Current Topics in Microbiology and Immunology

Stanton B. Gelvin *Editor*

Agrobacterium Biology

From Basic Science to Biotechnology

Springer

Current Topics in Microbiology and Immunology

Volume 418

More information about this series at http://www.springer.com/series/82

Stanton B. Gelvin
Editor

Agrobacterium Biology

From Basic Science to Biotechnology

Responsible series editor: Klaus Palme

Springer

Editor
Stanton B. Gelvin
Department of Biological Sciences
Purdue University
West Lafayette, IN, USA

ISSN 0070-217X ISSN 2196-9965 (electronic)
Current Topics in Microbiology and Immunology
ISBN 978-3-030-03256-2 ISBN 978-3-030-03257-9 (eBook)
https://doi.org/10.1007/978-3-030-03257-9

Library of Congress Control Number: 2018960380

This Springer imprint is published by the registered company Springer Nature Switzerland AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Preface

Agrobacterium may best be known as a genus of bacteria used for transferring genes to plants. Indeed, the use of *Agrobacterium* for plant genetic transformation forms the basis of the modern agricultural biotechnology industry, as well as a commonly used platform for plant basic research. However, the importance of *Agrobacterium* extends far beyond its use as a "gene jockeying" tool. *Agrobacterium* can be a serious pathogen of vineyards, orchards for apples, stonefruits, and nut trees, and nurseries, especially when cultivation requires grafting, pruning, or in vitro propagation. The appearance of crown galls, a tumorous disease caused by many strains of *Agrobacterium*, can seriously decrease yields or profits because of regulatory restrictions on the movement of infected plant material.

Agrobacterium is also a model for plant molecular, cellular, and evolutionary biology, as well as pathogenesis and bacterial ecology. The *Agrobacterium* Type IV Secretion System (T4SS) serves as the model for T4SSs of many human and animal pathogenic bacteria, and components of those cognate T4SSs often use the *Agrobacterium virB* nomenclature. The *Agrobacterium* Type VI Secretion System (T6SS) was among the first described, and serves as a model for inter-bacterial warfare. T-DNA transfer from *Agrobacterium* (and its close relatives) to plants and fungi represents the only example of extant horizontal gene transfer between prokaryotes and eukaryotes, and the presence of T-DNA in the genomes of several plant species suggests that ancient horizontal gene transfer events may have influenced plant growth and/or developmental processes. This volume describes in detail these and other aspects of *Agrobacterium* biology.

The study of *Agrobacterium* taxonomy and ecology has undergone a revitalization. As discussed in the chapter by Gan and Savka, whole genome sequencing of many members of the Rhizobeaceae has indicated that many alphaproteobacteria formerly designated *Agrobacterium* are more closely related to other genus groups, including *Rhizobium*. The close relationship among members of the *Agrobacterium* and *Ensifer*/*Sinorhizobium* groups may be reflected by the ability of many of the members, under the appropriate conditions, to transfer DNA to plants.

Agrobacterium species can live in many different environments, including the soil, the rhizosphere surrounding plant roots, and crown gall tumors. Two chapters, by Kuzmanovic et al., and Dessaux and Faure, describe how *Agrobacterium* can utilize various compounds, including opines synthesized by tumors, to create and enlarge their ecological niche.

A chapter by Figueroa-Cuilan and Brown describes the unusual mode of polar cellular growth *Agrobacterium* uses prior to cell division, and some of the proteins involved in the growth and division processes. Two chapters, by Matthysse and by Thompson et al., describe *Agrobacterium* motility and bacterial surface properties, including the synthesis and composition of polysaccharides that may mediate attachment of *Agrobacterium* to synthetic and plant surfaces. Within the bacterium, Biran et al. describe how *Agrobacterium* responds to heat stress, whereas Lee and Wang discuss *Agrobacterium* noncoding RNAs and how they may regulate bacterial gene expression, including the possible regulation of virulence gene induction preceding transformation.

Many *Agrobacterium* strains encode two important protein secretion systems: The Type VI Secretion System (T6SS) that is used for inter-bacterial "warfare", and the type IV secretion system that is used to conjugate plasmids between bacteria or to transfer T-DNA from *Agrobacterium* to host eukaryotic cells. Chapters by Wu et al., and Li and Christie present an overview of these two important systems used by *Agrobacterium* to interact with other bacteria in a niche or to genetically transform plants, respectively.

Agrobacterium is best known as an organism that can genetically transform plants. Although transformation-associated processes occurring in the bacterium are reasonably well understood, relatively little is known about events that occur in the plant host. Three chapters discuss recent work to understand these plant-associated events. Tu et al. describe some of the latest results regarding the trafficking of virulence proteins (and perhaps T-DNA) within the plant cytoplasm as they exit the bacterium and target the nucleus. Singer discusses models of T-DNA integration, and Willig et al. describe the plant transcriptional response to *Agrobacterium* infection. Hooykaas et al. compare and contrast *Agrobacterium*-mediated transformation of plants with that of yeast and fungi.

Although Genetically Modified Organisms (GMOs) remain controversial, some plants are naturally transgenic. Diverse members of the *Nicotiana* family, as well as *Ipomea* (sweet potato) and *Lineria*, contain T-DNA from what must have been an ancient *Agrobacterium* infection. Two chapters by Otten and Matveeva describe this phenomenon, and speculate on how such ancient transformation events may have affected plant evolution and development.

Agrobacterium-mediated transformation is the most commonly used platform for the generation of genetically modified plants, but other bacteria can also transform plants when supplied with the appropriate components, including virulence genes and T-DNA. Lacroix and Citovsky review what is known about interkingdom horizontal gene transfer by bacterial species other than *Agrobacterium* and speculate on how these other species can be used for biotechnology purposes.

Transgenic plants have been generated using *Agrobacterium*-mediated transformation since 1983. With continuing research, *Agrobacterium* has been harnessed for biotechnological purposes. Two chapters, by Sardesai and Subramanyam and by Anand and Jones, describe some of the history associated with the development of this gene transfer platform. These authors describe the increasing sophistication by which scientists have used *Agrobacterium* for generating transgenic plants, and how *Agrobacterium* is currently being used to deliver reagents for intricate and precise plant genome modification.

I am grateful to all the authors for the considerable work they put into writing these chapters, and to all my *Agrobacterium* colleagues and mentors who have, in the space of only about 40 years, brought the understanding of *Agrobacterium* biology so far.

West Lafayette, USA Stanton B. Gelvin

Contents

One More Decade of *Agrobacterium* Taxonomy 1
Han Ming Gan and Michael A. Savka

The Ecology of *Agrobacterium vitis* and Management of Crown Gall Disease in Vineyards 15
Nemanja Kuzmanović, Joanna Puławska, Lingyun Hao and Thomas J. Burr

Niche Construction and Exploitation by *Agrobacterium*: How to Survive and Face Competition in Soil and Plant Habitats 55
Y. Dessaux and D. Faure

Cell Wall Biogenesis During Elongation and Division in the Plant Pathogen *Agrobacterium tumefaciens* 87
Wanda M. Figueroa-Cuilan and Pamela J. B. Brown

Exopolysaccharides of *Agrobacterium tumefaciens* 111
Ann G. Matthysse

Function and Regulation of *Agrobacterium tumefaciens* Cell Surface Structures that Promote Attachment 143
Melene A. Thompson, Maureen C. Onyeziri and Clay Fuqua

Coping with High Temperature: A Unique Regulation in *A. tumefaciens* 185
Dvora Biran, Or Rotem, Ran Rosen and Eliora Z. Ron

Small Noncoding RNAs in *Agrobacterium tumefaciens* 195
Keunsub Lee and Kan Wang

The *Agrobacterium* Type VI Secretion System: A Contractile Nanomachine for Interbacterial Competition 215
Chih-Feng Wu, Delaney A. Smith, Erh-Min Lai and Jeff H. Chang

The *Agrobacterium* VirB/VirD4 T4SS: Mechanism and Architecture Defined Through In Vivo Mutagenesis and Chimeric Systems 233
Yang Grace Li and Peter J. Christie

Real-Time Trafficking of *Agrobacterium* Virulence Protein VirE2 Inside Host Cells 261
Haitao Tu, Xiaoyang Li, Qinghua Yang, Ling Peng and Shen Q. Pan

The Mechanism of T-DNA Integration: Some Major Unresolved Questions 287
Kamy Singer

Transcriptome Profiling of Plant Genes in Response to *Agrobacterium tumefaciens*-Mediated Transformation 319
Christopher J. Willig, Kaixuan Duan and Zhanyuan J. Zhang

***Agrobacterium*-Mediated Transformation of Yeast and Fungi** 349
Paul J. J. Hooykaas, G. Paul H. van Heusden, Xiaolei Niu, M. Reza Roushan, Jalal Soltani, Xiaorong Zhang and Bert J. van der Zaal

The *Agrobacterium* Phenotypic Plasticity (*Plast*) Genes 375
Léon Otten

***Agrobacterium*-Mediated Transformation in the Evolution of Plants** 421
Tatiana V. Matveeva

Beyond *Agrobacterium*-Mediated Transformation: Horizontal Gene Transfer from Bacteria to Eukaryotes 443
Benoît Lacroix and Vitaly Citovsky

***Agrobacterium*: A Genome-Editing Tool-Delivery System** 463
Nagesh Sardesai and Subhashree Subramanyam

Advancing *Agrobacterium*-Based Crop Transformation and Genome Modification Technology for Agricultural Biotechnology 489
Ajith Anand and Todd J. Jones

Correction to: *Agrobacterium*: A Genome-Editing Tool-Delivery System E1
Nagesh Sardesai and Subhashree Subramanyam

One More Decade of *Agrobacterium* Taxonomy

Han Ming Gan and Michael A. Savka

Contents

1 Introduction 2
2 Pre-2006 *Agrobacterium* Taxonomy 3
3 Alternative Views of the *Agrobacterium* Phylogeny 4
3.1 The *recA* Gene as an Alternative Genetic Marker to 16S rRNA for Inferring *Agrobacterium* Phylogeny 4
3.2 Four (or Six) Is Better Than One: Refining and Revising the *Agrobacterium* Genus Through Multilocus Sequence Analysis (MLSA) 4
4 *Agrobacterium* and the Genomic Era 5
4.1 Pre-next-Generation Sequencing Period 5
4.2 Next-Generation Sequencing and *Agrobacterium* 6
4.3 Updating the *Agrobacterium* Phylogeny in the Light of More Publicly Available Genomic Resources 7
5 Genomic Species Within *Agrobacterium* 9
6 Concluding Remarks 11
References 11

Abstract This chapter presents a historical overview of the development and changes in scientific approaches to classifying members of the *Agrobacterium* genus. We also describe the changes in the inference of evolutionary relationships among *Agrobacterium* biovars and *Agrobacterium* strains from using the 16S rRNA

H. M. Gan (✉)
Centre for Integrative Ecology, School of Life and Environmental Sciences, Deakin University, Geelong, VIC, Australia
e-mail: han.gan@deakin.edu.au

M. A. Savka (✉)
Thomas H. Gosnell School of Life Sciences, Rochester Institute of Technology, Rochester, NY, USA
e-mail: massbi@rit.edu

Current Topics in Microbiology and Immunology (2018) 418:1–14
DOI 10.1007/82_2018_81

Published Online: 3 March 2018

marker to *recA* genes and to the use of multilocus sequence analysis (MLSA). Further, the impacts of the genomic era enabling low cost and rapid whole genome sequencing on *Agrobacterium* phylogeny are reviewed with a focus on the use of new and sophisticated bioinformatics approaches to refine phylogenetic inferences. An updated genome-based phylogeny of ninety-seven *Agrobacterium tumefaciens* complex isolates representing ten known genomic species is presented, providing additional support to the monophyly of the *Agrobacterium* clade. Additional taxon sampling within *Agrobacterium* genomovar G3 indicates potential exceptions to interpretation of the concept of bacterial genomics species as ecological species because the genomovar G3 genomic cluster, which initially includes clinical strains, now also includes plant-associated and cave isolates.

1 Introduction

Since the first uses of DNA sequences to classify relationship among bacterial strains became routine (Janda and Abbott 2007; Stackebrandt and Goebel 1994), new and increasing amounts of single-copy protein-coding DNA markers have been employed to revaluate and revise the taxonomy of *Agrobacterium*. Phylogenetic analyses based on increased taxon and gene sampling have led to the reclassification of the traditional *Agrobacterium* biovars 1 and 3 to two new genera (Costechareyre et al. 2010; Mousavi et al. 2014). With the now common practice of sequencing whole bacterial genomes, large data sets are increasingly available, and these sequences have become linked to more sophisticated approaches to analyse data using multiple and linear bioinformatical approaches. These approaches have provided new and improved insight into the evolutionary relationships among *Agrobacterium* species. In this review chapter, we first provide a historical overview of the molecular systematics of the genus *Agrobacterium* which led to an intense debate among the scientific community during the 16S rRNA era. We next review changes to the *Agrobacterium* taxonomy which is gradually embraced by the scientific community in the light of more recent and refined phylogenetic analyses using improved gene and taxon sampling. The unprecedented genetic information about *Agrobacterium* derived from the advent of next-generation sequencing and its impacts on the inference and delineation of *Agrobacterium* at the strain level is summarized. We also provide a genome-based phylogeny of ninety-seven *Agrobacterium tumefaciens* complex isolates, representing a significant increase in taxon sampling compared to a previous phylogenomic study (Ormeno-Orrillo et al. 2015). The validity of bacterial genome species being ecological species (Lassalle et al. 2011) is briefly assessed and discussed in the light of new phylogenomic inferences and observed ecological niche diversity among recently sequenced strains belonging to *Agrobacterium* genomovar 3.

2 Pre-2006 *Agrobacterium* Taxonomy

The use of 16S rRNA sequence as a genetic marker for microbial taxonomy brought about both chaos and order within the taxonomy of *Agrobacterium*. The availability of universal 16S rRNA primers and the inherent high copy number of 16S rRNA in most bacterial genomes are two of the main attributes promoting the inclusion of the 16S rRNA sequence as part of the developed polyphasic taxonomy for bacteria (Janda and Abbott 2007; Woo et al. 2008). Furthermore, the high sequence conservation of the 16S rRNA gene makes it a very powerful genetic marker when inferring deep relationships. However, at the species or genus level, the use of the 16 s rRNA gene to discriminate among species tends to be modest if not inferior to other universal genetic markers (Kisand and Wikner 2003; Stackebrandt and Goebel 1994).

It is important to note that 16S rRNA gene substitution rates appear to vary among different groups of bacteria (Ochman et al. 1999; Smit et al. 2007). In other words, if the 16S rRNA gene substitution rates are lower in the family Rhizobiaceae, this will translate into low 16S rRNA gene nucleotide divergence and/or phylogenetic signals among members of the Rhizobiaceae. This may negatively affect phylogenetic interpretation, raising doubts about the veracity of their inferred evolutionary relationships. An initial proposal by Young et al. (2001) to incorporate all species of *Agrobacterium* and *Allorhizobium* into the genus *Rhizobium* due to the lack of concordance between DNA hybridization, biochemical traits, and fatty acid profiles among members of the described genera sparked an intense response from the scientific community (Farrand et al. 2003; Young et al. 2001). Farrand et al. (2003) claimed that members of the genus *Agrobacterium* and *Rhizobium* can be distinguished based on chromosomal structure and phenotype (as an individual species but not genera). Young et al. (2001) replied to Farrand et al. (2003) defending the initial proposal in addition to highlighting that the proposal is in accordance with the rules/codes set out by the International Code of Nomenclature of Bacteria. Young et al. (2001) further cautioned that bending the codes to retain the genus *Agrobacterium* may trigger a potential return to unregulated and chaotic bacterial nomenclature. The initial classification of *Agrobacterium* species based on their pathogenicity has been problematic, as it is now well established that the virulence factors are usually encoded on plasmids and some of these can even be lost relatively easily through growth at elevated temperature (Genetello et al. 1977). For further reading on the change and development in *Agrobacterium* taxonomy until 2006, we direct reader to a comprehensive review by Young (2008).

3 Alternative Views of the *Agrobacterium* Phylogeny

3.1 *The* recA *Gene as an Alternative Genetic Marker to 16S rRNA for Inferring* Agrobacterium *Phylogeny*

The *recA* gene encodes a multifunctional and important enzyme involved in homologous recombination and DNA repair (Kowalczykowski et al. 1994). A *recA* mutant is therefore characterized by its high sensitivity to UV light in addition to being recombination-deficient, a desirable trait for genetic studies involving trans-complementation of mutations located on a chromosome or plasmid (Kanie et al. 2007; Kuzminov and Stahl 1997). The importance of a *recA* mutant is well recognized among *Agrobacterium* geneticists, leading to the construction of strains LBA4301 and UIA143, *recA* mutants of *Agrobacterium tumefaciens* Ach5, and *Agrobacterium tumefaciens* C58, respectively (Farrand et al. 1989). Beyond molecular genetics, the *recA* gene is also well known in molecular systematics (Lloyd and Sharp 1993) and has been incorporated as one of the main genes for multilocus typing (MLSA) (Bennasar et al. 2010; Delamuta et al. 2012; Huo et al. 2017; Martens et al. 2008; Menna et al. 2009; Sakamoto and Ohkuma 2011). Phylogenetic analysis based on the *recA* gene of 138 strains from 13 genomic species of *Agrobacterium* lends support to the use of this marker gene for speciation of the genus *Agrobacterium* (Costechareyre et al. 2010). Genomic species is a concept of bacterial species based on similarities among bacterial chromosomal DNAs as determined by DNA–DNA hybridization or alternatively by in silico calculation of pair-wise average nucleotide identity (ANI) using whole genome sequences (Konstantinidis et al. 2006; Stackebrandt and Goebel 1994). A genomic species is defined as a group of bacterial strains with DNA–DNA reassociation values of more than 70%, which corresponds closely to ~95% ANI (Konstantinidis et al. 2006). A *recA*-based phylogenetic analysis indicates that *Agrobacterium* biovar 2, typically represented by *Agrobacterium rhizogenes*, and biovar 3 represented by *Agrobacterium vitis* are distantly related to *Agrobacterium* biovar 1. In addition, inclusion of *recA* sequences from several *Rhizobium* type strains in the analysis showed a stronger affiliation of *Agrobacterium rhizogenes* and *Agrobacterium vitis* to the *Rhizobium* clade.

3.2 *Four (or Six) Is Better Than One: Refining and Revising the* Agrobacterium *Genus Through Multilocus Sequence Analysis (MLSA)*

Phylogenetic tree construction based on six protein-coding housekeeping genes consisting of ATP synthase F1, beta subunit (*atpD*), glutamine synthetase type I (*glnA*), glutamine synthetase type II (*glnII*), recombinase A (*recA*), RNA polymerase beta subunit (*rpoB*), and threonine synthase (*thrC*) from 114 rhizo- and

agrobacteria reinforced the monophyly of the genus *Agrobacterium* which was previously reestablished based on the *recA* gene. In addition to resolving other pending taxonomic issues related to the family Rhizobiaceae, the substantial increase in gene and taxon sampling also lent support to the reclassification of *Agrobacterium vitis* to an existing genus *Allorhizobium* (Mousavi et al. 2014). Once belonging to three different biovars of the same genus, the phythopathogenic *Agrobacterium tumefaciens*, *Agrobacterium vitis* (now *Allorhizobium vitis*), and *Agrobacterium rhizogenes* (now *Rhizobium rhizogenes*) now belong to three different genera. Furthermore, with the creation of the genus *Neorhizobium* which is a sister group to *Agrobacterium*, *Agrobacterium* can now remain a suitable genus name for a monophyletic clade within the Rhizobiaceae family. A follow-up study based on three housekeeping genes and the 16S rRNA gene again supported the monophyly of the revised *Agrobacterium* clade in addition to expanding the membership of the genus *Allorhizobium* to include *R. taibanshenase*, *R. paknamense*, *R. oryzae*, *R.psuedoryzae*, *R. qilianshanense*, and *R. borbori*. However, in contrast to a previous study based on six housekeeping genes, a sister grouping of *Agrobacterium–Neorhizobium* was not observed. The *Agrobacterium* clade instead shared a sister grouping with the *R. aggregatum* complex (Mousavi et al. 2015). Mousavi et al., however, did not suggest the reclassification of members from the *R. aggregatum* complex to the genus *Agrobacterium* as members of this sister clade, citing the lack of *Agrobacterium*-specific genome architecture (linear chromosome and the presence of the protelomerase-coding gene, *telA* (Ramirez-Bahena et al. 2014).

4 *Agrobacterium* and the Genomic Era

4.1 Pre-next-Generation Sequencing Period

Whole genome sequencing provides an unprecedented view into the evolutionary relationships of microorganisms. With a repertoire of single-copy and near-universal genes, usually in the range of hundreds, that can be used for phylogenetic inference, there is no longer a limitation to gene sampling, one of the main requirements for accurate phylogenetic analysis (Hedges 2002; Rosenberg and Kumar 2003). *Agrobacterium tumefaciens* C58 (now *Agrobacterium fabrum* C58) is the first *Agrobacterium* strain to have its complete genome sequenced by two separate research groups using conventional Sanger sequencing (Goodner et al. 2001; Wood et al. 2001) and subsequently revised with improved annotation (Slater et al. 2013). Approximately nine years later, the complete genome for members from the remaining two biovars, e.g. *Agrobacterium vitis* (biovar 3, now *Allorhizobium vitis*) and *Agrobacterium radiobacter* (biovar 2, now *Rhizobium* sp.; Slater et al. 2009), was reported. In addition, for the first time a high-resolution phylogeny of *Agrobacterium* was constructed based on the concatenated protein

alignment of 507 single-copy orthologous gene families encoded on the primary chromosomes. Phylogenetic clustering patterns indicated that biovar 2 should be grouped to the genus *Rhizobium*, whereas biovar 3 and biovar 1 are still members of the *Agrobacterium* genus. The limited taxon sampling resulting from the high cost of whole genome sequencing at the time unfortunately prevented Slater et al. (2009) from inferring the delineation of biovar 3 and biovar 1 into two separate genera.

4.2 Next-Generation Sequencing and Agrobacterium

The advent of next-generation sequencing brought about a revolution in microbial genomics by enabling the whole genome sequence of a pure culture to be obtained at a small fraction of the cost and time initially required by Sanger sequencing (MacLean et al. 2009; Metzker 2010). Coupled with advances in algorithms for quick and accurate microbial genome assembly and annotation (Bankevich et al. 2012; Seemann 2014), the scientific community is now blessed with an explosion of publicly available microbial genomic resources which naturally invite a new investigation of the phylogeny of *Agrobacterium*. Ormeno-Orrillo and workers used a sophisticated and reproducible bioinformatics pipeline (Segata et al. 2013) to reconstruct the *Agrobacterium* phylogeny based on the concatenated alignment of 384 universal proteins identified from 113 sequenced strains from the family Rhizobiaceae (Ormeno-Orrillo et al. 2015). In contrast to the previously inferred whole genome phylogeny, *Agrobacterium vitis* S4 no longer formed a tight cluster with *Agrobacterium tumefaciens* C58. Instead, the increased taxon sampling supported previous *recA* and MLSA-based analyses indicating the monophyletic clustering of *Agrobacterium vitis* S4 with members of the genus *Allorhizobium* such as *Allorhizobium undicola* (de Lajudie et al. 1998), lending further support to the revival of *Allorhizobium* as a genus within the Rhizobiaceae (Mousavi et al. 2014). By reclassifying *Agrobacterium* biovars 2 and 3 into separate genera (Mousavi et al. 2014, 2015; Velázquez et al. 2010), a monophyletic cluster consisting solely of members from the genus *Agrobacterium* can be obtained with maximal support, indicating that at the genomic level, *Agrobacterium* is a definable genus of the family Rhizobiaceae (Ormeno-Orrillo et al. 2015). The author noted, however, the exclusion of an important Agrobacterium genome, e.g. *Agrobacterium radiobacter* NCPPB 3001 = DSM30147^T (accession number ASXY01, Bioproject PRJNA212112; Zhang et al. 2014) from their analysis, citing unusual genomic anomalies such as low sequence homology (<97%) to some of its published gene sequences. Leveraging the recent availability of key *Agrobacterium* species genomes, Kim and Gan (2017) performed a smaller scale phylogenomic analysis of the genus *Agrobacterium* showing the monophyletic clustering of *A. tumefaciens* B6 and *A. radiobacter* NCPPB 3001^T = DSM30147^T with high pair-wise ANI value (>95%), providing conclusive genomic evidence that both strains are identical species (Kim and Gan 2017).

4.3 *Updating the* Agrobacterium *Phylogeny in the Light of More Publicly Available Genomic Resources*

In this chapter, we present an updated phylogeny of *Agrobacterium* and more generally the Rhizobiaceae using a similar PhyloPhlAn approach implemented by Ormeno-Orrillo et al. (2015). PhyloPhlAn is a bioinformatics pipeline which takes the predicted proteomes from multiple microbial strains in fasta format as input and uses an ultra-fast protein similarity search (Edgar 2010) to identify more than 400 single-copy and conserved proteins within each predicted proteome. The identified proteins are aligned individually using MUSCLE (Edgar 2004), concatenated, and used for maximum likelihood tree reconstruction with FastTree2 (Price et al. 2010). Consistent with previous reports, a cluster consisting of mainly *Agrobacterium* strains could be recovered with maximal support, with *Agrobacterium rubi* and *Agrobacterium larrymoorei* being basal to the rest of *Agrobacterium* (Figs. 1 and 2). The presence of a substantial number of *Rhizobium* strains in the *Agrobacterium* clade (Fig. 2) is an aftermath of Young et al.'s initial proposal (2001) for merging *Agrobacterium* with the genus *Rhizobium*. In addition, the phylogenetic placement of *A. radiobacter* DSM 30147^T basal to the rest of agrobacteria genomovar 4, which now includes a more recent and improved genome of *A. radiobacter* DSM 30147^T

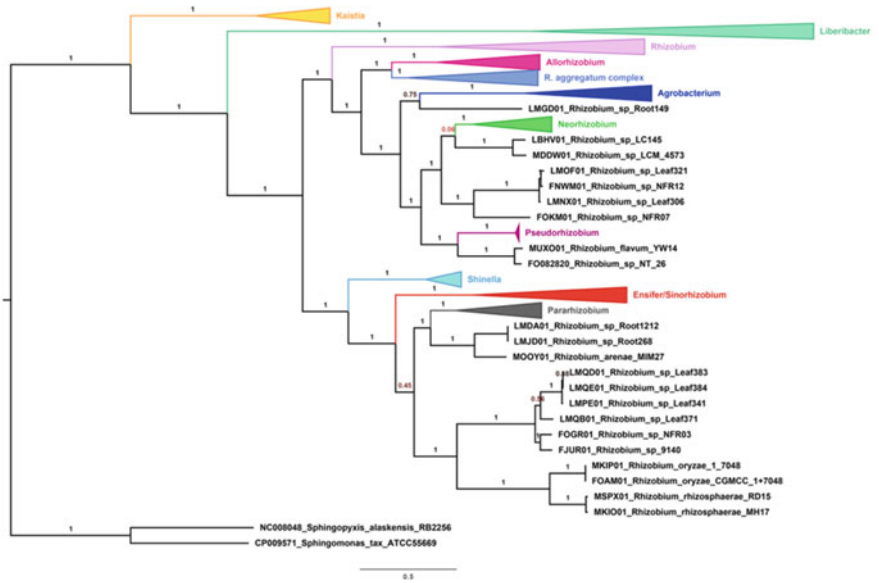
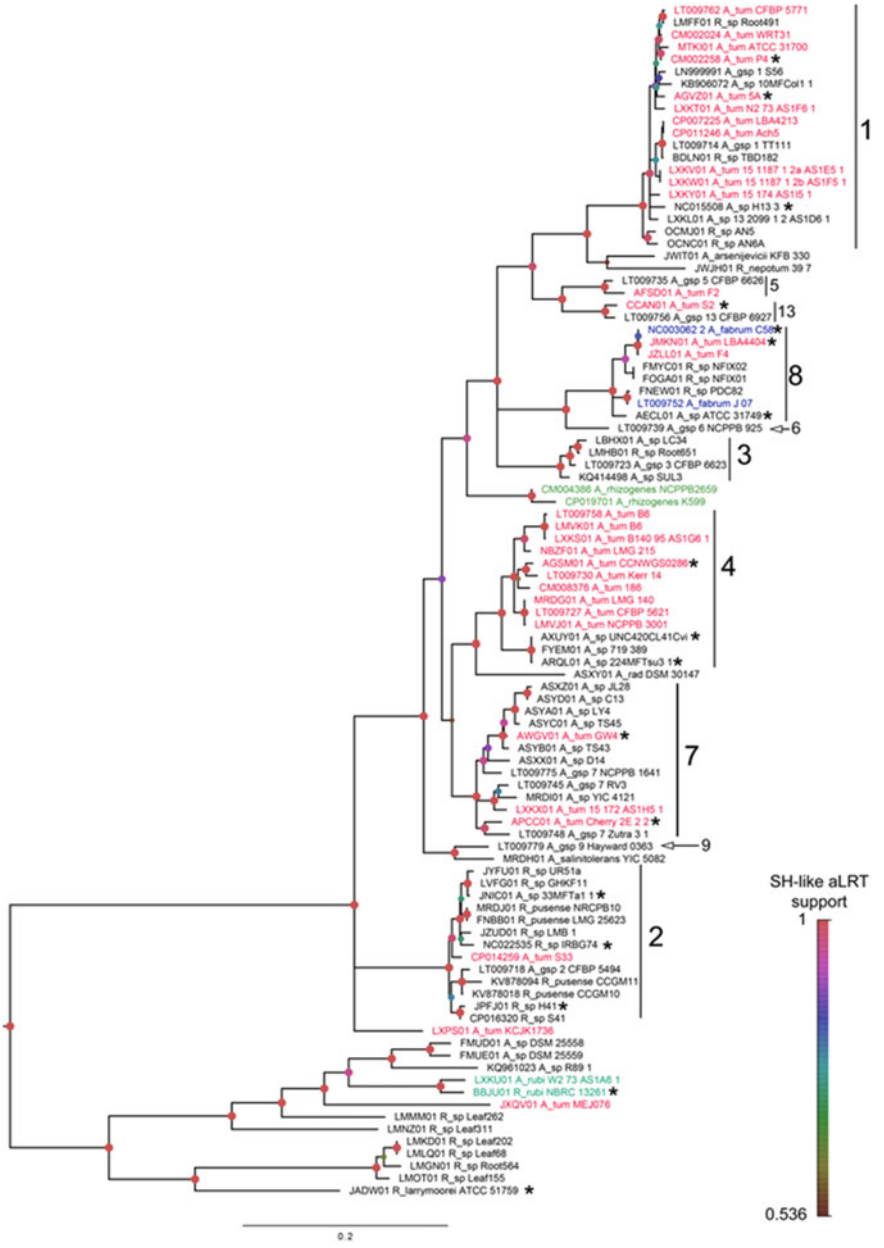


Fig. 1 Reconstruction of the Rhizobiaceae phylogeny using maximum likelihood inference based on the concatenated amino acid alignment of universal single-copy genes as implemented in the PhyloPhlAn pipeline (Segata et al. 2013). Members of the family Sphingomondaceae were rooted as the outgroup. Values along branch indicate SH-like aLRT support values (Shimodaira and Hasegawa 1999) calculated using FastTree2 (Price et al. 2010)



(=NCPPB3001^{T}; WGS Accession: LMVJ01; Lee et al., unpublished), is unusual, suggesting a genome assembly anomaly as previously noted (Ormeno-Orrillo et al. 2015). Another notable anomaly revealed by increased taxon sampling is the unexpected clustering of strain LBA4404, a disarmed derivative of the wild-type

◂**Fig. 2** Expanded *Agrobacterium* clade from Fig. 1 depicting the evolutionary relationships among *Agrobacterium* strains. First text strings are the WGS accession numbers, and the first letters after the strings represent the submitted genus name (R = *Rhizobium*; A = *Agrobacterium*). Taxon name is as per species name deposited into the NCBI whole genome shotgun database. Taxa coloured green: *Agrobacterium rubi*; taxa coloured red: *Agrobacterium tumefaciens*; taxa coloured blue: *Agrobacterium fabrum*. Nodes were coloured according to their SH-like local support values, and genomic species clusters were indicated by the vertical lines or arrows next to the tree. Asterisk signs indicate taxa that were included in a previous large-scale phylogenomic analysis by Ormeno-Orrillo et al. (2015). The tree was constructed using a whole genome-based (400 universal single-copy genes) approach

Ach5 Tn*904* mutant (strain LBA4213), with members from the genovar 8 containing *Agrobacterium fabrum* C58 (Ooms et al. 1982). Recently, both strains Ach5 and LBA4213 have been sequenced by two independent groups (Henkel et al. 2014; Huang et al. 2015) and in contrast to strain LBA4404, both strains resided in the genomovar 1 clade, forming a monophyletic group. Given the known divergence between strain Ach5 and strain C58, this strongly indicates that the currently deposited whole genome sequence of strain LBA4404 is incorrect and warrants future investigation. The abnormal phylogenetic placement of strain LBA4404 was similarly observed but not explicitly mentioned in a study by Ormeno-Orrillo et al. (2015). Clustering based on genospecies is apparent; albeit the relationships among some of the genospecies are not strongly supported, suggesting the limitation of amino acid-based phylogenomic analysis for fully resolving strain, subspecies, and/or species-level relationships similarly observed in a recent genome-based phylogeny of *Pseudomonas* (Tran et al. 2017). To infer accurately the phylogeny of the currently well-supported *Agrobacterium* clade, future work utilizing the newly published phylogenetic-aware pan-genome analysis tool (Ding et al. 2017) to improve the recovery of core *Agrobacterium* single-copy genes, coupled with complementary analysis based on pair-wise average nucleotide identity (ANI) (Richter et al. 2016), will be instructive.

5 Genomic Species Within *Agrobacterium*

Traditionally, genome–genome hybridization has been used to establish genomic relatedness among strains, and a hybridization ratio of approximately 70% between two strains usually indicates a species-level relationship (Wayne et al. 1987; Stackebrandt et al. 2002). Average nucleotide calculation (ANI) is becoming increasingly popular for in silico species delineation in the light of genomic data availability. An initial genomic comparison indicated 95% pair-wise ANI as correlated with 70% DNA–DNA hybridization (DDH), and this correlation was consistently observed in various subsequent studies (Auch et al. 2010; Colston et al. 2014). Using the established 70% DDH criterion in addition to a follow-up validation based on mathematical models and amplified fragment length polymorphism (AFLP) data, members within the *Agrobacterium tumefaciens* complex were

classified into ten distinct genomic species with a non-continuous genomovar numbering, e.g. G1–G9 followed by G13, as a consequence of the reclassification of some initially established genomovars to a different genus, e.g. *Agrobacterium rhizogenes* (genomovar 10) to *Rhizobium rhizogenes* or to a greater extent *Agrobacterium* clade, e.g. *Agrobacterium rubi* (genomovar 11; clade 2 in Fig. 2). To date, most of the genomovars have not received official Latin binomials due to the lack of differentiating biochemical features that are traditionally used to describe new bacterial species. Lassalle et al. (2011) took one of the first initiatives to differentiate the *Agrobacterium tumefaciens* species complex by identifying the gene repertoire specific to *Agrobacterium* genospecies 8 which includes strain C58, a widely used strain among *Agrobacterium* geneticists that has had its genome sequenced and annotated. By comparing the C58 genome against 25 strains from different *Agrobacterium* genospecies based on hybridization to DNA microarrays spanning the whole genome of strain C58, genes relevant to the speciation and ecological isolation of genomovar G8 were identified. Phenotypic traits specific to genomovar G8 initially inferred from microarray data, such as ferulic acid degradation and curdlan production, were subsequently validated using HPLC and Congo red assays, respectively. As a result, the species name *Agrobacterium fabrum* was suggested for strains of *Agrobacterium* genomovar G8, from the Latin plural genitive of *smith*, in reference to the pioneer isolator of an *Agrobacterium* strain (Smith and Townsend 1907).

Based on identification of a gene repertoire unique to genomovar G8 that is associated with commensal interactions with plants, and by citing several similar studies linking ecological niche and genomic species beyond the genus *Agrobacterium* (Cai et al. 2009; Johnson et al. 2006; Lefébure et al. 2010; Porwollik et al. 2002), Lassalle et al. (2011) suggested the generalization of the concept of bacterial genomic species as ecological species. A potential exception to this generalization is currently emerging within *Agrobacterium* genomovar G3. The *Agrobacterium* genomovar G3 initially consisted of strains isolated from clinical environments, e.g. human host and antiseptic flask (Popoff et al. 1984). However, based on the newly constructed phylogenomic tree, in addition to the classical *Agrobacterium* sp. CFBP 6623, the agrobacteria G3 clade now consists of strains LC34, SUL3, and Root651 which were isolated from a diverse and non-clinical environment. Notably, *Agrobacterium* sp. LC34 originated from the rock surface of the Lechuguilla Cave which has been isolated from humans for over four million years (Bhullar et al. 2012), an environment that substantially differs from that of strain CFBP 6623. On the contrary, *Agrobacterium* sp. Root651 may share a similar ecological niche with that of G8 agrobacteria given that it is a member of the *Arabidopsis* plant root microbiota (Bai et al. 2015). *Agrobacterium* sp. SUL3 was isolated from a laboratory culture of the hydrocarbon-producing *Botryococcus braunii*, a non-plant photosynthetic organism (green microalga; (Jones et al. 2016). Taken together, it will be hard to convince microbial ecologists that members of the *Agrobacterium* genomovar G3 are a single ecological species despite their high genomic relatedness.

6 Concluding Remarks

The progress of using whole genome sequence data for establishing relatedness among members of the Rhizobiaceae family is presented. As additional whole genome sequences of these members are elucidated, further insight into the complex phylogeny of *Agrobacterium* will become available. Further and rigorous analysis of large data sets will validate or further contest the concept of bacterial genomic species as ecological species.

Acknowledgements M.A.S. and H.M.G acknowledge funding for research provided by the Gosnell School of Life Science at Rochester Institute of Technology. H.M.G was partially supported by a research start-up grant from the Deakin University Centre of Integrative Ecology.

References

Auch AF, von Jan M, Klenk H-P, Göker M (2010) Digital DNA-DNA hybridization for microbial species delineation by means of genome-to-genome sequence comparison. Stand Genomic Sci 2:117–134

Bai Y et al (2015) Functional overlap of the *Arabidopsis* leaf and root microbiota. Nature 528:364–369

Bankevich A et al (2012) SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. J Comp Biol 19:455–477

Bennasar A, Mulet M, Lalucat J, García-Valdés E (2010) PseudoMLSA: a database for multigenic sequence analysis of *Pseudomonas* species. BMC Microbiol 10:118

Bhullar K et al (2012) Antibiotic resistance is prevalent in an isolated cave microbiome. PLoS ONE 7:e34953

Cai H, Thompson R, Budinich MF, Broadbent JR, Steele JL (2009) Genome sequence and comparative genome analysis of *Lactobacillus casei*: insights into their niche-associated evolution. Genome Bio Evol 1:239–257

Colston SM, Fullmer MS, Beka L, Lamy B, Gogarten JP, Graf J (2014) Bioinformatic genome comparisons for taxonomic and phylogenetic assignments using *Aeromonas* as a test Case. mBio 5

Costechareyre D et al (2010) Rapid and efficient identification of *Agrobacterium* species by *recA* allele analysis. Microb Ecol 60:862–872

de Lajudie P et al (1998) *Allorhizobium undicola* gen. nov., sp. nov., nitrogen-fixing bacteria that efficiently nodulate *Neptunia natans* in Senegal. Int J Syst Bacteriol 4:1277–1290

Delamuta JRM, Ribeiro RA, Menna P, Bangel EV, Hungria M (2012) Multilocus sequence analysis (MLSA) of *Bradyrhizobium* strains: revealing high diversity of tropical diazotrophic symbiotic bacteria. Braz J Microbiol 43:698–710

Ding W, Baumdicker F, Neher RA (2017) panX: pan-genome analysis and exploration. Nucleic Acids Res 25

Edgar RC (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res 32:1792–1797

Edgar RC (2010) Search and clustering orders of magnitude faster than BLAST. Bioinformatics 26:2460–2461

Farrand SK, O'Morchoe SP, McCutchan J (1989) Construction of an *Agrobacterium tumefaciens* C58 *recA* mutant. J Bacteriol 171:5314–5321

Farrand SK, van Berkum PB, Oger P (2003) *Agrobacterium* is a definable genus of the family Rhizobiaceae. Int J Syst Evol Microbiol 53:1681–1687

Genetello C, Van Larebeke N, Holsters M, De Picker A, Van Montagu M, Schell J (1977) Ti plasmids of *Agrobacterium* as conjugative plasmids. Nature 265:561–563

Goodner B et al (2001) Genome sequence of the plant pathogen and biotechnology agent *Agrobacterium tumefaciens* C58. Science 294:2323–2328

Hedges SB (2002) The origin and evolution of model organisms. Nat Rev Gen 3:838–849

Henkel CV, den Dulk-Ras A, Zhang X, Hooykaas PJJ (2014) Genome sequence of the octopine-type *Agrobacterium tumefaciens* strain Ach5. Genome Announcements 2:e00225–00214

Huang Y-Y, Cho S-T, Lo W-S, Wang Y-C, Lai E-M, Kuo C-H (2015) Complete genome sequence of *Agrobacterium tumefaciens* Ach5. Genome Announc 3:e00570–00515

Huo Y-B, Chan Y, Lacap-Bugler DC, Mo S, Woo PC, Leung WK, Watt RM (2017) Multilocus sequence analysis of phylogroup 1 and 2 oral treponeme strains. Appl Environ Microbiol 83: e02499–02416

Janda JM, Abbott SL (2007) 16S rRNA gene sequencing for bacterial identification in the diagnostic laboratory: pluses, perils, and pitfalls. J Clin Microbiol 45:2761–2764

Johnson ZI, Zinser ER, Coe A, McNulty NP, Woodward EMS, Chisholm SW (2006) Niche partitioning among *Prochlorococcus* ecotypes along ocean-scale environmental gradients. Science 311:1737–1740

Jones KJ, Moore K, Sambles C, Love J, Studholme DJ, Aves SJ (2016) Draft genome sequences of *Achromobacter piechaudii* GCS2, *Agrobacterium* sp. Strain SUL3, *Microbacterium* sp. Strain GCS4, *Shinella* sp. Strain GWS1, and *Shinella* sp. strain SUS2 isolated from consortium with the hydrocarbon-producing alga *Botryococcus braunii*. Genome Announc 4

Kanie S et al (2007) Roles of RecA protein in spontaneous mutagenesis in *Escherichia coli*. Genes Genet Syst 82:99–108

Kim K, Gan HM (2017) A glimpse into the genetic basis of symbiosis between *Hydrogenophaga* and their helper strains in the biodegradation of 4-aminobenzenesulfonate. J Genomics 5:77–82

Kisand V, Wikner J (2003) Limited resolution of 16S rDNA DGGE caused by melting properties and closely related DNA sequences. J Microbiol Meth 54:183–191

Konstantinidis KT, Ramette A, Tiedje JM (2006) The bacterial species definition in the genomic era. Philos Trans R Soc Lond B Biol Sci 361:1929–1940

Kowalczykowski SC, Dixon DA, Eggleston AK, Lauder SD, Rehrauer WM (1994) Biochemistry of homologous recombination in *Escherichia coli*. Microbiol Rev 58:401–465

Kuzminov A, Stahl FW (1997) Stability of linear DNA in *recA* mutant *Escherichia coli* cells reflects ongoing chromosomal DNA degradation. J Bacteriol 179:880–888

Lassalle F et al (2011) Genomic species are ecological species as revealed by comparative genomics in *Agrobacterium tumefaciens*. Genome Bio Evol 3:762–781

Lefébure T, Pavinski Bitar PD, Suzuki H, Stanhope MJ (2010) Evolutionary dynamics of complete *Campylobacter* pan-genomes and the bacterial species concept. Genome Bio Evol 2:646–655

Lloyd AT, Sharp PM (1993) Evolution of the *recA* gene and the molecular phylogeny of bacteria. J Mol Evol 37:399–407

MacLean D, Jones JD, Studholme DJ (2009) Application of next-generation sequencing technologies to microbial genetics. Nat Rev Gen 7:287–296

Martens M, Dawyndt P, Coopman R, Gillis M, De Vos P, Willems A (2008) Advantages of multilocus sequence analysis for taxonomic studies: a case study using 10 housekeeping genes in the genus *Ensifer* (including former *Sinorhizobium*). Int J Syst Evol Microbiol 58:200–214

Menna P, Barcellos FG, Hungria M (2009) Phylogeny and taxonomy of a diverse collection of *Bradyrhizobium* strains based on multilocus sequence analysis of the 16S rRNA gene, ITS region and *glnII*, *recA*, *atpD* and *dnaK* genes. Int J Syst Evol Microbiol 59:2934–2950

Metzker ML (2010) Sequencing technologies—the next generation. Nat Rev Gen 11:31–46

Mousavi SA et al (2014) Phylogeny of the *Rhizobium-Allorhizobium-Agrobacterium* clade supports the delineation of *Neorhizobium* gen. nov. Syst Appl Microbiol 37:208–215

Mousavi SA, Willems A, Nesme X, de Lajudie P, Lindstrom K (2015) Revised phylogeny of Rhizobiaceae: proposal of the delineation of *Pararhizobium* gen. nov., and 13 new species combinations. Syst Appl Microbiol 38:84–90

Ochman H, Elwyn S, Moran NA (1999) Calibrating bacterial evolution. PNAS 96:12638–12643

Ooms G, Hooykaas PJ, Van Veen RJ, Van Beelen P, Regensburg-Tuïnk TJ, Schilperoort RA (1982) Octopine Ti-plasmid deletion mutants of *Agrobacterium tumefaciens* with emphasis on the right side of the T-region. Plasmid 7:15–29

Ormeno-Orrillo E et al (2015) Taxonomy of rhizobia and agrobacteria from the Rhizobiaceae family in light of genomics. Syst Appl Microbiol 38:287–291

Popoff MY, Kersters K, Kiredjian M, Miras I, Coynault C (1984) Taxonomic position of *Agrobacterium* strains of hospital origin. Ann Microbiol 3:427–442

Porwollik S, Wong RM-Y, McClelland M (2002) Evolutionary genomics of *Salmonella*: gene acquisitions revealed by microarray analysis. PNAS 99:8956–8961

Price MN, Dehal PS, Arkin AP (2010) FastTree 2—approximately maximum-likelihood trees for large alignments. PLoS ONE 5:e9490

Ramirez-Bahena MH et al (2014) Single acquisition of protelomerase gave rise to speciation of a large and diverse clade within the *Agrobacterium/Rhizobium* supercluster characterized by the presence of a linear chromid. Mol Phylogenet Evol 73:202–207

Richter M, Rossello-Mora R, Oliver Glockner F, Peplies J (2016) JSpeciesWS: a web server for prokaryotic species circumscription based on pairwise genome comparison. Bioinformatics 32:929–931

Rosenberg MS, Kumar S (2003) Taxon sampling, bioinformatics, and phylogenomics. Syst Biol 52:119

Sakamoto M, Ohkuma M (2011) Identification and classification of the genus *Bacteroides* by multilocus sequence analysis. Microbiology 157:3388–3397

Seemann T (2014) Prokka: rapid prokaryotic genome annotation. Bioinformatics 30:2068–2069

Segata N, Bornigen D, Morgan XC, Huttenhower C (2013) PhyloPhlAn is a new method for improved phylogenetic and taxonomic placement of microbes. Nat Commun 4

Shimodaira H, Hasegawa M (1999) Multiple comparisons of log-likelihoods with applications to phylogenetic inference. Mol Biol Evol 16:1114–1114

Slater SC et al (2009) Genome sequences of three *Agrobacterium* biovars help elucidate the evolution of multichromosome genomes in bacteria. J Bacteriol 191:2501–2511

Slater S et al (2013) Reconciliation of sequence data and updated annotation of the genome of *Agrobacterium tumefaciens* C58, and distribution of a linear chromosome in the genus *Agrobacterium*. Appl Environ Microbiol 79:1414–1417

Smit S, Widmann J, Knight R (2007) Evolutionary rates vary among rRNA structural elements. Nucleic Acids Res 35:3339–3354

Smith EF, Townsend CO (1907) A plant-tumor of bacterial origin. Science 25:671–673

Stackebrandt E, Goebel B (1994) Taxonomic note: a place for DNA-DNA reassociation and 16S rRNA sequence analysis in the present species definition in bacteriology. Int J Syst Evol Microbiol 44:846–849

Stackebrandt E, Frederiksen W, Garrity GM, Grimont PAD, Kamper P, Maiden MCJ, Nesme X, Rossello-Mora R, Swings J et al (2002) Report of the ad hoc committee for the re-evaluation of the species definition in bacteriology. Int J Syst Evol Microbiol 52:1043–1047

Tran PN, Savka MA, Gan HM (2017) In-silico taxonomic classification of 373 genomes reveals species misidentification and new genospecies within the genus *Pseudomonas*. Front Microbiol 8:1296

Velázquez E et al (2010) Analysis of core genes supports the reclassification of strains *Agrobacterium radiobacter* K84 and *Agrobacterium tumefaciens* AKE10 into the species *Rhizobium rhizogenes*. Syst Appl Microbiol 33:247–251

Wayne LG, Brenner DJ, Colwell RR et al (1987) International committee on systematic bacteriology. Report of the ad hoc committee on reconciliation of approaches to bacterial systematics. Int J Syst Bacteriol 37:463–464

Woo P, Lau S, Teng J, Tse H, Yuen KY (2008) Then and now: use of 16S rDNA gene sequencing for bacterial identification and discovery of novel bacteria in clinical microbiology laboratories. Clin Microbiol Infect 14:908–934

Wood DW et al (2001) The genome of the natural genetic engineer *Agrobacterium tumefaciens* C58. Science 294:2317–2323

Young J, Kuykendall L, Martinez-Romero E, Kerr A, Sawada H (2001) A revision of *Rhizobium* Frank 1889, with an emended description of the genus, and the inclusion of all species of *Agrobacterium* Conn 1942 and *Allorhizobium undicola* de Lajudie et al. 1998 as new combinations: *Rhizobium radiobacter*, *R. rhizogenes*, *R. rubi*, *R. undicola* and *R. vitis*. Int J Syst Evol Microbiol 51:89–103

Young JM (2008) *Agrobacterium*—Taxonomy of plant-pathogenic *Rhizobium* species. In: Tzfira T, Citovsky V (eds) *Agrobacterium*: from biology to biotechnology. Springer, New York, pp 183–220. https://doi.org/10.1007/978-0-387-72290-0_5

Zhang L, Li X, Zhang F, Wang G (2014) Genomic analysis of *Agrobacterium radiobacter* DSM 30147(T) and emended description of *A. radiobacter* (Beijerinck and van Delden 1902) Conn 1942 (Approved Lists 1980) emend. Sawada et al. 1993. Stand Genomic Sci 9:574–584

The Ecology of *Agrobacterium vitis* and Management of Crown Gall Disease in Vineyards

Nemanja Kuzmanović, Joanna Puławska, Lingyun Hao and Thomas J. Burr

Contents

1 Introduction 16
2 Taxonomy of *A. vitis* 17
3 Geographic Distribution of *A. vitis* 18
4 Genetic Characteristics and Diversity of *A. vitis* 19
4.1 Genome Organization of *A. vitis* 19
4.2 Genetic Diversity of *A. vitis* 19
4.3 Ti Plasmids 21
4.4 Opine-Catabolic Plasmids 24
4.5 Tartrate Utilization Plasmids 25
5 Plant–Host Interactions 26
5.1 Tumor Induction and Opine Production 26
5.2 Root Necrosis and Associated Mechanisms 28
5.3 Host Range 30
6 Epidemiology of Grapevine Crown Gall 30
6.1 Survival in and on Grapevine—The Role in Pathogen Dissemination 30
6.2 Survival in Soil and Importance of Soil-Borne Infections 33
6.3 Microbial Community Associated with Grapevine Crown Gall 34

N. Kuzmanović (✉)
Institute for Epidemiology and Pathogen Diagnostics, Julius Kühn-Institut, Federal Research Centre for Cultivated Plants (JKI), Messeweg 11-12, 38104 Braunschweig, Germany
e-mail: kuzmanovic1306@gmail.com

J. Puławska
Research Institute of Horticulture, ul. Konstytucji 3 Maja 1/3, 96-100, Skierniewice, Poland

L. Hao
College of Life Sciences and Oceanography, Shenzhen University, Shenzhen, Guang Dong 518060, People's Republic of China

T. J. Burr
School of Integrative Plant Sciences, Section of Pathology and Plant-Microbe Biology, Cornell University, Geneva 14456, USA

Current Topics in Microbiology and Immunology (2018) 418:15–53
DOI 10.1007/82_2018_85

Published Online: 20 March 2018

7 Disease Management 36
7.1 Early Diagnosis—Detection of the Pathogen 36
7.2 Management Practices 37
7.3 Evaluation of Strategies for Producing *A. vitis*-Free Grapevines 39
7.4 Biological Control 40
References 42

Abstract *Agrobacterium vitis* is the primary causal agent of grapevine crown gall worldwide. Symptoms of grapevine crown gall disease include tumor formation on the aerial plant parts, whereas both tumorigenic and nontumorigenic strains of *A. vitis* cause root necrosis. Genetic and genomic analyses indicated that *A. vitis* is distinguishable from the members of the *Agrobacterium* genus and its transfer to the genus *Allorhizobium* was suggested. *A. vitis* is genetically diverse, with respect to both chromosomal and plasmid DNA. Its pathogenicity is mainly determined by a large conjugal tumor-inducing (Ti) plasmid characterized by a mosaic structure with conserved and variable regions. Traditionally, *A. vitis* Ti plasmids and host strains were differentiated into octopine/cucumopine, nopaline, and vitopine groups, based on opine markers. However, tumorigenic and nontumorigenic strains of *A. vitis* may carry other ecologically important plasmids, such as tartrate- and opine-catabolic plasmids. *A. vitis* colonizes vines endophytically. It is also able to survive epiphytically on grapevine plants and is detected in soil exclusively in association with grapevine plants. Because *A. vitis* persists systemically in symptomless grapevine plants, it can be efficiently disseminated to distant geographical areas via international trade of propagation material. The use of healthy planting material in areas with no history of the crown gall represents the crucial measure of disease management. Moreover, biological control and production of resistant grape varieties are encouraging as future control measures.

1 Introduction

Agrobacterium vitis (i.e., *Agrobacterium* biovar 3) is considered one of the most significant and destructive bacterial pathogens of grapevine worldwide. This bacterium is the primary causal agent of crown gall of grapevine (*Vitis vinifera*) (Burr et al. 1998; Burr and Otten 1999). However, tumorigenic strains belonging to the *Agrobacterium tumefaciens* complex (i.e., *Agrobacterium* biovar 1/*Agrobacterium tumefaciens*) and *Rhizobium rhizogenes* (i.e., *Agrobacterium* biovar 2/*Agrobacterium rhizogenes*) are occasionally associated with disease (see below Sect. 6.3).

Typical symptoms of grapevine crown gall disease include tumor formation on the aerial plant parts, unlike symptoms on most other hosts of tumorigenic agrobacteria. *A. vitis* also causes root necrosis on grapevines (Burr et al. 1987a).

To our knowledge, the first report on grapevine crown gall was made by Fabre and Dunal (1853) describing the disease in France. The infectious nature of this bacterium as a causal agent of grapevine crown gall was demonstrated by Cavara (1897) in Italy. However, at that time *A. vitis* was not established as a separate species. The evolution of classification and nomenclature of *A. vitis* will be reviewed in the next section referring to the taxonomy of *A. vitis* (see below Sect. 2).

Crown gall of grapevine is an economically important plant disease and is particularly serious in regions prone to temperatures that cause freeze injuries to dormant trunks and canes. Potential economic losses in replant and wine sales over a six-year period are calculated to be notably high (Stewart et al. 2013). The disease reduces the vigor and yield of infected plants by up to 40% depending on the extent of infection (Schroth et al. 1988). Severe infections can lead to dieback of canes or whole plants. Crown gall is especially destructive when the graft union is affected. Serious economic losses occur especially in nurseries because grafted grapevines with visible symptoms are unmarketable and are generally discarded. Moreover, both tumorigenic and nontumorigenic strains of *A. vitis* may negatively affect graft strength and subsequent nursery production and vineyard establishment (Hao et al. 2017).

2 Taxonomy of *A. vitis*

A. vitis belongs to the family *Rhizobiaceae*, order *Rhizobiales*, class *Alphaproteobacteria*. As with related species of the *Agrobacterium* and *Rhizobium* genera, *A. vitis* is an aerobic, nonsporeforming, Gram-negative, rod-shaped bacterium with peritrichous flagella (Young et al. 2005). In general, *A. vitis* grows optimally at 25–28 °C and produces copious amounts of extracellular polysaccharide slime on carbohydrate-rich media.

Initially, tumorigenic strains exclusively isolated from grapevine were defined as an atypical group that could not be classified to *Agrobacterium* biovar 1 or biovar 2 (Panagopoulos and Psallidas 1973). These atypical strains were subsequently classified to *Agrobacterium* biovar 3 (biotype 3) based on biochemical and physiological properties (Kerr and Panagopoulos 1977; Panagopoulos et al. 1978; Süle 1978). *Agrobacterium* biovar 3 could also be differentiated by serological analysis using monoclonal antibodies (Bishop et al. 1989). Finally, Ophel and Kerr (1990) formally described a new species *A. vitis* (vi'tis. L. n. *Vitis*, generic name of grapevines) based on polyphasic characterization of biovar 3 strains. However, it was recently shown that *A. vitis* is phylogenetically distinguishable from *Agrobacterium* spp. by using multilocus sequence analysis (MLSA), and its transfer to the revived genus *Allorhizobium* was suggested (Mousavi et al. 2014, 2015).

Moreover, *A. vitis* has a different organization of genetic material compared to that of the genus *Agrobacterium* (see below Sect. 4.1), which also supports its distinctiveness (Ramírez-Bahena et al. 2014).

In recent years, genomics has significantly impacted the taxonomic status of bacteria. In particular, by using quantitative whole-genome comparison methods which include calculation of average nucleotide identity (ANI; Richter and Rossello-Mora 2009) and in silico DNA–DNA hybridization (*is*DDH) values (Meier-Kolthoff et al. 2013), it is possible to delineate prokaryotic species boundaries. Moreover, whole-genome-based phylogeny allows generation of highly robust phylogenetic trees based on large nucleotide or amino acid datasets. In this regard, whole-genome-based phylogeny based on 384 protein sequences conserved in the chromosomes also suggested that *A. vitis* represents a species distinct from *Agrobacterium* spp. and is more similar to *Allorhizobium undicola* (Ormeño-Orrillo et al. 2015). Moreover, our results based on genome comparisons (ANI and *is*DDH) of different *A. vitis* strains indicate that *A. vitis* is not a homogenous species but a species complex comprising at least three genomic species (unpublished data). Type strain K309^{T} represents genomic species G1, whereas the well-known strains AB4 and S4 are representatives of genomic species G2 and G3, respectively. They all have similar average GC contents, ranging from 57.5 to 57.6%. Therefore, the taxonomy of *A. vitis* requires further elucidation.

3 Geographic Distribution of *A. vitis*

The geographic distribution of *A. vitis* generally reflects that of its host grapevine. So far, the presence of the pathogen has been reported in Australia (Ophel et al. 1988), Brazil (De Oliveira et al. 1994), Bulgaria (Genov et al. 2006a), Canada (Dhanvantari 1983), China (Ma et al. 1987), Egypt (Tolba and Zaki 2011), France (Ridé et al. 2000), Germany (Bien et al. 1990), Greece (Panagopoulos and Psallidas 1973), Hungary (Süle 1978), Italy (Bini et al. 2008b), Japan (Sawada et al. 1990), Jordan (Al-Momani et al. 2006), Iran (Mohammadi and Fatehi-Paykani 1999), Israel (Haas et al. 1991), Portugal (Nascimento et al. 1999), Russia (Ignatov et al. 2016), Serbia (Kuzmanović et al. 2014), Slovenia (Zidarič 2009), South Africa (Loubser 1978), South Korea (Lim et al. 2009), Spain (López et al. 1988), Tunisia (Chebil et al. 2013a), Turkey (Argun et al. 2002), and the USA (Burr and Hurwitz 1981) (Fig. 1). Although to the best of our knowledge, official reports have not been made, tumorigenic *A. vitis* strains originating from Afghanistan, Croatia, Moldova, Montenegro, Morocco, and Poland were included in some studies, confirming the presence of the pathogen in these countries (Kerr and Panagopoulos 1977; McGuire et al. 1991; Bini et al. 2008b; Kuzmanović et al. 2015). The pathogen is most likely present in more countries although, to the extent of our knowledge, this is not yet documented in the literature.

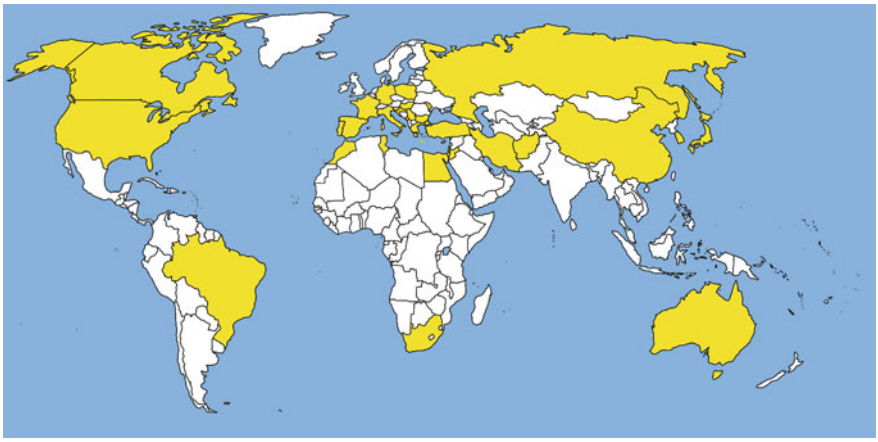


Fig. 1 Geographic distribution (marked in yellow) of *A. vitis* according to the data documented in the literature

4 Genetic Characteristics and Diversity of *A. vitis*

4.1 Genome Organization of A. vitis

A. vitis has genome architecture that includes two circular chromosomes and a variable number of plasmids (Jumas-Bilak et al. 1998; Tanaka et al. 2006; Slater et al. 2009). On the other hand, members of the genus *Agrobacterium* are characterized by the presence of a circular chromosome and a linear chromosome (Allardet-Servent et al. 1993; Jumas-Bilak et al. 1998; Slater et al. 2009; Slater et al. 2013; Ramírez-Bahena et al. 2014). Strain S4 was the first *A. vitis* strain with a fully sequenced genome and so far the only one with high-quality, complete, and published chromosome and plasmid sequences (Slater et al. 2009). The genome of strain S4 contains two circular chromosomes and five plasmids. Plasmids play a substantial role in the pathogenicity and ecology of agrobacteria and rhizobia. The number of plasmids in tumorigenic strains of *A. vitis* may range from 2 to 5 (Perry and Kado 1982; Albiach and Lopez 1992). Their pathogenicity is primarily determined by a large conjugal tumor-inducing (Ti) plasmid. Besides the Ti plasmid, *A. vitis* may also carry other ecologically important plasmids, such as tartrate-catabolic plasmids and opine-catabolic plasmids (see below Sects. 4.3, 4.4 and 4.5).

4.2 Genetic Diversity of A. vitis

Knowledge of genetic variation and the relatedness between strains are of crucial importance, particularly in epidemiological studies and for a better understanding of

the ecology and evolution of the pathogen. For this reason, the diversity of *A. vitis* was investigated both in terms of chromosomal and plasmid DNA. Here, we will focus on the genetic variation of chromosomal DNA, whereas the diversity of *A. vitis* plasmids will be discussed separately in the next section.

Overall, numerous studies demonstrated that *A. vitis* is genetically very diverse. A number of genetic groups were differentiated by analyzing populations of *A. vitis* in Australia (Gillings and Ophel-Keller 1995), Bulgaria (Genov et al. 2006b), Germany (Schulz et al. 1993), Iran (Rouhrazi and Rahimian 2012b), Japan (Kawaguchi et al. 2008b), Serbia (Kuzmanović et al. 2014, 2015), Spain (Palacio-Bielsa et al. 2009b), USA (Irelan and Meredith 1996; Otten et al. 1996b; Momol et al. 1998; Burr et al. 1999), and Turkey (Argun et al. 2002; Canik Orel et al. 2016). Moreover, strains isolated from the same locality and grapevine cultivar may belong to different genetic clusters (Kuzmanović et al. 2015).

Methods targeting sequences distributed throughout the whole genome, such as restriction analysis of total genomic DNA, as well as by PCR-based methods such as repetitive sequence-based PCR (rep-PCR) and randomly amplified polymorphic DNA (RAPD), allow comparison of closely related strains and assessment of clonality among them, which is particularly valuable in epidemiological studies. Additionally, the 16S-23S rRNA internal transcribed spacer (ITS) region is highly variable and therefore suitable for estimating epidemiological relationships among strains. Interestingly, Bautista-Zapanta et al. (2009) reported the existence of intercistronic heterogeneity of the 16S–23S rRNA ITS region among some *A. vitis* strains. However, the frequency of this phenomenon among a broader collection of *A. vitis* strains is unknown.

By applying the methodologies mentioned above, grouping of *A. vitis* strains isolated worldwide did not reflect the geographic origin of the strains. Some homogenous genetic groups were comprised of strains isolated in different countries and even different continents (Gillings and Ophel-Keller 1995; Otten et al. 1996b; Momol et al. 1998; Kuzmanović et al. 2014, 2015), which is consistent with the main means of spread of *A. vitis* via grapevine propagation material.

The 16S rDNA is a widely used phylogenetic marker for classification and discrimination of bacteria. However, it generally lacks the resolution power to discriminate among strains at the intraspecies or even intragenus level. Indeed, 16S rDNA sequences of *A. vitis*-type strain K309 and strain S4 differ by only one nucleotide (Otten et al. 1996a), although whole-genome comparisons suggest that they belong to different genomic species (Kuzmanović, unpublished). The 16S rDNA sequences of various *A. vitis* strains isolated in Spain also differed by one or two nucleotides at the same position (Palacio-Bielsa et al. 2009b). On the other hand, housekeeping genes, which are responsible for basic cellular functions and are relatively conserved on chromosomes, are particularly suitable markers for the assessment of phylogenetic relatedness among bacteria (Gevers et al. 2005). They were therefore exploited for analyzing *A. vitis* strains originating from various geographic locations (Kawaguchi et al. 2008b; Kuzmanović et al. 2015). Sequence analysis of *dnaK*, *gyrB*, and *recA* housekeeping genes was employed to characterize a representative collection of *A. vitis* strains originating from several European

countries, Africa, North America, and Australia (Kuzmanović et al. 2015). Nucleotide sequence analysis indicated high genetic diversity among the strains studied and suggested the presence of recombination events in *A. vitis*, particularly affecting the *dnaK* locus. Phylogeny based on *recA* gene sequences revealed four main phylogenetic groups among the *A. vitis* strains studied. Strains K309, AB4, and S4 were clustered in separate phylogenetic groups (Kuzmanović et al. 2015), which was in accordance with whole-genome comparisons mentioned above and suggested that they belong to three different genomic species within *A. vitis*.

4.3 Ti Plasmids

As for other agrobacteria, the Ti plasmid of *A. vitis* carries the primary genes required for pathogenicity and can determine the host range of the pathogen (host range is further discussed in Sect. 5.3). The size of the Ti plasmid is approximately 200 kbp (Table 1). The Ti plasmid consists of the following functional elements: transferred DNA (T-DNA), virulence (*vir*) region, opine catabolism genes, replication (*rep*) region, and conjugative transfer genes (*tra* and *trb* loci). T-DNA carries genes responsible for production of plant hormones and tumor induction (oncogenes), and genes encoding biosynthesis of low molecular weight molecules termed opines (more details are given in Sect. 5.1).

Ti plasmids belong to the *repABC* family of megaplasmids and encode two independent type IV secretion systems (T4SSs; Suzuki et al. 2009; Pappas and Cevallos 2011; Christie and Gordon 2014). The first T4SS system (*vir*) is responsible for processing and transfer of T-DNA from the bacterium to host plant cells, whereas the second (*tra*/*trb*) mediates conjugative transfer of the Ti plasmid (Christie and Gordon 2014). Conjugative transfer of the Ti plasmid is regulated by a quorum-sensing (QS) mechanism and is induced by opines produced in crown gall

Table 1 Characteristics of *A. vitis* Ti plasmids

Size	~160–260 kb
GC content (%)	~56–57%
Rep type[a]	*repABC*
Incompatibility groups	IncRh1, IncRh4
MOB families[b]	MOB_Q, MOB_P
T4SS types[c]	MPF_T (2)
Opine types	Nopaline (N) Vitopine (V) Octopine/cucumopine (O/C)

[a]Type of replication (rep) system
[b]Mobilization (MOB) family based on the amino acid sequence of conjugative relaxases (TraA and VirD2)
[c]Type IV secretion system (T4SS) type based on TrbE and VirB4 mating pair formation

tumors (Dessaux et al. 1998; Farrand 1998; White and Winans 2007; Faure and Lang 2014).

Ti plasmids have mosaic structures composed of conserved and variable regions, suggesting that they most likely evolved through horizontal gene transfer and recombination (Otten et al. 1992, 1993). This fact hinders their employment in the development of a satisfactory classification scheme. However, Ti plasmids can be grouped based on their backbone genes, such as those controlling plasmid replication and partitioning (*repABC* operon), as well as conjugative transfer. Thus, Ti plasmids and related root-inducing (Ri) plasmids have been classified into four incompatibility groups (Otten et al. 2008). *A. vitis* Ti plasmids pTiB10/7 and pTiAT66 were classified into the IncRh1 group, whereas pTiS4 belongs to the IncRh4 group (Table 1) (Szegedi et al. 1996). This classification scheme is based on their incompatibility (Inc) characteristics, referring to the inability of two related plasmids to be propagated stably in the same host cell line (Couturier et al. 1988). Incompatibility is mainly determined by functional elements located within the *repABC* cassette or in its proximity (Pappas and Cevallos 2011; Pinto et al. 2012). Therefore, the phylogeny of RepA, RepB, and RepC proteins has been used for grouping plasmids of agrobacteria (Wetzel et al. 2015). The Ti plasmid of *A. vitis* S4 was included in this study and clustered together with the OC plasmid pAoF64/95 of *R. rhizogenes* F64/95, the symbiotic plasmid pCB782 of *Rhizobium leguminosaurum* bv. *trifolii* CB782, and the cryptic plasmid pOV14c harbored by *Ensifer adhaerens* OV14, unlike other Ti plasmids investigated.

Classification of plasmids based on the amino acid sequence of conjugative relaxase and T4SS proteins has also been proposed (Garcillan-Barcia et al. 2009; Smillie et al. 2010; Garcillan-Barcia et al. 2011). In this respect, Ti plasmids are classified into the mobilization (MOB) families MOB_Q and MOB_P, based on their TraA and VirD2 relaxase protein sequences, respectively (Table 1) (Christie and Gordon 2014). Both T4SSs of Ti plasmids are classified into the MPF_T group based on TrbE and VirB4 mating pair formation (MPF) protein sequences (Table 1). The relaxase and T4SS proteins of the Ti plasmid harbored by *A. vitis* S4 are generally phylogenetically related to those of other Ti plasmids (Wetzel et al. 2015).

Plasmid backbone markers used for the classifications discussed above can be associated with different gene clusters having important roles in the ecology of the pathogen and epidemiology of crown gall. For example, Ti plasmids possess genes and gene clusters responsible for the synthesis and catabolism of opines, which are a specific class of compounds produced in crown gall tumors following genetic transformation by the pathogen. Opine markers were widely used for grouping of Ti plasmids and their host bacteria, and classification based on opines reflects different aspects involved in plant–pathogen and pathogen–pathogen interactions (see Sect. 5.1). Based on opine types, *A. vitis* Ti plasmids and host strains were differentiated into octopine/cucumopine (O/C), nopaline (N), and vitopine (V) groups (Table 1) (Szegedi et al. 1988; Paulus et al. 1989a). Additionally, an octopine-type Ti plasmid carried by atypical strain CG474 has been reported (Burr et al. 1998; Burr and Otten 1999). Interestingly, strains carrying both octopine and vitopine synthase genes were detected in Italy and Tunisia (Bini et al. 2008b; Chebil et al.

2013b). These strains might harbor both O/C- and V-type Ti plasmids, or a new Ti plasmid type with another arrangement of opine-related genes. In this regard, the existence of Ti plasmids having different combinations of genes encoding synthesis and catabolism of opines cannot be excluded, which complicates classification based on the presence of opine genes. Tumorigenic *Agrobacterium* biovar 1 strains isolated from grapevine harbor Ti plasmids typical for this species, although they may also carry O/C-type Ti plasmids characteristic of *A. vitis* (Szegedi et al. 2005).

The size of the O/C-type Ti plasmid of *A. vitis* AB3 (pTiAB3) is approximately 234 kb (Otten et al. 1995). On the other hand, the nopaline-type Ti plasmid of *A. vitis* AB4 (pTiAB4) is relatively small (∼157 kb) (Otten and De Ruffray 1994). The vitopine-type Ti plasmid of *A. vitis* S4 (pTiS4) is particularly large (258,824 bp). Its GC content is 56.7%, which is similar to that of other Ti plasmids (Suzuki et al. 2009).

O/C-type Ti plasmids are predominant within the population of *A. vitis* (60%), followed by N Ti plasmids (30%), whereas the V type is less abundant (10%) (Burr et al. 1998). Studies performed in Hungary (Szegedi 2003), France (Ridé et al. 2000), and Turkey (Canik Orel et al. 2016) were more or less in accordance with this proportion. Strains carrying an O/C-type Ti plasmid were also predominant in Bulgaria (Genov et al. 2006a, 2015), China (Ma et al. 1987), Spain (Palacio-Bielsa et al. 2009b), and Serbia (Kuzmanović et al. 2014). In Germany, more or less equal numbers of O/C and N strains of *A. vitis* were isolated from grapevine tumors, whereas the vitopine strains were not detected (Bien et al. 1990). On the contrary, vitopine strains were more abundant in Italy (Bini et al. 2008b) and Iran (Rouhrazi and Rahimian 2012b). The N strains were not identified within *A. vitis* populations in Bulgaria, Italy, and Serbia. It is unclear whether these differences in distribution of particular opine types are influenced by ecological factors or if they are a consequence of distribution of specific grape cultivars and/or rootstocks.

The O/C, N, and V types of Ti plasmids are characterized by a complex structure and gene arrangement of their T-DNA. Different variants of T-DNA structures have been described and thoroughly reviewed previously (Paulus et al. 1989a; Huss et al. 1990; Burr et al. 1998; Burr and Otten 1999). The O/C-type Ti plasmids possess two independent T-DNA fragments, TA-DNA and TB-DNA. Although at least six different O/C Ti plasmid structures have been characterized, based on the structure of the TA-DNA, they are divided into two main sub-groups OS and OL, having a small or large TA-DNA region, respectively. T-DNA of the atypical octopine strain CG474 has a unique T-DNA structure lacking TB-DNA, although it has some similarities with classical octopine TL-DNA and O/C TA-DNA (Burr and Otten 1999). The N-type Ti plasmids comprise a single T-DNA (Otten and De Ruffray 1994), whereas V types possess three independent T-DNAs (Canaday et al. 1992). The N- and V-type Ti plasmids are less variable than are the O/C plasmids. Reconstruction of evolutionary relationships among different T-DNA variants is hindered because it has been shown that Ti plasmids mainly evolved through horizontal gene transfer, insertions, and deletions (Otten et al. 1992; van Nuenen

et al. 1993; Otten and De Ruffray 1994). However, little is known about the ecological dynamics of different Ti plasmid variants.

In a study on distribution and localization of insertion elements in *A. vitis* strains, Paulus et al. (1989b) found a correlation between Ti plasmid genotype and the particular chromosomal background within an O/C group of strains. Later, Otten et al. (1996b) showed that most Ti plasmids in *A. vitis* are associated with a particular chromosomal background determined by restriction fragment length polymorphism (RFLP) analysis of the 16S-23S rRNA ITS region, which was further supported by Genov et al. (2006b). Similar studies reported a general coherence between 16S rRNA gene sequences, RFLP profiles of the 5′-end of the 23S rRNA gene, and RAPD fingerprints with a particular type of Ti plasmid (Momol et al. 1998; Palacio-Bielsa et al. 2009b). However, Turkish *A. vitis* strains were exceptions (Argun et al. 2002). Based on analysis of *A. vitis* strains isolated in Serbia, no strong correlation between 16S-23S rRNA ITS genotype of the strains and their type of Ti plasmid was found, although some 16S-23S rRNA ITS groups and clusters were composed of strains harboring a particular type of Ti plasmid (Kuzmanović et al. 2014). This plasmid–chromosome correlation found in *A. vitis* is unique among tumor-inducing *Agrobacterium* and *Rhizobium* species. For other agrobacteria and rhizobia causing crown gall, no correlation between chromosome and Ti plasmid type was found. Although it is assumed that Ti plasmid exchange or transfer occurs more or less frequently between *A. vitis* strains, certain Ti plasmids appear to be strongly and stably associated with a particular bacterial host. This may be influenced by the specific pathosystem of *A. vitis,* with the highly specialized pathogen persisting systemically in grapevines. This may limit the chance of its contact with diverse and large pools of tumorigenic strains in contrast to predominantly soil-inhabiting agrobacteria and rhizobia.

The diversity of *vir* region and opine catabolism gene clusters among *A. vitis* strains has not been studied extensively. However, O/C and N Ti plasmids have very similar or identical virulence regions (Otten and De Ruffray 1994). On the contrary, the vitopine Ti plasmid of *A. vitis* has a different organization of *vir* region (Gerard et al. 1992). Kawaguchi and Inoue (2009) analyzed the phylogeny of strains belonging to the *A. vitis* and *A. tumefaciens* complex, including *Agrobacterium* strains from other hosts, using the partial nucleotide sequences of the *virC* operon. The majority of *A. vitis* strains formed a separate cluster, whereas the remaining *A. vitis* strains formed additional monophyletic groups with strains of the *A. tumefaciens* complex isolated from grapevine. Strains from other hosts clustered separately on the phylogenetic tree.

4.4 Opine-Catabolic Plasmids

Opine-catabolic plasmids (pOC) are a specific group of replicons that, like Ti plasmids, carry genes encoding the uptake and catabolism of opines but do not contain *vir* genes and T-DNA required for pathogenicity. pOCs have been identified

both in tumorigenic and in nontumorigenic *Rhizobiaceae* strains isolated from tumors or soil around diseased plants (Merlo and Nester 1977; Sciaky et al. 1978; Wabiko et al. 1990; Szegedi et al. 1999; Wetzel et al. 2014; Puławska et al. 2016). Therefore, nonpathogenic strains carrying this plasmid are provided with the ability to metabolize opines and to proliferate inside or near tumors and diseased plants. Plasmids pAtK84b and pAoF64/95 carried by nonpathogenic *R. rhizogenes* strains K84 and F64/95, respectively, are the only pOCs that have been studied in detail (Clare et al. 1990; Wetzel et al. 2014). As with Ti plasmids, these two plasmids carry all genes required for conjugative transfer that is induced by opines and regulated by a QS system (Oger and Farrand 2002; Wetzel et al. 2014).

Data regarding the distribution of pOCs in strains associated with grapevine tumors are limited. Sciaky et al. (1978) characterized plasmids of various agrobacteria, including one avirulent *Agrobacterium* biovar 1 strain AG19 isolated from tumor tissue on grapevine in Greece. Although this strain did not confer tumorigenicity, it could utilize octopine and harbored a plasmid encoding this ability. Additional *Agrobacterium* biovar 1 strains from grapevine carrying large plasmids encoding catabolism of octopine were also described (Knauf et al. 1983). Later, Szegedi et al. (1999) reported the existence of pOC in nonpathogenic *A. vitis* strain F2/5 encoding the catabolism of octopine. Strain F2/5 is a potential biocontrol agent able to inhibit crown gall disease on grapevines. Similarly, pAtK84b conferring utilization of nopaline is harbored by the nonpathogenic biocontrol strain *R. rhizogenes* K84 (Clare et al. 1990).

4.5 Tartrate Utilization Plasmids

Tartrate utilization plasmids (pTrs) are a group of plasmids described in strains of *A. vitis* and related *Agrobacterium* species associated with grapevine. Gallie et al. (1984) first identified a plasmid encoding utilization of tartrate (named pTAR) in an atypical nonpathogenic *Agrobacterium* biovar 1 strain isolated from grapevine. The amino acid sequence of the replication protein RepA encoded by pTAR showed close homology to RepA protein of pAgK84, a plasmid that encodes synthesis of agrocin 84 in the biocontrol strain *R. rhizogenes* K84 (Gallie and Kado 1988; Kim et al. 2006). These two plasmids may belong to the same family and have a common evolutionary origin.

Szegedi et al. (1992) subsequently demonstrated that various tumorigenic *A. vitis* strains also carry large pTrs, some of which were self-conjugal. The pTr can be also carried by nonpathogenic *A. vitis* strains, for instance by the biocontrol strain F2/5 (Szegedi et al. 1999). The pTrs identified in tumorigenic *A. vitis* strains represent a diverse group of plasmids differing in size, transfer frequency, and stability in recipient *Agrobacterium* biovar 1 strains (Szegedi et al. 1992). They were also clearly different from the pTAR plasmid described by Gallie et al. (1984). pTrs also showed a high diversity in their incompatibility properties (Szegedi and Otten 1998). Therefore, they could coexist with pTis from different incompatibility groups.

The tartrate utilization system of the nopaline strain AB4 has been characterized and the corresponding genes identified as an operon (Crouzet and Otten 1995).

Tartrate utilization in this strain is encoded by the 170-kb conjugative plasmid pTrAB4. On the other hand, the O/C strain AB3 carries two independent tartrate utilization systems, one encoded by the Ti plasmid (pTiAB3) and another by pTrAB3 (245 kb; Otten et al. 1995). Regions encoding tartrate utilization genes are found in most of *A. vitis* strains, and they have been grouped into three different types based on analysis of their sequences (Salomone et al. 1996; Salomone and Otten 1999). Vitopine strains of *A. vitis* are also able to degrade tartrate, although their tartrate utilization system is different and has not yet been characterized (Salomone et al. 1996). In any case, tartrate degradation by vitopine strain S4 is encoded by a large plasmid pAtS4c (211,620 bp), which was initially named pTrS4 (Szegedi and Otten 1998). Interestingly, different genetic regions encoding tartrate utilization showed complex distribution patterns among various *A. vitis* strains that correlated with their chromosomal backgrounds (Salomone et al. 1996). Some pTrs encode production of putative signal molecules used in QS regulation; however, their regulatory role remains unknown (Lowe et al. 2009).

Because tartrate is an abundant compound in grapevine, pTrs may enhance host strain competitiveness on this plant species (Kado 1998; Salomone et al. 1998). Both *A. vitis* and *R. rhizogenes* use tartrate as a sole carbon source (Kerr and Panagopoulos 1977; Moore et al. 2001). However, the latter species prefers glucose to tartrate, unlike *A. vitis*, that utilizes tartrate more intensively than glucose (Szegedi 1985). Although it was shown that *Agrobacterium* biovar 1 strains generally do not utilize tartrate (Kerr and Panagopoulos 1977; Moore et al. 2001), some biovar 1 strains isolated from grapevine were able to catabolize this compound (Gallie et al. 1984; Szegedi et al. 2005). Interestingly, most of the nontumorigenic *A. vitis* strains isolated from roots of asymptomatic feral *Vitis riparia* vines in the USA were not able to utilize tartrate (Burr et al. 1999).

5 Plant–Host Interactions

5.1 Tumor Induction and Opine Production

A. vitis can be present in symptomless propagation material and thereby cause infections on young plants in newly established vineyards. In such cases, infection mostly occurs on aerial plant parts through wounds caused by abiotic and biotic factors, especially by freezing temperatures and cultural practices (Burr et al. 1998; Burr and Otten 1999). However, in warmer climates, such as that of Israel and South Africa, high temperatures and humidity can cause injuries and initiate infection (Burr et al. 1998). On the other hand, wounds made by disbudding and grafting are particularly important for triggering infection in nurseries. For soil-borne infections, injuries made by cultural practices or nematode wounds may also be conducive for infection by the pathogen (Süle et al. 1995; Burr et al. 1998). Interestingly, unlike tumorigenic strains belonging to the *A. tumefaciens* complex, *A. vitis* was unable to induce tumors on in vitro grown grapevine stem segments

(Szegedi et al. 2014). However, factors determining differences in the susceptibility of intact grapevines and explants to *A. vitis* remain to be elucidated.

Infection of plants by tumorigenic agrobacteria is a complex multistage process of natural genetic transformation of plants that includes DNA transfer from bacteria to plants (Hooykaas 2000; Zhu et al. 2000; Zupan et al. 2000; McCullen and Binns 2006; Pitzschke and Hirt 2010; Gelvin 2012). In brief, wounded plant tissue releases signal molecules that trigger infection through induction of *vir* genes of the bacteria. For a number of grapevine cultivars, syringic acid methyl ester was identified as a *vir*-inducing phenolic compound (Spencer et al. 1990). Vir proteins are involved in the processing and transfer of the T-DNA of the Ti plasmid and its stable integration into the plant host genome.

T-DNA genes are expressed in the host plant and encode biosynthesis of the phytohormones auxin and cytokinin (oncogenes) that lead to uncontrolled proliferation of plant cells, resulting in tumor formation. Tumors develop mainly on the lower trunks, graft unions, cordons, and canes (Fig. 2). Initial symptoms may be inconspicuous and remain unnoticed. However, as disease develops, tumor tissue can enlarge rapidly. Tumors can be localized or in the form of continuous proliferations that completely girdle the trunk (Fig. 2). Tumors are rarely observed on grapevine roots.

However, T-DNA genes also encode production of opines, as previously described in this chapter, which play important roles in the epidemiology of crown gall and the ecology of tumorigenic bacteria. Opines are typically conjugates of amino acids and α-ketoacids or sugars, and less frequently, they are sugar phosphodiesters (Dessaux et al. 1993, 1998; Chilton et al. 2001). So far, more than 20 different opine types belonging to different structural families have been described and characterized (Dessaux et al. 1993, 1998; Chilton et al. 2001).

In general, multiple opines belonging to different families may be produced in tumors as a consequence of plant genetic transformation. For example, O/C strains

Fig. 2 Symptoms of grapevine crown gall. Tumors on cordons and canes (left and middle), and graft unions (right)

of *A. vitis* are responsible for production of octopine and cucumopine in tumor tissue (Szegedi et al. 1988; Paulus et al. 1989a). Although additional opines are produced in tumors induced by various tumorigenic strains carrying well-studied octopine-type Ti plasmids (Dessaux et al. 1998; Zhu et al. 2000), their presence has not been investigated in grapevine tumors caused by *A. vitis* carrying O/C-type Ti plasmids. Nopaline was the only opine detected in tumors caused by *A. vitis* strains carrying an N type of Ti plasmid. Finally, V strains of *A. vitis* induced tumors in which vitopine and ridéopine were produced (Szegedi et al. 1988; Paulus et al. 1989a; Chilton et al. 2001). Chilton et al. (2001) suggested that vitopine is identical to heliopine. Heliopine is one of the opines detected in tumors caused by *Agrobacterium* strains carrying a classical octopine-type plasmid (e.g., *Agrobacterium* biovar 1 strain 15955), and its structure has been published (Chang et al. 1989).

Primarily, opines serve as selective nutrient sources for the pathogen because, as indicated above, genes responsible for uptake and catabolism of opines are located on the Ti plasmid, outside of the T-DNA region. The presence of opines is not limited to tumors, and opines can translocate to other plant parts and can also be secreted from roots as a component of root exudates (Savka et al. 1996).

Some opines can also induce conjugative transfer of Ti plasmids among agrobacteria (Dessaux et al. 1998; Farrand 1998). Opines therefore contribute to dissemination of Ti plasmids. Although transferred Ti plasmid genes may encode production of more than one opine type in a particular tumor, thus far it appears that only some opines serve as conjugal inducers (Farrand 1998). The conjugal opines induce the QS system that directly regulates conjugative transfer of the Ti plasmid. Taken together, the conjugative transfer of the Ti plasmid is dependent on pathogen population density and requires the presence of a conjugal opine. Conjugation mechanisms of plasmids harbored by *A. vitis* strains likely behave similarly, but this has thus far not been studied.

Unwounded tobacco seedlings can elicit *Agrobacterium vir* gene induction and T-DNA transfer (Brencic et al. 2005). Intriguingly, transformation did not lead to tumor formation, although plants produced opines. These results suggested that genetic transformation of plants by tumorigenic bacteria does not require wounding and that cell division during wound healing may play a role in tumor formation. However, such interaction between *A. vitis* and grapevine has not yet been studied.

5.2 Root Necrosis and Associated Mechanisms

Unlike other tumorigenic agrobacteria, *A. vitis* causes necrosis on roots of grapevine plants (Fig. 3; Burr et al. 1987a). Necrosis may provide a niche for the bacterium to persist in the soil, as *A. vitis* can persist in grapevine root debris for at least two years (Burr et al. 1995). Necrosis develops within 24 to 48 h after inoculation and is generally restricted to localized lesions from which *A. vitis* can be consistently isolated.

Fig. 3 Symptoms of root necrosis on grapevine. Necrosis on grapevine roots from a nursery (left), on an inoculated rooted cutting (middle), and on an inoculated seedling (right)

Interestingly, both tumorigenic and nontumorigenic strains of *A. vitis* cause root necrosis. Later studies indicated that the enzyme polygalacturonase encoded by the chromosomal *pehA* gene represents a virulence factor associated with grape root decay (McGuire et al. 1991; Rodriguez-Palenzuela et al. 1991). Nonetheless, a mutant strain lacking the polygalacturonase gene could still induce grape root necrosis when higher concentrations of bacteria were used, suggesting that additional factors are associated with this process (Herlache 1999).

Polygalacturonase may also play a role in tumorigenesis on grapevine, because *pehA* mutants were less pathogenic on this host (Rodriguez-Palenzuela et al. 1991). However, both wild-type and *pehA* mutants were equally tumorigenic on potato disks. Synthesis of polygalacturonase appears to be associated with host specialization of *A. vitis*, because it may affect attachment of bacteria to grape roots and their multiplication at wound sites (Brisset et al. 1991).

Polygalacturonase of *A. vitis* showed significant similarity to those of two other plant pathogens, *Pectobacterium carotovorum* and *Ralstonia solanacearum*, although differences in their soft rotting effects on potato tuber tissue were reported (Herlache et al. 1997).

Interestingly, *A. vitis* also causes a hypersensitive-like response (HR) on nonhost plants such as tobacco (Herlache et al. 2001). The underlying mechanism of grape necrosis and HR response may be related. In this respect, both grape necrosis and the HR are regulated by a complex QS regulatory system (Hao et al. 2005; Li et al. 2005; Hao and Burr 2006). More recently, the ability of *A. vitis* to cause the HR and necrosis was shown to be associated with a phosphopantetheinyl transferase (PPTase), and bacterial polyketide and nonribosomal peptide synthase-associated genes (Zheng and Burr 2013).

The QS system associated with induction of necrosis on grape and a HR response on tobacco also regulates characteristic surface motility (swarming) of *A. vitis*, which is associated with surfactant secretion (Süle et al. 2009). *A. vitis* was the only of the tested agrobacteria expressing swarming activity, and such behavior may facilitate colonization of grapevine by this pathogen.

5.3 Host Range

A. vitis has been detected in nature almost exclusively in association with grapevine, suggesting a high degree of natural host specialization. Nevertheless, in one exceptional case, *A. vitis* was isolated from galls on the roots of kiwi in Japan (Sawada and Ieki 1992). Interestingly, it was recently reported that *A. vitis* can cause banana leaf blight in China (Huang et al. 2015). Although strains isolated from banana were related to *A. vitis* based on 16S rDNA analysis, further study is required for reliable identification.

Inoculation of various test plants under greenhouse environment conditions revealed differences in host range among *A. vitis* strains and generally among tumorigenic strains isolated from grapevine (Panagopoulos et al. 1978; Knauf et al. 1982; Ma et al. 1987; Bien et al. 1990). Although initial studies suggested that *A. vitis* has a narrow host range limited to grapevine and a few other test plants, it later became evident that there are limited and wide host range strains. In particular, *A. vitis* strains harboring an O/C Ti plasmid (previously named octopine Ti plasmid) showed different host range patterns and were divided into limited and wide host range strains (Thomashow et al. 1980, 1981; Knauf et al. 1982, 1983). In this respect, the Ti plasmid largely determines host range of the bacterium (Thomashow et al. 1980; Knauf et al. 1982). Knauf et al. (1982) showed that different grapevine cultivars can respond differently, depending on the Ti plasmid carried by the inoculant strain. Further studies identified particular Ti plasmid genes located either in T-DNA or in the *vir* region that are associated with *A. vitis* host range (Buchholz and Thomashow 1984a, b; Hoekema et al. 1984; Yanofsky et al. 1985a, b; Bonnard et al. 1989). Moreover, Ti plasmids of limited host range strains may have evolved from ones carried by wide host range strains (Paulus et al. 1991a, b). Although previous studies clearly showed differences in host range of *A. vitis*, they do not answer the question as to why this pathogen is largely restricted to grapevines in natural environments.

6 Epidemiology of Grapevine Crown Gall

6.1 Survival in and on Grapevine—The Role in Pathogen Dissemination

A. vitis can systemically infect grapevine and spread through xylem sap (Fig. 4). The first direct evidence of this phenomenon was provided by Lehoczky (1968), who isolated the pathogen from the xylem bleeding sap of symptomatic vines. Moreover, development of secondary tumors was observed on experimentally injured canes, which provided further evidence of bacteria internally present in the vascular system. Systemic colonization of grapevine by *A. vitis* was confirmed in subsequent studies (Lehoczky 1971; Burr and Katz 1983; Tarbah and Goodman

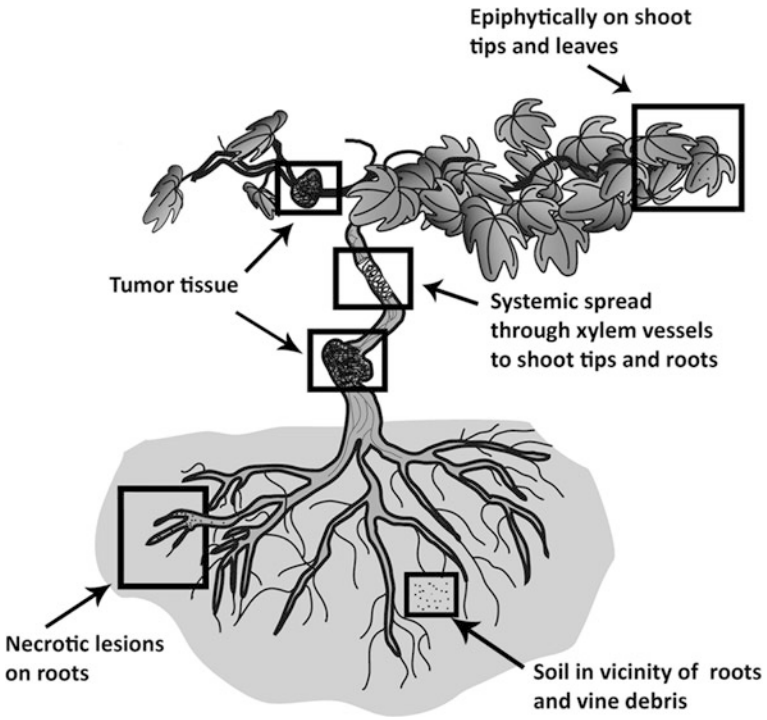


Fig. 4 Survival sites of *A. vitis* in and on grapevine and in soil

1987). *A. vitis* was also isolated from latently infected symptomless grapevine material (Lehoczky 1971; Burr and Katz 1984).

Although *A. vitis* rarely induces tumors on grapevine roots, the bacterium is able to persist in root tissues as it was consistently isolated from tumor-free grapevine roots (Lehoczky 1971, 1978; Süle 1986; Burr et al. 1987b, 1995; Thies et al. 1991). Lehoczky (1978) suggested that grapevine root systems represent a reservoir for *A. vitis* in which the pathogen can multiply and survive extreme environmental conditions. Additionally, *A. vitis* was frequently isolated from necrotic lesions on roots of infected vines (Fig. 4) (Burr et al. 1987b).

Tarbah and Goodman (1987) showed that *A. vitis* can move rapidly through xylem vessels of grapevine shoots. Within a period of 24 h, bacteria translocated 30 cm through the vascular system of shoots inoculated by dipping their freshly cut basal ends in bacterial suspension. Bauer et al. (1994) inoculated shoots of actively growing plants and demonstrated that migration of *A. vitis* from inoculation sites to the roots requires at least 15 weeks. It was later found that population dynamics of *A. vitis* in vines may vary with the cultivar and vegetation period (Burr et al. 1988, 1994; Stover et al. 1997b). Moreover, poor systemic movement of *A. vitis* and its presence in higher numbers mainly at inoculated sites were reported (Bauer et al.

1994; Stover et al. 1997b). Freezing may facilitate systemic movement of *A. vitis* in naturally infected and inoculated vines (Stover et al. 1997b).

Regarding seasonal fluctuations in pathogen population in grapevine, *A. vitis* was most abundant during the spring, decreased during the summer, and then in autumn returned to nearly the same levels as in spring (Bauer et al. 1994). In a recent study, Faist et al. (2016) compared the endophytic microbiota of organs from grapevine plants with or without crown gall disease symptoms using 16S rDNA amplicon sequencing. Their results also suggested that populations of *A. vitis* decreased in summer in comparison with spring and autumn, confirming the results of Bauer et al. (1994).

By testing plant material collected from a crown gall-infested vineyard, *A. vitis* could not be isolated from young green shoots until after the shoots became lignified (Burr et al. 1988). In addition, the pathogen was also absent from shoot tips. Similar results were reported by Bauer et al. (1994), who assayed experimentally inoculated grapevine plants. Nevertheless, Lehoczky (1989) infected healthy vines by grafting them with green shoots from infected plants, indirectly showing that *A. vitis* actually can be present in green shoots. Therefore, the detection methods used in previous studies may not have been sensitive enough for reliable detection of the pathogen. Indeed, using a PCR-based method, Poppenberger et al. (2002) showed that *A. vitis* can be translocated to shoot tips. Recently, by using a highly sensitive magnetic capture hybridization procedure followed by real-time polymerase chain reaction (real-time PCR), *A. vitis* was detected in shoot tips and meristem tissue of grapevine plants that were derived from symptomless cuttings grown in the greenhouse but latently infected with the pathogen (Johnson et al. 2016). The authors also revealed the frequent presence of tumorigenic *A. vitis* in dormant canes and green shoots. In addition, *A. vitis* is irregularly present in tested tissue. Intriguingly, Johnson et al. (2016) suggested that tumorigenic *A. vitis* is able to survive epiphytically on shoot tips. In a related study, Canik Orel et al. (2017) detected *A. vitis* in dormant grape buds and on surfaces of grapevine leaves collected from commercial vineyards (Fig. 4).

Considering its ability to survive in and on grapevine, *A. vitis* can be efficiently disseminated from nurseries to distant geographical areas via symptomless propagation material, which is regarded as the most important means of pathogen spread. Hence, *A. vitis* can be latently present in vines until favorable conditions for disease development arise. In this respect, some studies attempted to resolve the possible introduction pathway and evaluate epidemiological relationships among *A. vitis* strains. Gillings and Ophel-Keller (1995) suggested that *A. vitis* was introduced to South Australia in grapevine cuttings imported from California. Furthermore, a majority of *A. vitis* strains isolated predominantly from young commercial vineyards over the period of eight years (from 2003 to 2011) in six European countries were genotypically related, suggesting that they most likely had a common origin and were distributed following the movement of infected grapevine planting material (Kuzmanović et al. 2015).

The presence of *A. vitis* was also monitored in feral and wild grapevines. In this respect, nontumorigenic *A. vitis* strains were isolated from roots of asymptomatic

feral *V. riparia* vines collected in different locations in the USA (Burr et al. 1999). In contrast, Genov et al. (2006a) isolated tumorigenic *A. vitis* strains from stem samples of wild *V. vinifera* ssp. *sylvestris* collected in natural forests in Bulgaria. Tumorigenic *A. vitis* was also detected in dormant canes collected from symptomless wild grapevines (*V. riparia*) in New York and feral grapevines (including *Vitis californica*) in California, which were either proximal to or distant from commercial vineyards (Johnson et al. 2016; Canik Orel et al. 2017). Therefore, wild grapevines may serve as a significant reservoir of inoculum.

6.2 Survival in Soil and Importance of Soil-Borne Infections

Members of the family *Rhizobiaceae* are generally soil-inhabiting and plant-associated bacteria. Accordingly, *A. vitis* strains are able to survive in soil. However, *A. vitis* has been detected in soil exclusively in association with grapevine plants and, to the best of our knowledge, has never been isolated from non vineyard soil (Fig. 4). Tumorigenic *A. vitis*, including tumorigenic *Agrobacterium* biovar 1 strains, are thus isolated from soil samples taken from the root zone of vines, only from samples collected during the period when fresh galls were present on trunks (Burr and Katz 1983). However, the ratio of pathogenic to nonpathogenic strains was remarkably low. In this respect, Bien et al. (1990) could not isolate pathogenic *A. vitis* strains from soil samples taken from the root zone of infected vines; only nonpathogenic *Agrobacterium* biovar 1 strains were isolated from this source. As determined by isolation of bacteria, *A. vitis* was also detected in nonrhizosphere soil of infected vineyards, although the pathogen may survive to a lesser extent in this environment and may preferentially inhabit the grapevine rhizosphere (Burr et al. 1987b). In their greenhouse study, Bishop et al. (1988) compared the population dynamics of *A. vitis* in grapevine and oat rhizosphere to that of fallow soil. Unlike fallow soil, both host and nonhost rhizospheres enhanced the survival of *A. vitis*. However, the pathogen population was greater in the grapevine rhizosphere without a population decline over the course of the study. As shown before for some other plant species (Guyon et al. 1993; Oger et al. 1997; Savka and Farrand 1997; Mansouri et al. 2002), opines exuded from roots of transformed grapevine might affect the composition of bacterial populations in the rhizosphere and promote the growth of opine-degrading bacteria, primarily *A. vitis*.

Although crown gall outbreaks are primarily associated with diseased or systemically infected propagation material, soil-borne *A. vitis* can be responsible for infections of grapevine plants grown in greenhouse conditions and vineyards (Bishop et al. 1988; Pu and Goodman 1993). However, planting of grapevine in soil containing lower levels of *A. vitis* ($\leq 10^4$ cfu/g) did not result in systemic infection within 10 weeks (Bishop et al. 1988). Soil-borne populations of *A. vitis* represent an important source of inoculum at higher levels of soil infestation (10^6 cfu/g). On

the other hand, nematodes can enhance infection from soil. Combined inoculation of grapevine roots with *Meloidogyne hapla* and *A. vitis* at the level of 10^{2-3} cfu/g of soil resulted in root infestation by *A. vitis* and subsequent systemic colonization of grapevine by the pathogen (Süle et al. 1995). It is still unclear if nematodes could be vectors of *A. vitis*, although nematodes have been used as vectors to transfer *A. tumefaciens* into plant roots in order to transform *Arabidopsis* plants (Karimi et al. 2000). *A. vitis* can also be spread over short or long distances in rhizosphere soil or on the rhizoplane of apparently healthy propagation material (Burr et al. 1987b).

A. vitis can survive in association with decaying grape roots and canes in soil for at least two years after plants were artificially inoculated with the pathogen (Fig. 4) (Burr et al. 1995). However, *A. vitis* can most likely survive longer in decaying grape and cane tissue as long as grapevine residues are present in soil, which could serve as significant reservoirs of inoculum.

Overall, the methods used in the studies described above may not be sensitive enough to detect *A. vitis* in soil at low population densities. Further investigations on the survival of *A. vitis* in soil are therefore needed. In this respect, a magnetic capture hybridization procedure followed by real-time PCR might be the method of choice for such studies.

6.3 Microbial Community Associated with Grapevine Crown Gall

Infected grapevine plants and especially crown gall tumors are dynamic ecological niches inhabited by diverse microorganisms, both pathogenic and nonpathogenic, and their genetic diversity remains largely unexplored. The possible roles of different members of the microbial community associated with grapevine tumors in relation to the ecology and epidemiology of crown gall disease are yet to be explored.

A. vitis is the predominant species identified as a causal agent of grapevine crown gall worldwide. Tumorigenic strains belonging to *Agrobacterium* biovar 1 and *R. rhizogenes* were, however, sporadically isolated from infected grapevine plants (Panagopoulos and Psallidas 1973; Panagopoulos et al. 1978; Süle 1978; Burr and Katz 1984; Ma et al. 1987; Thies et al. 1991; Ridé et al. 2000; Argun et al. 2002; Bini et al. 2008b; Kawaguchi and Inoue 2009; Palacio-Bielsa et al. 2009b; Rouhrazi and Rahimian 2012a; Abdellatif et al. 2013; Genov et al. 2015; Perović et al. 2015). In addition, tumorigenic strain 0 belonging to the newly described species *Agrobacterium nepotum* was isolated from a crown gall tumor on grapevine in Hungary (Süle and Kado 1980; Puławska et al. 2012; Mousavi et al. 2015).

Nontumorigenic *A. vitis* strains were isolated as cohabitants with tumorigenic *A. vitis* strains from grapevine tumors, roots, and sap (Panagopoulos and Psallidas 1973; Burr and Katz 1983; Staphorst et al. 1985; Burr et al. 1987a; Bien et al. 1990; Genov et al. 2006a; Bini et al. 2008b; Kawaguchi et al. 2008b; Rouhrazi and Rahimian 2012a; Canik Orel et al. 2017; Kuzmanović et al., unpublished). Their

occurrence is most likely more frequent in diseased plants; however, they have not been studied in more detail. Nevertheless, it was determined that some non-pathogenic *A. vitis* strains isolated from callus tissue on dormant scion cuttings can catabolize octopine and nopaline, although pathogenic strains were not detected in the same plant material (Ophel et al. 1988). Most likely, these strains carry pOC encoding catabolism of opines.

Grapevine tumors were also inhabited by nontumorigenic strains belonging to novel phylogenetic groups within the genus *Agrobacterium*. In this respect, strains related to *Agrobacterium rubi*, but phylogenetically clearly different, were isolated from grapevine tumors (Kuzmanović, unpublished). As determined by PCR, these strains carried an *ooxA* gene-encoding oxidoreductase for conversion of octopine-type opines to pyruvate and corresponding basic amino acid, suggesting that they are able to catabolize octopine and likely harbor pOC. One of these strains (strain 384) forms a novel *Agrobacterium* species, *Agrobacterium rosae*, together with the atypical tumorigenic strain NCPPB 1650 isolated from hybrid tea rose (*Rosa* x *hybrida*) in South Africa, and three nonpathogenic strains isolated from tumors on raspberry and blueberry (Kuzmanović et al. 2018).

Opine utilization is not restricted to *Agrobacterium* spp. and related organisms. Several reports described that other microorganisms isolated from tumors, soil, and the rhizosphere, including fluorescent and nonfluorescent *Pseudomonas* spp., *Arthrobacter* spp., coryneforms, and several fungal species, are able to utilize opines (Beaulieu et al. 1983; Dahl et al. 1983; Bouzar and Moore 1987; Tremblay et al. 1987; Beauchamp et al. 1990; Bergeron et al. 1990; Nautiyal and Dion 1990; Beauchamp et al. 1991; Canfield and Moore 1991; Nautiyal et al. 1991; Moore et al. 1997). Moore et al. (1997) analyzed opine-catabolizing bacteria in tumors on several hosts, including grapevine. Interestingly, besides one nonpathogenic *Agrobacterium* strain utilizing octopine, they identified various fluorescent *Pseudomonas* strains having the ability to catabolize either octopine or nopaline, or both opines. Genes encoding opine catabolism in some non-*Agrobacterium* species were found on the chromosome, but not on plasmids (Watanabe et al. 2015).

Faist et al. (2016) investigated the bacterial endophytic community associated with grapevine plants with and without crown gall disease, including the surrounding vineyard soil over one year. The authors used cultivation-independent 16S rDNA-based analysis in preference to traditional isolation techniques. Taken together, they found the highest diversity of bacterial taxa in soil; the diversity decreased with the distance the soil was from roots and the graft union, and the cane. Crown gall disease affected the makeup of the bacterial community only on graft unions with visible tumors. Compared to graft unions on healthy plants, galls possessed higher species richness with a more stable bacterial population structure over time and shared more bacterial species with the soil microbial population. Besides *A. vitis*, the most abundant bacteria in graft unions of diseased plants were *Pseudomonas* sp. and *Enterobacteriaceae* sp. However, the reasons for the higher abundance of bacteria belonging to these genera are unknown.

7 Disease Management

7.1 Early Diagnosis—Detection of the Pathogen

The use of healthy planting material in areas with no history of the crown gall is crucial because, once established in a vineyard, *A. vitis* may be impossible to eliminate. Therefore, analysis of the grapevine propagation material for the presence of the pathogen is important for disease control. Additionally, it is also important to test the soil prior to planting.

Despite its destructiveness, *A. vitis* and other species causing grapevine crown gall are not considered quarantine pathogens in many countries. They are commonly regarded as harmful, widespread pathogens that can reduce the value of propagation material (quality pathogens), and grapevine material exchanges are not subject to strict phytosanitary control. Moreover, there is a lack of standardized protocols for pathogen detection and identification.

Testing of infected plants is mainly based on pathogen isolation on semiselective and/or differential media, followed by analysis of isolated strains using biochemical tests and pathogenicity assays (Moore et al. 2001). Although isolation of *A. vitis* from tumor tissue of infected vines is relatively straightforward, low numbers of bacteria in asymptomatic samples and their irregular distribution limit the efficiency of detection methods. A method involving callusing of dormant cuttings and isolation of bacteria from callus tissue on semiselective medium was initially established and used by some laboratories (Lehoczky 1971). Because of its systemic nature, the pathogen can be isolated from grapevine vascular sap by flushing water or buffer through dormant grapevine cuttings using a vacuum pump (Tarbah and Goodman 1986; Bazzi et al. 1987). In this respect, a greater number of *A. vitis* cells has been recovered when dormant cuttings were frozen prior to vacuum flushing (Stover et al. 1997b). However, these procedures are time-consuming and laborious and are therefore not suitable for routine analysis of large numbers of samples.

PCR-based techniques are the method of choice for rapid pathogen detection in plant material and soil samples (Burr et al. 2017). Although a number of different PCR primers have been reported (Palacio-Bielsa et al. 2009a), remarkable genetic variations in *A. vitis* and its Ti plasmids may limit the specificity of described protocols (Kuzmanović et al. 2016). Moreover, most PCR protocols developed so far are generally suitable for testing bacteria from pure culture and may not be sensitive enough for pathogen detection in plant material.

Efficient DNA extraction followed by highly sensitive real-time PCR currently represents the most promising tool for early pathogen detection in plant and soil samples. So far, two real-time PCR protocols for detection of *A. vitis* in grapevine samples have been reported, both based on SYBR Green I dye chemistry (Bini et al. 2008a; Johnson et al. 2013). The protocol developed by Johnson et al. (2013) involves sample enrichment followed by efficient DNA extraction via magnetic

capture hybridization (MCH) and detection of tumorigenic strains by using *virD*2 gene-specific primers. This assay allowed detection of 10^1 CFU/ml and was able to detect *A. vitis* in asymptomatic grapevine material (Johnson et al. 2013; Johnson et al. 2016; Canik Orel et al. 2017).

7.2 Management Practices

Because *A. vitis* persists systemically in grapevines and there are no effective chemical controls for crown gall on grapes, the disease is especially challenging to manage. As covered in Sect. 7.4 in this chapter, biological control is encouraging as a future commercial control. The development of *A. vitis*-free propagation material is also a viable consideration which is covered in this chapter (see Sect. 7.3). However, the uncertainty of material being completely free of the pathogen and the potential sources of *A. vitis* in the environment that may infect "clean" vines will ultimately affect the effectiveness of this management strategy.

Currently in commercial vineyards, crown gall is managed primarily through the use of cultural practices that aim to reduce injuries to vines which may serve as infection sites for *A. vitis* (Moyer 2013). These practices include selecting vineyard sites that have well-drained soils and which are geographically located to have good air drainage and thus are not as prone to freezing temperatures. Crown gall is often most prevalent in low-lying regions of vineyards where standing water may accumulate and cold air may settle in frost pockets. Such wet soil conditions can affect late-season acclimation of vines, making them more prone to injury from sudden freezing temperatures. Where possible, management of irrigation water is another practice employed for slowing vine growth to facilitate hardening-off and making them more tolerant to winter freezes. Other factors that affect vine sensitivity to freezing temperatures and injury include excessive fertilization and over-cropping which can stimulate late growth of vines and affect the onset of dormancy.

In cold regions, the practice of hilling soil around grafts of vines in the fall is employed as a means to preserve vines and as a crown gall management tool. In this case, should winter freezes severely damage or kill vine trunks it becomes possible to train a new trunk the following year that is generated from the remaining living scion wood that was buried by soil and protected from the freeze. The training of multiple trunks per vine is often implemented as well. In this case, once a trunk becomes severely injured and diseased with crown gall, it can be removed and the remaining trunk or trunks will allow crop production as new trunks are trained.

There are no effective chemical controls for grape crown gall. Although antibacterial compounds, such as copper products, are lethal to *A. vitis*, topical treatments to vines have limited value considering the bacterium is systemic in the vine. Moreover, the plants affected by crown gall are genetically transformed and stay permanently infected.

Several papers have been published to demonstrate that grape species and varieties (scion and rootstock) differ in their susceptibility to infection by *A. vitis* (Ferreira and van Zyl 1986; Szegedi et al. 1989; Goodman et al. 1993; Stover et al. 1997a; Roh et al. 2003; Mahmoodzadeh et al. 2004; Jung et al. 2016). In general, *V. vinifera* is most susceptible, whereas *V. riparia* and hybrids of *Vitis* species used for scion and rootstock varieties are most resistant to infection, producing fewer and smaller galls. Among grape rootstocks, Courderc 3309 and Riparia Gloire are generally viewed as resistant. Differences within specific accessions have been noted among genotypes of these and other *Vitis* spp. Regardless, even those considered "resistant," such as *V. riparia*, may carry internal populations of *A. vitis* with the possibility of spreading it to more susceptible grapevines.

One study compared the relative levels of crown gall susceptibility of 43 *Vitis* genotypes following inoculation with a diverse set of *A. vitis* strains, followed by measuring gall size and the proportion of inoculation sites with galls (Stover et al. 1997a). None of the genotypes were immune and, depending on genotype, galls formed at 10–100% of the inoculated sites; the mean gall size ranged from 1.0 to 12 mm when averaged across *A. vitis* strains. Significant strain by genotype interactions was observed. For example, *Vitis amurensis* was most susceptible to a limited host range *A. vitis* strain, AG57. Commonly used rootstocks, 3309C, T5C, Riparia Gloire, and 101–14 Mgt, were among the most crown gall-resistant genotypes (Stover et al. 1997a). *Vitis flexuosa*, *Vitis piasezkii*, and *V. amurensis* that had been reported as resistant previously developed some galls. Rootstocks 110R, 420A, and Dogridge were categorized as highly susceptible. In another study, Szegedi et al. (1989) inoculated various grapevine cultivars with *A. vitis* strains belonging to different opine groups. Grapevine varieties were separated into four groups based on their susceptibility or resistance. Both host and bacterial factors likely contribute to the susceptibility/resistance of grapevine to *A. vitis*.

A limited amount of research has been done on the genetics of crown gall resistance in grape. Szegedi and Kozma (1984) tested seedlings from 27 hybrid families by inoculating with *A. vitis* At-1. Crown gall resistance in these crosses originated from *V. amurensis*. Their results showed a segregation of 1:1 following crosses of resistant and susceptible phenotypes and 3:1 (resistant to susceptible) following selfing of resistant parents. Therefore, a Mendelian-dominant inheritance of crown gall resistance to strain AT-1 was proposed. Subsequent research (Kuczmog et al. 2012) developed molecular markers linked to the *Rcg1* crown gall resistance gene from *V. amurensis*. This technology holds promise for future marker-assisted breeding of high-quality crown gall-resistant grape varieties.

Attempts to control grapevine crown gall by developing transgenic disease-resistant plants have been also made (Vidal et al. 2006; Krastanova et al. 2010; Galambos et al. 2013). In one such study, grapevine rootstock (*Vitis berlandieri* × *V. rupestris* cv. "Richter 110") plants were transformed with an oncogene-silencing transgene based on *iaaM* and *ipt* oncogene sequences (Galambos et al. 2013). However, oncogene silencing in grapevine is highly strain-specific and thus has limited effectiveness in disease control.

7.3 *Evaluation of Strategies for Producing* A. vitis*-Free Grapevines*

The discovery that *A. vitis* survives systemically and distributes randomly in grapevine propagation material stimulated research to determine if and how vines free of the pathogen could be produced. One approach used hot water treatments of dormant cuttings (Burr et al. 1989; Bazzi et al. 1991). In this case, dormant cuttings were submersed in water at 50 °C for 30 min. Although populations of the pathogen could be greatly reduced with this treatment, the bacteria were not eliminated (Burr et al. 1996). When temperatures higher than 50 °C were tested, the potential for increased bud mortality became apparent. Subsequently, Wample et al. (1991) showed that dormant cuttings collected in Washington State could withstand higher temperatures without enduring bud kill. Therefore, factors (temperatures, cutting hardiness, etc.) prior to the treatment appear to have a significant effect on grape bud tolerance to heat and should be explored further to determine the effectiveness of such treatments on internal *A. vitis* populations. Despite the fact that hot water treatment may result in a certain amount of bud kill, the practice (50 °C for 30 min) is still used in some regions of the world and felt to be of benefit for disease management.

Plant tissue culture has also been used to eliminate viral and bacterial pathogens from plants, including grapes (Dula et al. 2007; Cassells 2012). Explants from shoot tips or meristems are cultured in specific media to facilitate plant development (Sim and Golino 2010). Shoot tip culture was previously tested and shown to be effective for elimination of *A. vitis*. However, the detection method for evaluating its effectiveness was much less sensitive than is magnetic capture hybridization in conjunction with real-time PCR (MCH real-time PCR) technology currently in use (Johnson et al. 2013).

More recently, the MCH real-time PCR method was used in multiple experiments to assay tissue culture plants that were propagated from vines collected from a commercial vineyard that had severe crown gall. Shoot tips and meristems from the plants were assayed in 2013. Eighteen of the first 29 plants propagated tested positive for *A. vitis*. These included meristems, shoot tips with meristems extracted, and shoot tips with meristems (Johnson et al. 2016). When the same plants were cut back and regrown, only 4 of the 29 were positive. These results indicate the irregular distribution of *A. vitis* in the tissues as well as the uncertainty of a negative result.

Similar experiments were done in 2014 using 31 plants that were grown from cuttings taken from Riesling vines that were heavily infected with crown gall. In this case, *A. vitis* was not detected in any of the meristems for the two repetitions of the experiment (second repetition involved evaluating the regrowth of the plants after they were cut back following the first set of assays). For shoot tips with the meristems removed, four tested positive in the first repetition, but none were positive in the second (Johnson et al. 2016). Detecting *A. vitis* associated with grape tissue culture plants is not totally unexpected and has been reported previously

(Poppenberger et al. 2002). *A. vitis* can persist on surfaces of grape leaves, thus epiphytically on grapevines (Canik Orel et al. 2017). Additional research on the use of tissue culture is underway to determine if in fact vines free of the pathogen can be generated. However, from work concluded thus far it is apparent that shoot tips do not necessarily comprise tissue that is free of the pathogen. Another consideration is the sensitivity of the MCH real-time PCR assay, which was found to be about 10 bacterial cells per sample. Therefore, if populations lower than 10 cells are present in a sample they may not be detected. From this research, we conclude that tissue culture alone may not eliminate *A. vitis* from grape explants and that additional practices such as incorporation of effective antibiotics that do not inhibit plant growth and/or the use of heat therapy should be evaluated as a component of the tissue culture propagation.

7.4 Biological Control

Biological control of crown gall disease caused by tumorigenic agrobacteria represents a major success story in the field of plant pathology, resulting from the discovery of the nontumorigenic *R. rhizogenes* (former name *Agrobacterium radiobacter*) strain K84 by Kerr (1972). Control of crown gall by K84 has been implemented on different plant hosts in many regions worldwide. The primary mode of action of K84 is by antibiosis through the production of agrocin 84, which is encoded by the conjugative plasmid pAgK84 (Kim et al. 2006). Subsequently, a genetically modified form of K84 (strain K1026) was developed having a deleted fragment of a *tra* gene responsible for plasmid transfer, thereby preventing its transfer to pathogenic strains (Jones et al. 1988). Such transfer can result in strains becoming resistant to agrocin 84. Nevertheless, K84 and K1026 are not effective in preventing crown gall caused by some tumorigenic agrobacteria including *A. vitis*, the primary cause of the grapevine crown gall disease.

The impressive success of K84 and K1026 together with their ineffectiveness against *A. vitis* on grape led several researchers to search for bacterial strains that may control crown gall on grape. A number of strains have shown an ability to inhibit growth of *A vitis* and tumor formation, including *A. vitis* strains E26 (Yang et al. 2009) and VAR03-1 and ARK-1 (Kawaguchi et al. 2005, 2007, 2008a, 2014, 2017; Kawaguchi and Inoue 2012; Kawaguchi 2013, 2014, 2015; Saito et al. 2018), the *Agrobacterium* biovar 1 strain HLB2 (Pu and Goodman 1993), and the *R. rhizogenes* strain J73 (Webster et al. 1986). The potential of strains belonging to other bacterial genera to control grapevine crown gall has been also tested. These investigations included various endophytic strains isolated from grapevine (Bell et al. 1995; Ferrigo et al. 2017), various *Pseudomonas* spp. strains (Khmel et al. 1998; Eastwell et al. 2006; Biondi et al. 2009), and *Rahnella aquatilis* strain HX2 (Chen et al. 2007). These strains have shown a range of effects on pathogen growth as well as on various levels of disease suppression in experiments done under laboratory and greenhouse conditions. In addition, preliminary screens of bacteria

in vitro and on different indicator plants in the greenhouse have revealed bacterial strains belonging to different genera that have activity against *A. vitis* (Habbadi et al. 2017a). Interestingly, one fungal isolate belonging to the genus *Acremonium* and commercial biological control agents *Bacillus subtilis* SR63 and *Trichoderma asperellum* T1 were also effective against *A. vitis* (Ferrigo et al. 2017). A thorough review of bacterial strains that have been tested for activity against *A. vitis* has been published (Filo et al. 2013).

The nontumorigenic *A. vitis* strain F2/5, which was originally isolated from grape and shown to inhibit grape crown gall in South Africa (Staphorst et al. 1985), has been further studied by several laboratories (Burr and Reid 1994; Bazzi et al. 1999; Zäuner et al. 2006). Strain F2/5 inhibits tumor formation by diverse *A. vitis* strains on different grape varieties (Burr and Reid 1994). Although tumor formation by most strains of *A. vitis* is greatly inhibited by F2/5, a few strains appear to be unaffected (Staphorst et al. 1985; Burr and Reid 1994). As with other *A. vitis* strains, F2/5 also causes necrosis on grape roots (Burr et al. 1987a) which was recently shown to have a negative impact on graft wound healing and on plant growth but was not required for grape tumor inhibition (GTI; Hao et al. 2017). The molecular mechanism involved in necrosis is not fully understood but includes QS regulation and the involvement of specific polyketide and nonribosomal peptide synthases (Zheng et al. 2003; Hao et al. 2005; Zheng et al. 2012; Zheng and Burr 2013). Our current understanding of the biochemical pathways associated with necrosis and tumor inhibition by F2/5 shows overlap, but also that they are distinct processes (Zheng and Burr 2016).

Additional research has focused on characterizing GTI and identifying factors associated with the ability of F2/5 to inhibit crown gall. For example, although F2/5 inhibits *A. vitis* from causing tumors on grapevine, it does not block tumor formation on most other plants such as tobacco. An exception to this rule was reported to be inhibition of tumors on *Ricinus* (Zäuner et al. 2006). In addition, for F2/5 to inhibit tumor formation it must be applied to wounds at the same time or prior to the tumorigenic strain, and usually at cell numbers equal to or greater (Burr and Reid 1994). Through mutational analyses, it was shown that tumor inhibition was not associated with antibiosis even though an antibiosis phenotype could be demonstrated in vitro (Burr et al. 1997). Additionally, this study also demonstrated that tumor inhibition was not the result of competition for attachment sites on plant wounds. More recently, it was shown that F2/5 does not reduce populations of the tumorigenic strain at grape wound surfaces but, by an unknown mechanism, inhibits *A. vitis* from causing crown gall on grape (Kaewnum et al. 2013). This study also demonstrated that the genetic mechanism of gall inhibition is associated with at least two regulatory mechanisms that include QS and the involvement of *clp* protease genes.

Essential oils of *Origanum compactum* and *Thymus vulgaris* showed in vitro and in planta antibacterial activity against *A. vitis* (Habbadi et al. 2017b). Therefore, the use of essential oils of medicinal and aromatic plants could be a valuable alternative strategy in the control of grapevine crown gall. Moreover, in one recent study, a specific phage display-selected peptide displayed inhibitory effect toward *A. vitis*

polygalacturonase, which could be a promising approach in disease control strategy (Warren et al. 2016).

Acknowledgements The research of T. J. Burr presented in this paper was partially funded by USDA Federal Capacity Fund grants, USDA-APHIS National Clean Plant Network, and by the New York Wine and Grape Foundation. N. Kuzmanović was supported by the Georg Forster Fellowship for postdoctoral researchers from the Alexander von Humboldt-Foundation, Bonn, Germany. We kindly thank Dr. Ernő Szegedi for his critical reading of the manuscript.

References

Abdellatif E, Valentini F, Janse JD et al (2013) Occurrence of crown gall of the grapevine in Tunisia and characterization of Tunisian *Agrobacterium vitis* and *A. tumefaciens* strains. J Plant Pathol 95:115–126

Albiach MR, Lopez MM (1992) Plasmid heterogeneity in Spanish isolates of *Agrobacterium tumefaciens* from thirteen different hosts. Appl Environ Microbiol 58:2683–2687

Allardet-Servent A, Michaux-Charachon S, Jumas-Bilak E et al (1993) Presence of one linear and one circular chromosome in the *Agrobacterium tumefaciens* C58 genome. J Bacteriol 175:7869–7874

Al-Momani F, Albasheer S, Saadoun I (2006) Distribution of *Agrobacterium tumefaciens* biovars in Jordan and variation of virulence. Plant Pathol J 22:318–322

Argun N, Momol MT, Maden S et al (2002) Characterization of *Agrobacterium vitis* strains isolated from Turkish grape cultivars in the Central Anatolia region. Plant Dis 86:162–166

Bauer C, Schulz TF, Lorenz D et al (1994) Population dynamics of *Agrobacterium vitis* in two grapevine varieties during the vegetation period. Vitis 33:25–29

Bautista-Zapanta JN, Arafat HH, Tanaka K et al (2009) Variation of 16S-23S internally transcribed spacer sequence and intervening sequence in rDNA among the three major *Agrobacterium* species. Microbiol Res 164:604–612

Bazzi C, Piazza C, Burr TJ (1987) Detection of *Agrobacterium tumefaciens* in grapevine cuttings. EPPO Bulletin 17:105–112

Bazzi C, Stefani E, Gozzi R et al (1991) Hot-water treatment of dormant grape cuttings: its effects on *Agrobacterium tumefaciens* and on grafting and growth of vine. Vitis 30:177–187

Bazzi C, Alexandrova M, Stefani E et al (1999) Biological control of *Agrobacterium vitis* using non-tumorigenic agrobacteria. Vitis 38:31–35

Beauchamp CJ, Chilton WS, Dion P et al (1990) Fungal catabolism of crown gall opines. Appl Environ Microbiol 56:150–155

Beauchamp CJ, Kloepper JW, Lifshitz R et al (1991) Frequent occurrence of the ability to utilize octopine in rhizobacteria. Can J Microbiol 37:158–164

Beaulieu C, Coulombe LJ, Granger RL et al (1983) Characterization of opine-utilizing bacteria isolated from Quebec. Phytoprotect 64:61–68

Bell CR, Dickie GA, Chan JWYF (1995) Variable response of bacteria isolated from grapevine xylem to control grape crown gall disease in planta. Am J Enol Vitic 46:499–508

Bergeron J, Macleod RA, Dion P (1990) Specificity of octopine uptake by *Rhizobium* and *Pseudomonas* strains. Appl Environ Microbiol 56:1453–1458

Bien E, Lorenz D, Eichhorn K et al (1990) Isolation and characterization of *Agrobacterium tumefaciens* from the German vineregion Rheinpfalz. J Plant Dis Protect 97:313–322

Bini F, Geider K, Bazzi C (2008a) Detection of *Agrobacterium vitis* by PCR using novel *virD2* gene-specific primers that discriminate two subgroups. Eur J Plant Pathol 122:403–411

Bini F, Kuczmog A, Putnoky P et al (2008b) Novel pathogen-specific primers for the detection of *Agrobacterium vitis* and *Agrobacterium tumefaciens*. Vitis 47:181–189

Biondi E, Bini F, Anaclerio F et al (2009) Effect of bioagents and resistance inducers on grapevine crown gall. Phytopathol Mediter 48:379–384

Bishop AL, Katz BH, Burr TJ (1988) Infection of grapevines by soilborne *Agrobacterium tumefaciens* biovar 3 and population dynamics in host and nonhost rhizospheres. Phytopathol 78:945–948

Bishop AL, Burr TJ, Mittak VL et al (1989) A monoclonal antibody specific to *Agrobacterium tumefaciens* biovar 3 and its utilization for indexing grapevine propagation material. Phytopathol 79:995–998

Bonnard G, Tinland B, Paulus F et al (1989) Nucleotide sequence, evolutionary origin and biological role of a rearranged cytokinin gene isolated from a wide host range biotype III *Agrobacterium* strain. Mol Gen Genet 216:428–438

Bouzar H, Moore LW (1987) Isolation of different *Agrobacterium* biovars from a natural oak savanna and tallgrass prairie. Appl Environ Microbiol 53:717–721

Brencic A, Angert ER, Winans SC (2005) Unwounded plants elicit *Agrobacterium vir* gene induction and T-DNA transfer: transformed plant cells produce opines yet are tumour free. Mol Microbiol 57:1522–1531

Brisset M-N, Rodriguez-Palenzuela P, Burr TJ et al (1991) Attachment, chemotaxis, and multiplication of *Agrobacterium tumefaciens* biovar 1 and biovar 3 on grapevine and pea. Appl Environ Microbiol 57:3178–3182

Buchholz WG, Thomashow MF (1984a) Comparison of T-DNA oncogene complements of *Agrobacterium tumefaciens* tumor-inducing plasmids with limited and wide host ranges. J Bacteriol 160:319–326

Buchholz WG, Thomashow MF (1984b) Host range encoded by the *Agrobacterium tumefaciens* tumor-inducing plasmid pTiAg63 can be expanded by modification of its T-DNA oncogene complement. J Bacteriol 160:327–332

Burr TJ, Hurwitz B (1981) Occurrence of *Agrobacterium radiobacter* pv. *tumefaciens* (Smith & Townsend) biotype 3 on grapevine in New York State (Abstr.). Phytopathol 71:26

Burr TJ, Katz BH (1983) Isolation of *Agrobacterium tumefaciens* biovar 3 from grapevine galls and sap, and from vineyard soil. Phytopathol 73:163–165

Burr TJ, Katz BH (1984) Grapevine cuttings as potential sites of survival and means of dissemination of *Agrobacterium tumefaciens*. Plant Dis 68:976–978

Burr TJ, Otten L (1999) Crown gall of grape: Biology and disease management. Annu Rev Phytopathol 37:53–80

Burr TJ, Reid CL (1994) Biological control of grape crown gall with non-tumorigenic *Agrobacterium vitis* strain F2/5. Am J Enol Vitic 45:213–219

Burr TJ, Bishop AL, Katz BH et al (1987a) A root-specific decay of grapevine caused by *Agrobacterium tumefaciens* and *A. radiobacter* biovar 3. Phytopathol 77:1424–1427

Burr TJ, Katz BH, Bishop AL (1987b) Populations of *Agrobacterium* in vineyard and non vineyard soils and grape roots in vineyards and nurseries. Plant Dis 71:617–620

Burr TJ, Katz BH, Bishop AL et al (1988) Effect of shoot age and tip culture propagation of grapes on systemic infestations by *Agrobacterium tumefaciens* biovar 3. Am J Enol Vitic 39:67–70

Burr TJ, Ophel K, Kerr A (1989) Effect of hot water treatment on systemic *Agrobacterium tumefaciens* biovar 3 in dormant grape cuttings. Plant Dis 73:242–245

Burr TJ, Reid CL, Tagliatti E et al (1995) Survival and tumorigenicity of *Agrobacterium vitis* in living and decaying grape roots and canes in soil. Plant Dis 79:677–682

Burr TJ, Reid CL, Splittstoesser DF et al (1996) Effect of heat treatments on grape bud mortality and survival of *Agrobacterium vitis* in vitro and in dormant grape cuttings. Am J Enol Vitic 47:119–123

Burr TJ, Reid CL, Tagliati E et al (1997) Biological control of grape crown gall by strain F2/5 is not associated with agrocin production or competition for attachment sites on grape cells. Phytopathol 87:706–711

Burr TJ, Bazzi C, Süle S et al (1998) Crown gall of grape: Biology of *Agrobacterium vitis* and the development of disease control strategies. Plant Dis 82:1288–1297

Burr TJ, Reid CL, Adams CE et al (1999) Characterization of *Agrobacterium vitis* strains isolated from feral *Vitis riparia*. Plant Dis 83:102–107

Burr TJ, Johnson KL, Kaewnum S et al (2017) Detection of *Agrobacterium* spp. in grapevines (Chapter 43). In: Fatmi MB, Walcott RR, Schaad NW (eds) Detection of plant-pathogenic bacteria in seed and other planting material. Second edn. APS Press, Minneapolis, MN, USA

Canaday J, Gerad JC, Crouzet P et al (1992) Organization and functional analysis of three T-DNAs from the vitopine Ti plasmid pTiS4. Mol Gen Genet 235:292–303

Canfield ML, Moore LW (1991) Isolation and characterization of opine-utilizing strains of *Agrobacterium tumefaciens* and fluorescent strains of *Pseudomonas* spp. from rootstocks of *Malus*. Phytopathol 81:440–443

Canik Orel D, Karagoz A, Durmaz R et al (2016) Phenotypic and molecular characterization of *Rhizobium vitis* strains from vineyards in Turkey. Phytopathol Mediter 55

Canik Orel D, Reid CL, Fuchs M et al (2017) Identifying environmental sources of *Agrobacterium vitis* in vineyards and wild grapevines. Am J Enol Vitic 68:213–217

Cassells AC (2012) Pathogen and biological contamination management in plant tissue culture: phytopathogens, vitro pathogens, and vitro pests. Methods in molecular biology (Clifton, NJ) 877:57–80

Cavara F (1897) Tubercolosi della vite. Intorno alla eziologia de alcune malattie di piante coltivate. Stazoni Sperimentali Agrarie Italiane 30:483–487

Chang CC, Jayaswal RK, Chen CM et al (1989) Altered imino diacid synthesis and transcription in crown gall tumors with transposon Tn*5* insertions in the 3′ end of the octopine synthase gene. J Bacteriol 171:5922–5927

Chebil S, Fersi R, Chenenaoui S et al (2013a) First report of *Agrobacterium vitis* as causal agent of crown gall disease of grapevine in Tunisia. Plant Dis 97:836

Chebil S, Fersi R, Chenenaoui S et al (2013b) Occurrence of *Agrobacterium vitis* carrying two opine-type plasmids in Tunisian vineyards. J Plant Pathol Microbiol 4:175

Chen F, Guo YB, Wang JH et al (2007) Biological control of grape crown gall by *Rahnella aquatilis* HX2. Plant Dis 91:957–963

Chilton WS, Petit A, Chilton MD et al (2001) Structure and characterization of the crown gall opines heliopine, vitopine and rideopine. Phytochem 58:137–142

Christie PJ, Gordon JE (2014) The *Agrobacterium* Ti Plasmids. Microbiology spectrum 2. https://doi.org/10.1128/microbiolspec.plas-0010-2013

Clare BG, Kerr A, Jones DA (1990) Characteristics of the nopaline catabolic plasmid in *Agrobacterium* strains K84 and K1026 used for biological control of crown gall disease. Plasmid 23:126–137

Couturier MA, Bex F, Bergquist PL, Maas WK (1988) Identification and classification of bacterial plasmids. Microbiol rev 52(3):375–395 http://www.ncbi.nlm.nih.gov/pmc/articles/PMC373151/

Crouzet P, Otten L (1995) Sequence and mutational analysis of a tartrate utilization operon from *Agrobacterium vitis*. J Bacteriol 177:6518–6526

Dahl GA, Guyon P, Petit A et al (1983) Characterization of opine utilizing bacteria isolated from Quebec Canada. Phytoprotect 64:61–68

De Oliveira JR, Da Silva Romeiro R, De Souza Leäo Lacerda B (1994) Occurrence of *Agrobacterium tumefaciens* biovar 3 on grapevine in Brazil. J Phytopathol 140:363–366

Dessaux Y, Petit A, Tempe J (1993) Chemistry and biochemistry of opines, chemical mediators of parasitism. Phytochem 34:31–38

Dessaux Y, Petit A, Farrand SK et al (1998) Opines and opine-like molecules involved in plant-*Rhizobiaceae* interactions. In: Spaink HP, Kondorosi A, Hooykaas PJJ (eds) The *Rhizobiaceae*, molecular biology of model plant-associated bacteria. Kluwer Academic Publishers, Dordrecht, The Netherlands, pp 173–197

Dhanvantari BN (1983) Etiology of grape crown gall in Ontario. Can J Bot 61:2641–2646

Dula T, Kolber M, Lazar J et al (2007) Production of healthy grapevine propagating material: Pathogens and methods. Online publication http://oiv2007hu/documents/viticulture/Szegedi_OIV_2007_textpdf

Eastwell KC, Sholberg PL, Sayler RJ (2006) Characterizing potential bacterial biocontrol agents for suppression of *Rhizobium vitis*, causal agent of crown gall disease in grapevines. Crop Protect 25:1191–1200

Fabre E, Dunal MF (1853) Observations sur les maladies régnantes de la vigne. Bulletin de la Société d'Agriculture du Département de l'Hérault 40:46

Faist H, Keller A, Hentschel U et al (2016) Grapevine (*Vitis vinifera*) crown galls host distinct microbiota. Appl Environ Microbiol 82:5542–5552

Farrand SK (1998) Conjugal plasmids and their transfer. In: Spaink HP, Kondorosi A, Hooykaas PJJ (eds) The rhizobiaceae: molecular biology of model plant-associated bacteria. Springer, Netherlands, Dordrecht, pp 199–233

Faure D, Lang J (2014) Functions and regulation of quorum-sensing in *Agrobacterium tumefaciens*. Front Plant Sci 5

Ferreira JHS, van Zyl FGH (1986) Susceptibility of grapevine rootstocks to strains of *Agrobacterium tumefaciens* biovar 3. S Afr J Enol Vitic 7:101–104

Ferrigo D, Causin R, Raiola A (2017) Effect of potential biocontrol agents selected among grapevine endophytes and commercial products on crown gall disease. Biocontrol 62:821–833

Filo A, Sabbatini P, Sundin GW et al (2013) Grapevine crown gall suppression using biological control and genetic engineering: a review of recent research. Am J Enol Vitic 64:1–14

Galambos A, Zok A, Kuczmog A et al (2013) Silencing *Agrobacterium* oncogenes in transgenic grapevine results in strain-specific crown gall resistance. Plant Cell Rep 32:1751–1757

Gallie DR, Kado CI (1988) Minimal region necessary for autonomous replication of pTAR. J Bacteriol 170:3170–3176

Gallie DR, Zaitlin D, Perry KL et al (1984) Characterization of the replication and stability regions of *Agrobacterium tumefaciens* plasmid pTAR. J Bacteriol 157:739–745

Garcillan-Barcia MP, Francia MV, de la Cruz F (2009) The diversity of conjugative relaxases and its application in plasmid classification. FEMS Microbiol Rev 33:657–687

Garcillan-Barcia MP, Alvarado A, de la Cruz F (2011) Identification of bacterial plasmids based on mobility and plasmid population biology. FEMS Microbiol Rev 35:936–956

Gelvin SB (2012) Traversing the cell: *Agrobacterium* T-DNA's journey to the host genome. Front Plant Sci 3:52

Genov I, Atanassov I, Tsvetkov I et al (2006a) Isolation and characterization of *Agrobacterium* strains from grapevines in Bulgarian vineyards and wild grapes, *V. vinifera* ssp. *silvestris*. Vitis 45:97–101

Genov I, Atanassov I, Yordanov Y et al (2006b) Genetic diversity of *Agrobacterium vitis* strains, isolated from grapevines and wild grapes in Bulgaria, assessed by cleaved amplified polymorphic sequences analysis of 16S-23S rDNA. Vitis 45:125–130

Genov N, Llop P, Lopez MM et al (2015) Molecular and phenotypic characterization of *Agrobacterium* species from vineyards allows identification of typical *Agrobacterium vitis* and atypical biovar 1 strains. J Appl Microbiol 118:1465–1477

Gerard JC, Canaday J, Szegedi E et al (1992) Physical map of the vitopine Ti plasmid pTiS4. Plasmid 28:146–156

Gevers D, Cohan FM, Lawrence JG et al (2005) Re-evaluating prokaryotic species. Nat Rev. Micro 3:733–739

Gillings M, Ophel-Keller K (1995) Comparison of strains of *Agrobacterium vitis* from grapevine source areas in Australia. Austral Plant Pathol 24:29–37

Goodman RN, Grimm R, Frank M (1993) The influence of grape rootstocks on the crown gall infection process and on tumor development. Am J Enol Vitic 44:22–26

Guyon P, Petit A, Tempe J et al (1993) Transformed plants producing opines specifically promote growth of opine-degrading agrobacteria. Mol Plant-Microbe Interact 6:92–98

Haas JH, Zveibil A, Zutra D et al (1991) The presence of crown gall of grape incited by *Agrobacterium tumefaciens* biovar 3 in Israel. Phytoparasitica 19:311–318

Habbadi K, Benkirane R, Benbouazza A et al (2017a) Biological control of grapevine crown gall caused by *Allorhizobium vitis* using bacterial antagonists. Int J Science Res 6:1390–1397

Habbadi K, Meyer T, Vial L et al (2017b) Essential oils of *Origanum compactum* and *Thymus vulgaris* exert a protective effect against the phytopathogen *Allorhizobium vitis*. Environ Sci Pollut Res Int

Hao G, Burr TJ (2006) Regulation of long-chain *N*-acyl-homoserine lactones in *Agrobacterium vitis*. J Bacteriol 188:2173–2183

Hao G, Zhang H, Zheng D et al (2005) *luxR* homolog *avhR* in *Agrobacterium vitis* affects the development of a grape-specific necrosis and a tobacco hypersensitive response. J Bacteriol 187:185–192

Hao L, Kemmenoe DJ, Orel DC et al (2017) The impacts of tumorigenic and non-tumorigenic *Agrobacterium vitis* strains on graft strength and growth of grapevines. Plant Dis

Herlache TC (1999) Biochemical and molecular genetic investigations of the *Agrobacterium vitis*–grapevine interaction. Ph.D. Thesis, Cornell University, Ithaca, NY

Herlache TC, Hotchkiss AT, Burr TJ et al (1997) Characterization of the *Agrobacterium vitis pehA* gene and comparison of the encoded polygalacturonase with the homologous enzymes from *Erwinia carotovora* and *Ralstonia solanacearum*. Appl Environ Microbiol 63:338–346

Herlache TC, Zhang HS, Ried CL et al (2001) Mutations that affect *Agrobacterium vitis*-induced grape necrosis also alter its ability to cause a hypersensitive response on tobacco. Phytopathol 91:966–972

Hoekema A, de Pater BS, Fellinger AJ et al (1984) The limited host range of an *Agrobacterium tumefaciens* strain extended by a cytokinin gene from a wide host range T-region. EMBO 3:3043–3047

Hooykaas PJJ (2000) *Agrobacterium,* a natural metabolic engineer of plants. In: Verpoorte R, Alfermann AW (eds) Metabolic engineering of plant secondary metabolism. Springer, The Netherlands, Dordrecht, pp 51–67

Huang S, Long M, Fu G et al (2015) Characterization of a new pathovar of *Agrobacterium vitis* causing banana leaf blight in China. J Basic Microbiol 55:129–134

Huss B, Tinland B, Paulus F et al (1990) Functional analysis of a complex oncogene arrangement in biotype III *Agrobacterium tumefacien*s strains. Plant Mol Biol 14:173–186

Ignatov AN, Khodykina MV, Vinogradova SV et al (2016) First report of *Agrobacterium vitis* causing crown galls of wine grape in Russia. Plant Dis 100:853

Irelan NA, Meredith CP (1996) Genetic analysis of *Agrobacterium tumefaciens* and *A. vitis* using randomly amplified polymorphic DNA. Am J Enol Vitic 47:145–151

Johnson KL, Zheng D, Kaewnum S et al (2013) Development of a magnetic capture hybridization real-time PCR assay for detection of tumorigenic *Agrobacterium vitis* in grapevines. Phytopathol 103:633–640

Johnson KL, Cronin H, Reid CL et al (2016) Distribution of *Agrobacterium vitis* in grapevines and its relevance to pathogen elimination. Plant Dis 100:791–796

Jones DA, Ryder MH, Clare BG et al (1988) Construction of a Tra^- deletion mutant of pAgK84 to safeguard the biological control of crown gall. Mol Gen Genet 212:207–214

Jumas-Bilak E, Michaux-Charachon S, Bourg G et al (1998) Unconventional genomic organization in the alpha subgroup of the Proteobacteria. J Bacteriol 180:2749–2755

Jung S-M, Hur Y-Y, Preece JE et al (2016) Profiling of disease-related metabolites in grapevine internode tissues infected with *Agrobacterium vitis*. Plant Pathol J 32:489–499

Kado CI (1998) Origin and evolution of plasmids. Antonie Van Leeuwenhoek 73:117–126

Kaewnum S, Zheng D, Reid CL et al (2013) A host-specific biological control of grape crown gall by *Agrobacterium vitis* strain F2/5: its regulation and population dynamics. Phytopathol 103:427–435

Karimi M, Van Montagu M, Gheysen G (2000) Nematodes as vectors to introduce *Agrobacterium* into plant roots. Mol Plant Pathol 1:383–387

Kawaguchi A (2013) Biological control of crown gall on grapevine and root colonization by nonpathogenic *Rhizobium vitis* strain ARK-1. Microbes Environ 28:306–311

Kawaguchi A (2014) Reduction in pathogen populations at grapevine wound sites is associated with the mechanism underlying the biological control of crown gall by *Rhizobium vitis* strain ARK-1. Microbes Environ 29:296–302

Kawaguchi A (2015) Biological control agent *Agrobacterium vitis* strain ARK-1 suppresses expression of the *virD2* and *virE2* genes in tumorigenic *A. vitis*. Eur J Plant Pathol 143:789–799

Kawaguchi A, Inoue K (2009) Grapevine crown gall caused by *Rhizobium radiobacter* (Ti) in Japan. J Gen Plant Pathol 75:205–212

Kawaguchi A, Inoue K (2012) New antagonistic strains of non-pathogenic *Agrobacterium vitis* to control grapevine crown gall. J Phytopathol 160:509–518

Kawaguchi A, Inoue K, Nasu H (2005) Inhibition of crown gall formation by *Agrobacterium radiobacter* biovar 3 strains isolated from grapevine. J Gen Plant Pathol 71:422–430

Kawaguchi A, Inoue K, Nasu H (2007) Biological control of grapevine crown gall by nonpathogenic *Agrobacterium vitis* strain VAR03-1. J Gen Plant Pathol 73:133–138

Kawaguchi A, Inoue K, Ichinose Y (2008a) Biological control of crown gall of grapevine, rose, and tomato by nonpathogenic *Agrobacterium vitis* strain VAR03-1. Phytopathol 98:1218–1225

Kawaguchi A, Sawada H, Ichinose Y (2008b) Phylogenetic and serological analyses reveal genetic diversity of *Agrobacterium vitis* strains in Japan. Plant Pathol 57:747–753

Kawaguchi A, Inoue K, Tanina K (2014) Evaluation of the nonpathogenic *Agrobacterium vitis* strain ARK-1 for crown gall control in diverse plant species. Plant Dis 99:409–414

Kawaguchi A, Inoue K, Tanina K et al (2017) Biological control for grapevine crown gall using nonpathogenic *Rhizobium vitis* strain ARK-1. Proc Jpn Acad Ser B Phys Biol Sci 93:547–560

Kerr A (1972) Biological control of crown gall: seed inoculation. J Appl Bacteriol 35:493–497

Kerr A, Panagopoulos CG (1977) Biotypes of *Agrobacterium radiobacter* var. *tumefaciens* and their biological control. J Phytopathol 90:172–179

Khmel IA, Sorokina TA, Lemanova NB et al (1998) Biological control of crown gall in grapevine and raspberry by two *Pseudomonas* spp. with a wide spectrum of antagonistic activity. Biocontrol Sci Technol 8:45–57

Kim JG, Park BK, Kim SU et al (2006) Bases of biocontrol: sequence predicts synthesis and mode of action of agrocin 84, the Trojan horse antibiotic that controls crown gall. Proc Natl Acad Sci USA 103:8846–8851

Knauf VC, Panagopoulos CG, Nester EW (1982) Genetic factors controlling the host range of *Agrobacterium tumefaciens*. Phytopathol 72:1545–1549

Knauf VC, Panagopoulos CG, Nester EW (1983) Comparison of Ti plasmids from three different biotypes of *Agrobacterium tumefaciens* isolated from grapevines. J Bacteriol 153:1535–1542

Krastanova SV, Balaji V, Holden MR et al (2010) Resistance to crown gall disease in transgenic grapevine rootstocks containing truncated *virE2* of *Agrobacterium*. Transgenic Res 19:949–958

Kuczmog A, Galambos A, Horvath S et al (2012) Mapping of crown gall resistance locus *Rcg1* in grapevine. Theor Appl Genet 125:1565–1574

Kuzmanović N, Ivanović M, Prokić A et al (2014) Characterization and phylogenetic diversity of *Agrobacterium vitis* from Serbia based on sequence analysis of 16S-23S rRNA internal transcribed spacer (ITS) region. Eur J Plant Pathol 140:757–768

Kuzmanović N, Biondi E, Bertaccini A et al (2015) Genetic relatedness and recombination analysis of *Allorhizobium vitis* strains associated with grapevine crown gall outbreaks in Europe. J Appl Microbiol 119:786–796

Kuzmanović N, Biondi E, Ivanović M et al (2016) Evaluation of different PCR primers for identification of tumorigenic bacteria associated with grapevine crown gall. J Plant Pathol 98:311–319

Kuzmanović N, Puławska J, Smalla K et al (2018) *Agrobacterium rosae* sp. nov., isolated from galls on different agricultural crops. Syst Appl Microbiol

Lehoczky J (1968) Spread of *Agrobacterium tumefaciens* in the vessels of the grapevine, after natural infection. J Phytopathol 63:239–246

Lehoczky J (1971) Further evidences concerning the systemic spreading of *Agrobacterium tumefaciens* in the vascular system of grapevines. Vitis 10:215–221

Lehoczky J (1978) Root-system of the grapevine as a reservoir of *Agrobacterium tumefaciens* cells. In: 4th international conference on plant pathogenic bacteria, Angers, France, pp 239–243

Lehoczky J (1989) Inoculation experiment on trellised grapevines with *Agrobacterium tumefaciens* to study the process of crown gall disease. Acta Phytopathol Entomol Hung 24:283–291

Li Y, Gronquist MR, Hao G et al (2005) Chromosome and plasmid-encoded *N*-acyl homoserine lactones produced by *Agrobacterium vitis* wildtype and mutants that differ in their interactions with grape and tobacco. Physiol Mol Plant Pathol 67:284–290

Lim SH, Kim JG, Kang HW (2009) Novel SCAR primers for specific and sensitive detection of *Agrobacterium vitis* strains. Microbiol Res 164:451–460

López MM, Gorris MT, Montojo AM (1988) Opine utilization by Spanish isolates of *Agrobacterium tumefaciens*. Plant Pathol 37:565–572

Loubser JT (1978) Identification of *Agrobacterium tumefaciens* biotype 3 on grapevine in South Africa. Plant Dis Rep 62:730–731

Lowe N, Gan HM, Chakravartty V et al (2009) Quorum-sensing signal production by *Agrobacterium vitis* strains and their tumor-inducing and tartrate-catabolic plasmids. FEMS Microbiol Lett 296:102–109

Ma DQ, Yanofsky MF, Gordon MP et al (1987) Characterization of *Agrobacterium tumefaciens* strains isolated from grapevine tumors in China. Appl Environ Microbiol 53:1338–1343

Mahmoodzadeh H, Nazemieh A, Majidi I et al (2004) Evaluation of crown gall resistance in *Vitis vinifera* and hybrids of *Vitis* spp. Vitis 43:75–79

Mansouri H, Petit A, Oger P et al (2002) Engineered rhizosphere: the trophic bias generated by opine-producing plants is independent of the opine type, the soil origin, and the plant species. Appl Environ Microbiol 68:2562–2566

McCullen CA, Binns AN (2006) *Agrobacterium tumefaciens* and plant cell interactions and activities required for interkingdom macromolecular transfer. Annu Rev Cell Dev Biol 22:101–127

McGuire RG, Rodriguez-Palenzuela P, Collmer A et al (1991) Polygalacturonase production by *Agrobacterium tumefaciens* biovar 3. Appl Environ Microbiol 57:660–664

Meier-Kolthoff JP, Auch AF, Klenk H-P et al (2013) Genome sequence-based species delimitation with confidence intervals and improved distance functions. BMC Bioinform 14:1–14

Merlo DJ, Nester EW (1977) Plasmids in avirulent strains of *Agrobacterium*. J Bacteriol 129:76–80

Mohammadi M, Fatehi-Paykani R (1999) Phenotypical characterization of Iranian isolates of *Agrobacterium vitis*, the causal agent of crown gall disease of grapevine. Vitis 38:115–121

Momol EA, Burr TJ, Reid CL et al (1998) Genetic diversity of *Agrobacterium vitis* as determined by DNA fingerprints of the 5′-end of the 23S rRNA gene and random amplified polymorphic DNA. J Appl Microbiol 85:685–692

Moore LW, Chilton WS, Canfield ML (1997) Diversity of opines and opine-catabolizing bacteria isolated from naturally occurring crown gall tumors. Appl Environ Microbiol 63:201–207

Moore LW, Bouzar H, Burr TJ (2001) *Agrobacterium*. In: Schaad NW, Jones JB, Chun W (eds) Laboratory guide for identification of plant pathogenic bacteria, 3rd edn. APS Press, St Paul, Minnesota, pp 17–35

Mousavi SA, Osterman J, Wahlberg N et al (2014) Phylogeny of the *Rhizobium-Allorhizobium-Agrobacterium* clade supports the delineation of *Neorhizobium* gen. nov. Syst Appl Microbiol 37:208–215

Mousavi SA, Willems A, Nesme X et al (2015) Revised phylogeny of *Rhizobiaceae*: proposal of the delineation of *Pararhizobium* gen. nov., and 13 new species combinations. Syst Appl Microbiol 38:84–90

Moyer M (2013) Grapevine crown gall, disease management white paper. http://winewsuedu/wp-content/uploads/sites/66/2013/05/2013-CrownGallWhiteSheetpdf

Nascimento T, Oliveira H, Schulz T (1999) Identification of *Agrobacterium vitis* by amplification of molecular markers. Petria 9:81–84

Nautiyal CS, Dion P (1990) Characterization of the opine-utilizing microflora associated with samples of soil and plants. Appl Environ Microbiol 56:2576–2579
Nautiyal CS, Dion P, Chilton WS (1991) Mannopine and mannopinic acid as substrates for *Arthrobacter* sp. strain MBA209 and *Pseudomonas putida* NA513. J Bacteriol 173:2833–2841
Oger P, Farrand SK (2002) Two opines control conjugal transfer of an *Agrobacterium* plasmid by regulating expression of separate copies of the quorum-sensing activator gene *traR*. J Bacteriol 184:1121–1131
Oger P, Petit A, Dessaux Y (1997) Genetically engineered plants producing opines alter their biological environment. Nat Biotechnol 15:369–372
Ophel K, Kerr A (1990) *Agrobacterium vitis* sp. nov. for strains of *Agrobacterium* biovar 3 from grapevines. Int J Syst Bacteriol 40:236–241
Ophel K, Burr TJ, Magarey PA et al (1988) Detection of *Agrobacterium tumefaciens* biovar 3 in South Australian grapevine propagation material. Austral Plant Pathol 17:61–66
Ormeño-Orrillo E, Servín-Garcidueñas LE, Rogel MA et al (2015) Taxonomy of rhizobia and agrobacteria from the *Rhizobiaceae* family in light of genomics. Syst Appl Microbiol 38:287–291
Otten L, De Ruffray P (1994) *Agrobacterium vitis* nopaline Ti plasmid pTiAB4: relationship to other Ti plasmids and T-DNA structure. Mol Gen Genet 245:493–505
Otten L, Canaday J, Gerard JC et al (1992) Evolution of agrobacteria and their Ti plasmids—a review. Mol Plant Microbe Interact 5:279–287
Otten L, Gerard JC, De Ruffray P (1993) The Ti plasmid from the wide host range *Agrobacterium vitis* strain Tm4: Map and homology with other Ti plasmids. Plasmid 29:154–159
Otten L, Crouzet P, Salomone JY et al (1995) *Agrobacterium vitis* strain AB3 harbors two independent tartrate utilization systems, one of which is encoded by the Ti plasmid. Mol Plant Microbe Interact 8:138–146
Otten L, De Ruffray P, de Lajudie P et al (1996a) Sequence and characterisation of a ribosomal RNA operon from *Agrobacterium vitis*. Mol Gen Genet 251:99–107
Otten L, de Ruffray P, Momol EA et al (1996b) Phylogenetic relationship between *Agrobacterium vitis* isolates and their Ti plasmids. Mol Plant Microbe Interact 9:782–786
Otten L, Burr T, Szegedi E (2008) *Agrobacterium*: a disease-causing bacterium. In: Tzfira T, Citovsky V (eds) *Agrobacterium*: from biology to biotechnology. Springer, New York, USA, pp 1–46
Palacio-Bielsa A, Cambra MA, López MM (2009a) PCR detection and identification of plant-pathogenic bacteria: updated review of protocols (1989–2007). J Plant Pathol 91:249–297
Palacio-Bielsa A, González-Abolafio R, Álvarez B et al (2009b) Chromosomal and Ti plasmid characterization of tumorigenic strains of three *Agrobacterium* species isolated from grapevine tumours. Plant Pathol 58:584–593
Panagopoulos CG, Psallidas PG (1973) Characteristics of Greek isolates of *Agrobacterium tumefaciens* (E.F. Smith & Townsend). Conn J Appl Bacteriol 36:233–240
Panagopoulos CG, Psallidas PG, Alivizatos AS (1978) Studies on biotype 3 of *Agrobacterium radiobacter* var. *tumefaciens*. In: 4th international conference on plant pathogenic bacteria, Angers, France, pp 221–228
Pappas KM, Cevallos MA (2011) Plasmids of the *Rhizobiaceae* and their role in interbacterial and transkingdom interactions. In: Witzany G (ed) Biocommunication in soil microorganisms. Springer, Berlin Heidelberg, Berlin, Heidelberg, pp 295–337
Paulus F, Huss B, Bonnard G et al (1989a) Molecular systematics of biotype III Ti plasmids of *Agrobacterium tumefaciens*. Mol Plant-Microbe Interact 2:64–74
Paulus F, Ride M, Otten L (1989b) Distribution of two *Agrobacterium tumefaciens* insertion elements in natural isolates: evidence for stable association between Ti plasmids and their bacterial hosts. Mol Gen Genet 219:145–152
Paulus F, Canaday J, Otten L (1991a) Limited host range Ti plasmids: recent origin from wide host range Ti plasmids and involvement of a novel IS element, IS868. Mol Plant Microbe Interact 4:190–197

Paulus F, Canaday J, Vincent F et al (1991b) Sequence of the *iaa* and *ipt* region of different *Agrobacterium tumefaciens* biotype III octopine strains: reconstruction of octopine Ti plasmid evolution. Plant Mol Biol 16:601–614

Perović T, Renzi M, Mazzaglia A et al (2015) First report of *Agrobacterium tumefaciens* as a causal agent of crown gall on grapevine in Montenegro. Plant Dis 100:515

Perry KL, Kado CI (1982) Characteristics of Ti plasmids from broad-host-range and ecologically specific biotype 2 and 3 strains of *Agrobacterium tumefaciens*. J Bacteriol 151:343–350

Pinto UM, Pappas KM, Winans SC (2012) The ABCs of plasmid replication and segregation. Nat Rev Microbiol 10:755–765

Pitzschke A, Hirt H (2010) New insights into an old story: *Agrobacterium*-induced tumour formation in plants by plant transformation. EMBO J 29:1021–1032

Poppenberger B, Leonhardt W, Redl H (2002) Latent persistence of *Agrobacterium vitis* in micropropagated *Vitis vinifera*. Vitis 41:113–114

Pu X-A, Goodman RN (1993) Effects of fumigation and biological control on infection of indexed crown gall free grape plants. Am J Enol Vitic 44:241–248

Puławska J, Willems A, De Meyer SE et al (2012) *Rhizobium nepotum* sp. nov. isolated from tumors on different plant species. Syst Appl Microbiol 35:215–220

Puławska J, Kuzmanović N, Willems A et al (2016) *Pararhizobium polonicum* sp. nov. isolated from tumors on stone fruit rootstocks. Syst Appl Microbiol 39:164–169

Ramírez-Bahena MH, Vial L, Lassalle F et al (2014) Single acquisition of protelomerase gave rise to speciation of a large and diverse clade within the *Agrobacterium*/*Rhizobium* supercluster characterized by the presence of a linear chromid. Mol Phylogenet Evol 73:202–207

Richter M, Rossello-Mora R (2009) Shifting the genomic gold standard for the prokaryotic species definition. Proc Natl Acad Sci USA 106:19126–19131

Ridé M, Ridé S, Petit A et al (2000) Characterization of plasmid-borne and chromosome-encoded traits of *Agrobacterium* biovar 1, 2, and 3 strains from France. Appl Environ Microbiol 66:1818–1825

Rodriguez-Palenzuela P, Burr TJ, Collmer A (1991) Polygalacturonase is a virulence factor in *Agrobacterium tumefaciens* biovar 3. J Bacteriol 173:6547–6552

Roh JH, Yun HK, Park KS et al (2003) In vivo evaluation of resistance of grape varieties to crown gall disease. Plant Pathol J 19:235–238

Rouhrazi K, Rahimian H (2012a) Characterization of Iranian grapevine isolates of *Rhizobium* (*Agrobacterium*) spp. J Plant Pathol 94:555–560

Rouhrazi K, Rahimian H (2012b) Genetic diversity of Iranian *Agrobacterium* strains from grapevine. Ann Microbiol 62:1661–1667

Saito K, Watanabe M, Matsui H et al (2018) Characterization of the suppressive effects of the biological control strain VAR03-1 of *Rhizobium vitis* on the virulence of tumorigenic *R. vitis*. J Gen Plant Pathol 84:58–64

Salomone JY, Otten L (1999) Structure and function of a conserved DNA region coding for tartrate utilization in *Agrobacterium vitis*. FEMS Microbiol Lett 174:333–337

Salomone JY, Crouzet P, De Ruffray P et al (1996) Characterization and distribution of tartrate utilization genes in the grapevine pathogen *Agrobacterium vitis*. Mol Plant Microbe Interact 9:401–408

Salomone JY, Szegedi E, Cobanov P et al (1998) Tartrate utilization genes promote growth of *Agrobacterium* spp. on grapevine. Mol Plant Microbe Interact 11:836–838

Savka MA, Farrand SK (1997) Modification of rhizobacterial populations by engineering bacterium utilization of a novel plant-produced resource. Nat Biotechnol 15:363–368

Savka MA, Black RC, Binns AN et al (1996) Translocation and exudation of tumor metabolites in crown galled plants. Mol Plant Microbe Interact 9:310–313

Sawada H, Ieki H (1992) Crown gall of kiwi caused by *Agrobacterium tumefaciens* in Japan. Plant Dis 76:212

Sawada H, Ieki H, Takikawa Y (1990) Identification of grapevine crown gall bacteria isolated in Japan. Jap J Phytopathol 56:199–206

Schroth MN, McCain AH, Foott JH et al (1988) Reduction in yield and vigor of grapevine caused by crown gall disease. Plant Dis 72:241–246
Schulz TF, Bauer C, Lorenz D et al (1993) Studies on the evolution of *Agrobacterium vitis* as based on genomic fingerprinting and IS element analysis. Syst Appl Microbiol 16:322–329
Sciaky D, Montoya AL, Chilton M-D (1978) Fingerprints of *Agrobacterium* Ti plasmids. Plasmid 1:238–253
Sim ST, Golino D (2010) Micro- vs. macroshoot tip culture therapy for disease elimination in grapevines. FPS Grape Program Newsl Foundation Plant Services, University of California, Davis
Slater SC, Goldman BS, Goodner B et al (2009) Genome sequences of three *Agrobacterium* biovars help elucidate the evolution of multichromosome genomes in bacteria. J Bacteriol 191:2501–2511
Slater S, Setubal JC, Goodner B et al (2013) Reconciliation of sequence data and updated annotation of the genome of *Agrobacterium tumefaciens* C58, and distribution of a linear chromosome in the genus *Agrobacterium*. Appl Environ Microbiol 79:1414–1417
Smillie C, Garcillan-Barcia MP, Francia MV et al (2010) Mobility of plasmids. Microbiol Mol Biol Rev 74:434–452
Spencer PA, Tanaka A, Towers GHN (1990) An *Agrobacterium* signal compound from grapevine cultivars. Phytochem 29:3785–3788
Staphorst JL, van Zyl FGH, Strijdom BW et al (1985) Agrocin-producing pathogenic and nonpathogenic biotype-3 strains of *Agrobacterium tumefaciens* active against biotype-3 pathogens. Curr Microbiol 12:45–52
Stewart E, Wenner N, Long L et al (2013) Crown gall of grape: understanding the disease, prevention and management. Online Publication. http://plantpath.psu.edu/research/labs/grapes/publications/disease-fact-sheets/Crown-gall-grape.pdf
Stover EW, Swartz HJ, Burr TJ (1997a) Crown gall formation in a diverse collection of *Vitis* genotypes inoculated with *Agrobacterium vitis*. Am J Enol Vitic 48:26–32
Stover EW, Swartz HJ, Burr TJ (1997b) Endophytic *Agrobacterium* in crown gall-resistant and -susceptible *Vitis* genotypes. Vitis 36:21–26
Süle S (1978) Biotypes of *Agrobacterium tumefaciens* in Hungary. J Appl Bacteriol 44:207–213
Süle S (1986) Survival of *Agrobacterium tumefaciens* in *Berlandieri* x *Riparia* grapevine rootstock. Acta Phytopathol Entomol Hung 21:203–206
Süle S, Kado CI (1980) Agrocin resistance in virulent derivatives of *Agrobacterium tumefaciens* harboring the pTi plasmid. Physiol Plant Pathol 17:347–356
Süle S, Lehoczky J, Jenser G et al (1995) Infection of grapevine roots by *Agrobacterium vitis* and *Meloidogyne hapla*. J Phytopathol 143:169–171
Süle S, Cursino L, Zheng D et al (2009) Surface motility and associated surfactant production in *Agrobacterium vitis*. Lett Appl Microbiol 49:596–601
Suzuki K, Tanaka K, Yamamoto S et al (2009) Ti and Ri Plasmids. In: Schwartz E (ed) Microbial megaplasmids, vol 11. Microbiology monographs. Springer, Berlin, Heidelberg, pp 133–147
Szegedi E (1985) Host range and specific L(+)tartrate utilization of biotype 3 of *Agrobacterium tumefaciens*. Acta Phytopathol Acad Sci Hung 20:17–22
Szegedi E (2003) Opines in naturally infected grapevine crown gall tumors. Vitis 42:39–41
Szegedi E, Kozma P (1984) Studies on the inheritance of resistance to crown gall disease of grapevine. Vitis 23:121–126
Szegedi E, Otten L (1998) Incompatibility properties of tartrate utilization plasmids derived from *Agrobacterium vitis* strains. Plasmid 39:35–40
Szegedi E, Czakó M, Otten L et al (1988) Opines in crown gall tumours induced by biotype 3 isolates of *Agrobacterium tumefaciens*. Physiol Mol Plant Pathol 32:237–247
Szegedi E, Korbuly J, Otten L (1989) Types of resistance of grapevine varieties to isolates of *Agrobacterium tumefaciens* biotype 3. Physiol Mol Plant Pathol 35:35–43
Szegedi E, Otten L, Czakó M (1992) Diverse types of tartrate plasmids in *Agrobacterium tumefaciens* biotype III strains. Mol Plant Microbe Interact 5:435–438

Szegedi E, Czakó M, Otten L (1996) Further evidence that the vitopine-type pTi's of *Agrobacterium vitis* represent a novel group of Ti plasmids. Mol Plant Microbe Interact 9:139–143

Szegedi E, Süle S, Burr TJ (1999) *Agrobacterium vitis* strain F2/5 contains tartrate and octopine utilization plasmids which do not encode functions for tumor inhibition on grapevine. J Phytopathol 147:554–558

Szegedi E, Bottka S, Mikulas J et al (2005) Characterization of *Agrobacterium tumefaciens* strains isolated from grapevine. Vitis 44:49–54

Szegedi E, Deak T, Forgács I et al (2014) *Agrobacterium vitis* strains lack tumorigenic ability on in vitro grown grapevine stem segments. Vitis 53:147–154

Tanaka K, Urbanczyk H, Matsui H et al (2006) Construction of physical map and mapping of chromosomal virulence genes of the biovar 3 *Agrobacterium* (*Rhizobium vitis*) strain K-Ag-1. Genes Genet Syst 81:373–380

Tarbah FA, Goodman RN (1986) Rapid detection of *Agrobacterium tumefaciens* in grapevine propagating material and the basis for an efficient indexing system. Plant Dis 70(70):566–568

Tarbah FA, Goodman RN (1987) Systemic spread of *Agrobacterium tumefaciens* biovar 3 in the vascular system of grapes. Phytopathol 77:915–920

Thies KL, Griffin DE, Graves CH Jr et al (1991) Characterization of *Agrobacterium* isolates from muscadine grape. Plant Dis 75:634–637

Thomashow MF, Panagopoulos CG, Gordon MP et al (1980) Host range of *Agrobacterium tumefaciens* is determined by the Ti plasmid. Nature 283:794–796

Thomashow MF, Knauf VC, Nester EW (1981) Relationship between the limited and wide host range octopine-type Ti plasmids of *Agrobacterium tumefaciens*. J Bacteriol 146:484–493

Tolba IH, Zaki MF (2011) Characterization of *Agrobacterium vitis* isolates obtained from galled grapevine plants in Egypt. Annals of Agricultural Sciences 56:113–119

Tremblay G, Gagliardo R, Chilton WS et al (1987) Diversity among opine-utilizing bacteria: identification of coryneform isolates. Appl Environ Microbiol 53:1519–1524

van Nuenen M, de Ruffray P, Otten L (1993) Rapid divergence of *Agrobacterium vitis* octopine-cucumopine Ti plasmids from a recent common ancestor. Mol Gen Genet 240:49–57

Vidal JR, Kikkert JR, Malnoy MA et al (2006) Evaluation of transgenic 'Chardonnay' (*Vitis vinifera*) containing magainin genes for resistance to crown gall and powdery mildew. Transgenic Res 15:69–82

Wabiko H, Kagaya M, Sano H (1990) Various nopaline catabolism genes located outside the Ti-plasmids in *Agrobacterium tumefaciens* strains. Microbiol 136:97–103

Wample RL, Bary A, Burr TJ (1991) Heat tolerance of dormant *Vitis vinifera* cuttings. Am J Enol Vitic 42:67–72

Warren JG, Kasun GW, Leonard T et al (2016) A phage display-selected peptide inhibitor of *Agrobacterium vitis* polygalacturonase. Mol Plant Pathol 17:480–486

Watanabe S, Sueda R, Fukumori F et al (2015) Characterization of flavin-containing opine dehydrogenase from bacteria. PLoS ONE 10:e0138434

Webster J, Dos Santos M, Thomson JA (1986) Agrocin-producing *Agrobacterium tumefaciens* strain active against grapevine isolates. Appl Environ Microbiol 52:217–219

Wetzel ME, Kim KS, Miller M et al (2014) Quorum-dependent mannopine-inducible conjugative transfer of an *Agrobacterium* opine-catabolic plasmid. J Bacteriol 196:1031–1044

Wetzel ME, Olsen GJ, Chakravartty V et al (2015) The *repABC* plasmids with quorum-regulated transfer systems in members of the rhizobiales divide into two structurally and separately evolving groups. Genome Biol Evol 7:3337–3357

White CE, Winans SC (2007) Cell–cell communication in the plant pathogen *Agrobacterium tumefaciens*. Philos Trans. R Soc B: Biol Sci 362:1135–1148

Yang YL, Li JY, Wang JH et al (2009) Mutations affecting chemotaxis of *Agrobacterium vitis* strain E26 also alter attachment to grapevine roots and biocontrol of crown gall disease. Plant Pathol 58:594–605

Yanofsky M, Lowe B, Montoya A et al (1985a) Molecular and genetic analysis of factors controlling host range in *Agrobacterium tumefaciens*. Mol Gen Genet 201:237–246

Yanofsky M, Montoya A, Knauf V et al (1985b) Limited-host-range plasmid of *Agrobacterium tumefaciens*: molecular and genetic analyses of transferred DNA. J Bacteriol 163:341–348

Young JM, Kerr A, Sawada H (2005) Genus *Agrobacterium* Conn 1942, 359^{AL}. In: Brenner DJ, Krieg NR, Staley JT, Garrity GM (eds) Bergey's manual of systematics of archaea and bacteria. The Proteobacteria, vol 2C, 2ed edn. Springer, New York, pp 340–345

Zäuner S, Crespan JE, Burr TJ et al (2006) Inhibition of crown gall induction by Agrobacterium vitis strain F2/5 in grapevine and Ricinus. Vitis 45:131–139

Zheng D, Burr TJ (2013) An Sfp-type PPTase and associated polyketide and nonribosomal peptide synthases in *Agrobacterium vitis* are essential for induction of tobacco hypersensitive response and grape necrosis. Mol Plant Microbe Interact 26:812–822

Zheng D, Burr TJ (2016) Inhibition of grape crown gall by *Agrobacterium vitis* F2/5 requires two nonribosomal peptide synthetases and one polyketide synthase. Mol Plant Microbe Interact 29:109–118

Zheng D, Zhang H, Carle S et al (2003) A *luxR* homolog, *aviR*, in *Agrobacterium vitis* is associated with induction of necrosis on grape and a hypersensitive response on tobacco. Mol Plant Microbe Interact 16:650–658

Zheng D, Hao G, Cursino L et al (2012) LhnR and upstream operon LhnABC in *Agrobacterium vitis* regulate the induction of tobacco hypersensitive responses, grape necrosis and swarming motility. Mol Plant Pathol 13:641–652

Zhu J, Oger PM, Schrammeijer B et al (2000) The bases of crown gall tumorigenesis. J Bacteriol 182:3885–3895

Zidarič I (2009) Monitoring of bacteria of the genus *Agrobacterium* on grapevine in 2006 and 2007. In: 9th Slovenian conference on plant protection, Nova Gorica, Slovenia. Book of lectures and papers, pp 225–229

Zupan J, Muth TR, Draper O et al (2000) The transfer of DNA from *Agrobacterium tumefaciens* into plants: a feast of fundamental insights. Plant J 23:11–28

Niche Construction and Exploitation by *Agrobacterium*: How to Survive and Face Competition in Soil and Plant Habitats

Y. Dessaux and D. Faure

Contents

1 Introduction 56
2 *Agrobacterium*: A Soil-Adapted Bacterium 57
2.1 Exploiting Soil Resources 57
2.2 Facing and Sustaining Competition 58
3 The Plant Tumor: A Niche Extension for Agrobacteria 59
3.1 An Instance of Natural Genetic Engineering 61
3.2 The Opine Concept 64
3.3 Opine Metabolism Genes 64
3.4 Opines as Growth Substrates 65
3.5 Opines as Inducers of Ti Plasmid Horizontal Transfer 66
3.6 Cost and Control of Opine-Niche Construction and Exploitation 67
3.7 Competition for the Opine Niche by the Plant Microbiota 68
3.8 Exploitation of Other Plant Tumor Resources 69
4 Niche Construction and Exploitation by *Agrobacterium*-Related Genera 70
5 Unsolved Mysteries in *Agrobacterium* Ecology 72
5.1 Where Do Pathogenic Agrobacteria Hide in Nature? 72
5.2 Origin of T-DNA, Origin of Opines 73
5.3 Is *Agrobacterium*'s Ability to Transfer DNA to Organisms Belonging to Other Kingdoms Unique? 75
References 76

Y. Dessaux (✉) · D. Faure
Institute for Integrative Biology of the Cell (I2BC), CEA - CNRS - Univ. Paris-Sud, Université Paris-Saclay, Avenue de la Terrasse, 91198 Gif-sur-Yvette, France
e-mail: yves.dessaux@i2bc.paris-saclay.fr

D. Faure
e-mail: denis.faure@i2bc.paris-saclay.fr

Current Topics in Microbiology and Immunology (2018) 418:55–86
DOI 10.1007/82_2018_83

Published Online: 20 March 2018

Abstract *Agrobacterium* populations live in different habitats (bare soil, rhizosphere, host plants), and hence face different environmental constraints. They have evolved the capacity to exploit diverse resources and to escape plant defense and competition from other microbiota. By modifying the genome of their host, *Agrobacterium* populations exhibit the remarkable ability to construct and exploit the ecological niche of the plant tumors that they incite. This niche is characterized by the accumulation of specific, low molecular weight compounds termed opines that play a critical role in *Agrobacterium*'s lifestyle. We present and discuss the functions, advantages, and costs associated with this niche construction and exploitation.

1 Introduction

Agrobacterium is known among microbiologists, geneticists, and biotechnologists as a robust and versatile tool used to introduce foreign genes into plants or fungi (for reviews, see Vain 2007; Idnurm et al. 2017). However, most members of this genus are primarily plant pathogens that induce galls on dicotyledonous plants. Formerly, the *Agrobacterium* genus encompassed various species such as *A. rubi*, *A. larrymoorei*, *A. vitis*, *and A. tumefaciens*. The latter species is now recognized as a complex of several species including *A. fabrum* to which belongs *A. fabrum* C58, whose genome was the first sequenced in *Agrobacterium* (for more on *Agrobacterium* taxonomy see, *e.g.,* Mousavi et al. 2014; Kuzmanović et al. 2015; De Lajudie and Young 2017). In this chapter, we deal with members of the *Agrobacterium* genus and related genera, irrespective of their species designation, but the most abundant literature is associated with the *A. tumefaciens* species complex, and especially with the strain C58. For consistency, we will retain the ancient name *A. tumefaciens* to designate this strain.

Collectively, agrobacteria belong to the family Rhizobiaceae of the class alpha-proteobacteria, members of which are often found in soils of various origins and appear to be among the most common inhabitants of these environments (*e.g.*, Bouzar and Moore 1987; Nüsslein and Tiedje 1998; Texeira et al. 2010; Inceoglu et al. 2011; Lundgerg et al., 2012; Bulgarelli et al. 2012). Interestingly, agrobacteria isolated from soils, including rhizospheric soils, are most often avirulent (Bouzar and Moore 1987; Burr et al. 1987), *i.e.*, they do not harbor a Ti plasmid, the key replicon that determines virulence, unless the soil has an history of crown gall disease (Bouzar et al. 1993; Krimi et al. 2002). These findings suggest that agrobacteria are soil- and rhizosphere-adapted bacteria. As expected, agrobacteria exhibit several traits to exploit soil and rhizosphere resources and to survive under competition with other micro- and macro-organisms. Aside from these adaptive traits, the acquisition of a Ti plasmid that confers pathogenicity can be considered as a process leading to the construction of a more specific and less competitive ecological niche on plant hosts. Data to support these views on niche exploitation and construction by agrobacteria in the soil and plant habitats are presented below.

2 *Agrobacterium*: A Soil-Adapted Bacterium

Depending on the soil type, agrobacteria members can be either rare or relatively abundant among cultivatable bacteria, with concentrations ranging from 10^3 to 10^7 CFU/g (Bouzar and Moore 1987; Bouzar et al. 1993; Krimi et al. 2002). Agrobacterial traits that favor the adaptation to the soil environment remain largely unidentified, as they do for most soil bacteria. However, analysis of the metabolic properties of the bacteria and recent genomic data revealed several interesting features that may allow *Agrobacterium* to colonize the highly competitive soil environment.

2.1 Exploiting Soil Resources

Agrobacteria may survive for weeks and months under oligotrophic conditions, including pure water (Iacobellis and Devay 1986). Surface waters and aerosols could therefore contribute to dissemination of *Agrobacterium* populations. Members of this genus are also resistant to osmotic stress, both by taking up osmoprotectants (Nobile and Deshusses 1986; Boncompagni et al. 1999) or by synthesizing them (Smith et al. 1990).

However, bare soils are rare. Most often they are covered by plants that decompose in fall and winter to form humic acids in which agrobacteria can survive for months (Süle 1978). Plants also release at their root system a mixture of carbon compounds known as rhizodeposits. The rhizodeposits consist mainly of root cell debris and exudates, these later originating from plant photosynthesis and metabolism (for reviews, see Hinsinger et al. 2009; Jones et al. 2009; Sánchez-Cañizares et al. 2017). In possible relation with the supply of diverse carbon sources in the rhizosphere, agrobacteria have evolved a wide metabolic capability. For instance and with some variations from one strain to another, agrobacteria can degrade a large range of oses, polyols, and sugar derivatives often from plant origin, including cellobiose, trehalose, maltitol (Marasco et al. 1995; Ampomah et al. 2013), altritol and galactitol (Wichelecki et al. 2015), xylose and glucosamine (Zhao and Binns 2014), melezibiose, raffinose, gentobiose, turanose, lyxose, tagatose, D- and L-fucose, aldonitol, D- and L-arabitol, dulcitol, inositol, sorbitol, xylitol, gluconic acid, keto-gluconic acid, arbutin, esculin, and salicin (Dessaux, unpublished). Agrobacteria can also utilize a wide range of nitrogen-containing compounds as nitrogen sources such as urea (Riley and Weaver 1977), amino-valerate, amino benzoate, ethanolamine, tryptamine (Dessaux, unpublished), and gamma amino-butyrate (Chevrot et al. 2006). In relation to these potential nutrients, agrobacteria exhibit potent urease (Dessaux et al. 1986a, b) and transaminase activities (Sukanya and Vaidyanathan 1964) and a putative nitrilase that permits the scavenging of nitrogen from the plant glycoside amygdalin (Dessaux et al. 1989) and possibly from other cyanogenic compounds. In agreement with the above

catabolites, agrobacteria also encode a large number of diverse transporters likely used to take up various potential nutrients.

2.2 *Facing and Sustaining Competition*

In the soil, agrobacteria are armed to face microbial competitors. Indeed, agrobacteria benefit from a set of potent siderophores that permit an efficient recovery of iron in iron-deprived environments. Several types of siderophores have been identified. The first of these discovered is agrobactin, a derivative of 2,3-dihydroxybenzoic acid, spermidine, and threonine (Ong et al. 1979). The second one is a hydroxamate (Penyalver et al. 2001). The third one, detected in strain C58, remains unidentified (Rondon et al. 2004) but may be specific for this strain (Baude et al. 2016). In addition, with respect to microbial competitors, agrobacteria appear to be partly resistant to antibiotics such as chloramphenicol (Tennigkeit and Matzura 1991), penicillin, erythromycin, streptomycin, and moderately to tetracycline (Khanaka et al. 1981). Aside from these traits, some agrobacteria also express a type VI secretion system (T6SS; for review, see Ryu 2015) that drives the injection of at least three effectors with enzymatic activities (DNase and putative peptidoglycan amidase) into neighboring, competing bacteria (Wu et al. 2008; Ma et al. 2014).

When *Agrobacterium* colonizes a plant habitat, it can resist adverse antimicrobial compounds such as phenolics produced by plants upon wounding or biotic stress (reviews: Kefeli et al. 2003; Bhattacharya et al. 2010; Caretto et al. 2015). Phenolics play multiple roles in plant protection. With respect to the microflora, phenolics can be potent growth inhibitors of fungi and antibacterial agents (for reviews, see Cushnie and Lamb 2005; Lattanzio et al. 2008). However, nonpathogenic *Agrobacterium* strains possess an efflux pump active on a group of phenolics, the isoflavonoids that include medicarpin and coumestrol (Palumbo et al. 1998). Other phenolics such as vanillyl alcohol, vanillin, coniferyl alcohol, coniferyl aldehyde, sinapyl alcohol, sinapinaldehyde, and syringaldehyde can also be degraded by nonpathogenic agrobacteria (Brencic et al. 2004). Recently, ferulic acid was also shown to be degraded by *Agrobacterium* strain C58 (Baude et al. 2016). In addition, pathogenic agrobacteria can detoxify other phenolics via the products of two Ti plasmid genes, *virH1* and *virH2*, located in the virulence region. The VirH1 and VirH2 proteins share sequence homology with cytochrome P450-like enzymes (Kanemoto et al. 1989), and VirH2 appears to be an O-demethylase that is active on over 15 phenolic substrates such as sinapinic acid and acetosyringone. VirH2 can also convert vanillic acid to protocatechuate, which can be further metabolized via the β-ketoadipate pathway (Brencic et al. 2004). Taken together, these data indicate that pathogenic agrobacteria are more resistant to phenolics than are nonpathogenic ones, a result confirmed by the analysis of a *virH2* mutant (Brencic et al. 2004). Remarkably, many of the above-mentioned phenolics are inducers of the virulence genes of *Agrobacterium* (Bolton et al. 1986;

Engström et al. 1987) and a few may also be chemoattractant (Parke et al. 1987), a feature that could allow agrobacteria to move upward the concentration gradient toward the wounded plant cells (for review, see Shaw 1991). The route to the plant is also traced by root exudates that are also chemoattractant for agrobacteria (Hawes and Pueppke 1987; Hawes and Smith 1989).

3 The Plant Tumor: A Niche Extension for Agrobacteria

The above data indicate that *Agrobacterium* is well-equipped to survive in the soil and the plant rhizosphere. However, these environments remain quite competitive. The ability of *Agrobacterium* to generate a plant tumor can therefore be seen as a "coup de génie" that permits these bacteria to benefit from a much more private habitat, *i.e.,* a quasi-specific niche (Fig. 1). *Agrobacterium* takes a triple ecological advantage from tumor-niche construction: (i) an increase of resources supporting its proliferation to a high population level; (ii) a decrease of plant defense response in

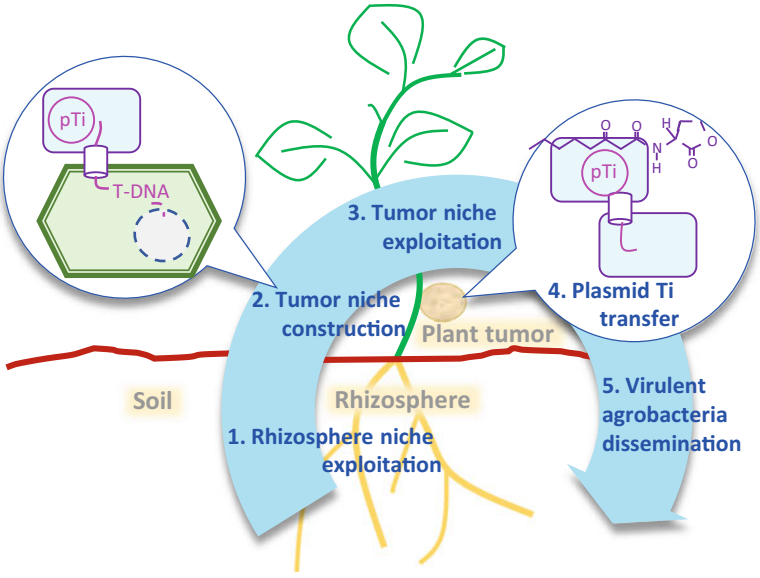
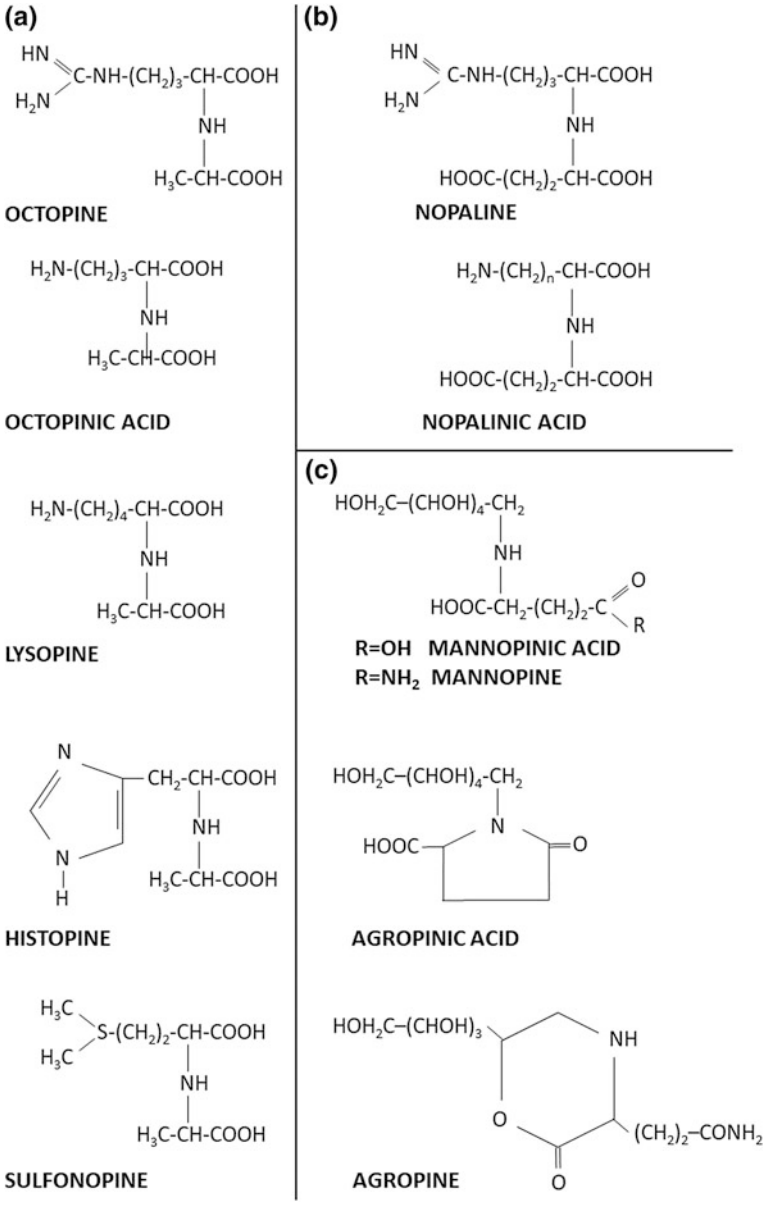


Fig. 1 Ecological niches of *Agrobacterium.* Saprophytic and pathogenic (carrying Ti plasmid) *Agrobacterium* populations efficiently colonize the rhizosphere of host and non-host plants. Upon permissive conditions (wounding), virulent agrobacteria construct a novel ecological niche that is the plant tumor, as the result of the transfer and expression of the T-DNA in plant host genome. *Agrobacterium* populations exploit the tumor resources, including opines which confer a selective advantage to *Agrobacterium* pathogens. Opines also activate quorum-sensing pathways that promote Ti plasmid conjugative transfer, hence contributing to the maintenance and propagation of the virulence genes. The high abundance of virulent *Agrobacterium* in plant tumor facilitates the dissemination to a new host as well as the maintenance of populations in the rhizosphere and soil



plant tumor tissues; and, (iii) a decrease of competition with resident microbiota, especially through the exploitation of specific growth substrates known as opines. The first point is still poorly understood but could be hypothesized from the high abundance of organic and mineral nutrients that accumulate in plant tumors (Deeken et al. 2006; Lang et al. 2016), whereas tumor development represents a

◂**Fig. 2** Structural formulas of opines **a** octopine family, **b** nopaline family, **c** agropine family, **d** agrocinopines family, **e** cucumopine family, **f** succinamopine and leucinopine families, **g** chrysopine family, **h** ridéopine and heliopine families. Octopine family opines are all synthesized by the enzyme octopine synthase and derive from various proteinous and nonproteinous amino acids, and pyruvate. They include the most recently discovered opine sulfonopine (Flores-Mireles et al. 2012). Nopaline and nopalinic acid synthesized by nopaline synthase derive from alpha-ketoglutarate and, respectively, arginine and ornithine. Succinamopine, leucinopine, cucumopine (and its diastereomer mikimopine) are also alpha-ketoglutarate condensates and exhibit asparagine, leucine, and histidine moieties, respectively. Heliopine (also termed vitopine) is a condensation product of pyruvate and glutamine. The mannityl opines are sugar and glutamate or glutamine-containing compounds as are the closely related opines of the chrysopine family. Other sugar opines include the agrocinopines A and B that are the only phosphorus-containing opines

metabolic sink from the plant host; this process makes diversified and abundant resources available to the pathogen. The second point was revealed by transcriptomic and genetic analyses of plant defense pathways (Gohlke and Deeken 2014). Tumor tissue development not only results in abnormally proliferating cells, but also causes differentiation and serves as a mechanism to balance pathogen defense, thereby contributing to the long-term coexistence of agrobacteria and the host plant. The third point, *i.e.*, the opine contribution to *Agrobacterium* lifestyle in plant tumors, is detailed below.

3.1 An Instance of Natural Genetic Engineering

Agrobacterium'*s* ability to incite a plant tumor, known as crown gall, depends upon the presence in the bacteria of a large plasmid termed the tumor-inducing (Ti) plasmid. During the infection process, a portion of this plasmid, T-DNA, is transferred via a type IV secretion system ($T4SS_{T\text{-}DNA}$) as a single-stranded DNA linked with proteins with plant nuclear localization signals. These proteins and T-DNA localize to the nucleus of the plant where T-DNA is eventually integrated into the genome and expressed. These proteins and the $T4SS_{T\text{-}DNA}$ are encoded by the non-transferred virulence (*vir*) genes also located on the Ti plasmid (for reviews and more details on the transfer machinery and genetic transformation formation process, see Pitzschke and Hirt 2010; Gelvin 2012; Lacroix and Citovsky 2013; Subramoni et al. 2014; Nester 2015; Christie 2016; Gelvin 2017). Two major sets of genes are borne on T-DNA. The first set, the oncogenes, is responsible for the synthesis of the plant hormones auxin and cytokinin by the transformed host cells, a feature that triggers their proliferation to form a tumor (Ooms et al. 1981; Akiyoshi et al. 1983; Ream et al. 1983). The second set is responsible for the synthesis, of low molecular weight compounds collectively termed opines (see Fig. 2) at the expense of the metabolite pool of the plant. Opines play key ecological roles in the *Agrobacterium*/plant interaction (for reviews, see Dessaux et al. 1998; Subramoni et al. 2014).

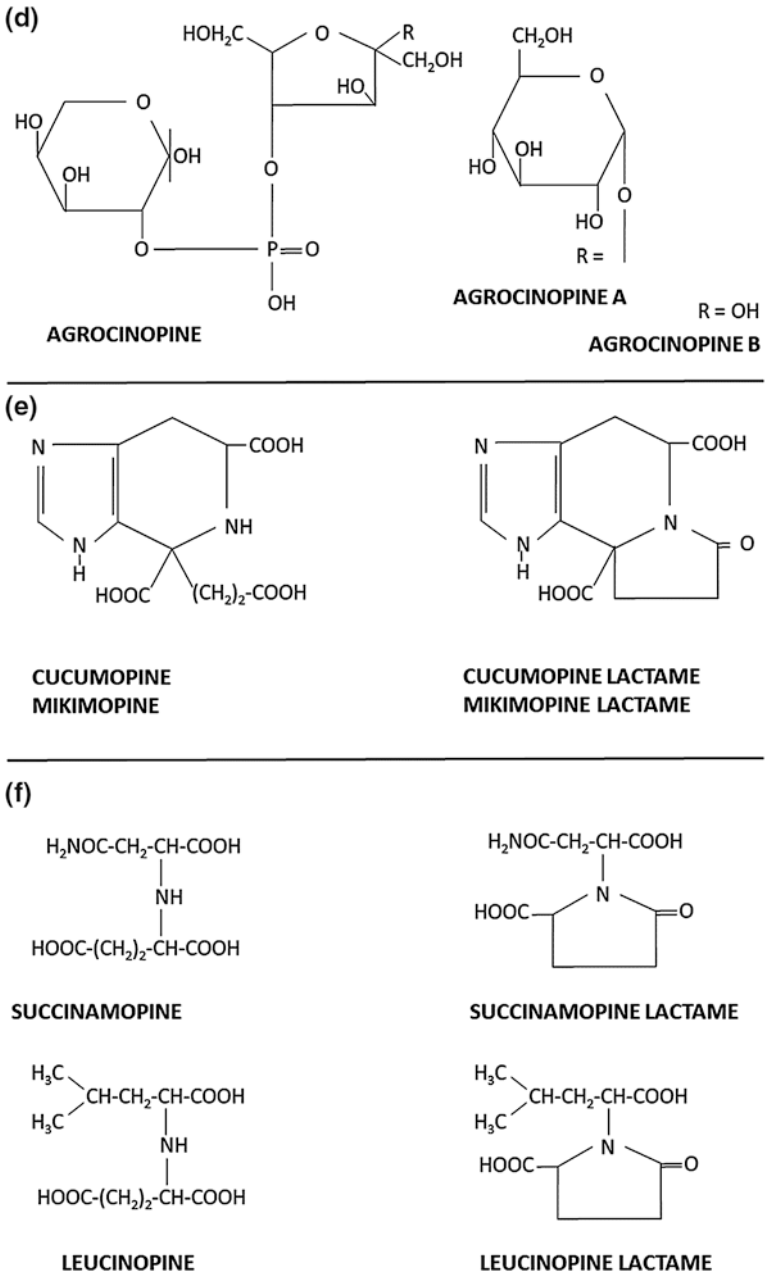


Fig. 2 (continued)

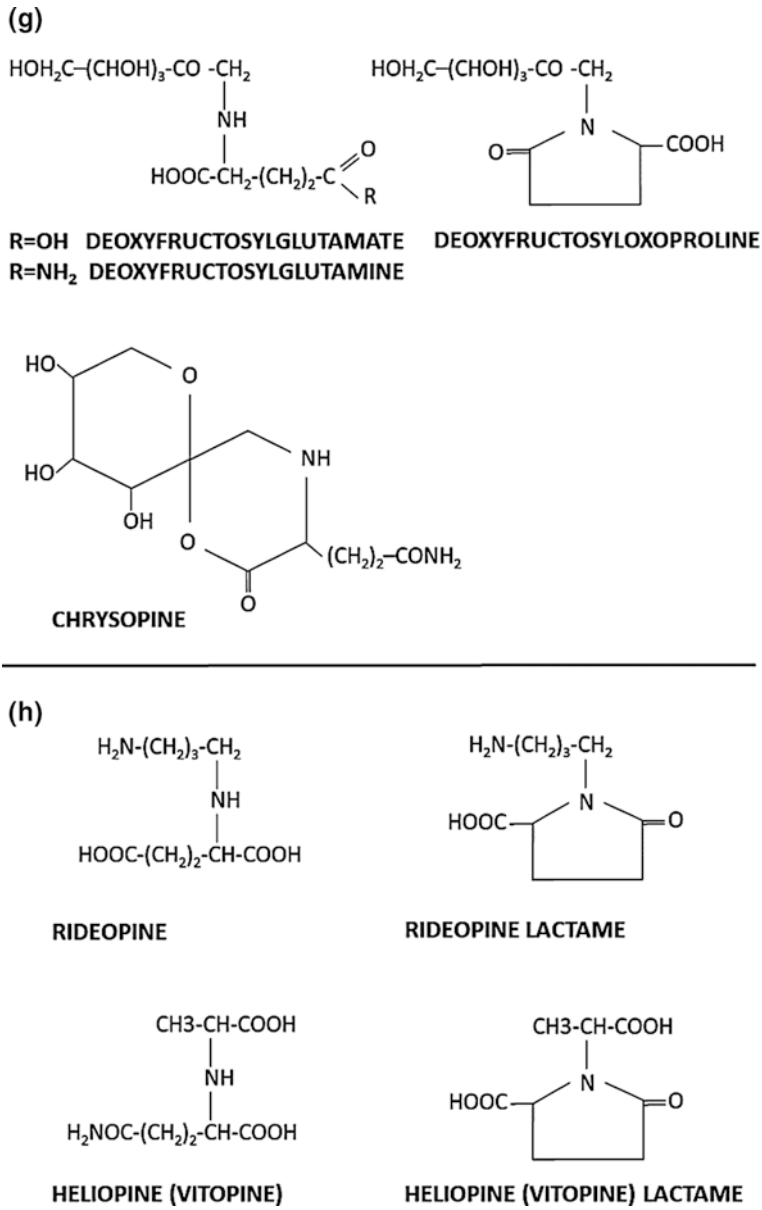


Fig. 2 (continued)

3.2 The Opine Concept

Opine synthesis by crown gall tumors and their assimilation by agrobacteria represents an archetype of ecological niche construction and exploitation processes by a pathogen. Opines are secreted by transformed plant cells into the intercellular spaces in the tumor, and to a lesser extent the whole plant (Savka and Farrand 1992; Savka et al. 1996). Opines play two major roles in niche construction for agrobacteria. First, they serve as growth substrates for the tumor-inciting strain and, second, they stimulate the conjugative transfer of the Ti plasmid from pathogenic *Agrobacterium* to other *Agrobacterium* cells (for a review, see Dessaux et al. 1998). These features are at the origin of the opine concept that describes opines as chemical mediators of parasitism. Synthesis of opines is induced by the pathogen, thus providing an environment favorable to the growth of the bacteria and dissemination of its pathogenicity (Schell et al. 1979; Tempé and Petit 1983).

The opine concept was formulated years before the discovery of plants that naturally harbor in their genomes DNA regions highly homologous to *Agrobacterium* T-DNA. Among these species are members of the genera *Nicotiana*, *Linaria,* and *Ipomoea* (White et al. 1983, Aoki et al. 1994; Suzuki et al. 2002; Matveeva et al. 2012; Kyndt et al. 2015; Quispe-Huamanquispe et al. 2017). Interestingly, some of these plants produce detectable amounts of opines (Chen and Otten 2017). The opine concept could therefore incorporate both the tumorous temporary niche and the permanent niche that naturally genetically modified plants and their offspring represent. However, no clear demonstration of a stimulation of the community of opine-degrading bacteria at the root system of these naturally transformed plants has yet been reported.

3.3 Opine Metabolism Genes

Opines are most often synthesized from common molecules such as amino acids, alpha-ketoacids, and sugars. Over 20 opine molecules are known (Fig. 2a–h). They are not all present at the same time in a tumor and some opines are specific for a given agrobacteria species. Indeed, the type opine synthesized by plant cells and degraded by agrobacteria depends upon the type of Ti plasmid, a feature that has been used to classify agrobacterial Ti plasmids (for a review, see Dessaux et al. 1998). The current list of agrobacterial opines is likely to be near complete. Indeed, over the last 15 years no novel opine has been discovered except sulfonopine, a sulfur-containing molecule detected in tumors induced by a single octopine-type *Agrobacterium* strain (Flores-Mireles et al. 2012).

Genes involved in the biosynthesis and catabolism of opines are known for several opine systems. Generally, opines derived from amino acids and alpha-ketoacids (such as octopine or nopaline; Fig. 2a, b) are synthesized in one step by a protein encoded by a single gene located on T-DNA (De Greve et al.

1982; Koncz et al. 1983). The same is true for phospho-sugar opines of the agrocinopine family (Joos et al. 1983; Fig. 2d). On the contrary, opines derived from condensation of sugars and amino acids, the mannityl opines or the chrysopine family opines (Fig. 2c, g), are synthesized in one, two, or three steps by the corresponding number of enzymes encoded by one, two, or three genes. These are most often located on a T-DNA separate from that which carries the oncogenes (Hood et al. 1986; Palanichelvam et al. 2000).

Opine catabolic genes are generally clustered in operons and regulons in delineated regions of the Ti plasmids, and their expression is inducible by the degraded opines (Klapwik et al. 1977; Klapwik et al. 1978; Chilton and Chilton 1984; Dessaux et al. 1988). Two sets of genes are present in the catabolic region. The first encodes the transport system (*e.g.,* Klapwik et al. 1977; Zanker et al. 1992) that often consists of an ABC transporter and its cognate, high affinity (nM–μM range) periplasmic-binding protein (Lang et al. 2014; El Sahili et al. 2015; Marty et al. 2016; Vigouroux et al. 2017). The second encodes the enzymes involved in the degradation of the opines to molecules that belong to central bacterial metabolism. For example, octopine and nopaline are degraded into arginine, ornithine, and glutamate, and pyruvate or alpha-ketoglutarate, respectively (Montoya et al. 1977; Ellis et al. 1979; Dessaux et al. 1986a, b). Remarkably, for some opines such as the mannityl opines, genes, and functions involved in the synthesis and degradation are closely related, suggesting that duplication events occurred in the course of the evolution of the Ti plasmids (Kim et al. 1996; Hong et al. 1997; Kim and Farrand 1996). A similar duplication also occurred with respect to genes involved in the synthesis and degradation of the phospho-sugar opines agrocinopines A and B (Kim and Farrand 1997).

3.4 Opines as Growth Substrates

The opine concept has been elaborated from the observation that all crown gall tumors, including those initially reported not to contain any opine (*i.e.,* the so-called null type tumors), indeed contain such compounds (Guyon et al. 1980). The opine hypothesis later received experimental validation. The first support for the opine concept came from comparison of the growth of two closely related *Agrobacterium* strains, one capable of degrading opines, the other not, at the root system of transformed plants producing opines. The experiment revealed that plants producing opines preferentially promote the growth of opine-degrading agrobacteria (Guyon et al. 1993). A second set of experiments involved transformed plants producing opines and two closely related *Pseudomonas* strains, one engineered—via the introduction of an opine catabolic plasmid—to degrade opines, the other not. The experiment demonstrated that the growth of the opine-degrading pseudomonad was favored at the root and leaf surface of opine-producing plants (Wilson et al. 1995; Savka and Farrand 1997). A recent experiment (Lang et al.

2014) involved the wild-type (WT) *Agrobacterium* strain C58 and a mutant unable to degrade nopaline, the major opine found in the tumors incited by this strain. When both strains were inoculated separately onto a plant, they multiplied in the tumor to reach a similar bacterial concentration. However, when co-inoculated the WT opine-degrading bacteria outcompeted the mutant. This observation formally demonstrated that the presence of the opine does not increase the carrying capacity of the tumor habitat for *Agrobacterium* but "*selects for those able to assimilate it*" (Lang et al. 2014). A similar study extended this paradigm to the octopine-niche (submitted by Vigouroux et al. 2017).

3.5 *Opines as Inducers of Ti Plasmid Horizontal Transfer*

The discovery of the Ti plasmid as key pathogenic element for *Agrobacterium* (Van Larebeke et al. 1974; Watson et al. 1975) was rapidly followed by the demonstration that these plasmids can be transferred by conjugation between bacteria; a phenomenon also regulated by opines (Kerr et al. 1977; Genetello et al. 1977). The nature of the opines that induce conjugation varied as a function of the opine-type of the plasmid. Thus, octopine induces the transfer of octopine-type plasmids, whereas agrocinopines A and B induce transfer of nopaline-type plasmids, and agrocinopines C and D the conjugation of agropine-type plasmids (Klapwijk et al. 1978; Petit et al. 1978; Ellis et al. 1982).

Ti plasmid transfer is also regulated by quorum sensing (QS; Piper et al. 1993; Zhang et al. 1993). QS is a widely occurring regulatory process that couples gene expression (in a positive or negative way) with bacterial cell concentration. It relies upon the production and sensing by a bacterial population of diffusible signal(s), the concentration of which indicates that of the microbial cells. Once a threshold concentration of signal is reached in the environment, the presence of the signal is sensed by receptors and translated into activation or repression of the expression of the genes regulated by QS (for recent reviews on QS, see Garg et al. 2014; Grandclément et al. 2016; Papenfort and Bassler 2016).

In the reference *Agrobacterium* strain C58, the presence of agrocinopines A and B triggers the expression of the *acc* operon of the Ti plasmid that encodes agrocinopine degradation, and that of the adjacent *arc* operon by releasing the repression exerted by the master regulator AccR (Beck von Bodman et al. 1992). Agrocinopine A can be cleaved into arabinose-2-P and sucrose by AccF, because only arabinose-2-phosphate (and not agrocinopine A) interacts with AccR (El Sahili et al. 2015). One of the genes of the *arc* operon is *traR*. It encodes the regulatory protein TraR that, once bound to the QS signal, dimerizes and activates the transcription of the *traAFB*, *traCDG*, and *trb* operons (Piper et al. 1999). The *tra* operons encode components of the DNA transfer and replication (DTR) system that recognizes and cleaves plasmid DNA at the origin of transfer (oriT) located between the two *tra* operons (Farrand et al. 1996; Zechner et al. 2001). The *trb*

operon encodes components of a type IV secretion system ($T4SS_{pTi}$) that permits the transfer of the plasmid DNA and associated proteins to recipient bacteria (Li et al. 1999). Interestingly, the first gene of the *trb* operon is *traI*. The eponym protein TraI is responsible for the synthesis of a diffusible QS signal that belongs to the widely distributed *N*-acyl homoserine lactone (AHL) class of signals (Hwang et al. 1994). In the presence of agrocinopines but at low cell concentration, the *trb* operon—hence, the TrbI QS signal synthase—is very weakly expressed and only low amounts of QS signals accumulate in the environment. In the presence of agrocinopines and at high cell concentrations, the QS signal concentration increases and its presence is sensed by TraR that becomes activated and induces the full expression of the $T4SS_{pTi}$ and DTR system, stimulating the transfer of the Ti plasmid (Li et al. 1999; Li and Farrand 2000).

3.6 *Cost and Control of Opine-Niche Construction and Exploitation*

As indicated above, the key step of opine-niche construction is the transfer of T-DNA to plant cells via a dedicated T4SS ($T4SS_{T\text{-}DNA}$) that imposes a fitness cost to agrobacteria (Platt et al. 2012). In a competitive arena, individuals expressing the $T4SS_{T\text{-}DNA}$ are disadvantaged compared to those impaired for $T4SS_{T\text{-}DNA}$ or defective for a Ti plasmid. Indeed, in short-term experimental evolution cultures in the presence of acetosyringone (an inducer of $T4SS_{T\text{-}DNA}$ expression) and in plant tumors, spontaneous mutants arose in the progeny of a virulent *Agrobacterium* ancestor. These mutants were altered in virulence because of alteration or loss of the Ti plasmid (Bélanger et al. 1995; Fortin et al. 1992, 1993; Llop et al. 2009). Virulent agrobacteria exhibit three potential mechanisms to balance the fitness cost and damage imposed by T-DNA transfer: (i) a tight control of *vir* gene expression by phenolics, acidic pH, and sugars contributes to optimize the cost/benefit of $T4SS_{T\text{-}DNA}$ expression, hence the success of T-DNA transfer into plant cells (Nair et al. 2011; He et al. 2009); (ii) Ti plasmid horizontal transfer that may re-introduce the Ti plasmid into those cells which have lost it (Lang et al. 2013); (iii) a fitness gain to agrobacteria—that have kept or acquired a Ti plasmid—because of opine-niche exploitation (Lang et al. 2014). Conditioning the transfer of the Ti plasmid to the tumor environment (opine as ecological proxy) ensures that the Ti plasmid-carrying *Agrobacterium* individuals will gain a selective advantage in the most compatible ecological niche.

In nature, the Ti plasmid may be transferred to other agrobacteria (other species or clonal lineages) or non-agrobacteria that is free of a Ti plasmid, whereas this transfer could be considered as advantageous for the Ti plasmid per se (selfish gene and reservoir hypotheses), and it could be disadvantageous for the Ti plasmid donor lineage because potential bacterial competitors could acquire the opine-niche exploitation trait. Another important consideration in Ti plasmid transfer is its cost

as the process uses a second $T4SS_{pTi}$. An experimental evolution experiment conducted with an *A. tumefaciens* C58 derivative expressing QS and $T4SS_{pTi}$ revealed the emergence of mutants defective for QS signal synthesis (mutations in *traR*) or exhibiting a QS-hijacking behavior or defective for the presence of a Ti plasmid (Tannières et al. 2017). *Agrobacterium* Ti plasmid donors exhibit several mechanisms to finely control QS, and hence Ti plasmid transfer. QS relates Ti plasmid transfer to a high population level of donors. This major requirement allows a virulent population to become dominant in a plant tumor habitat before activating Ti plasmid transfer, which is costly (growth slowdown) and hazardous (increase of opine-assimilating competitors). Additional mechanisms which are not present in all agrobacteria also contribute to delay QS signaling, therefore leaving time for donors to proliferate before transferring a Ti plasmid. First, the TraM protein encoded by *traM* on the Ti plasmid interacts with TraR and blocks the formation of an active TraR homodimer at low QS signal concentrations (Khan et al. 2008; Qin et al. 2007). Second, the lactonases BlcC and AiiB open the gamma-butyrolactone ring of AHLs (Haudecoeur et al. 2009). The *traM* and *aiiB* genes are encoded by the Ti plasmid and are expressed in the presence of agrocinopines in strain C58. The *blcC* gene (formerly *attM*) belongs to the *blcABC* (formerly the *attKLM*) operon located on pAt plasmids. *BlcC* is activated in the presence of gamma-butyrolactone (GBL), gamma-hydroxybutyric acid (GHB), and succinic semialdehyde (SSA), which are activated and repressed by a high and low gamma-aminobutyric acid (GABA)/proline ratio, respectively (Carlier et al. 2004; Lang et al. 2016), whereas TraM titrates TraR and prevents its early production, intracellular lactonases constrain the level of AHLs in the intra- and extracellular environments, hence their binding to TraR. Both these QS-delaying mechanisms are bypassed when TraR and AHLs are produced at a high level (Khan et al. 2008; Haudecoeur et al. 2009).

3.7 Competition for the Opine Niche by the Plant Microbiota

Although engineered by *Agrobacterium* as a niche, the tumor can be colonized by other opine-degrading microorganisms, including bacteria such as pseudomonads, *Sinorhizobium meliloti*, *Arthrobacter* sp., coryneform isolates (Tremblay et al. 1987; Nautiyal and Dion 1990; Nautiyal et al. 1991; Moore et al. 1997; Faist et al. 2016), or by fungal strains (*Cylindrocarpon heteronema* and *Fusarium solani*; Beauchamp et al. 1990). These microorganisms are naturally present in soils of diverse origins, and their growth can be stimulated by opines produced by the tumor and released at the root system of the plant independently of the soil and plant considered (Oger et al. 1997; Mansouri et al. 2002, Mondy et al. 2014; Faist et al. 2016). Interestingly, as indicated earlier, opines are chemoattractants for *Agrobacterium* (Kim and Farrand 1988). This feature may provide a way for

agrobacteria that migrate from the tumor to return to the opine-rich niche of the crown gall. A possibility exists that opines could also attract non-agrobacterial organisms, but to the best of our knowledge, this has not yet been investigated.

3.8 Exploitation of Other Plant Tumor Resources

Besides opines, a wide variety of organic (amino acids, organic acids, oses, polyols, etc.) and mineral compounds, which are potential resources for agrobacteria, accumulate in plant tumors (Deeken et al. 2006; Lang et al. 2016). Unlike opines, these compounds are not specific to tumor tissues and Ti plasmid type. To be considered part of the niche construction process, these compounds should not only accumulate in plant tumors, and their exploitation should also confer a selective advantage to agrobacteria for colonizing this habitat. Numerous traits are potentially consistent with this definition but experimental evidence is missing. A chromosomal locus *picA*, which may be involved in the degradation of plant polymers and whose expression is induced in the presence of plant tissues, may be such a candidate (Rong et al. 1991).

Numerous *Agrobacterium* isolates (carrying or lacking a Ti plasmid) harbor larger plasmids known as pAt plasmids (Merlo and Nester 1977; Rosenberg and Huguet 1984; Hynes et al. 1985). pAt plasmids can be very different from one strain to another, whereas they may comprise up to 10% of the agrobacterial genome, only a limited number of pAt functions are known. In *A. tumefaciens* C58, aside from utilization of GBL, GHB, and SSA (a by-product of GABA) as nutrients (Carlier et al. 2004), the plasmid pAtC58 encodes degradation of the Amadori compound deoxy-fructosyl-glutamine (Vaudequin-Dransart et al. 1995; Baek et al. 2003). Exploitation of these plant compounds could contribute to tumor colonization by virulent (carrying Ti plasmid) and avirulent (free of Ti plasmid) agrobacteria.

The question about the cost associated with At plasmid maintenance has been investigated in *Agrobacterium* strain C58 by comparing different derivatives carrying two, only one, or none of the plasmids pAtC58 or pTiC58. In culture medium when the $T4SS_{T\text{-}DNA}$ and $T4SS_{pTi}$ are not expressed, the cost of carrying the pAt plasmid was higher than that of the Ti plasmid (Morton et al. 2014). This may be related to the large size of the pAt plasmid as well as to the constitutive expression of the $T4SS_{pAt}$ that promotes its conjugative transfer (Chen et al. 2002). A fitness gain associated with the pAt plasmid was reported in the rhizosphere of *Helianthus annuus* (Morton et al. 2014), but this question remains unsolved in plant tumors. Interestingly, in *Agrobacterium* strain C58 the transfer of the pAt plasmid is co-regulated with that of the Ti plasmid and strongly depends upon the activity of the master regulatory protein AccR encoded by a Ti plasmid gene, the transcription of which is induced in the presence of agrocinopines A and B (Lang et al. 2013). This observation suggests that the tumor habitat stimulates a simultaneous propagation of both the pAt and pTi plasmids, probably meaning that a selective

advantage could be conferred by the acquisition of the two plasmids. In some *Agrobacterium* strains, pTi and pAt plasmids can cointegrate and cooperate for opine degradation (Vaudequin-Dransart et al. 1998). This cooperation has also been observed for some Ri plasmids (Costantino et al. 1980; Petit et al. 1983). In strains devoid of a Ti plasmid, the transfer of the pAt plasmid may also be regulated, by QS, by-products of genes located on this plasmid. In this case, the existence of one or more possible inducers of conjugation has not been demonstrated (Mhedbi-Hajri et al. 2016).

4 Niche Construction and Exploitation by *Agrobacterium*-Related Genera

All the findings described above paved the path to investigate whether the opine concept can be extended outside the *Agrobacterium* clade. Experiments performed with closely related Rhizobiaceae revealed that transformed plant roots induced by *Rhizobium rhizogenes* (formerly *Agrobacterium rhizogenes*) also contain opines (Petit et al. 1983). These two pathogens are closely related. Indeed, as in *Agrobacterium* spp., pathogenic strains of *R. rhizogenes* harbor large plasmids known as Ri plasmids (Moore et al. 1979). A portion of these plasmids, T-DNA, is transferred to the nucleus of the plant cells where it integrates into the genome upon infection (Chilton et al. 1982; Willmitzer et al. 1982; White et al. 1982). *R. rhizogenes* T-DNA harbors oncogenes that for the most part differ from those of *A. tumefaciens* and trigger the formation of transformed roots (*e.g.,* Durand-Tardif et al. 1985; Slightom et al. 1986; Cardarelli et al. 1987; Spena et al. 1987). However, the genes involved in opine biosynthesis are often highly related to those of *Agrobacterium* Ti plasmids, and several of them direct the synthesis of opines, such as cucumopine or mannityl opines (Fig. 2d, e), that are also found in crown gall tumors (Tepfer and Tempé 1981; Jouanin 1984; De Paolis et al. 1985; Petit and Tempé 1985).

A further extension dealt with nitrogen-fixing nodules incited by *Sinorhizobium meliloti* and *Rhizobium leguminosarum* strains on leguminous plants. Some of these nodules contain opine-like molecules, identified as *scyllo*-inosamine (SI) and 3-*O*-methyl-*scyllo*-inosamine (3OSI; Murphy et al. 1987; Saint et al. 1993) and collectively termed rhizopines (Fig. 3). However, only a limited number of strains of these species (ca. 11–12% of assayed clones) were able to produce and degrade rhizopines, independent of their geographical origin (Rossbach et al. 1995; Wexler et al. 1995). Genes involved in both the synthesis and degradation of SI and 3OSI have been identified. They are adjacent on the symbiotic plasmid of the bacteria (Murphy et al. 1987). In contrast to the *Agrobacterium* system, these biosynthetic genes are not transferred to plant cells but are expressed by the bacteria itself in the nodule context only. As with *Agrobacterium* opines, rhizopines provide a selective advantage for rhizopine utilizers in the plant environment, possibly by providing a

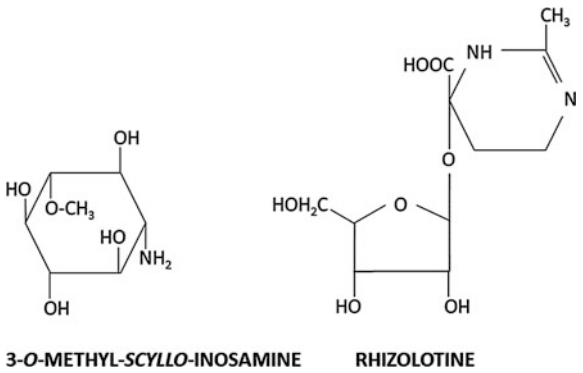


Fig. 3 Structural formulas of opine-like molecules found in nodules. The opine-like compound 3-*O*-methyl-*scyllo*-inosamine (as well as *scyllo*-inosamine, not shown) is opine-like molecules detected in nodules incited in alfalfa (*Medicago sativa*) by some strains of *Rhizobium meliloti*. Rhizolotine is an opine-like compound found in *Lotus* spp. nodules incited by *Mesorhizobium loti* strain NZP2037. This riboside molecule exhibits a tetrahydropyrimidine ring

selective nutrient to members of the population living around the nodules. This selective advantage has been demonstrated by competition experiments that involved a wild-type *S. meliloti* strain and a mutant unable to degrade rhizopines (Gordon et al. 1996; Heinrich et al. 1999). For a recent review on rhizopines and more data on genes involved in biosynthesis and degradation, the reader can refer to Savka et al. (2013).

Two other opine-like molecules have been detected in the nitrogen-fixing nodules induced on *Lotus* spp. by *Mesorhizobium loti*. One was identified as the riboside of an alpha-hydroxy-imino acid and named rhizolotine (Fig. 3). The second is an unidentified ninhydrin-positive compound (Shaw et al. 1986; Scott et al. 1987). No indication of the competitive advantage given to the rhizolotine-degrading strains in nature is available.

Aside from the above-described interactions, other interactions between bacteria and their hosts involve a trophic link. This is the case, for instance, for rhizobia that induce nodules on mimosa (*Acacia dealbata*) or *Leucaena* spp. plants. Plants of both genera produce large amount of mimosine, a toxic amino acid that only rhizobia nodulating these plants can degrade (Soerdajo et al. 1994), providing them with a selective advantage (Soedarjoa et al. 1998). Also the alkaloids calystegins present in the roots and exudates of morning glory (*Convolvulus arvensis*), hedge bindweed (*Calystegia sepium*), and belladonna (*Atropa belladonna*) can be efficiently degraded by *Sinorhizobium meliloti* strain Rm 41, a strain that is frequently detected in the root system of these plants, though they are not members of the legume clade and not hosts for symbiotic nitrogen fixation (Tepfer et al. 1988). None of these interactions, however, fits the description of the opine concept that remains limited to agrobacteria and—to a certain extent—to some rhizobia. As most rhizobia are symbionts, the opine concept should therefore be reformulated as "opines are chemical mediators of plant-microbe interactions, the synthesis of

which is induced by the micro-organism, thus providing an environment favorable to its growth and dissemination of its plant-interacting capacity."

All the above data prompted scientists to propose that the growth of beneficial microbial populations in the rhizosphere could be engineered and favored by establishing an opine-based, trophic link between the plant to protect and selected microbial population (Savka et al. 2002: Dessaux et al. 2016). Though elegant, and in spite of encouraging preliminary results obtained for some plant growth promoting rhizobacteria (Dessaux et al. 1987; Guyon et al. 1993; Wilson et al. 1995; Savka and Farrand 1997; Oger et al. 1997), this concept has not yet received definitive experimental validation.

5 Unsolved Mysteries in *Agrobacterium* Ecology

5.1 Where Do Pathogenic Agrobacteria Hide in Nature?

Though some pathogenic *Agrobacterium* strains can be isolated from uncultivated pasture soil (Schroth et al. 1971), natural soil, and plant rhizospheres, agrobacteria isolates are most often nonpathogenic unless the place of isolation has a history of crown gall contamination (Bouzar et al. 1993; Krimi et al. 2002; Dessaux, unpublished). This feature led scientists to wonder whether pathogenic agrobacteria can be isolated from some nursery soils because plants are contaminated and therefore provide the source of bacteria, or whether the plants are contaminated because virulent agrobacteria are present in these soils. This "chicken or egg" causal dilemma cannot be definitively resolved at this time, but some factual and speculative elements can be proposed. First, it is clear that exchange of contaminated plant material between various locations and countries could be at the origin of crown gall outbursts (Pionnat et al. 1999). Once contaminated, and in spite of seasonal fluctuations, soils can host agrobacterial populations and maintain them for years (Bouzar et al. 1993; Krimi et al. 2002). Second, it cannot be excluded that pathogenic agrobacteria can "hide" in the rhizosphere of non-host plants (*i.e.,* plants that do not develop crown gall symptoms) and, consequently, in places where they will not be searched for. In agreement with this proposal, agrobacteria have been detected at the root system of maize (Gomes et al. 2001) and wheat (Bednárová et al. 1979).

An alternative or complementary hypothesis is that *Agrobacterium* do not hide, but Ti plasmids do. It could be speculated that Ti plasmids could conjugate in tumors to other, non-agrobacterial isolates where they could replicate. In support of this model, Ti plasmids could conjugate to *E. coli* under laboratory conditions but they do not replicate in this host (Holsters et al. 1978). They can also be transferred to rhizobia that possess genetic backgrounds in which Ti plasmids can replicate but do not always express their tumorigenic functions (Hooykaas et al. 1977; van Veen et al. 1989; Teyssier-Cuvelle et al. 1999). Though attractive, this later hypothesis is

not really supported by probabilistic elements. First, the conjugative transfer frequency of the Ti plasmid in vitro reaches at best 10^3 per donor (Lang et al. 2013). Second, to generate a pathogenic *Agrobacterium*, the Ti plasmid would have to conjugate from the replicative bacteria back to an agrobacterial isolate in environments deprived of opines that are precisely the inducers of this conjugative transfer.

Clearly, the question of the reservoirs of Ti plasmids in nature remains an open but critical one. Further investigations are necessary to identify such reservoirs and complement our understanding of the ecology of both agrobacteria and Ti plasmid. Studies that aim to elucidate the genes and functions that contribute to bacterial fitness in tumors, the rhizosphere, bare soil, and possibly surface waters could contribute to reach these objectives.

5.2 *Origin of T-DNA, Origin of Opines*

Analyses of Ti plasmids revealed that they exhibit a chimeric structure (Otten et al. 1993; Otten and De Ruffray 1994) composed mostly of four key clusters of genes: the T-DNA, the virulence region that encodes the $T4SS_{T\text{-}DNA}$ involved in T-DNA transfer, the opine catabolic region, and the conjugative transfer regions that includes the $T4SS_{pTi}$. Interestingly, in *A. tumefaciens* Ti plasmids, T-DNA, the $T4SS_{T\text{-}DNA}$, and the conjugative region(s) are highly related, whereas the opine catabolic regions differ from one plasmid type to another. The homology of several genes that encode the $T4SS_{T\text{-}DNA}$ and the $T4SS_{pTi}$ (Chen et al. 2002) clearly suggests that both may derive from a common ancestral protein secretion system. Similarly, T-DNA genes responsible for the production of the plant hormone auxin, namely *iaaM* and *iaaH*, are orthologues of genes found in members of the *Pseudomonas savastanoi* species (Yamada et al. 1985). The cytokinin biosynthetic gene, *ipt* or *tmr*, is also related to the cytokinin biosynthetic gene *ptz* of *P. savastanoi* (Powell and Morris 1986).

As indicated above, *A. tumefaciens* T-DNAs differ from one another mostly by the nature of the opine anabolic genes. A parsimonious hypothesis therefore implies that T-DNA and the $T4SS_{T\text{-}DNA}$ have been acquired before the genes involved in opine metabolism in the evolutionary history of the plasmids, possibly as a way to reduce plant defense (Dunoyer et al. 2006; Gohlke and Deeken 2014). Furthermore, opine metabolic genes could have different origins. Some of these opine metabolic genes have evolved by duplication from common ancestor(s). This is the case of the genes involved in the synthesis and degradation of the mannityl opines. The synthesis of mannopine and mannopinic acid proceeds in two steps: (i) the condensation of glucose with glutamine or glutamate, respectively, to Schiff bases and their Amadori rearrangement compounds to form deoxy-fructosyl-glutamine (dfgln) and deoxy-fructosyl-glutamate (dfglu; Fig. 2g); (ii) the reduction of dfgln and dfglu to mannopine or mannopinic acid, respectively (Ellis et al. 1984). Mannopine can be dehydrated to yield the cognate lactone agropine, another mannityl opine (Dessaux

et al. 1986a, b). Degradation proceeds almost in the reverse way. Agropine undergoes a lactonolysis to mannopine that is in turn degraded to dfgln and mannose and glutamine. Dfglu is degraded to mannose and glutamate. In this scheme, dfgln appears to play a central role. First, it is also an opine found in the tumors induced by strains of *Agrobacterium* that harbor a chrysopine-type Ti plasmid; second, and in contrast to most opines, dfgln can be degraded by numerous strains of *Agrobacterium* irrespective of their virulence (Bouzar et al. 1995; Chilton et al. 1995; Vaudequin-Dransart et al. 1995). Accordingly, the Ti plasmid-free derivative of the reference strain C58 can metabolize dfgln via the product of genes located on the At plasmid that are highly homologous to genes found on the Ti plasmids (Baek et al. 2003). Remarkably, contrary to the situation with other opines, dfgln and dfglu occur widely in nature, *i.e.,* outside *Agrobacterium*-induced tumors. As with numerous Amadori compounds, dfgln and dfglu form spontaneously in decaying plant material (Anet 1957; Anet and Reynolds 1957). It is tempting to speculate that their common occurrence in nature provides a sufficient selective pressure to account for the emergence and selection of degradative functions, as seen in nonpathogenic strains of *Agrobacterium*. The duplication of the degradative opine genes and their incorporation into a "proto T-DNA" could have provided *Agrobacterium* with a way to force living plant cells to produce dfgln and dfglu. A further step in the evolution of the Ti plasmid could be the acquisition of opine anabolic and catabolic functions to allow the conversion of the two Amadori compounds to mannopine and mannopinic acid and later agropine (and vice versa) that are less accessible to competing organisms. Though entirely speculative, this model is nevertheless consistent with the physiological, biochemical, and molecular data available today.

The origins of other opine metabolic functions are even more speculative than those of the dfgln and mannityl opines. Octopine is synthesized in the muscle of marine animals such as octopus and squid during anaerobic muscle contraction (Thoai and Robin 1959) as a way to re-oxidize NADH, regenerate ATP, and reduce the concentrations of both arginine and pyruvate that accumulate under this condition (Grieshaber and Gäde 1976). Because marine agrobacterial isolates have been obtained (Rüger and Höfle 1992), a possibility exists that octopine degradation in these bacteria arose in relation with the presence of octopine in marine animals. In agreement, octopine-degrading bacteria have been isolated from mussels and oysters (Dion 1986). The related structures and sequence homologies of both the catabolic and anabolic genes for octopine and nopaline (Zanker et al. 1992, 1994) also suggest that these two opine systems may have evolved from a common ancestor.

The origin of sugar opines, such as the agrocinopines, is also unclear. Agrocinopine A is composed of sucrose linked to L-arabinose by a phosphodiester bond, whereas in Agrocinopine C, a D-glucose is present instead of the L-arabinose in agrocinopine A. Agrocinopines B and D differ from A and C, respectively, by lacking one sugar from the sucrose moiety (Ellis and Murphy 1981). In *Agrobacterium* strain C58, agrocinopine A is cleaved into arabinose-2-phosphate that is able to interact with AccR for activating quorum-sensing and Ti plasmid conjugation (El Sahili et al. 2015). Noticeably, arabinose-2-phosphate is uncommon

(unique until now) in the living world due to the unusual phosphate linkage on the C2 atom of the pyranose. This exemplifies the capacity of *Agrobacterium* to innovate by the use of a signal that is discriminable among the diverse sugars in plant hosts.

The occurrence of various opine anabolic and degradative systems may appear puzzling at first glance. However, the occurrence of multiple opine systems could indeed allow the diversification and coexistence of various agrobacterial populations in the same plant environment. These populations can therefore be considered as sympatric, and may eventually evolve novel species in further evolutionary steps (Lassalle et al. 2015). In agreement, whereas octopine or heliopine can be found in tumors incited by numerous *Agrobacterium* species, a number of opines such as cucumopine or ridéopine have been found only in grapevine tumors induced by members of the *A. vitis* species (Chilton et al. 2001). Similarly, cucumopine (or mikimopine) are detected only in hairy roots induced by *R. rhizogenes* (Davioud et al. 1988).

5.3 *Is* Agrobacterium*'s Ability to Transfer DNA to Organisms Belonging to Other Kingdoms Unique?*

Agrobacterium spp. and *R. rhizogenes*, due to the presence of Ti and highly related Ri plasmids, are to the best of our knowledge rare examples of bacteria naturally capable of transferring DNA to members of the eukaryote kingdom (Lacroix and Citovsky 2016). However, *Ensifer adhaerens*, a related bacterium, has recently been reported to transform plant cells when equipped with an appropriate plant transformation plasmid vector of the pCAMBIA series (Wendt et al. 2012). Aside from *Agrobacterium*, the only transkingdom DNA transfer that has been reported under laboratory conditions is between the pathogen *Bartonella henselae* and a human endothelial cell line (Schröder et al. 2011). *B. henselae* is not a major human pathogen except in immune-compromised patients where it may trigger a disease known as bacillary angiomatosis (Dehio 2005). A mobilisable cryptic plasmid from another *Bartonella* species was tagged with a fluorescent protein gene expressed only in eukaryotic backgrounds and introduced into a *B. henselae* strain that was used to infect endothelial cells. Post-infection, a low numbers of cells were fluorescent, indicating a T4SS-mediated transfer frequency of the plasmid of $\sim 2 \times 10^{-4}$. There is, however, no direct evidence that such a transfer may occur in animals, and no indication that such a transfer may lead to a permanent transformation of the recipient eukaryotic cells.

A recurring question is why no other bacteria have evolved comparable host transformation systems? First, to inquire whether other systems comparable to the Ti and Ri plasmids exist, 21 bp DNA border sequences have been compared to sequences of bacterial genomes or soil microbial metagenomes in data banks. The only hits identified were members of the two former genera (*Agrobacterium* and *Rhizobium*; Dessaux, unpublished). Second, the uniqueness of the *Agrobacterium*

spp. and *R. rhizogenes* transformation machinery could be explained by some of the evolutionary elements presented above which indicate that the occurrence of Ti and Ri plasmids proceeded in several steps, under selective pressures that may have rarely encountered in the living world. In other words, acquisition of Ti and Ri plasmids was quite an exceptional event.

In addition, once Ti and Ri plasmids evolved, it appears that their propagation in other bacteria was restricted by various factors. For instance, Ti plasmids do not replicate in firmicutes and actinobacteria, nor do they in beta- and gamma-proteobacteria such as *E. coli* or pseudomonads (Holsters et al. 1978; Dessaux, unpublished). Also, cloned Ti plasmid genes such as the opine catabolic genes are generally not expressed in other bacteria, including proteobacteria (Dessaux et al. 1987). Even in the related *Rhizobium* group where Ti plasmids replicate, tumor-inducing functions may or may not be expressed (Klein and Klein 1953; Hooykaas et al. 1977; van Veen et al. 1989) possibly because chromosomal genes involved in this process (see for instance Douglas et al. 1985; Close et al. 1985; Thomashow et al. 1987) may be missing. These data imply that transfer of the Ti plasmid outside the *Agrobacterium* genera, the *R. rhizogenes* species, and some *Rhizobium* species may be an evolutionary cul-de-sac either because the plasmid does not replicate or because the plasmid functions are not expressed. To a certain extent, and from an anthropomorphic point of view, *Agrobacterium* drastically protects the invention of the Ti plasmids that allow it to shift from a generalist behavior in the soil and the rhizosphere to a specialist behavior in the tumor where it escapes most microbial competitors and a part of plant defense.

References

Akiyoshi DE, Morris RO, Hinz R et al (1983) Cytokinin/auxin balance in crown gall tumors is regulated by specific loci in the T-DNA. Proc Natl Acad Sci USA 80:407–411

Ampomah OY, Avetisyan A, Hansen E et al (2013) The *thuEFGKAB* operon of rhizobia and *Agrobacterium tumefaciens* codes for transport of trehalose, maltitol, and isomers of sucrose and their assimilation through the formation of their 3-keto derivatives. J Bacteriol 195: 3797–3807

Anet EFLJ (1957) Chemistry of non-enzymic browning. II. Some crystalline amino acid-deoxy-sugars. Austral J Chem 10:193–197

Anet EFLJ, Reynolds TM (1957) Chemistry of non-enzymic browning. II. Reactions between amino acids, organic acids, and sugars in freeze-dried apricots and peaches. Austral J Chem 10:182–192

Aoki S, Kawaoka A, Sekine M et al (1994) Sequence of the cellular T-DNA in the untransformed genome of *Nicotiana glauca* that is homologous to ORFs 13 and 14 of the Ri plasmid and analysis of its expression in genetic tumours of *N. glauca* x *N. langsdorffii*. Mol Gen Genet 243:706–710

Baek CH, Farrand SK, Lee KE et al (2003) Convergent evolution of Amadori opine catabolic systems in plasmids of *Agrobacterium tumefaciens*. J Bacteriol 185:513–524

Baude J, Vial L, Villard C et al (2016) Coordinated regulation of species-specific hydroxycinnamic acid degradation and siderophore biosynthesis pathways in *Agrobacterium fabrum*. Appl Environ Microbiol 82:3515–3524

Beauchamp CJ, Chilton WS, Dion P et al (1990) Fungal catabolism of crown gall opines. Appl Environ Microbiol 56:150–155

Beck von Bodman S, Hayman GT, Farrand SK (1992) Opine catabolism and conjugal transfer of the nopaline Ti plasmid pTiC58 are coordinately regulated by a single repressor. Proc Natl Acad Sci USA 89:643–647

Bednárová M, Staněk M, Vancura V et al (1979) Microorganisms in the rhizosphere of wheat colonized by the fungus *Gaeumannomyces graminis* var. *tritici*. Folia Microbiol (Praha) 24:253–261

Bélanger C, Canfield ML, Moore LW et al (1995) Genetic analysis of nonpathogenic *Agrobacterium tumefaciens* mutants arising in crown gall tumors. J Bacteriol 177:3752–3757

Bhattacharya A, Sood P, Citovsky V (2010) The roles of plant phenolics in defence and communication during *Agrobacterium* and *Rhizobium* infection. Mol Plant Pathol 11:705–719

Brencic A, Eberhard A, Winans SC (2004) Signal quenching, detoxification and mineralization of *vir* gene-inducing phenolics by the VirH2 protein of *Agrobacterium tumefaciens*. Mol Microbiol 51:1103–1115

Bolton GW, Nester EW, Gordon MP (1986) Plant phenolic compounds induce expression of the *Agrobacterium tumefaciens* loci needed for virulence. Science 232:983–985

Boncompagni E, Osteras M, Poggi MC, le Rudulier D (1999) Occurrence of choline and glycine betaine uptake and metabolism in the family rhizobiaceae and their roles in osmoprotection. Appl Environ Microbiol 65:2072–2077

Bouzar H, Chilton WS, Nesme X, Dessaux Y, Vaudequin V, Petit A, Jones JB, Hodge NC (1995) A new *Agrobacterium* strain isolated from aerial tumors on *Ficus benjamina* L. Appl Environ Microbiol 61:65–73

Bouzar H, Moore LW (1987) Isolation of different *Agrobacterium* biovars from a natural oak savanna and tallgrass prairie. Appl Environ Microbiol 53:717–721

Bouzar H, Ouadah D, Krimi K et al (1993) Correlative association between resident plasmids and the host chromosome in a diverse *Agrobacterium* soil population. Appl Environ Microbiol 59:1310–1317

Bulgarelli D, Rott M, Schlaeppi K et al (2012) Revealing structure and assembly cues for *Arabidopsis* root-inhabiting bacterial microbiota. Nature 48:91–95

Burr TJ, Katz BH, Bishop AL (1987) Populations of *Agrobacterium* in vineyard and nonvineyard soils and grape roots in vineyards and nurseries. Plant Dis 71:617–620

Cardarelli M, Mariotti D, Pomponi M et al (1987) *Agrobacterium rhizogenes* T-DNA genes capable of inducing hairy root phenotype. Mol Gen Genet 209:475–480

Caretto S, Linsalata V, Colella G et al (2015) Carbon fluxes between primary metabolism and phenolic pathway in plant tissues under stress. Int J Mol Sci 16:26378–26394. https://doi.org/10.3390/ijms161125967

Carlier A, Chevrot R, Dessaux Y et al (2004) The assimilation of gamma-butyrolactone in *Agrobacterium tumefaciens* C58 interferes with the accumulation of the N-acyl-homoserine lactone signal. Mol Plant Microbe Interact 17:951–957

Chen K, Otten L (2017) Natural *Agrobacterium* transformants: recent results and some theoretical considerations. Front Plant Sci 8:1600. https://doi.org/10.3389/fpls.2017.01600

Chen L, Chen Y, Wood DW et al (2002) A new type IV secretion system promotes conjugal transfer in *Agrobacterium tumefaciens*. J Bacteriol 184:4838–4845

Chevrot R, Rosen R, Haudecoeur E et al (2006) GABA controls the level of quorum-sensing signal in *Agrobacterium tumefaciens*. Proc Natl Acad Sci USA 103:7460–7464

Chilton MD, Tepfer DA, Petit A et al (1982) *Agrobacterium rhizogenes* inserts T-DNA into the genomes of the host plant root cells. Nature 295:432–434

Chilton WS, Chilton MD (1984) Mannityl opine analogs allow isolation of catabolic pathway regulatory mutants. J Bacteriol 158:650–658

Chilton WS, Petit A, Chilton MD, Dessaux Y (2001) Structure and characterization of the crown gall opines heliopine, vitopine and ridéopine. Phytochemistry 58:137–142

Chilton WS, Stomp AM, Beringue V et al (1995) The chrysopine family of amadori-type crown gall opines. Phytochemistry 40:619–628

Christie PJ (2016) The mosaic Type IV secretion systems. EcoSal Plus 7(1). https://doi.org/10.1128/ecosalplus.esp-0020-2015

Close TJ, Tait RC, Kado CI (1985) Regulation of Ti plasmid virulence genes by a chromosomal locus of *Agrobacterium tumefaciens*. J Bacteriol 164:774–781

Costantino P, Hooykaas PJJ, Dendulk-Ras H et al (1980) Tumor formation and rhizogenicity of *Agrobacterium rhizogenes* carrying Ti plasmids. Gene 11:79–87

Cushnie TTP, Lamb AJ (2005) Antimicrobial activity of flavonoids. Int J Antimicrob Agents 26:342–356

Davioud E, Petit A, Tate ME et al (1988) Cucumopine—a new T-DNA-encoded opine in hairy root and crown gall. Phytochemistry 27:2429–2433

Deeken R, Engelmann JC, Efetova M et al (2006) An integrated view of gene expression and solute profiles of *Arabidopsis* tumors: a genome-wide approach. Plant Cell 18:3617–3634

De Greve H, Dhaese P, Seurinck J et al (1982) Nucleotide sequence and transcript map of the *Agrobacterium tumefaciens* Ti plasmid-encoded octopine synthase gene. J Mol Appl Genet 1:499–511

Dehio C (2005) *Bartonella*-host-cell interactions and vascular tumour formation. Nat Rev Microbiol 3:621–631

De Lajudie PM, Young JPW (2017) International Committee on Systematics of Prokaryotes Subcommittee for the Taxonomy of *Rhizobium* and *Agrobacterium*—minutes of the meeting, Budapest, 25 August 2016. Int J Syst Evol Microbiol. https://doi.org/10.1099/ijsem.0.002144

De Paolis A, Mauro ML, Pomponi M et al (1985) Localization of agropine-synthesizing functions in the TR region of the root-inducing plasmid of *Agrobacterium rhizogenes* 1855. Plasmid 13:1–7

Dessaux Y, Grandclément C, Faure D (2016) Engineering the rhizosphere. Trends Plant Sci 21:266–278

Dessaux Y, Guyon P, Farrand SK et al (1986a) *Agrobacterium* Ti and Ri plasmids specify enzymic lactonization of mannopine to agropine. J Gen Microbiol 132:2549–2559

Dessaux Y, Guyon P, Petit A et al (1988) Opine utilization by *Agrobacterium* spp.: octopine-type Ti plasmids encode two pathways for mannopinic acid degradation. J Bacteriol 170:2939–2946

Dessaux Y, Petit A, Ellis JG et al (1989) Ti plasmid-controlled chromosome transfer in *Agrobacterium tumefaciens*. J Bacteriol 171:6363–6366

Dessaux Y, Petit A, Farrand SK et al (1998) Opine and opine-like molecules involved in plant-Rhizobiaceaea interactions. In: Spaink HP, Kondorosi A, Hooykaas P (eds) The Rhizobiaceae. Kluwer Academic Publishers, Dordrecht, pp 173–197

Dessaux Y, Petit A, Tempé J et al (1986b) Arginine catabolism in *Agrobacterium* strains: role of the Ti plasmid. J Bacteriol 166:44–50

Dessaux Y, Tempé J, Farrand SK (1987) Genetic analysis of mannityl opine catabolism in octopine-type *Agrobacterium tumefaciens* strain 15955. Mol Gen Genet 208:301–308

Dion P (1986) Utilization of octopine by marine bacteria isolated from mollusks. Can J Microbiol 32:959–963

Douglas CJ, Staneloni RJ, Rubin RA et al (1985) Identification and genetic analysis of an *Agrobacterium tumefaciens* chromosomal virulence region. J Bacteriol 161:850–860

Durand-Tardif M, Broglie R, Slightom J (1985) Structure and expression of Ri T-DNA from *Agrobacterium rhizogenes* in *Nicotiana tabacum*. Organ and phenotypic specificity. J Mol Biol 186:557–564

Dunoyer P, Himber C, Voinnet O (2006) Induction, suppression and requirement of RNA silencing pathways in virulent *Agrobacterium tumefaciens* infections. Nat Genet 38:258–263

Ellis JG, Kerr A, Tempé J et al (1979) Arginine catabolism: a new function of both octopine and nopaline Ti-plasmids of *Agrobacterium*. Mol Gen Genet 173:263–269

Ellis JG, Kerr A, Petit A et al (1982) Conjugal transfer of nopaline and agropine Ti-plasmids—the role of agrocinopines. Mol Gen Genet 186:269–274

Ellis JG, Murphy PJ (1981) Four new opines from crown gall tumours—their detection and properties. Mol Gen Genet 181:36–43

Ellis JG, Ryder MH, Tate ME (1984) *Agrobacterium tumefaciens* TR-DNA encodes a pathway for agropine biosynthesis. Mol Gen Genet 195:466–473

El Sahili A, Li SZ, Lang J et al (2015) A pyranose-2-phosphate motif is responsible for both antibiotic import and quorum-sensing regulation in *Agrobacterium tumefaciens*. PLoS Pathog 11(8):e1005071. https://doi.org/10.1371/journal.ppat.1005071

Engström P, Zambryski P, Van Montagu M et al (1987) Characterization of *Agrobacterium tumefaciens* virulence proteins induced by the plant factor acetosyringone. J Mol Biol 197:635–645

Farrand SK, Hwang I, Cook DM (1996) The *tra* region of the nopaline-type Ti plasmid is a chimera with elements related to the transfer systems of RSF1010, RP4, and F. J Bacteriol 178:4233–4247

Flores-Mireles AL, Eberhard A, Winans SC (2012) *Agrobacterium tumefaciens* can obtain sulphur from an opine that is synthesized by octopine synthase using S-methylmethionine as a substrate. Mol Microbiol 84:845–856

Fortin C, Nester EW, Dion P (1992) Growth inhibition and loss of virulence in cultures of *Agrobacterium* tumefaciens treated with acetosyringone. J Bacteriol 174:5676–5685

Fortin C, Marquis C, Nester EW et al (1993) Dynamic structure of *Agrobacterium tumefaciens* Ti plasmids. J Bacteriol 175:4790–4799

Garg N, Manchanda G, Kumar A (2014) Bacterial quorum sensing: circuits and applications. Antonie Van Leeuwenhoek 105:289–305

Gelvin SB (2012) Traversing the Cell: *Agrobacterium* T-DNA's journey to the host genome. Front Plant Sci 3:52. https://doi.org/10.3389/fpls.2012.00052

Gelvin SB (2017) Integration of *Agrobacterium* T-DNA into the plant genome. Annu Rev Genet. https://doi.org/10.1146/annurev-genet-120215-035320

Genetello C, Van Larebeke N, Holsters M et al (1977) Ti plasmids of *Agrobacterium* as conjugative plasmids. Nature 265:561–563

Gohlke J, Deeken R (2014) Plant responses to *Agrobacterium tumefaciens* and crown gall development. Front Plant Sci 5:155. https://doi.org/10.3389/fpls.2014.00155

Gomes NCM, Heuer H, Schönfeld J et al (2001) Bacterial diversity of the rhizosphere of maize (*Zea mays*) grown in tropical soil studied by temperature gradient gel electrophoresis. Plant Soil 232:167–180

Gordon DM, Ryder MH, Heinrich K et al (1996) An experimental test of the rhizopine concept in *Rhizobium meliloti*. Appl Environ Microbiol 62:3991–3996

Grandclément C, Tannières M, Moréra S et al (2016) Quorum quenching: Role in nature and applied developments. FEMS Microbiol Rev 40:86–116

Grieshaber M, Gäde G (1976) The biological role of octopine in the squid, *Loligo vulgaris* (Lamarck). J Comp Physiol 108:225–232

Guyon P, Chilton MD, Petit A et al (1980) Agropine in "null-type" crown gall tumors: Evidence for generality of the opine concept. Proc Natl Acad Sci USA 77:2693–2697

Guyon P, Petit A, Tempé J et al (1993) Transformed plants producing opines specifically promote growth of opine-degrading agrobacteria. Mol Plant-Microbe Interact 6:92–98

Haudecoeur E, Tannières M, Cirou A et al (2009) Different regulation and roles of lactonases AiiB and AttM in *Agrobacterium tumefaciens* C58. Mol Plant Microbe Interact 22:529–537

Hawes MC, Smith LY (1989) Requirement for chemotaxis in pathogenicity of *Agrobacterium tumefaciens* on roots of soil grown pea plants. J Bacteriol 171:5668–5671

Hawes MC, Pueppke SG (1987) Correlation between binding of *Agrobacterium tumefaciens* by root cap cells and susceptibility of plants to crown gall. Plant Cell Rep 6:287–290

Hood EE, Chilton WS, Chilton MD et al (1986) T-DNA and opine synthetic loci in tumors incited by *Agrobacterium tumefaciens* A281 on soybean and alfalfa plants. J Bacteriol 168:1283–1290

He F, Nair GR, Soto CS et al (2009) Molecular basis of ChvE function in sugar binding, sugar utilization, and virulence in *Agrobacterium tumefaciens*. J Bacteriol 191:5802–5813

Heinrich K, Gordon DM, Ryder MH et al (1999) A rhizopine strain of *Sinorhizobium meliloti* remains at a competitive nodulation advantage after an extended period in the soil. Soil Biol Biochem 31:1063–1065

Hinsinger P, Bengough AG, VetterleinIain D et al (2009) Rhizosphere: biophysics, biogeochemistry and ecological relevance. Plant Soil 321:117–152

Holsters M, Silva B, Van Vliet F et al (1978) In vivo transfer of the Ti-plasmid of *Agrobacterium tumefaciens* to *Escherichia coli*. Mol Gen Genet 163:335–358

Hong SB, Hwang I, Dessaux Y (1997) A T-DNA gene required for agropine biosynthesis by transformed plants is functionally and evolutionarily related to a Ti plasmid gene required for catabolism of agropine by *Agrobacterium* strains. J Bacteriol 179:4831–4840

Hooykaas PJJ, Klapwijk PM, Nuti MP et al (1977) Transfer of the *Agrobacterium tumefaciens* Ti-plasmid to avirulent agrobacteria and to *Rhizobium ex planta*. J Gen Microbiol 98:477–484

Hwang I, Li PL, Zhang L, Piper KR et al (1994) TraI, a LuxI homologue, is responsible for production of conjugation factor, the Ti plasmid N-acylhomoserine lactone autoinducer. Proc Natl Acad Sci USA 91:4639–4643

Hynes MF, Simon R, Pühler A (1985) The development of plasmid-free strains of *Agrobacterium tumefaciens* by using incompatibility with a *Rhizobium meliloti* plasmid to eliminate pAtC58. Plasmid 13:99–105

Iacobellis NS, Devay JE (1986) Long-term storage of plant-pathogenic bacteria in sterile distilled water. Appl Environ Microbiol 52:388–389

Idnurm A, Bailey AM, Cairns TC et al (2017) A silver bullet in a golden age of functional genomics: the impact of *Agrobacterium*-mediated transformation of fungi. Fungal Biol Biotechnol 4:6. https://doi.org/10.1186/s40694-017-0035-0

Inceoglu Ö, Al-Soud WA, Salles JF et al (2011) Comparative analysis of bacterial communities in a potato field as determined by pyrosequencing. PLoS One 63(2):460–470

Jones DL, Nguyen C, Finlay RD (2009) Carbon flow in the rhizosphere: carbon trading at the soil–root interface. Plant Soil 321:5–33

Joos H, Inzé D, Caplan A (1983) Genetic analysis of T-DNA transcripts in nopaline crown galls. Cell 32:1057–1067

Jouanin L (1984) Restriction map of an agropine-type Ri plasmid and its homologies with Ti plasmids. Plasmid 12:91–102

Kanemoto RH, Powell AT, Akiyoshi DE et al (1989) Nucleotide sequence and analysis of the plant-inducible locus *pinF* from *Agrobacterium tumefaciens*. J Bacteriol 171:2506–2512

Kefeli VI, Kalevitch MV, Borsari B (2003) Phenolic cycle in plants and environment. J Cell Mol Biol 2:13–18

Kerr A, Manigault P, Tempé J (1977) Transfer of virulence in vivo and in vitro in *Agrobacterium*. Nature 265:560–561

Khan SR, Gaines J, Roop RM et al (2008) Broad-host-range expression vectors with tightly regulated promoters and their use to examine the influence of TraR and TraM expression on Ti plasmid quorum sensing. Appl Environ Microbiol 74:5053–5062

Khanaka H, Catteau M, Tailliez R (1981) Antibiotic sensitivity in *Rhizobium* and *Agrobacterium*. Zentralbl Bakt Mikrobiol Hygiene I Abt Orig C2:282–288

Kim H, Farrand SK (1997) Characterization of the *acc* operon from the nopaline-type Ti plasmid pTiC58, which encodes utilization of agrocinopines A and B and susceptibility to agrocin 84. J Bacteriol 179:7559–7572

Kim H, Farrand SK (1988) Opine catabolic loci from *Agrobacterium* plasmids confer chemotaxis to their cognate substrates. Mol Plant Microbe Interact 11:131–143

Kim KS, Chilton WS, Farrand SK (1996) A Ti plasmid-encoded enzyme required for degradation of mannopine is functionally homologous to the T-region-encoded enzyme required for synthesis of this opine in crown gall tumors. J Bacteriol 178:3285–3292

Kim KS, Farrand SK (1996) Ti plasmid-encoded genes responsible for catabolism of the crown gall opine mannopine by *Agrobacterium tumefaciens* are homologs of the T-region genes responsible for synthesis of this opine by the plant tumor. J Bacteriol 178:3275–3284

Klapwijk PM, Oudshoorn M, Schilperoort RA (1977) Inducible permease involved in the uptake of octopine, lysopine and octopinic acid by *Agrobacterium tumefaciens* strains carrying virulence-associated plasmids. J Gen Microbiol 102:1–11

Klapwijk PM, Scheulderman T, Schilperoort RA (1978) Coordinated regulation of octopine degradation and conjugative transfer of Ti plasmids in *Agrobacterium tumefaciens*: evidence for a common regulatory gene and separate operons. J Bacteriol 136:775–785

Klein DT, Klein RM (1953) Transmittance of tumor-inducing ability to avirulent crown-gall and related bacteria. J Bacteriol 66:220–228

Koncz C, De Greve H, André D et al (1983) The opine synthase genes carried by Ti plasmids contain all signals necessary for expression in plants. EMBO J 2:1597–1603

Krimi Z, Petit A, Mougel C et al (2002) Seasonal fluctuations and long-term persistence of pathogenic populations of *Agrobacterium* spp. in soils. Appl Environ Microbiol 68:3358–3365

Kuzmanović N, Puławska J, Prokić A et al (2015) *Agrobacterium arsenijevicii* sp. nov., isolated from crown gall tumors on raspberry and cherry plum. Syst Appl Microbiol 38:373–378

Kyndt T, Quispe D, Zhai H et al (2015) The genome of cultivated sweet potato contains *Agrobacterium* T-DNAs with expressed genes: an example of a naturally transgenic food crop. Proc Natl Acad Sci USA 112:5844–5849

Lacroix B, Citovsky V (2013) The roles of bacterial and host plant factors in *Agrobacterium*-mediated genetic transformation. Int J Dev Biol 57:467–481

Lacroix B, Citovsky V (2016) Transfer of DNA from bacteria to eukaryotes. MBio 7(4). pii: e00863-16

Lang J, Gonzalez-Mula A, Taconnat L et al (2016) The plant GABA signaling downregulates horizontal transfer of the *Agrobacterium tumefaciens* virulence plasmid. New Phytol 210: 974–983

Lang J, Planamente S, Mondy S et al (2013) Concerted transfer of the virulence Ti plasmid and companion At plasmid in the *Agrobacterium tumefaciens*-induced plant tumour. Mol Microbiol 90:1178–1189

Lang J, Vigouroux A, Planamente S et al (2014) *Agrobacterium* uses a unique ligand-binding mode for trapping opines and acquiring a competitive advantage in the niche construction on plant host. PLoS Pathog 10(10):e1004444

Lassalle F, Muller D, Nesme X (2015) Ecological speciation in bacteria: reverse ecology approaches reveal the adaptive part of bacterial cladogenesis. Res Microbiol 166:729–741

Lattanzio V, Kroon PA, Quideau S et al (2008) Plant phenolics—secondary metabolites with diverse functions. In: Daayf F, Lattanzio V (eds) Recent advances in polyphenol research. Wiley, Oxford, pp 1–35

Li PL, Hwang I, Miyagi H et al (1999) Essential components of the Ti plasmid trb system, a type IV macromolecular transporter. J Bacteriol 181:5033–5041

Li PL, Farrand SK (2000) The replicator of the nopaline-type Ti plasmid pTiC58 is a member of the *repABC* family and is influenced by the TraR-dependent quorum-sensing regulatory system. J Bacteriol 182:179–188

Llop P, Murillo J, Lastra B et al (2009) Recovery of nonpathogenic mutant bacteria from tumors caused by several *Agrobacterium tumefaciens* strains: a frequent event? Appl Environ Microbiol 75:6504–6514

Lundberg DS, Lebeis SL, Paredes SH et al (2012) Defining the core *Arabidopsis thaliana* root microbiome. Nature 488:86–90

Ma LS, Hachani A, Lin JS et al (2014) *Agrobacterium tumefaciens* deploys a superfamily of type VI secretion DNase effectors as weapons for interbacterial competition in planta. Cell Host Microbe 16:94–104

Mansouri H, Petit A, Oger P et al (2002) Engineered rhizosphere: the trophic bias generated by opine-producing plants is independent of the opine type, the soil origin, and the plant species. Appl Environ Microbiol 68:2562–2566

Marasco R, Lago CT, De Felice M (1995) Utilization of cellobiose and other beta-D-glucosides in *Agrobacterium tumefaciens*. Res Microbiol 146:485–492

Marty L, Vigouroux A, Aumont-Nicaise M et al (2016) Structural basis for high specificity of Amadori compound and mannopine opine binding in bacterial pathogens. J Biol Chem 291:22638–22649

Matveeva TV, Bogomaz DI, Pavlova OA et al (2012) Horizontal gene transfer from genus *Agrobacterium* to the plant *Linaria* in nature. Mol Plant Microbe Interact 25:1542–1551

Merlo DJ, Nester EW (1977) Plasmids in avirulent strains of *Agrobacterium*. J Bacteriol 129: 76–80

Mhedbi-Hajri N, Yahiaoui N, Mondy S et al (2016) Transcriptome analysis revealed that a quorum sensing system regulates the transfer of the pAt megaplasmid in *Agrobacterium tumefaciens*. BMC Genomics 17:661

Mondy S, Lenglet A, Beury-Cirou A et al (2014) An increasing opine carbon bias in artificial exudation systems and genetically modified plant rhizospheres leads to an increasing reshaping of bacterial populations. Mol Ecol 23:4846–4861

Montoya AL, Chilton MD, Gordon MP et al (1977) Octopine and nopaline metabolism in *Agrobacterium tumefaciens* and crown gall tumor cells: role of plasmid genes. J Bacteriol 129:101–107

Moore LW, Chilton WS, Canfield ML (1997) Diversity of opines and opine-catabolizing bacteria isolated from naturally occurring crown gall tumors. Appl Environ Microbiol 63:201–207

Moore L, Warren G, Strobel G (1979) Involvement of a plasmid in the hairy root disease of plants caused by *Agrobacterium rhizogenes*. Plasmid 2:617–626

Morton ER, Platt TG, Fuqua C et al (2014) Non-additive costs and interactions alter the competitive dynamics of co-occurring ecologically distinct plasmids. Proc Biol Sci 281 (1779):20132173

Mousavi SA, Österman J, Wahlberg N et al (2014) Phylogeny of the *Rhizobium-Allorhizobium-Agrobacterium* clade supports the delineation of *Neorhizobium* gen. nov. Syst Appl Microbiol 37:208–215

Murphy PJ, Heycke N, Banfalvi Z et al (1987) Genes for the catabolism and synthesis of an opine-like compound in *Rhizobium meliloti* are closely linked and on the Sym plasmid. Proc Natl Acad Sci USA 84:493–497

Nair GR, Lai X, Wise AA et al (2011) The integrity of the periplasmic domain of the VirA sensor kinase is critical for optimal coordination of the virulence signal response in *Agrobacterium tumefaciens*. J Bacteriol 193:1436–1448

Nautiyal CS, Dion P (1990) Characterization of the opine-utilizing microflora associated with samples of soil and plants. Appl Environ Microbiol 56:2576–2579

Nautiyal CS, Dion P, Chilton WS (1991) Mannopine and mannopinic acid as substrates for *Arthrobacter* sp. strain MBA209 and *Pseudomonas putida* NA513. J Bacteriol 173:2833–2841

Nester EW (2015) *Agrobacterium*: nature's genetic engineer. Front Plant Sci 5:730

Nobile S, Deshusses J (1986) Transport of gamma-butyrobetaine in an *Agrobacterium* species isolated from soil. J Bacteriol 168:780–784

Nüsslein K, Tiedje JM (1998) Characterization of the dominant and rare members of a young Hawaiian soil bacterial community with small-subunit ribosomal DNA amplified from DNA fractionated on the basis of its guanine and cytosine composition. Appl Environ Microbiol 64:1283–1289

Ooms G, Hooykaas PJ, Moolenaar G et al (1981) Grown gall plant tumors of abnormal morphology, induced by *Agrobacterium tumefaciens* carrying mutated octopine Ti plasmids; analysis of T-DNA functions. Gene 14:33–50

Ong SA, Peterson T, Neilands JB (1979) Agrobactin, a siderophore from *Agrobacterium tumefaciens*. J Biol Chem 254:1860–1865

Oger P, Petit A, Dessaux Y (1997) Genetically engineered plants producing opines alter their biological environment. Nat Biotechnol 15:369–372

Otten L, De Ruffray P (1994) *Agrobacterium vitis* nopaline Ti plasmid pTiAB4: relationship to other Ti plasmids and T-DNA structure. Mol Gen Genet 245:493–505

Otten L, Gérard JC, De Ruffray P (1993) The Ti plasmid from the wide host range *Agrobacterium vitis* strain Tm4: map and homology with other Ti plasmids. Plasmid 29:154–159

Parke D, Ornston LN, Nester EW (1987) Chemotaxis to plant phenolic inducers of virulence genes is constitutively expressed in the absence of the Ti plasmid in *Agrobacterium tumefaciens*. J Bacteriol 169:5336–5338

Palanichelvam K, Oger P, Clough SJ et al (2000) A second T-region of the soybean-supervirulent chrysopine-type Ti plasmid pTiChry5, and construction of a fully disarmed vir helper plasmid. Mol Plant Microbe Interact 13:1081–1091

Palumbo JD, Kado CI, Phillips DA (1998) An isoflavonoid-inducible efflux pump in *Agrobacterium tumefaciens* is involved in competitive colonization of roots. J Bacteriol 180:3107–3113

Papenfort K, Bassler BL (2016) Quorum sensing signal-response systems in Gram-negative bacteria. Nat Rev Microbiol 14:576–588

Penyalver R, Oger P, López MM et al (2001) Iron-binding compounds from *Agrobacterium* spp.: biological control strain *Agrobacterium rhizogenes* K84 produces a hydroxamate siderophore. Appl Environ Microbiol 67:654–664

Petit A, David C, Dahl G (1983) Further extension of the opine concept: plasmids in *Agrobacterium rhizogenes* cooperate for opine degradation. Mol Gen Genet 190:204–214

Petit A, Tempé J (1985) The function of T-DNA in nature. In: Van Vloten-Doting L, Groot GSP, Hall TC (eds) Molecular form and function of the plant genome. Plenum Press, New York/ London, pp 625–636

Petit A, Tempé J, Kerr A et al (1978) Substrate induction of conjugative activity of *Agrobacterium tumefaciens* Ti plasmids. Nature 271:570–572

Pionnat S, Keller H, Héricher D et al (1999) Ti plasmids from *Agrobacterium* characterize rootstock clones that initiated a spread of crown gall disease in Mediterranean countries. Appl Environ Microbiol 65:4197–4206

Piper KR, Beck von Bodman S, Farrand SK (1993) Conjugation factor of *Agrobacterium tumefaciens* regulates Ti plasmid transfer by autoinduction. Nature 362:448–450

Piper KR, Beck Von Bodman S, Hwang I (1999) Hierarchical gene regulatory systems arising from fortuitous gene associations: controlling quorum sensing by the opine regulon in *Agrobacterium*. Mol Microbiol 32:1077–1089

Pitzschke A, Hirt H (2010) New insights into an old story: *Agrobacterium*-induced tumour formation in plants by plant transformation. EMBO J 29:1021–1032

Platt TG, Bever JD, Fuqua C (2012) A cooperative virulence plasmid imposes a high fitness cost under conditions that induce pathogenesis. Proc Biol Sci 279:1691–1699

Powell GK, Morris RO (1986) Nucleotide sequence and expression of *a Pseudomonas savastanoi* cytokinin biosynthetic gene: homology with *Agrobacterium tumefaciens tmr* and *tzs* loci. Nucleic Acids Res 14:2555–2565

Qin Y, Su S, Farrand SK (2007) Molecular basis of transcriptional antiactivation. TraM disrupts the TraR-DNA complex through stepwise interactions. J Biol Chem 282:19979–19991

Ream LW, Gordon MP, Nester EW (1983) Multiple mutations in the T region of the *Agrobacterium tumefaciens* tumor-inducing plasmid. Proc Natl Acad Sci USA 80:1660–1664

Riley PS, Weaver RE (1977) Comparison of thirty-seven strains of Vd-3 bacteria with *Agrobacterium radiobacter*: morphological and physiological observations. J Clin Microbiol 5:172–177

Rondon MR, Ballering KS, Thomas MG (2004) Identification and analysis of a siderophore biosynthetic gene cluster from *Agrobacterium tumefaciens* C58. Microbiology 150:3857–3866

Rong LJ, Karcher SJ, Gelvin SB (1991) Genetic and molecular analyses of *picA*, a plant-inducible locus on the *Agrobacterium tumefaciens* chromosome. J Bacteriol 173:5110–5120

Rosenberg S, Huguet T (1984) The pAtC58 plasmid of *Agrobacterium tumefaciens* is not essential for tumor induction. Mol Gen Genet 196:533–536

Rossbach S, Rasul G, Schneider M et al (1995) Structural and functional conservation of the rhizopine catabolism (*moc*) locus is limited to selected *Rhizobium meliloti* strains and unrelated to their geographical origin. Mol Plant Microbe Interact 8:549–559

Rüger HJ, Höfle MG (1992) Marine star-shaped-aggregate-forming bacteria: *Agrobacterium atlanticum* sp. nov.; *Agrobacterium meteori* sp. nov.; *Agrobacterium ferrugineum* sp. nov.,

nom. rev.; *Agrobacterium gelatinovorum* sp. nov., nom. rev.; and *Agrobacterium stellulatum* sp. nov., nom. rev. Int J Syst Bacteriol 42:133–143

Ryu CM (2015) Against friend and foe: type 6 effectors in plant-associated bacteria. J Microbiol 53:201–208

Saint CP, Wexler M, Murphy PJ et al (1993) Characterization of genes for synthesis and catabolism of a new rhizopine induced in nodules by *Rhizobium meliloti* Rm220-3: extension of the rhizopine concept. J Bacteriol 175:5205–5215

Sánchez-Cañizares C, Jorrín B et al (2017) Understanding the holobiont: the interdependence of plants and their microbiome. Curr Opin Microbiol 38:188–196

Savka MA, Black RC, Binns AN et al (1996) Translocation and exudation of tumor metabolites in crown galled plants. Mol Plant Microbe Interact 9:310–313

Savka MA, Dessaux Y, McSpadden-Gardener BB et al (2013) The "biased rhizosphere" concept and advances in the omics era to study bacterial competitiveness and persistence in the phytosphere. In: de Bruijn FJ (ed) Molecular microbial ecology of the rhizosphere, vol 2. Wiley, Hoboken, pp 1147–1161

Savka MA, Dessaux Y, Oger P et al (2002) Engineering bacterial competitiveness and persistence in the phytosphere. Mol Plant Microbe Interact 15:866–874

Savka MA, Farrand SK (1992) Mannityl opine accumulation and exudation by transgenic tobacco. Plant Physiol 98:784–789

Savka MA, Farrand SK (1997) Modification of rhizobacterial populations by engineering bacterium utilization of a novel plant-produced resource. Nat Biotechnol 15:363–368

Schell J, Van Montagu M, De Beuckeleer M et al (1979) Interactions and DNA transfer between *Agrobacterium tumefaciens*, the Ti-plasmid and the plant host. Proc R Soc Lond B Biol Sci 204:251–266

Schröder G, Schuelein R, Quebatte M, Dehio C (2011) Conjugative DNA transfer into human cells by the VirB/VirD4 type IV secretion system of the bacterial pathogen *Bartonella henselae*. Proc Natl Acad Sci USA 108:14643–14648

Schroth MN, Weinhold AR, McCain AH (1971) Biology and control of *Agrobacterium tumefaciens*. Hilgardia 40:537–552

Scott DB, Wilson R, Shaw GJ (1987) Biosynthesis and degradation of nodule-specific *Rhizobium loti* compounds in Lotus nodules. J Bacteriol 169:278–282

Shaw CH (1991) Swimming against the tide: chemotaxis in *Agrobacterium*. Bioessays 13:25–29

Shaw GJ, Wilson RD, Lane GA (1986) Structure of rhizolotine, a novel opine-like metabolite from *Lotus tenuis* nodules. J Chem Soc Chem Commun 2:180–181

Slightom JL, Durand-Tardif M (1986) Nucleotide sequence analysis of TL-DNA of *Agrobacterium rhizogenes* agropine type plasmid. Identification of open reading frames. J Biol Chem 261:108–121

Smith LT, Smith GM, Madkour MA (1990) Osmoregulation in *Agrobacterium tumefaciens*: accumulation of a novel disaccharide is controlled by osmotic strength and glycine betaine. J Bacteriol 172:6849–6855

Soedarjo M, Hemscheidt TK, Borthakur D (1994) Mimosine, a toxin present in leguminous trees (*Leucaena* spp.), induces a mimosine-degrading enzyme activity in some *Rhizobium* strains. Appl Environ Microbiol 60:4268–4272

Soedarjo M, Borthakur D (1998) Mimosine, a toxin produced by the tree-legume *Leucaena* provides a nodulation competition advantage to mimosine-degrading *Rhizobium* strains. Soil Biol Biochem 30:1605–1613

Spena A, Schmülling T, Koncz C et al (1987) Independent and synergistic activity of *rolA*, *B* and *C* loci in stimulating abnormal growth in plants. EMBO J 6:3891–3899

Subramoni S, Nathoo N, Klimov E (2014) *Agrobacterium tumefaciens* responses to plant-derived signaling molecules. Front Plant Sci 5:322

Sukanya NK, Vaidyanathan CS (1964) Aminotransferases of *Agrobacterium tumefaciens*: transamination between tryptophan and phenylpyruvate. Biochem J 92:594–598

Süle S (1978) Biological control of crown gall by a peat cultured antagonist. Phytopath Z 91: 273–275

Suzuki K, Yamashita I, Tanaka N (2002) Tobacco plants were transformed by *Agrobacterium rhizogenes* infection during their evolution. Plant J 32:775–787

Tannières M, Lang J, Barnier C et al (2017) Quorum-quenching limits quorum-sensing exploitation by signal-negative invaders. Sci Rep 7:40126

Teixeira LC, Peixoto RS, Cury JC et al (2010) Bacterial diversity in rhizosphere soil from Antarctic vascular plants of Admiralty Bay, maritime Antarctica. ISME J 4(2010):989–1001

Tempé J, Petit A (1983) La piste des opines. In: Pühler A (ed) Molecular genetics of the bacteria-plant interaction. Springer, Berlin-Heidelberg, pp 14–32

Tennigkeit J, Matzura H (1991) Nucleotide sequence analysis of a chloramphenicol-resistance determinant from *Agrobacterium tumefaciens* and identification of its gene product. Gene 8:113–116

Tepfer D, Goldmann A, Pamboukdjian N et al (1988) A plasmid of *Rhizobium meliloti* 41 encodes catabolism of two compounds from root exudate of *Calystegium sepium*. J Bacteriol 170: 1153–1161

Tepfer DA, Tempé J (1981) Production d'agropine par des racines formées sous l'action d'*Agrobacterium rhizogenes* souche A4. C R Hebd Séances Acad Sci Ser III 292:153–156

Teyssier-Cuvelle S, Mougel C, Nesme X (1999) Direct conjugal transfers of Ti plasmid to soil microflora. Mol Ecol 8:1273–1284

Thoai NV, Robin Y (1959) Métabolisme des dérivés guanidylés: VIII. Biosynthèse de l'octopine et répartition de l'enzyme l'opérant chez les invertebrés. Biochimica Biophysica Acta 35: 446–453

Thomashow MF, Karlinsey JE, Marks JR et al (1987) Identification of a new virulence locus in *Agrobacterium tumefaciens* that affects polysaccharide composition and plant cell attachment. J Bacteriol 169:3209–3216

Tremblay G, Gagliardo R, Chilton WS et al (1987) Diversity among opine-utilizing bacteria: identification of coryneform isolates. Appl Environ Microbiol 53:1519–1524

Quispe-Huamanquispe DG, Gheysen G, Kreuze JF (2017) Horizontal gene transfer contributes to plant evolution: the case of *Agrobacterium T-DNAs*. Front Plant Sci 8:2015

Vain P (2007) Thirty years of plant transformation technology development. Plant Biotechnol J 5:221–229

Van Larebeke N, Engler G, Holsters M et al (1974) Large plasmid in *Agrobacterium tumefaciens* essential for crown gall-inducing ability. Nature 252:169–170

van Veen RJM, den Dulk-Ras H, Schilperoort RA et al (1989) Ti plasmid containing *Rhizobium meliloti* are non-tumorigenic on plants, despite proper virulence gene induction and T-strand formation. Arch Microbiol 153:85–89

Vaudequin-Dransart V, Petit A, Poncet C et al (1995) Novel Ti plasmids in *Agrobacterium* strains isolated from fig tree and chrysanthemum tumors and their opinelike molecules. Mol Plant Microbe Interact 8:311–321

Vaudequin-Dransart V, Petit A, Chilton WS et al (1998) The cryptic plasmid of *Agrobacterium tumefaciens* cointegrates with the Ti plasmid and cooperates for opine degradation. Mol Plant Microbe Interact 11:583–591

Vigouroux A, El Sahili A, Lang J et al (2017) Structural basis for high specificity of octopine binding to exploit the constructed octopine-niche on plant host. Sci Rep 7:18033

Watson B, Currier TC, Gordon MP, Chilton MD, Nester EW (1975) Plasmid required for virulence of *Agrobacterium tumefaciens*. J Bacteriol 123:255–264

Wendt T, Doohan F, Mullins E (2012) Production of *Phytophthora infestans*-resistant potato (*Solanum tuberosum*) utilising *Ensifer adhaerens* OV14. Transgenic Res 21:567–578

Wexler M, Gordon D, Murphy PJ (1995) The distribution of inositol rhizopine genes in *Rhizobium* populations. Soil Biol Biochem 27:531–537

White FF, Garfinkel DJ, Huffman GA et al (1983) Sequences homologous to *Agrobacterium rhizogenes* T-DNA in the genomes of uninfected plants. Nature 301:348–350

White FF, Ghidossi G, Gordon MP et al (1982) Tumor induction by *Agrobacterium rhizogenes* involves the transfer of plasmid DNA to the plant genome. Proc Natl Acad Sci USA 79: 3193–3197

Wichelecki DJ, Vetting MW, Chou L et al (2015) ATP-binding cassette (ABC) transport system solute-binding protein-guided identification of novel d-altritol and galactitol catabolic pathways in *Agrobacterium tumefaciens* C58. J Biol Chem 290:28963–28976

Willmitzer L, Sanchez-Serramo J, Duschelder E et al (1982) DNA from *Agrobacterium rhizogenes* is transferred to and expressed in axemic hairy root plant tissue. Mol Gen Genet 186:16–22

Wilson M, Savka MA, Hwang I et al (1995) Altered epiphytic colonization of mannityl opine-producing transgenic tobacco plants by a mannityl opine-catabolizing strain of *Pseudomonas syringae*. Appl Environ Microbiol 61:2151–2158

Wu HY, Chung PC, Shih HW et al (2008) Secretome analysis uncovers an Hcp-family protein secreted via a type VI secretion system in *Agrobacterium tumefaciens*. J Bacteriol 190: 2841–2850

Yamada T, Palm CJ, Brooks B et al (1985) Nucleotide sequences of the *Pseudomonas savastanoi* indoleacetic acid genes show homology with *Agrobacterium tumefaciens* T-DNA. Proc Natl Acad Sci USA 82:6522–6526

Zanker H, Lurz G, Langridge U et al (1994) Octopine and nopaline oxidases from Ti plasmids of *Agrobacterium tumefaciens*: molecular analysis, relationship, and functional characterization. J Bacteriol 176:4511–4517

Zanker H, von Lintig J, Schröder J (1992) Opine transport genes in the octopine (*occ*) and nopaline (*noc*) catabolic regions in Ti plasmids of *Agrobacterium tumefaciens*. J Bacteriol 174:841–849

Zechner EL, de la Cruz F, Eisenbrandt R et al (2001) Conjugative-DNA transfer processes. In: Thomas CM (ed) The horizontal gene pool: bacterial plasmids and gene spread. Harwood Academic Publishers, Amsterdam, p 87

Zhang L, Murphy PJ, Kerr A et al (1993) *Agrobacterium* conjugation and gene regulation by *N*-acyl-L-homoserine lactones. Nature 362:446–448

Zhao J, Binns AN (2014) GxySBA ABC transporter of *Agrobacterium tumefaciens* and its role in sugar utilization and *vir* gene expression. J Bacteriol 196:3150–3159

Cell Wall Biogenesis During Elongation and Division in the Plant Pathogen *Agrobacterium tumefaciens*

Wanda M. Figueroa-Cuilan and Pamela J. B. Brown

Contents

1 Introduction 88
2 Cell Growth and Division of Rod-Shaped Bacteria: A Historical Perspective 89
3 Cell Elongation in *A. tumefaciens* 91
3.1 *A. tumefaciens* Elongates Unipolarly 93
3.2 How Is the PG Synthesis Machinery Targeted to a Cell Pole During Elongation? 94
3.3 PG Biosynthesis at a Pole: Candidate Enzymes Contributing to Cell Wall Expansion in *A. tumefaciens* 95
3.4 Release from the Pole: Transition from Polar Growth to Mid-Cell Growth 97
4 Cell Division in *A. tumefaciens* 98
4.1 Establishment of the FtsZ-Ring 99
4.2 Divisome Maturation and Initiation of Septal PG Biosynthesis 101
4.3 Pole Remodeling and Cell Separation 102
5 Concluding Remarks and Outstanding Questions 103
References 105

Abstract A great diversity of bacterial cell shapes can be found in nature, suggesting that cell wall biogenesis is regulated both spatially and temporally. Although *Agrobacterium tumefaciens* has a rod-shaped morphology, the mechanisms underlying cell growth are strikingly different than other well-studied rod-shaped bacteria including *Escherichia coli*. Technological advances, such as the ability to deplete essential genes and the development of fluorescent D-amino acids, have enabled recent advances in our understanding of cell wall biogenesis during cell elongation and division of *A. tumefaciens*. In this review, we address how the field has evolved

W. M. Figueroa-Cuilan · P. J. B. Brown (✉)
Division of Biological Sciences, University of Missouri, Columbia, MO, USA
e-mail: brownpb@missouri.edu

Current Topics in Microbiology and Immunology (2018) 418:87–110
DOI 10.1007/82_2018_92

Published Online: 29 May 2018

over the years by providing a historical overview of cell elongation and division in rod-shaped bacteria. Next, we summarize the current understanding of cell growth and cell division processes in *A. tumefaciens*. Finally, we highlight the need for further research to answer key questions related to the regulation of cell wall biogenesis in *A. tumefaciens*.

1 Introduction

Hidden in the microscopic world, a great diversity of bacterial morphologies can be found. From simple rods and spheres, to very complex star- and helical-shaped bacteria, the mechanisms by which these fascinating organisms achieve, maintain, or evolve cell shapes are currently under investigation. The bacterial peptidoglycan (PG) cell wall is one of the fundamental determinants of cell shape, and research has revealed that carefully regulated enzymatic processes are used to achieve and maintain a characteristic morphology. Thus, the constant targeting of biosynthetic machineries and remodeling enzymes to specific subcellular locations ensure that a specific morphology can be maintained for many generations. The bacterial cell wall biosynthetic machinery is comprised of high molecular weight penicillin-binding proteins (HMW-PBPs) (Typas et al. 2011) and glycosyltransferases such as shape, elongation, division, and sporulation proteins (SEDS) (Meeske et al. 2016). A second subset of enzymes, including endopeptidases, carboxypeptidases, and L,D-transpeptidases (LDTs), is involved in remodeling of the PG (Typas et al. 2011). Finally, amidases and lytic transglycosylases (LTGs) are required for the hydrolysis of the PG mesh, enabling cell separation (van Heijenoort 2011; Uehara and Bernhardt 2011). Accordingly, PBPs, SEDS proteins, remodeling enzymes, and autolysins work in concert to modify continuously and expand the PG, allowing cell elongation and division to proceed. In recent years, technological advances such as the development of genetic tools (Morton and Fuqua 2012; Figueroa-Cuilan et al. 2016; Grangeon et al. 2017) and use of fluorescent D-amino acids (Kuru et al. 2012, 2015; Siegrist et al. 2015; Howell et al. 2017b) have contributed to our understanding of cell growth patterning and the underlying mechanisms in diverse bacteria, including *Agrobacterium tumefaciens*.

A. tumefaciens is a rod-shaped, gram-negative soil bacterium and the causative agent of crown gall disease in flowering plants (Escobar and Dandekar 2003; Nester 2014). Studies using amine-reactive dyes, non-canonical D-amino acids, and fluorescent D-amino acids demonstrated that *A. tumefaciens* and other members of the class Rhizobiales exhibit unipolar elongation (Brown et al. 2012; Kuru et al. 2012). In contrast, in *Escherichia coli*, a classical model for studies of bacterial cell growth and division, PG insertion is dispersed along the existing lateral cell wall during elongation (Cava et al. 2013).

2 Cell Growth and Division of Rod-Shaped Bacteria: A Historical Perspective

Here, we provide a historical perspective of the advances in our understanding of cell growth and division of rod-shaped bacteria including *E. coli*, *A. tumefaciens*, and *Rhizobium meliloti*. These early studies laid the groundwork for more recent studies which highlight key mechanistic differences in cell elongation among rod-shaped bacteria.

In *E. coli*, disruption of chromosome replication or cell division causes cells to elongate dramatically. Gamma irradiation (Lea et al. 1937), ultraviolet irradiation (Barner and Cohen 1956), and treatments with antibiotics that cause DNA damage or target components of the DNA replication machinery such as mitomycin C and nalidixic acid (Latch and Margolin 1997; Kilgore and Greenberg 1961; Helmstetter and Pierucci 1968; Goss et al. 1964) cause *E. coli* cells to form long, smooth filaments (Fig. 1a:1–2). Similarly, treatment with antibiotics, including cephalexin and carbenicillin, that inhibit mid-cell PG biogenesis, also causes *E. coli* cells to become filamentous (Fig. 1a:3–4) (Latch and Margolin 1997; Rolinson 1980; Zupan et al. 2013). In addition, numerous temperature-sensitive (TS) mutants with filamentous growth at elevated temperature were isolated (Fig. 1a:5) (Van De Putte et al. 1964). Mapping of genes with mutations responsible for cell filamentation led to the identification of the *fts* (filamentous growth is thermosensitive) genes (Ricard and Hirota 1973; Lutkenhaus and Donachie 1979; Lutkenhaus et al. 1980). *fts* genes encode proteins belonging to a core complex of highly conserved proteins, termed the divisome, which functions in cell division (See Sect. 4).

The first indication that *A. tumefaciens* may have a different growth pattern than *E. coli* emerged when isolation of TS mutants with a block in cell division failed to produce filamentous cells (Fig. 1b:5) (Fujiwara and Fukui 1972). Among 17 temperature-sensitive mutants which do not divide at elevated temperatures, three exhibited an atypical branched morphology, while the remaining mutants exhibited a spherical morphology (Fig. 1b:5). In the branching mutants, cell growth occurred exclusively at one pole of a cell and branches were formed from splitting of the growth-active poles (Fujiwara and Fukui 1974b). While the growth mode of individual wild-type cells could not be deduced, the pattern of micro-colony formation led to the hypothesis that *A. tumefaciens* growth occurs at a single pole (Fujiwara and Fukui 1974b). These results were supported by the observation that mitomycin C and nalidixic acid block cell division and cause branching of *R. meliloti* and *A. tumefaciens* cells (Fig. 1b:1–2) (Latch and Margolin 1997; Fujiwara and Fukui 1974a). Furthermore, treating *R. meliloti* and *A. tumefaciens* with cephalexin or carbenicillin also causes a branched morphology to emerge (Fig. 1b:3–4) (Latch and Margolin 1997; Zupan et al. 2013).

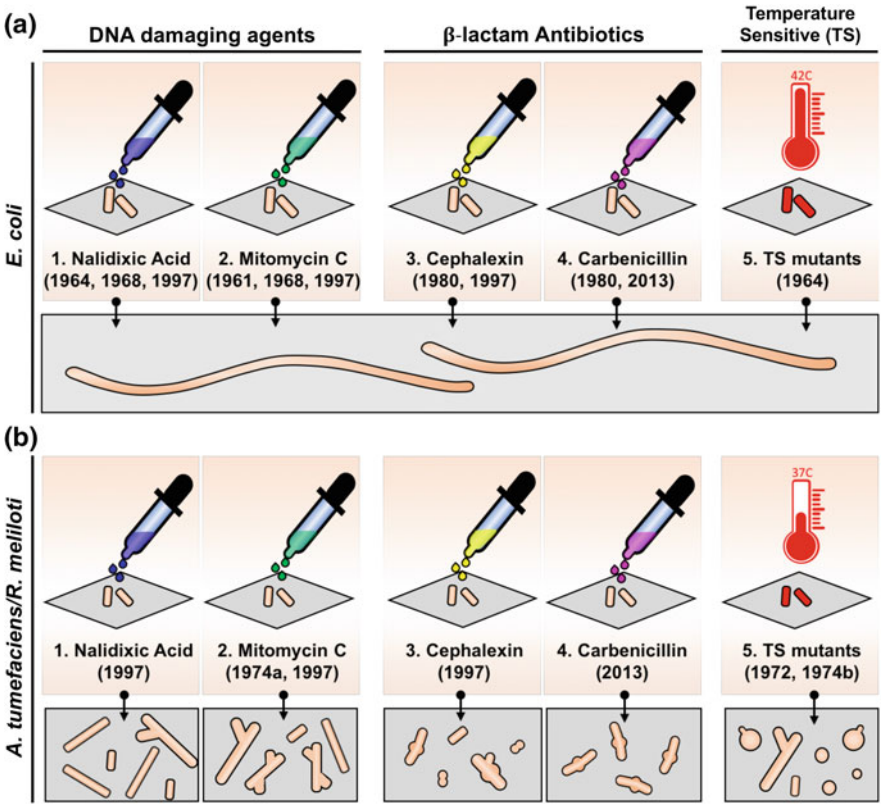


Fig. 1 Morphological changes induced by inhibition of cell division. Treatments include DNA damaging agents (1–2) and beta-lactam antibiotics (3–4). Cell division is also blocked in temperature-sensitive (TS) mutants (5) when incubated at high temperatures. **a** In *E. coli*, when cell division is inhibited, a filamentous phenotype is consistently observed. **b** In the Rhizobiales, phenotypes induced by cell division blocks are more variable and induce branching, bulging, elongated, and spherical cells. Schematics depicted in this figure show the impact of these treatments as observed in the indicated references (Latch and Margolin 1997; Zupan et al. 2013; Fujiwara and Fukui 1972, 1974a, b; Van De Putte et al. 1964, Kilgore and Greenberg 1961; Helmstetter and Pierucci 1968; Goss et al. 1964)

Despite the compelling evidence that blocking cell division causes branching, consistent with a model of polar elongation, these observations remained largely unrecognized. Indeed, Bergey's Manual of Systematic Bacteriology indicates that *Agrobacterium* does not reproduce by budding at one pole of the cell (Kuykendall 2005). More recently, the growth pattern of bacteria belonging to the order Rhizobiales, including *A. tumefaciens*, has been described as unipolar elongation (Fig. 2) (Brown et al. 2012; Kuru et al. 2012).

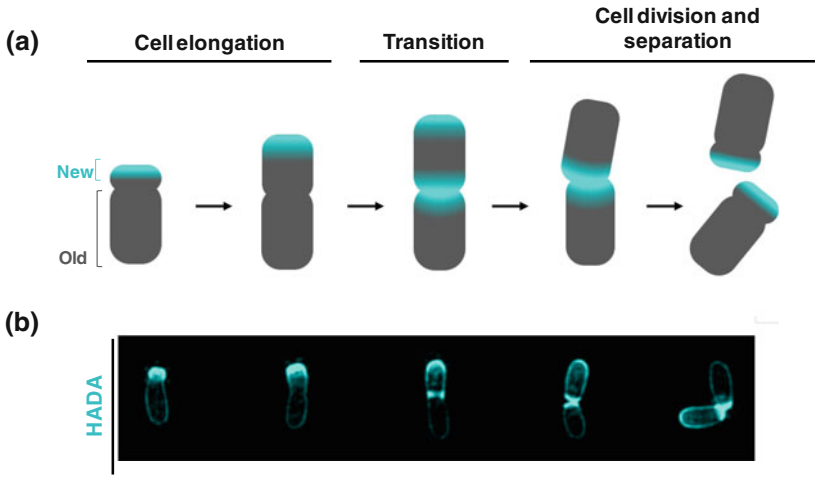
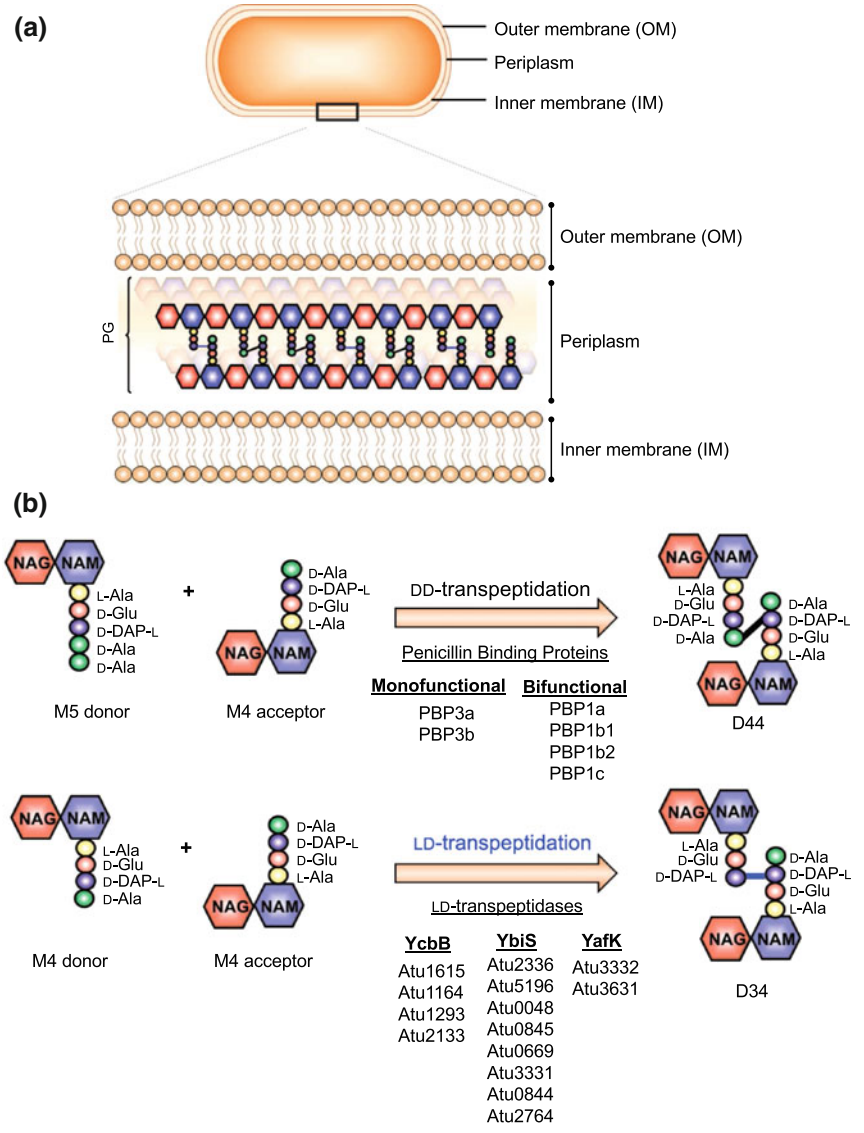


Fig. 2 Schematic and fluorescence images illustrating the *A. tumefaciens* growth pattern. The growth pattern of *A. tumefaciens* includes three stages: (1) polar growth during cell elongation, (2) a transition when polar growth terminates and mid-cell PG synthesis is initiated, and (3) mid-cell PG synthesis during cell division. **a** The schematic indicates areas of active PG synthesis (showed in cyan) throughout the cell cycle. **b** Fluorescence images of individual *A. tumefaciens* cells following a short pulse label with a fluorescent D-amino acid (HADA) reveal sites of PG synthesis throughout the cell cycle. These images highlight the three different stages of PG synthesis

3 Cell Elongation in *A. tumefaciens*

In rod-shaped bacteria, cell growth is defined as the insertion of peptidoglycan (PG) resulting in the elongation of the cell. In gram-negative bacteria, the PG cell wall is located in the periplasmic space between the outer and inner membranes (Fig. 3a). The PG cell wall is a net-like structure that aids bacteria to withstand osmotic pressure conferred by the environment and is comprised of a polysaccharide containing alternating *N*-acetylglucosamine (NAG) and *N*-acetylmuramic acid (NAM) sugars. The NAM sugar is decorated with a pentapeptide stem, composed of the following D-amino acids: L-Ala, D-Glu, D-DAP-L, D-Ala, D-Ala (Fig. 3b). The pentapeptide stem enables NAG-NAM disaccharides to be cross-linked to one another through peptide bridges (Fig. 3b). Although *A. tumefaciens* PG resembles the canonical structure of PG for a gram-negative bacterium (Erbs et al. 2008), a number of distinctive features were revealed by compositional analysis of the PG (Brown et al. 2012). These features include an increased abundance of overall muropeptide crosslinkage and an enrichment in L,D-Dap-Dap crosslinks formed by LDTs. Interestingly, increased abundance of L,D-cross-linkages is a common feature among polar growing bacteria such as *Mycobacterium* (Lavollay et al. 2008, 2011), indicating that LDTs may play an important role during polar elongation. Another feature of *A. tumefaciens* PG is the


(a)
Outer membrane (OM)
Periplasm
Inner membrane (IM)
PG
Outer membrane (OM)
Periplasm
Inner membrane (IM)
(b)
NAG NAM
L-Ala
D-Glu
D-DAP-L
D-Ala
D-Ala
M5 donor
+
NAG NAM
D-Ala
D-DAP-L
D-Glu
L-Ala
M4 acceptor
DD-transpeptidation
Penicillin Binding Proteins
Monofunctional
PBP3a
PBP3b
Bifunctional
PBP1a
PBP1b1
PBP1b2
PBP1c
NAG NAM
L-Ala
D-Glu
D-DAP-L
D-Ala
D-Ala
D-DAP-L
D-Glu
L-Ala
D44
NAG NAM
L-Ala
D-Glu
D-DAP-L
D-Ala
M4 donor
+
NAG NAM
D-Ala
D-DAP-L
D-Glu
L-Ala
M4 acceptor
LD-transpeptidation
LD-transpeptidases
YcbB
Atu1615
Atu1164
Atu1293
Atu2133
YbiS
Atu2336
Atu5196
Atu0048
Atu0845
Atu0669
Atu3331
Atu0844
Atu2764
YafK
Atu3332
Atu3631
NAG NAM
L-Ala
D-Glu
D-DAP-L
D-Ala
D-DAP-L
D-Glu
L-Ala
D34

◂**Fig. 3** Peptidoglycan structure and crosslinking reactions catalyzed by D,D-transpeptidases and L,D-transpeptidases of *A. tumefaciens*. **a** Schematic illustrating the cell envelope of *A. tumefaciens* which includes the outer membrane, inner membrane, and peptidoglycan (PG) in the periplasmic space. PG is composed of glycan strands containing alternating *N*-acetylglucosamine (NAG) and *N*-acetylmuramic acid (NAM) sugars. The NAM sugar is decorated with a pentapeptide stem, enabling adjacent glycan strands to be linked through peptide bridges to stabilize the PG. **b** Transpeptidation reactions crosslink peptidoglycan. A representative D,D-transpeptidation reaction is illustrated in the top panel. A monomeric donor with five peptides in the stem (M5) is crosslinked to a monomeric acceptor with four peptides in the stem (M4). This transpeptidation reaction produces a dimeric muropeptide (D44). The catalysis of D,D-transpeptidation reactions is performed by bifunctional and monofunctional PBPs. A representative L,D-transpeptidation reaction is shown in the bottom panel. Crosslinking of an M4 donor and M4 acceptor produces a dimeric muropeptide (D34). L,D-transpeptidation reactions are catalyzed by L,D-transpeptidases (LDTs). In *A. tumefaciens*, three different subfamilies of LDTs can be found: YcbB, YbiS, and YafK. With the exception of Atu2764, the proteins belonging to the YbiS subfamily are specific to the Rhizobiales

absence of detectable anhydromuropeptides. Anhydromuropeptides occupy the terminal position in glycan strands of PG from gram-negative bacteria, and thus the abundance of anhydromuropeptides provides an indirect measure of average glycan strand length. This inability to detect anhydromuropeptides in *A. tumefaciens*, PG could be explained by: (1) modification of anhydromuropeptides masking their detection, (2) an unusually long glycan strand length, or (3) a novel mechanism for glycan strand termination. Together, these observations suggest that the composition of PG in *A. tumefaciens* differs from that of other well-studied gram-negative bacteria such as *E. coli*.

In this section, we first highlight the unipolar growth pattern during elongation of *A. tumefaciens*. Next, potential mechanisms for targeting PG biosynthesis to the growth pole and the role of enzymes in polar synthesis of PG are discussed. Finally, we consider how PG biosynthesis is terminated at the growth pole, allowing establishment of a new growth zone near mid-cell for cell division.

3.1 A. tumefaciens *Elongates Unipolarly*

The growth of several rod-shaped bacteria is characterized by insertion of new PG alongside the lateral walls of the cell, a process mediated in part by the actin-like scaffolding protein MreB (for review, see (Errington 2015). MreB interacts with multiple components of the cell wall machinery and functions as a regulator to maintain proper rod-cell shape during cell elongation. Remarkably, the genomes of representative members of the order Rhizobiales lack the core components of the elongasome, the protein complex responsible for insertion of new PG into the existing wall material during cell elongation, including MreB (Errington 2015; Margolin 2009). This observation, coupled with the branched morphologies resulting from treatment of *A. tumefaciens* with DNA damaging agents or antibiotics which inhibit mid-cell PG biosynthesis (Fig. 1b), provides further evidence for

an alternative mode of growth during cell elongation. To compare the growth pattern of *A. tumefaciens* and *E. coli*, cells were transiently treated with the amine-reactive dye Texas Red-X succinimidyl ester (TRSE) to label the outer membrane proteins. During dispersed growth of *E. coli* cells in the absence of TRSE, incorporation of new unlabeled material along the lateral walls of the cell results in dilution of the TRSE signal. In contrast, the TRSE signal remains fixed in the cell body and old pole compartment of elongating *A. tumefaciens* because growth occurs at the new cell pole (Brown et al. 2012). Similar growth patterns were observed using methods that directly label the peptidoglycan, including pulse-chase labeling of PG with D-cysteine (Brown et al. 2012) and short pulse labeling of PG with fluorescent D-amino acids (FDAAs) (Fig. 2) (Kuru et al. 2012). Together, these experiments led to the elucidation of the growth patterning of *A. tumefaciens*, which consists of three main growth stages: (1) polar growth during cell elongation, (2) termination of polar growth and initiation of mid-cell PG synthesis, and (3) PG insertion at mid-cell to promote cell division (Fig. 2) (Cameron et al. 2015; Howell and Brown 2016; Kysela et al. 2013; Yang et al. 2016). During elongation, the PG biosynthesis machinery is strictly targeted to the new cell pole. The cells continue to grow unipolarly for about a third of a cell cycle by addition of new PG at the cell pole, displacing older PG. Next, termination of polar growth coincides with recruitment of FtsZ to mid-cell and subsequent PG insertion at mid-cell. Insertion of PG at mid-cell continues until the septum is formed and the daughter cells separate. Whereas the growth pattern of *A. tumefaciens* has been clearly established, we are just beginning to understand the mechanisms underlying unipolar elongation.

3.2 How Is the PG Synthesis Machinery Targeted to a Cell Pole During Elongation?

In *A. tumefaciens*, the strict polar targeting of PG biosynthesis machinery during cell elongation suggests that an underlying mechanism restricts growth to one pole. How can protein complexes be targeted to polar locations within bacterial cells? Several polar targeting principles have been described, including the ability of proteins to recognize negative membrane curvature and the accumulation of polymer-forming proteins in DNA-free regions such as poles (Laloux and Jacobs-Wagner 2014). Polymer-forming landmark proteins function to recruit additional proteins to the pole, enabling polar protein complexes to form. For example, a polar hub established through interactions with the polar organizing protein PopZ enables proper chromosome segregation in *Caulobacter crescentus* (Bowman et al. 2008; Ebersbach et al. 2008; Bowman et al. 2010), and DivIVA is necessary for polar growth in *Actinobacteria* (Flärdh 2003; Kang et al. 2008; Letek et al. 2009; Fuchino et al. 2013; Hempel et al. 2008). Thus, it is likely that establishment of a polar protein complex to regulate polar growth in *A. tumefaciens* may also rely on a

landmark protein. In *A. tumefaciens*, only two candidate polar landmark proteins, PopZ and PodJ, have been characterized. Surprisingly, neither PopZ nor PodJ are strictly required for polar growth, although both proteins contribute to the transition from polar growth to mid-cell growth (see Sect. 3.4) (Grangeon et al. 2015; Howell et al. 2017a; Ehrle et al. 2017; Grangeon et al. 2017; Anderson-Furgeson et al. 2016). Thus, dissection of the molecular composition of *A. tumefaciens* growth poles will be necessary to determine whether novel polar organizing proteins are present and able to target the PG biosynthetic machinery to the new pole.

An alternative mechanism of polar targeting that may contribute to regulation of PG biosynthesis during elongation in *A. tumefaciens* is inheritance from the division site (Laloux and Jacobs-Wagner 2014). When rod-shaped bacteria undergo cell division, each newborn cell inherits a new pole formed during cell division. Thus, proteins localizing to mid-cell during cell division will be present at the new pole of the new daughter cells. The absence of the canonical rod-shaped cell elongation machinery in *A. tumefaciens* (Brown et al. 2012; Cameron et al. 2014) indicates that the conserved cell division machinery may influence where new growth resumes after cell division. In *A. tumefaciens*, the cell division proteins FtsZ and FtsA display polar localization during elongation and then localize to mid-cell before cell division (Brown et al. 2012; Zupan et al. 2013; Grangeon et al. 2015; Howell et al. 2017a). The importance of the FtsZ and FtsA foci at the new cell pole during cell elongation is still largely unknown; however, the absence of FtsZ causes tip splitting to occur at growth poles, suggesting that FtsZ may be required for proper termination of polar growth (Howell and Brown 2016). It remains to be determined whether any of the divisome proteins are retained at the new poles following cell division function in polar growth.

3.3 *PG Biosynthesis at a Pole: Candidate Enzymes Contributing to Cell Wall Expansion in* A. tumefaciens

The enzymes contributing to PG biosynthesis are well-conserved across bacteria and are likely to contribute to polar PG biosynthesis in *A. tumefaciens* (Cameron et al. 2014; Howell and Brown 2016). The first group of enzymes consists of the high molecular weight penicillin-binding proteins (HMW-PBPs) (Fig. 3b, top panel) (Scheffers and Pinho 2005; Egan et al. 2015; Sauvage et al. 2008; Typas et al. 2011). HMW-PBPs are subdivided into two groups: bifunctional PBPs, which possess transglycosylase and DD-transpeptidase activities (PBP1), and monofunctional PBPs, which have DD-transpeptidase activity only (PBP2, PBP3). The transglycosylase activity of PBPs is responsible for the addition of the nascent NAG-NAM disaccharide to the existing glycan strand, whereas the DD-transpeptidase links adjacent peptide stems. A representative peptide bond formed by DD-transpeptidases activity occurs between the D-Ala at the fourth position of the peptide stem (M5 donor) and the D-stereogenic center of DAP at the third position

of the peptide stem (M4 acceptor) in two adjacent monomeric muropeptides to form a dimeric muropeptide (Fig. 3b, top panel). In addition to PBPs, another enzyme that may be contributing to PG biosynthesis in *A. tumefaciens* is the monofunctional peptidoglycan glycosyltransferase MtgA, which is responsible for elongation of glycan strands during cell division in *E. coli* (Derouaux et al. 2008).

The genome of *A. tumefaciens* contains genes predicted to encode six HMW-PBPs; four bifunctional (PBP1a, PBP1b1, PBP1b2, and PBP1c), and two monofunctional (PBP3a and PBP3b) (Fig. 3b, top panel) (Cameron et al. 2014). *A. tumefaciens* PBP1a localizes to the growth pole during most of the cell cycle (Cameron et al. 2014) and is essential for cell survival (Curtis and Brun 2014), suggesting that PBP1a may have an important function during polar elongation. In contrast, PBP3a and PBP3b localize at mid-cell during cell division (Cameron et al. 2014). PBP3a is essential for viability (Curtis and Brun 2014) whereas no changes in growth rates or phenotype were observed in the absence of PBP3b (Cameron et al. 2014). These results suggest that PBP3a may be a major contributor to PG biosynthesis during cell division; however, other PBPs may also function during cell division because labeling PBPs with a fluorescent penicillin derivative, Bocillin-FL (Zhao et al. 1999), results in a strong mid-cell signal (Cameron et al. 2014). Furthermore, treatment of WT cells with carbenicillin, a penicillin derivative that blocks the DD-transpeptidase activity of PBPs, leads to mid-cell bulges with no obvious effects on the cell poles (Fig. 1b:4)(Zupan et al. 2013). Together, these results indicate that PBP1a may contribute to polar growth, whereas PBP3a and possibly other PBPs are involved in septal PG synthesis during cell division. There are many open questions regarding the activities of HMW-PBPs in *A. tumefaciens*. If PBP1a functions in polar elongation, how its activity is restricted to the new pole during elongation and how its activity is terminated prior to cell division? In *E. coli*, the activity of PBP1A and PBP1B is dependent on the cognate lipoproteins LpoA and LpoB, respectively (Paradis-Bleau et al. 2010; Typas et al. 2010). Although Lpo-like proteins are not encoded by the genome of *A. tumefaciens*, the presence of a domain of unknown function between the enzymatic domains of PBP1a suggests that its activity may be regulated. Are both PBP3a and PBP3b core components of the division machinery? What is the function of the remaining HMW-PBPs? As answers to these questions emerge, we expect to gain insights into the molecular mechanism underlying the cell growth pattern.

The next set of enzymes hypothesized to contribute to the polar elongation of *A. tumefaciens* are the L,D-Transpeptidases (LDTs) (Cameron et al. 2014; Howell and Brown 2016; Brown et al. 2012). LDTs are penicillin-insensitive enzymes involved in the catalysis of direct crosslinks between two D-DAP-L peptides at the third position of adjacent muropeptide stems (Fig. 3B, bottom panel). Muropeptide composition analyses indicate that *A. tumefaciens* PG is enriched in muropeptides containing LD-crosslinks (Brown et al. 2012). PG from other polar growing bacteria such as *Mycobacterium* also contains muropeptides with a high percentage (60–80%) of LD-crosslinks, suggesting an important role of LDTs in polar growth (Lavollay et al. 2008, 2011). Moreover, a loss of several LDTs from *Mycobacterium tuberculosis* leads to defects in cell morphology, such as cell

rounding (Sanders et al. 2014), and β-lactam antibiotic sensitivity (Gupta et al. 2010). Consistent with the high abundance of LD-crosslinks in the PG, the genome of *A. tumefaciens* reveals an enrichment in genes encoding LDT enzymes (Cameron et al. 2014). Genes encoding a total of 14 candidate LDTs have been identified in the *A. tumefaciens* genome. Interestingly, the YbiS subfamily of LDTs consists of seven Rhizobiales-specific LDTs and one commonly found in other bacterial classes. The remaining six LDTs belong to two subfamilies called YcbB and YafK (Fig. 3b, bottom panel). Two of the Rhizobiales-specific LDTs (ATU0845, ATU0669) localize to the growth pole during most of the cell cycle (Cameron et al. 2014). These results indicate that LDTs may play a crucial, and perhaps essential, role during polar growth. More studies will be required to dissect the precise role and functional redundancy of LDTs in *A. tumefaciens*.

3.4 Release from the Pole: Transition from Polar Growth to Mid-Cell Growth

Before termination of polar elongation, a transition of growth from the new cell pole to mid-cell occurs (Fig. 2). During this transition, the septal PG biosynthetic machinery is recruited to mid-cell to form the septum of the incipient daughter cells. In *A. tumefaciens*, so far three proteins have been found to contribute to this transition.

First, the polar organizing protein (PopZ) is a polymer-forming landmark protein that serves as a hub for polar proteins. In *C. crescentus*, PopZ is required to tether the chromosomal origin to the cell pole and for assembly of proteins required for stalk development (Bowman et al. 2008, 2010, 2013; Ebersbach et al. 2008; Laloux and Jacobs-Wagner 2013; Holmes et al. 2016). In *A. tumefaciens*, PopZ strictly localizes to the new cell pole during cell elongation and then to mid-cell during late stages of cell division (Grangeon et al. 2015). The absence of PopZ results in cell division defects such as ectopic pole formation, tip splitting, bulging, and formation of small cells that lack DNA (Howell et al. 2017a; Grangeon et al. 2017). The chromosome partitioning protein, ParB, and cell division proteins, FtsZ and FtsA, are mislocalized in cells lacking PopZ (Howell et al. 2017a; Ehrle et al. 2017). Together, these observations suggest that PopZ participates in multiple processes including chromosome segregation and the transition from polar growth to mid-cell PG biosynthesis during cell division. It remains to be determined whether PopZ has a direct role in regulation of PG biosynthesis at the growth pole or during cell division. Alternatively, the observed defects in cell division could be secondary to the defect in chromosome segregation caused by the absence of PopZ.

Another candidate polymer-forming landmark protein is PodJ. In *C. crescentus*, PodJ is located at the old pole, recruits cell cycle regulators to the pole, and enables development of polar structures including the flagellum and pili (Hinz et al. 2003; Viollier et al. 2002). In *A. tumefaciens*, PodJ initially localizes to the old pole but slowly accumulates at the new pole indicating a possible role in the transition from a new, growth-active pole to an old, growth-inactive pole (Grangeon et al. 2015).

The absence of PodJ causes the formation of elongated cells with multiple constrictions and branching, suggesting that polar growth is not terminated efficiently leading to subsequent defects in cell division (Anderson-Furgeson et al. 2016). Remarkably, FtsZ and FtsA have atypical localization patterns and FtsZ-rings and FtsA-rings often form at failed sites of cell division in the absence of PodJ. These data suggest that PodJ contributes to the transitioning of PG biosynthesis machinery from the cell pole to mid-cell. One possibility is that the transition of the growth-active pole to a growth-inactive pole mediated by PodJ is coordinated with onset of PG biosynthesis at mid-cell for cell division. Further mechanistic studies will be necessary to determine why the sites of cell division that are established in the absence of PodJ often fail to septate.

Because proper establishment of future sites of cell division appears to be important for the transition of polar growth to mid-cell PG biosynthesis, it is logical to consider the role of cell division proteins in these processes. Both FtsZ and FtsA remain at the pole for a portion of the cell cycle before localizing near mid-cell to mark the future site of cell division (Brown et al. 2012; Cameron et al. 2014; Zupan et al. 2013). FtsZ-depleted cells exhibit a complete block of cell division leading to cell branching, multipolar elongation, and tip splitting (Howell and Brown 2016). This observation suggests that FtsZ is required not only for the establishment of PG biosynthesis at mid-cell, but also for the termination of polar PG biosynthesis. While it is clear that the absence of FtsZ prevents the transition of the growth-active pole to a growth-inactive pole, additional studies are needed to determine the precise functions of FtsZ and the proteins it recruits to mid-cell during the inactivation of the growth pole.

4 Cell Division in *A. tumefaciens*

The onset of cell division is characterized by a burst of septal PG synthesis at mid-cell (Fig. 2). Remarkably, septal growth must be regulated and the PG must be remodeled to generate the incipient bacterial cell poles. Indeed, precise coordination of these processes depends on the divisome, a multiprotein complex dedicated to orchestrating cell division. The divisome consists of over 30 proteins which are highly conserved among bacteria (Haeusser and Margolin 2016). Placement of FtsZ at mid-cell leads to the recruitment and assembly of the other cell division proteins to the site of cell division. In rod-shaped bacteria such as *E. coli* and *Bacillus subtilis*, FtsZ and its membrane anchor FtsA move the septal PG biosynthesis machinery around the circumference of the site of cell division using a treadmilling motion (Bisson-Filho et al. 2017; Yang et al. 2017).

Here, we speculate on the mechanism of cell division in *A. tumefaciens* using the linear hierarchical model for cell division of *E. coli* (Du and Lutkenhaus 2017) as a reference. We also discuss what little is known about the function of the proteins contributing to the cell division of *A. tumefaciens* and highlight questions for future work.

4.1 Establishment of the FtsZ-Ring

The typical assembly pathway for the divisome is comprised of two stages: establishment of the Z-ring and recruitment of late cell division proteins for PG biosynthesis (Du and Lutkenhaus 2017). During the first stage, GTP-dependent interactions between FtsZ monomers allow the formation of FtsZ protofilaments at mid-cell, where the concentration of proteins that inhibit FtsZ-polymer formation is the lowest (Oliva et al. 2004; Rowlett and Margolin 2013; Wu and Errington 2011). FtsA and ZipA are recruited to mid-cell where they bind the conserved C-terminal peptide (CTP) domain of FtsZ to anchor it to the inner membrane, allowing the Z-ring to coalesce at mid-cell (Pichoff and Lutkenhaus 2005; Haney et al. 2001). FtsA is also required for the recruitment and regulation of other downstream divisome proteins such as FtsN (Pichoff et al. 2015; Liu et al. 2015). FtsEX are the final core divisome proteins recruited during the first stage. They regulate key cell wall events such as PG septal synthesis and hydrolysis (Yang et al. 2011) and stimulate the FtsA-mediated recruitment of downstream divisome proteins to mid-cell (Du et al. 2016).

The predicted divisome of *A. tumefaciens* consists of the putative early proteins FtsZ, FtsA, and FtsEX, but lacks the FtsZ membrane anchor ZipA (Fig. 4a, b) (Cameron et al. 2014). Remarkably, the genome of *A. tumefaciens* contains three homologs of *ftsZ* (Atu2086, Atu4673, and Atu4215). Atu2086 (hereafter termed FtsZ) is syntenic with conserved genes that encode essential members of the cell division machinery, including FtsA (Atu2087) and FtsQ (Atu2088) (Cameron et al. 2014), suggesting that FtsZ encodes the major cell division scaffolding protein of *A. tumefaciens*. FtsZ is a tubulin homolog composed of a short N-terminal domain of unknown function, a GTPase domain necessary for the hydrolysis of GTP, a C-terminal linker (CTL) which influences polymer structure and dynamics, and a conserved C-terminal peptide (CTP) required for the membrane tethering of FtsZ (Zupan et al. 2013). The role of the two additional homologs of FtsZ in *A. tumefaciens* has not been determined; however, one of these copies lacks the both the CTL and CTP domains and the other is truncated in the GTPase domain (Zupan et al. 2013), suggesting that they may have distinct functions from FtsZ. Indeed, a saturating transposon mutagenesis screen suggests that only FtsZ is essential for viability (Curtis and Brun 2014). FtsZ localizes at the new pole during elongation and at mid-cell during cell division (Brown et al. 2012; Zupan et al. 2013; Grangeon et al. 2015; Cameron et al. 2015; Howell et al. 2017a). Moreover, recruitment of FtsZ to mid-cell is coordinated with the initiation of constriction at the future site of cell division (Brown et al. 2012). Time-lapse microscopy of FtsZ-depleted cells exhibits gross morphological defects such as branching, multipolar elongation, and tip splitting, suggesting that FtsZ is essential for cell division, viability, and establishment of cell shape and length (Howell and Brown 2016). Overall, these results suggest that FtsZ is the major cell division scaffolding protein necessary for the recruitment of the divisome to mid-cell in *A. tumefaciens* (Fig. 4a).

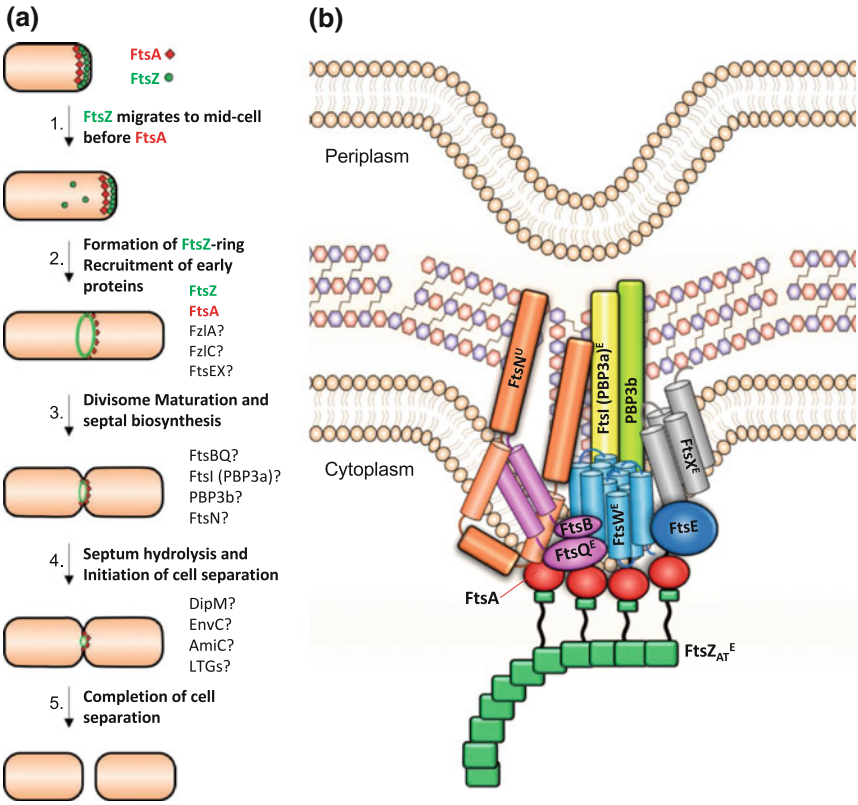


Fig. 4 Predicted divisome components of *A. tumefaciens*. **a** Predicted temporal order of assembly of candidate proteins involved in cell division and cell separation in *A. tumefaciens*. (1) $FtsZ_{AT}$ monomers migrate to mid-cell to form a Z-ring at the future site of cell division. (2) Recruitment of early divisome proteins such as FtsA and FtsEX to mid-cell. At this time, the Z-ring is tethered to the inner membrane by membrane-associated proteins such as FtsA, FzlA, or FzlC. (3) The late divisome proteins (FtsQB, FtsI, and FtsN) are recruited to mid-cell leading to the maturation of the divisome and activation septal PG biosynthesis. (4) PG hydrolases and their regulators are recruited to mid-cell to hydrolyze septal PG. (5) Finally, cell separation is completed upon the action of PG hydrolases. The presence of a question mark (?) indicates uncertainty in the localization or function of the protein during cell division. **b** Predicted spatial organization of the mature core *A. tumefaciens* divisome. Essentiality of the core divisome proteins is shown: ^{E} indicates essential for survival, ^{U} indicates unresolved essentiality, and the absence of superscript indicates that the protein is not essential (Curtis and Brun 2014)

In *A. tumefaciens*, localization studies place FtsZ at the site for cell division before the arrival of FtsA to mid-cell (Grangeon et al. 2015; Cameron et al. 2014) (Fig. 4a), indicating that FtsA may not be the only anchor of FtsZ. One candidate FtsZ anchor is FzlC. In *C. crescentus*, $FzlC_{CC}$ localizes to the incipient cell division site with the early wave of cell division proteins (Goley et al. 2010) and interacts with the CTP of FtsZ to anchor it to the inner membrane (Meier et al. 2016).

In *A. tumefaciens*, $FzlC_{AT}$ is present (Cameron et al. 2014) and shares 56% similarity with $FzlC_{CC}$, suggesting that $FzlC_{AT}$ may anchor FtsZ at mid-cell. Another candidate for tethering FtsZ is FzlA, a putative glutathione S-transferase (GST) protein. In *C. crescentus*, $FzlA_{CC}$ is an essential protein that localizes to mid-cell in an FtsZ-dependent manner, directly interacts with FtsZ in vitro, and regulates FtsZ protofilament curvature (Lariviere et al. 2018; Goley et al. 2011). In *A. tumefaciens*, $FzlA_{AT}$ is predicted to be an essential protein (Curtis and Brun 2014) and shares 67% similarity to $FzlA_{CC}$, suggesting that $FzlA_{AT}$ could interact with FtsZ. The final candidate for the tethering of FtsZ to the membrane is FtsE. FtsE is part of the early wave of divisome proteins and interacts with FtsZ (Huang et al. 2013; Goley et al. 2011). In addition to possible roles in stabilization of the FtsZ-rings, FtsA and FtsEX likely contribute to the proper regulation of PG biosynthesis at mid-cell; however, the contributions of these proteins to cell division have not yet been explored in *A. tumefaciens*.

4.2 *Divisome Maturation and Initiation of Septal PG Biosynthesis*

The second stage of divisome assembly involves the recruitment of late cell division proteins to regulate PG biosynthesis at mid-cell. In *E. coli*, the proteins sequentially recruited to mid-cell are FtsK, FtsQ, FtsL, FtsB, FtsW, FtsI, and FtsN (Du and Lutkenhaus 2017). FtsK has multiple functions, including the transport and decatenation of DNA, and the localization of FtsQLB to mid-cell (Massey et al. 2006). FtsQ, FtsL, and FtsB are membrane proteins that serve as scaffolds for other divisome proteins (Gonzalez et al. 2010). The complex formed by FtsQLB is activated by divisome maturation and regulates PG biosynthesis by modulating the activities of FtsI and FtsW (Tsang and Bernhardt 2015). The precise mechanism by which FtsQLB is activated is unknown, but it is hypothesized that FtsN may trigger the activation of the complex. Next, FtsW, a polytopic membrane protein with glycosyltransferase activity, is recruited to mid-cell and subsequently recruits FtsI (PBP3) to mid-cell (Cho et al. 2016). Together, FtsW and PBP3 are major contributors to PG biosynthesis at mid-cell. FtsN is the last essential divisome protein recruited to mid-cell (Weiss 2015). The arrival of FtsN to mid-cell signals the completion of divisome maturation, and septal PG biosynthesis is initiated. FtsN also recruits enzymes involved in the hydrolysis of septal PG to mid-cell, enabling pole remodeling and cell separation (Weiss 2015).

The late cell division proteins predicted to be encoded in the *A. tumefaciens* genome include two copies of FtsK, FtsQ, FtsB, FtsW, two copies of FtsI (PBP3a and PBP3b), and FtsN (Fig. 4a, b) (Cameron et al. 2014). Neither FtsK homolog is predicted to be essential (Curtis and Brun 2014) although the potential for redundancy means that one or both copies of FtsK could play important roles in chromosome segregation, cell division, or both processes. The functions of the FtsK proteins have not yet been explored in *A. tumefaciens*. The FtsQLB complex plays

an important role in cell division of *E. coli* (Gonzalez et al. 2010; Tsang and Bernhardt 2015); however, in *A. tumefaciens* only FtsQ is predicted to be essential, and a homolog of FtsL cannot be readily identified in the genome (Cameron et al. 2014; Curtis and Brun 2014). Thus, it will be of interest to determine whether FtsQ and FtsB form a complex important for recruitment of other divisome proteins and contribute to regulation of septal PG biosynthesis. PBP3a (Atu2100) and PBP3b (Atu1067) localize to mid-cell, consistent with a role in cell division (Cameron et al. 2014), and PBP3a is essential for cell survival (Curtis and Brun 2014). A deletion stain for PBP3b was constructed, but no phenotypic changes were observed (Cameron et al. 2014). Together, these results suggest that in *A. tumefaciens*, PBP3a is the major septal biosynthesis PBP. The predicted FtsN in *A. tumefaciens* has an extended N-terminal region giving rise to an atypical protein topology when compared to *E. coli* FtsN. The N-terminus is predicted to be periplasmic, followed by a cytoplasmic loop, and a periplasmic C-terminus (Fig. 4b). Only the C-terminal region (amino acids 675-1008) of the *A. tumefaciens* FtsN, which contains a 3-helix region, a glutamine-rich linker, and a SPOR domain, shares similarity with *E. coli* FtsN. In *E. coli*, the cytoplasmic tail of FtsN interacts with FtsA, the 3-helix region may interact with the FtsQLB complex, and the SPOR domain binds PG and is important for recruiting FtsN to mid-cell (Weiss 2015). Although the function of FtsN is unknown in *A. tumefaciens*, it is exciting to consider how the altered protein structure may affect the functions and regulatory abilities of FtsN during cell division and cell separation.

4.3 Pole Remodeling and Cell Separation

In *E. coli*, the cell separation process is mediated by PG hydrolases including lytic transglycosylases (LTGs) and amidases. LTGs are enzymes that cleave β-1,4 glycosidic linkages in the glycan strands, leading to the formation of terminating sugars with 1,6-anhydromuramic acid rings. The degradation of the PG by LTGs initiates the early steps in cell wall recycling as the product of LTG activity is internalized and converted to lipid II in the cytoplasm (Johnson et al. 2013; Park and Uehara 2008; Vollmer et al. 2008). Most bacteria possess multiple LTGs, and inactivation of all seven in *E. coli* is not tolerated, suggesting that at least some of the enzymes are redundant and have critical functions (Heidrich et al. 2002). *A. tumefaciens* has at least five LTGs (Cameron et al. 2014), and individual LTGs are not required for cell viability (Curtis and Brun 2014); however, the function of LTGs in cell division and peptidoglycan recycling remains to be explored.

Amidases are hydrolytic enzymes that specifically cleave the amide bond that links the NAM sugar with the peptide stem (van Heijenoort 2011). LytM-containing proteins typically function as endopeptidases; however, in *E. coli* catalytically inactive LytM proteins (dLytM) function as regulators of amidase activity. dLytMs are recruited to mid-cell by FtsN and FtsEX and activate amidases (Peters et al. 2011; Yang et al. 2011; Uehara et al. 2009, 2010). In *E. coli*, a block of

septal PG synthesis using cephalexin affects the localization of amidases but not dLytM factors, suggesting that ongoing PG synthesis may be required to modulate amidase activity (Peters et al. 2011). *A. tumefaciens* contains a single LytC-type amidase (AmiC) and two dLytM proteins (EnvC and DipM) (Cameron et al. 2014) (Fig. 4). The predictions that AmiC will function during septum cleavage and that its activity is regulated by DipM remain to be tested experimentally. In addition, genes encoding the targets of EnvC regulation, AmiA and AmiB, are not present in the *A. tumefaciens* genome, suggesting that EnvC may regulate other enzymes involved in cell separation. During the final step of cell separation, the Tol–Pal complex constricts the outer membrane as the septum is cleaved (van Heijenoort 2011; Egan 2018). The Tol–Pal system connects the bacterial outer membrane, periplasm, and cytoplasm through a series of protein–protein interactions. During cell division, this system is thought to provide the energy required for constricting the outer membrane as septum synthesis and hydrolysis shape the new daughter cell poles (Egan 2018). In *A. tumefaciens*, the Tol–Pal complex is present (Cameron et al. 2014) (Fig. 4b) and likely essential for cell viability (Curtis and Brun 2014), suggesting that these proteins may have an important function in outer membrane constriction and cell separation. Whereas the presence of genes encoding PG biosynthetic enzymes, PG hydrolases, and putative regulators allows a preliminary model of the *A. tumefaciens* divisome to emerge (Fig. 4b), the functions of these proteins need to be resolved to paint a clear picture of the processes of mid-cell PG biosynthesis, septum cleavage, pole remodeling, and cell separation in *A. tumefaciens*.

5 Concluding Remarks and Outstanding Questions

Although the mechanisms by which polar growth and cell division take place in *A. tumefaciens* are largely unknown, significant advances in the field have recently been accomplished. *A. tumefaciens* has become a model organism for the study of polar growth, and the development of new genetic tools to deplete essential proteins in *A. tumefaciens* (Figueroa-Cuilan et al. 2016; Grangeon et al. 2017) should enable mechanistic studies in the future. These genetic tools, coupled with advances in PG compositional analysis (Cava and de Pedro 2014) and the ability to visualize sites of PG synthesis (Kuru et al. 2012, 2015), provide the resources needed to tackle outstanding questions about the spatial and temporal regulation of PG biosynthesis in *A. tumefaciens*. Just a few of the many remaining questions are summarized below in hopes of encouraging future investigations to improve our understanding of mechanisms underlying the coordination of polar growth and cell division in *A. tumefaciens*.

First, is the PG biosynthesis machinery conserved among polar growing bacteria? Characteristic features have been identified among polar growing bacteria, including an increased abundance of overall muropeptide crosslinking and an enrichment in of LD-crosslinks (Brown et al. 2012). These observations suggest that the PG biosynthesis machinery of polar growing bacteria may require a conserved

subset of enzymes to direct polar elongation. Second, what proteins are required for polar elongation? In *A. tumefaciens*, the predicted elongasome complex is only composed of PBP1a, a bifunctional PBP which is predicted to be essential (Curtis and Brun 2014) and has a polar localization pattern (Cameron et al. 2014). Whereas PBPs are expected to be important contributors to polar PG biosynthesis, other enzymes are also likely contributors. Consistent with the increased abundance of LD-crosslinks, the genomes of polar growing bacteria are enriched in genes which are predicted to encode LDTs. Interestingly, a subfamily of LDTs, YbiS is Rhizobiales-specific and two members of this subfamily of LDTs exhibit polar localization (Cameron et al. 2014); however, the contribution of LDTs to polar elongation is still unknown. Furthermore, additional proteins are likely necessary either to target the enzymes to the growth pole or restrict their activity to the growth pole. Third, how does FtsZ find the middle in *A. tumefaciens*? In *E. coli*, FtsZ negative regulators prevent the polymerization of FtsZ at the cell poles and over the nucleoid, allowing the Z-ring to form only at mid-cell. In bacteria, two well-known systems prevent FtsZ polymerization: the Min system and Nucleoid Occlusion (NO) (Rowlett and Margolin 2013; Wu and Errington 2011). In *A. tumefaciens*, only the Min system is readily identifiable in the genome sequence; however, the Min system is predicted to be dispensable for viability (Curtis and Brun 2014), and absence of the Min system does not cause significant cell division defects (Flores et al. 2018) suggesting that another mechanism must contribute to proper FtsZ positioning. In addition to understanding how FtsZ is properly positioned, it is necessary to understand how the placement of FtsZ at mid-cell is coordinated with the termination of polar growth. Fourth, how is the divisome assembled in *A. tumefaciens*? To begin to address this question, it will be necessary to determine how FtsZ is tethered to the membrane at mid-cell. FtsA arrives at mid-cell after FtsZ (Cameron et al. 2014; Grangeon et al. 2015), indicating that FtsA may not be the only anchor of FtsZ. Other candidates for tethering FtsZ to the membrane include FzlA and FzlC. Next, it will be necessary to confirm and identify the components of the divisome and characterize the contribution of each protein component to septum formation at mid-cell. Finally, how is the activity of autolysins, including amidases and LTGs, coordinated to enable pole remodeling and cell separation? The use of new and existing genetic and biochemical tools for *A. tumefaciens* should help answer these questions and will reveal how essential processes including DNA replication, PG biosynthesis, chromosome segregation, cell division, and cell separation are coordinated during the cell cycle.

Acknowledgements Research in the Brown lab on *A. tumefaciens* cell growth and division is supported by the National Science Foundation (IOS1557806) and WFC is supported by a Gus T. Ridgel Fellowship. We thank Michael VanNieuwenhze (Indiana University) for the gift of the FDAA used in Fig. 2.

References

Anderson-Furgeson JC, Zupan JR, Grangeon R, Zambryski PC (2016) Loss of PodJ in *Agrobacterium tumefaciens* leads to ectopic polar growth, branching, and reduced cell division. J Bacteriol 198(13):1883–1891. https://doi.org/10.1128/JB.00198-16

Barner HD, Cohen SS (1956) The relation of growth to the lethal damage induced by ultraviolet irradiation in *Escherichia coli*. J Bacteriol 71(2):149–157

Bisson-Filho AW, Hsu YP, Squyres GR, Kuru E, Wu F, Jukes C, Sun Y, Dekker C, Holden S, VanNieuwenhze MS, Brun YV, Garner EC (2017) Treadmilling by FtsZ filaments drives peptidoglycan synthesis and bacterial cell division. Science 355(6326):739–743. https://doi.org/10.1126/science.aak9973

Bowman GR, Comolli LR, Zhu J, Eckart M, Koenig M, Downing KH, Moerner WE, Earnest T, Shapiro L (2008) A polymeric protein anchors the chromosomal origin/ParB complex at a bacterial cell pole. Cell 134(6):945–955. https://doi.org/10.1016/j.cell.2008.07.015

Bowman GR, Comolli LR, Gaietta GM, Fero M, Hong SH, Jones Y, Lee JH, Downing KH, Ellisman MH, McAdams HH, Shapiro L (2010) *Caulobacter* PopZ forms a polar subdomain dictating sequential changes in pole composition and function. Mol Microbiol 76(1):173–189. https://doi.org/10.1111/j.1365-2958.2010.07088.x

Bowman GR, Perez AM, Ptacin JL, Ighodaro E, Folta-Stogniew E, Comolli LR, Shapiro L (2013) Oligomerization and higher-order assembly contribute to sub-cellular localization of a bacterial scaffold. Mol Microbiol 90(4):776–795. https://doi.org/10.1111/mmi.12398

Brown PJB, de Pedro MA, Kysela DT, Van der Henst C, Kim J, De Bolle X, Fuqua C, Brun YV (2012) Polar growth in the Alphaproteobacterial order Rhizobiales. Proc Natl Acad Sci U S A 109(5):1697–1701. https://doi.org/10.1073/pnas.1114476109

Cameron TA, Anderson-Furgeson J, Zupan JR, Zik JJ, Zambryski PC (2014) Peptidoglycan synthesis machinery in *Agrobacterium tumefaciens* during unipolar growth and cell division. MBio 5(3):e01219–e01214. https://doi.org/10.1128/mbio.01219-14

Cameron TA, Zupan JR, Zambryski PC (2015) The essential features and modes of bacterial polar growth. Trends Microbiol 23(6):347–353. https://doi.org/10.1016/j.tim.2015.01.003

Cava F, de Pedro MA (2014) Peptidoglycan plasticity in bacteria: emerging variability of the murein sacculus and their associated biological functions. Curr Opin Microbiol 18:46–53. https://doi.org/10.1016/j.mib.2014.01.004

Cava F, Kuru E, Brun YV, de Pedro MA (2013) Modes of cell wall growth differentiation in rod-shaped bacteria. Curr Opin Microbiol 16(6):731–737. https://doi.org/10.1016/j.mib.2013.09.004

Cho H, Wivagg CN, Kapoor M, Barry Z, Rohs PD, Suh H, Marto JA, Garner EC, Bernhardt TG (2016) Bacterial cell wall biogenesis is mediated by SEDS and PBP polymerase families functioning semi-autonomously. Nat Microbiol:16172. https://doi.org/10.1038/nmicrobiol.2016.172

Curtis PD, Brun YV (2014) Identification of essential alphaproteobacterial genes reveals operational variability in conserved developmental and cell cycle systems. Mol Microbiol 93(4):713–735. https://doi.org/10.1111/mmi.12686

Derouaux A, Wolf B, Fraipont C, Breukink E, Nguyen-Disteche M, Terrak M (2008) The monofunctional glycosyltransferase of *Escherichia coli* localizes to the cell division site and interacts with penicillin-binding protein 3, FtsW, and FtsN. J Bacteriol 190(5):1831–1834. https://doi.org/10.1128/JB.01377-07

Du S, Lutkenhaus J (2017) Assembly and activation of the *Escherichia coli* divisome. Mol Microbiol 105(2):177–187. https://doi.org/10.1111/mmi.13696

Du S, Pichoff S, Lutkenhaus J (2016) FtsEX acts on FtsA to regulate divisome assembly and activity. Proc Natl Acad Sci U S A 113(34):E5052–E5061. https://doi.org/10.1073/pnas.1606656113

Ebersbach G, Briegel A, Jensen GJ, Jacobs-Wagner C (2008) A self-associating protein critical for chromosome attachment, division, and polar organization in *Caulobacter*. Cell 134(6):956–968. https://doi.org/10.1016/j.cell.2008.07.016

Egan AJF (2018) Bacterial outer membrane constriction. Mol Microbiol. https://doi.org/10.1111/mmi.13908

Egan AJ, Biboy J, van't Veer I, Breukink E, Vollmer W (2015) Activities and regulation of peptidoglycan synthases. Philos Trans R Soc Lond B Biol Sci 370(1679). https://doi.org/10.1098/rstb.2015.0031

Ehrle HM, Guidry JT, Iacovetto R, Salisbury AK, Sandidge DJ, Bowman GR (2017) Polar organizing protein PopZ is required for chromosome segregation in *Agrobacterium tumefaciens*. J Bacteriol 199(17). https://doi.org/10.1128/jb.00111-17

Erbs G, Silipo A, Aslam S, De Castro C, Liparoti V, Flagiello A, Pucci P, Lanzetta R, Parrilli M, Molinaro A, Newman MA, Cooper RM (2008) Peptidoglycan and muropeptides from pathogens *Agrobacterium* and *Xanthomonas* elicit plant innate immunity: structure and activity. Chem Biol 15(5):438–448. https://doi.org/10.1016/j.chembiol.2008.03.017

Errington J (2015) Bacterial morphogenesis and the enigmatic MreB helix. Nat Rev Microbiol 13(4):241–248. https://doi.org/10.1038/nrmicro3398

Escobar MA, Dandekar AM (2003) *Agrobacterium tumefaciens* as an agent of disease. Trends Plant Sci 8(8):380–386. https://doi.org/10.1016/S1360-1385(03)00162-6

Figueroa-Cuilan W, Daniel JJ, Howell M, Sulaiman A, Brown PJB (2016) Mini-Tn7 insertion in an artificial *att*Tn7 site enables depletion of the essential master regulator CtrA in the phytopathogen *Agrobacterium tumefaciens*. Appl Environ Microbiol 82(16):5015–5025. https://doi.org/10.1128/AEM.01392-16

Flärdh K (2003) Essential role of DivIVA in polar growth and morphogenesis in *Streptomyces coelicolor* A3(2). Mol Microbiol 49(6):1523–1536. https://doi.org/10.1046/j.1365-2958.2003.03660.x

Flores SA, Howell M, Daniel JJ, Piccolo R, Brown PJB (2018) Absence of the Min system does not cause major cell division defects in Agrobacterium tumefaciens. Front Microbiol 9:681. https://doi.org/10.3389/fmicb.2018.00681

Fuchino K, Bagchi S, Cantlay S, Sandblad L, Wu D, Bergman J, Kamali-Moghaddam M, Flardh K, Ausmees N (2013) Dynamic gradients of an intermediate filament-like cytoskeleton are recruited by a polarity landmark during apical growth. Proc Natl Acad Sci U S A 110(21):E1889–E1897. https://doi.org/10.1073/pnas.1305358110

Fujiwara T, Fukui S (1972) Isolation of morphological mutants of *Agrobacterium tumefaciens*. J Bacteriol 110(2):743–746

Fujiwara T, Fukui S (1974a) Effect of D-alanine and mitomycin-C on cell morphology of *Agrobacterium tumefaciens*. J Gen Appl Microbiol 20(6):345–349. https://doi.org/10.2323/jgam.20.345

Fujiwara T, Fukui S (1974b) Unidirectional growth and branch formation of a morphological mutant, Agrobacterium tumefaciens. J Bacteriol 120(2):583–589

Goley ED, Comolli LR, Fero KE, Downing KH, Shapiro L (2010) DipM links peptidoglycan remodelling to outer membrane organization in *Caulobacter*. Mol Microbiol 77(1):56–73. https://doi.org/10.1111/j.1365-2958.2010.07222.x

Goley ED, Yeh YC, Hong SH, Fero MJ, Abeliuk E, McAdams HH, Shapiro L (2011) Assembly of the *Caulobacter* cell division machine. Mol Microbiol 80(6):1680–1698. https://doi.org/10.1111/j.1365-2958.2011.07677.x

Gonzalez MD, Akbay EA, Boyd D, Beckwith J (2010) Multiple interaction domains in FtsL, a protein component of the widely conserved bacterial FtsLBQ cell division complex. J Bacteriol 192(11):2757–2768. https://doi.org/10.1128/JB.01609-09

Goss WA, Deitz WH, Cook TM (1964) Mechanism of action of nalidixic acid on *Escherichia coli*. J Bacteriol 88:1112–1118

Grangeon R, Zupan JR, Anderson-Furgeson J, Zambryski PC (2015) PopZ identifies the new pole, and PodJ identifies the old pole during polar growth in *Agrobacterium tumefaciens*. Proc Natl Acad Sci U S A 112(37):11666–11671. https://doi.org/10.1073/pnas.1515544112

Grangeon R, Zupan J, Jeon Y, Zambryski PC (2017) Loss of PopZ At activity in *Agrobacterium tumefaciens* by deletion or depletion leads to multiple growth poles, minicells, and growth defects. MBio 8(6). https://doi.org/10.1128/mbio.01881-17

Gupta R, Lavollay M, Mainardi JL, Arthur M, Bishai WR, Lamichhane G (2010) The *Mycobacterium tuberculosis* protein LdtMt2 is a nonclassical transpeptidase required for virulence and resistance to amoxicillin. Nat Med 16(4):466–469. https://doi.org/10.1038/nm.2120

Haeusser DP, Margolin W (2016) Splitsville: structural and functional insights into the dynamic bacterial Z ring. Nat Rev Microbiol 14(5):305–319. https://doi.org/10.1038/nrmicro.2016.26

Haney SA, Glasfeld E, Hale C, Keeney D, He Z, de Boer P (2001) Genetic analysis of the *Escherichia coli* FtsZ.ZipA interaction in the yeast two-hybrid system. Characterization of FtsZ residues essential for the interactions with ZipA and with FtsA. J Biol Chem 276(15):11980–11987. https://doi.org/10.1074/jbc.M009810200

Heidrich C, Ursinus A, Berger J, Schwarz H, Holtje JV (2002) Effects of multiple deletions of murein hydrolases on viability, septum cleavage, and sensitivity to large toxic molecules in *Escherichia coli*. J Bacteriol 184(22):6093–6099

Helmstetter CE, Pierucci O (1968) Cell division during inhibition of deoxyribonucleic acid synthesis in *Escherichia coli*. J Bacteriol 95(5):1627–1633

Hempel AM, Wang SB, Letek M, Gil JA, Flardh K (2008) Assemblies of DivIVA mark sites for hyphal branching and can establish new zones of cell wall growth in *Streptomyces coelicolor*. J Bacteriol 190(22):7579–7583. https://doi.org/10.1128/JB.00839-08

Hinz AJ, Larson DE, Smith CS, Brun YV (2003) The *Caulobacter crescentus* polar organelle development protein PodJ is differentially localized and is required for polar targeting of the PleC development regulator. Mol Microbiol 47(4):929–941

Holmes JA, Follett SE, Wang H, Meadows CP, Varga K, Bowman GR (2016) *Caulobacter* PopZ forms an intrinsically disordered hub in organizing bacterial cell poles. Proc Natl Acad Sci U S A 113(44):12490–12495. https://doi.org/10.1073/pnas.1602380113

Howell M, Brown PJB (2016) Building the bacterial cell wall at the pole. Curr Opin Microbiol 34:53–59. https://doi.org/10.1016/j.mib.2016.07.021

Howell M, Aliashkevich A, Salisbury AK, Cava F, Bowman GR, Brown PJB (2017a) Absence of the polar organizing protein PopZ results in reduced and asymmetric cell division in *Agrobacterium tumefaciens*. J Bacteriol 199(17). https://doi.org/10.1128/jb.00101-17

Howell M, Daniel JJ, Brown PJB (2017b) Live cell fluorescence microscopy to observe essential processes during microbial cell growth. J Vis Exp 129. https://doi.org/10.3791/56497

Huang KH, Durand-Heredia J, Janakiraman A (2013) FtsZ ring stability: of bundles, tubules, crosslinks, and curves. J Bacteriol 195(9):1859–1868. https://doi.org/10.1128/JB.02157-12

Johnson JW, Fisher JF, Mobashery S (2013) Bacterial cell-wall recycling. Ann N Y Acad Sci 1277:54–75. https://doi.org/10.1111/j.1749-6632.2012.06813.x

Kang CM, Nyayapathy S, Lee JY, Suh JW, Husson RN (2008) Wag31, a homologue of the cell division protein DivIVA, regulates growth, morphology and polar cell wall synthesis in mycobacteria. Microbiology 154(Pt 3):725–735. https://doi.org/10.1099/mic.0.2007/014076-0

Kilgore WW, Greenberg J (1961) Filament formation and resistane to 1-methyl-3-nitro-1-nitrosoguanidine and other radiomimetic compounds in *Escherichia coli*. J Bacteriol 81

Kuru E, Hughes HV, Brown PJB, Hall E, Tekkam S, Cava F, de Pedro MA, Brun YV, VanNieuwenhze MS (2012) *In situ* probing of newly synthesized peptidoglycan in live bacteria with fluorescent D-amino acids. Angew Chem Int Ed Engl 51(50):12519–12523. https://doi.org/10.1002/anie.201206749

Kuru E, Tekkam S, Hall E, Brun YV, Van Nieuwenhze MS (2015) Synthesis of fluorescent D-amino acids and their use for probing peptidoglycan synthesis and bacterial growth in situ. Nat Protoc 10(1):33–52. https://doi.org/10.1038/nprot.2014.197

Kuykendall LD (2005) Family Rhizobiaceae. In: Brenner DJ, Krieg NR, Stanley JT, Garrity GM (eds) Bergey's manual of systematic bacteriology, vol 2: The Proteobacteria (Part C). Springer, New York, p 324

Kysela DT, Brown PJB, Huang KC, Brun YV (2013) Biological consequences and advantages of asymmetric bacterial growth. Annu Rev Microbiol 67:417–435. https://doi.org/10.1146/annurev-micro-092412-155622

Laloux G, Jacobs-Wagner C (2013) Spatiotemporal control of PopZ localization through cell cycle-coupled multimerization. J Cell Biol 201(6):827–841. https://doi.org/10.1083/jcb.201303036

Laloux G, Jacobs-Wagner C (2014) How do bacteria localize proteins to the cell pole? J Cell Sci 127(Pt 1):11–19. https://doi.org/10.1242/jcs.138628

Lariviere PJ, Szwedziak P, Mahone CR, Lowe J, Goley ED (2018) FzlA, an essential regulator of FtsZ filament curvature, controls constriction rate during *Caulobacter* division. Mol Microbiol 107(2):180–197. https://doi.org/10.1111/mmi.13876

Latch JN, Margolin W (1997) Generation of buds, swellings, and branches instead of filaments after blocking the cell cycle of *Rhizobium meliloti*. J Bacteriol 179(7):2372–2381

Lavollay M, Arthur M, Fourgeaud M, Dubost L, Marie A, Veziris N, Blanot D, Gutmann L, Mainardi JL (2008) The peptidoglycan of stationary-phase *Mycobacterium tuberculosis* predominantly contains cross-links generated by L,D-transpeptidation. J Bacteriol 190 (12):4360–4366. https://doi.org/10.1128/JB.00239-08

Lavollay M, Fourgeaud M, Herrmann JL, Dubost L, Marie A, Gutmann L, Arthur M, Mainardi JL (2011) The peptidoglycan of *Mycobacterium abscessus* is predominantly cross-linked by L,D-transpeptidases. J Bacteriol 193(3):778–782. https://doi.org/10.1128/JB.00606-10

Lea DE, Haines RB, Coulson CA (1937) The action of radiations on bacteria III—γ-rays on growing and on non-proliferating bacteria. Proc R Soc Lond B Biol Sci 123:1–21

Letek M, Fiuza M, Ordonez E, Villadangos AF, Flardh K, Mateos LM, Gil JA (2009) DivIVA uses an N-terminal conserved region and two coiled-coil domains to localize and sustain the polar growth in *Corynebacterium glutamicum*. FEMS Microbiol Lett 297(1):110–116. https://doi.org/10.1111/j.1574-6968.2009.01679.x

Liu B, Persons L, Lee L, de Boer PA (2015) Roles for both FtsA and the FtsBLQ subcomplex in FtsN-stimulated cell constriction in *Escherichia coli*. Mol Microbiol 95(6):945–970. https://doi.org/10.1111/mmi.12906

Lutkenhaus J, Donachie WD (1979) Identification of the *ftsA* gene product. J Bacteriol 137 (3):1088–1094

Lutkenhaus JF, Wolf-Watz H, Donachie WD (1980) Organization of genes in the *ftsA-envA* region of the *Escherichia coli* genetic map and identification of a new *fts* locus (*ftsZ*). J Bacteriol 142 (2):615–620

Margolin W (2009) Sculpting the bacterial cell. Curr Biol 19(17):R812–R822. https://doi.org/10.1016/j.cub.2009.06.033

Massey TH, Mercogliano CP, Yates J, Sherratt DJ, Lowe J (2006) Double-stranded DNA translocation: structure and mechanism of hexameric FtsK. Mol Cell 23(4):457–469. https://doi.org/10.1016/j.molcel.2006.06.019

Meeske AJ, Riley EP, Robins WP, Uehara T, Mekalanos JJ, Kahne D, Walker S, Kruse AC, Bernhardt TG, Rudner DZ (2016) SEDS proteins are a widespread family of bacterial cell wall polymerases. Nature 537(7622):634–638. https://doi.org/10.1038/nature19331

Meier EL, Razavi S, Inoue T, Goley ED (2016) A novel membrane anchor for FtsZ is linked to cell wall hydrolysis in *Caulobacter crescentus*. Mol Microbiol 101(2):265–280. https://doi.org/10.1111/mmi.13388

Morton ER, Fuqua C (2012) Genetic manipulation of *Agrobacterium*. Curr Protoc Microbiol 25:3D.2.1–3D.2.15. https://doi.org/10.1002/9780471729259.mc03d02s25

Nester E (2014) *Agrobacterium:* nature's genetic engineer. Front Plant Sci. https://doi.org/10.3389/fpls.2014.00730

Oliva MA, Cordell SC, Lowe J (2004) Structural insights into FtsZ protofilament formation. Nat Struct Mol Biol 11(12):1243–1250. https://doi.org/10.1038/nsmb855

Paradis-Bleau C, Markovski M, Uehara T, Lupoli TJ, Walker S, Kahne DE, Bernhardt TG (2010) Lipoprotein cofactors located in the outer membrane activate bacterial cell wall polymerases. Cell 143(7):1110–1120. https://doi.org/10.1016/j.cell.2010.11.037

Park JT, Uehara T (2008) How bacteria consume their own exoskeletons (turnover and recycling of cell wall peptidoglycan). Microbiol Mol Biol Rev 72(2):211–227. https://doi.org/10.1128/MMBR.00027-07

Peters NT, Dinh T, Bernhardt TG (2011) A fail-safe mechanism in the septal ring assembly pathway generated by the sequential recruitment of cell separation amidases and their activators. J Bacteriol 193(18):4973–4983. https://doi.org/10.1128/JB.00316-11

Pichoff S, Lutkenhaus J (2005) Tethering the Z ring to the membrane through a conserved membrane targeting sequence in FtsA. Mol Microbiol 55(6):1722–1734. https://doi.org/10.1111/j.1365-2958.2005.04522.x

Pichoff S, Du S, Lutkenhaus J (2015) The bypass of ZipA by overexpression of FtsN requires a previously unknown conserved FtsN motif essential for FtsA-FtsN interaction supporting a model in which FtsA monomers recruit late cell division proteins to the Z ring. Mol Microbiol 95(6):971–987. https://doi.org/10.1111/mmi.12907

Ricard M, Hirota Y (1973) Process of cellular division in *Escherichia coli*: physiological study on thermosensitive mutants defective in cell division. J Bacteriol 116(1):314–322

Rolinson GN (1980) Effect of beta-lactam antibiotics on bacterial cell growth rate. J Gen Microbiol 120(2):317–323

Rowlett VW, Margolin W (2013) The bacterial Min system. Curr Biol 23(13):R553–R556. https://doi.org/10.1016/j.cub.2013.05.024

Sanders AN, Wright LF, Pavelka MS Jr (2014) Genetic characterization of mycobacterial L,D-transpeptidases. Microbiology 160(Pt 8):1795–1806. https://doi.org/10.1099/mic.0.078980-0

Sauvage E, Kerff F, Terrak M, Ayala JA, Charlier P (2008) The penicillin-binding proteins: structure and role in peptidoglycan biosynthesis. FEMS Microbiol Rev 32(2):234–258. https://doi.org/10.1111/j.1574-6976.2008.00105.x

Scheffers DJ, Pinho MG (2005) Bacterial cell wall synthesis: new insights from localization studies. Microbiol Mol Biol Rev 69(4):585–607. https://doi.org/10.1128/MMBR.69.4.585-607.2005

Siegrist MS, Swarts BM, Fox DM, Lim SA, Bertozzi CR (2015) Illumination of growth, division and secretion by metabolic labeling of the bacterial cell surface. FEMS Microbiol Rev 39 (2):184–202. https://doi.org/10.1093/femsre/fuu012

Tsang MJ, Bernhardt TG (2015) A role for the FtsQLB complex in cytokinetic ring activation revealed by an *ftsL* allele that accelerates division. Mol Microbiol 95(6):925–944. https://doi.org/10.1111/mmi.12905

Typas A, Banzhaf M, van den van Saparoea Berg B, Verheul J, Biboy J, Nichols RJ, Zietek M, Beilharz K, Kannenberg K, von Rechenberg M, Breukink E, den Blaauwen T, Gross CA, Vollmer W (2010) Regulation of peptidoglycan synthesis by outer-membrane proteins. Cell 143(7):1097–1109. https://doi.org/10.1016/j.cell.2010.11.038

Typas A, Banzhaf M, Gross CA, Vollmer W (2011) From the regulation of peptidoglycan synthesis to bacterial growth and morphology. Nat Rev Microbiol 10(2):123–136. https://doi.org/10.1038/nrmicro2677

Uehara T, Bernhardt TG (2011) More than just lysins: peptidoglycan hydrolases tailor the cell wall. Curr Opin Microbiol 14(6):698–703. https://doi.org/10.1016/j.mib.2011.10.003

Uehara T, Dinh T, Bernhardt TG (2009) LytM-domain factors are required for daughter cell separation and rapid ampicillin-induced lysis in *Escherichia coli*. J Bacteriol 191(16):5094–5107. https://doi.org/10.1128/JB.00505-09

Uehara T, Parzych KR, Dinh T, Bernhardt TG (2010) Daughter cell separation is controlled by cytokinetic ring-activated cell wall hydrolysis. EMBO J 29(8):1412–1422. https://doi.org/10.1038/emboj.2010.36

Van De Putte P, Van D, Roersch A (1964) The selection of mutants of *Escherichia coli* with impaired cell division at elevated tempratures. Mutat Res 106:121–128

van Heijenoort J (2011) Peptidoglycan hydrolases of *Escherichia coli*. Microbiol Mol Biol Rev 75 (4):636–663. https://doi.org/10.1128/MMBR.00022-11

Viollier PH, Sternheim N, Shapiro L (2002) Identification of a localization factor for the polar positioning of bacterial structural and regulatory proteins. Proc Natl Acad Sci U S A 99 (21):13831–13836. https://doi.org/10.1073/pnas.182411999

Vollmer W, Joris B, Charlier P, Foster S (2008) Bacterial peptidoglycan (murein) hydrolases. FEMS Microbiol Rev 32(2):259–286. https://doi.org/10.1111/j.1574-6976.2007.00099.x

Weiss DS (2015) Last but not least: new insights into how FtsN triggers constriction during *Escherichia coli* cell division. Mol Microbiol 95(6):903–909. https://doi.org/10.1111/mmi.12925

Wu LJ, Errington J (2011) Nucleoid occlusion and bacterial cell division. Nat Rev Microbiol 10(1):8–12. https://doi.org/10.1038/nrmicro2671

Yang DC, Peters NT, Parzych KR, Uehara T, Markovski M, Bernhardt TG (2011) An ATP-binding cassette transporter-like complex governs cell-wall hydrolysis at the bacterial cytokinetic ring. Proc Natl Acad Sci U S A 108(45):E1052–E1060. https://doi.org/10.1073/pnas.1107780108

Yang DC, Blair KM, Salama NR (2016) Staying in shape: the impact of cell shape on bacterial survival in diverse environments. Microbiol Mol Biol Rev 80(1):187–203. https://doi.org/10.1128/Mmbr.00031-15

Yang X, Lyu Z, Miguel A, McQuillen R, Huang KC, Xiao J (2017) GTPase activity-coupled treadmilling of the bacterial tubulin FtsZ organizes septal cell wall synthesis. Science 355 (6326):744–747. https://doi.org/10.1126/science.aak9995

Zhao G, Meier TI, Kahl SD, Gee KR, Blaszczak LC (1999) BOCILLIN FL, a sensitive and commercially available reagent for detection of penicillin-binding proteins. Antimicrob Agents Chemother 43(5):1124–1128

Zupan JR, Cameron TA, Anderson-Furgeson J, Zambryski PC (2013) Dynamic FtsA and FtsZ localization and outer membrane alterations during polar growth and cell division in *Agrobacterium tumefaciens*. Proc Natl Acad Sci U S A 110(22):9060–9065. https://doi.org/10.1073/pnas.1307241110

Exopolysaccharides of *Agrobacterium tumefaciens*

Ann G. Matthysse

Contents

1 Introduction 112
2 Cyclic β-(1, 2)-Glucan 114
2.1 Structure and Biosynthesis 114
2.2 Regulation 118
2.3 Function 119
3 Succinoglycan 120
3.1 Structure and Biosynthesis 120
3.2 Regulation 121
3.3 Function 122
4 Cellulose 122
4.1 Structure and Biosynthesis 122
4.2 Regulation 125
4.3 Function 127
5 Curdlan 127
5.1 Structure and Biosynthesis 127
5.2 Regulation 128
5.3 Function 130
6 The Unipolar Polysaccharide (UPP) 131
6.1 Structure and Biosynthesis 131
6.2 Regulation 131
6.3 Function 133
7 Summary and Conclusions 134
7.1 Structure and Biosynthesis 134
7.2 Regulation 135
7.3 Function 136
References 137

Abstract Agrobacterium exopolysaccharides play a major role in the life of the cell. Exopolysaccharides are required for bacterial growth as a biofilm and they protect the bacteria against environmental stresses. Five of the exopolysaccharides

A. G. Matthysse (✉)
Department of Biology, University of North Caroline, Chapel Hill, NC, USA
e-mail: ann_matthysse@unc.edu

Current Topics in Microbiology and Immunology (2018) 418:111–141
DOI 10.1007/82_2018_100

Published Online: 11 July 2018

made by *A. tumefaciens* have been characterized extensively with respect to their structure, synthesis, regulation, and role in the life of the bacteria. These are cyclic-β-(1, 2)-glucan, cellulose, curdlan, succinoglycan, and the unipolar polysaccharide (UPP). This chapter describes the structure, synthesis, regulation, and function of these five exopolysaccharides.

1 Introduction

Agrobacterium exopolysaccharides play a major role in the life of the cell. Exopolysaccharides are required for bacterial growth as a biofilm and they protect the bacteria against environmental stresses (Weiner et al. 1995). Under some growth conditions, exopolysaccharides make up 20% or more of the dry weight of the bacterial culture (Breedveld and Miller 1994). Five of the exopolysaccharides made by *A. tumefaciens* have been characterized extensively with respect to their structure, synthesis, regulation, and role in the life of the bacteria. These are cyclic-β-(1, 2)-glucan, cellulose, curdlan, succinoglycan, and the unipolar polysaccharide (UPP). *A. tumefaciens* and its close relatives make additional exopolysaccharides including a polymer of glucose, galactose, and pyruvic acid with a 17-mer repeat unit (O'Neill et al. 1992); a polymer of glucose, galactose, mannose, and pyruvic acid (Hou et al. 1996); a gel-forming polymer of galactose, arabinose, and aminogalactose named PGHX (Liu et al. 2016); and an acetylated polymer-containing glucose, galactose and an unidentified deoxy-sugar which is capable of interfering with the binding of *A. tumefaciens* to plants (Reuhs et al. 1997). There is little additional information on the synthesis, regulation, or role of these less characterized exopolysaccharides. Inspection of the genome of *A. tumefaciens* C58 suggests that the bacteria can produce additional exopolysaccharides, which have not yet been described.

This chapter will only discuss the five exopolysaccharides which have been well characterized. For each of these more is known about some aspects of their biology than others. The genes required for the synthesis of each of these exopolysaccharides have largely been identified. The structure is known for all but the unipolar polysaccharide (UPP). Much is known about the regulation of the synthesis of the UPP and of cellulose. There is less information about the regulation of cyclic β-(1, 2) glucans and succinoglycan. The information available on the regulation of curdlan synthesis exclusively derives from an overproducing industrial strain with an unknown relationship to wild-type strains such as C58. The roles of each of these exopolysaccharides in the life of the bacteria are still being explored. Tables listing the genes known to be involved in the synthesis (Table 1) and regulation (Table 2) of these exopolysaccharides are included in the chapter.

Table 1 Genes Involved in the Synthesis of Exopolysaccharides in *A. tumefaciens*

Gene name	Gene number	Function	References
Cyclic-beta-1,2-glucan			
chvA	*Atu2728*	Polysaccharide transport	Cangelosi et al. (1989)
chvB	*Atu2730*	Glycosyltransferase	Zorreguieta et al. (1985)
Cellulose			
celA	*Atu3309*	Cellulose synthase	Matthysse et al. (1995)
celB	*Atu3308*	Membrane protein	Matthysse et al. (1995)
celC	*Atu3307*	Endoglucanase	Matthysse et al. (1995)
celD	*Atu3302*	Unknown	Matthysse et al. (1995)
celE	*Atu3303*	Unknown	Matthysse et al. (1995)
Curdlan			
crdA	*Atu3057*	Membrane protein	Stasinopoulos et al. (1999)
crdS	*Atu3056*	Curdlan synthase	Stasinopoulos et al. (1999)
crdC	*Atu3055*	Polysaccharide transport	McIntosh et al. (2005)
Succinoglycan			
exoA	*Atu4053*	Glycosyltransferase	Reuber and Walker (1993)
exoB	*Atu4166*	UDP-glucose-epimerase	Reuber and Walker (1993)
exoC (pgm)	*Atu4074*	Phosphoglucomutase	Reuber and Walker (1993)
exoF	*Atu3326*	Polysaccharide transport	Reuber and Walker (1993)
exoH	*Atu4056*	Succinyltransferase	Reuber and Walker (1993)
exoK	*Atu4055*	Endo-glycanase	York and Walker (1998)
exoL	*Atu4054*	Glycosyltransferase	Reuber and Walker (1993)
exoM	*Atu4052*	Glycosyltransferase	Reuber and Walker (1993)
exoN	*Atu4050*	UTP-glucose-1-phosphate uridyltransferase	Reuber and Walker (1993)
exoO	*Atu4051*	Glycosyltransferase	Reuber and Walker (1993)
exoP	*Atu4049*	Wzz homologue, polysaccharide transport	Reuber and Walker (1993)
exoQ	*Atu3325*	Polysaccharide transport	Reuber and Walker (1993)

(continued)

Table 1 (continued)

Gene name	Gene number	Function	References
exoT	*Atu4057*	Polysaccharide transport	Reuber and Walker (1993)
exoU	*Atu4060*	Glycosyltransferase	Reuber and Walker (1993)
exoV	*Atu4059*	Polysaccharide pyruvyltransferase	Reuber and Walker (1993)
exoW	*Atu4058*	Glycosyltransferase	Reuber and Walker (1993)
exoY	*Atu3327*	Polysaccharide transport	Reuber and Walker (1993)
Unipolar polysaccharide			
uppA	*Atu1235*	Membrane protein, may transfer a sugar to bactoprenol	Xu et al. (2012)
uppB	*Atu1236*	Glycosyltransferase type 1	Xu et al. (2013)
uppC	*Atu1237*	Glycosyltransferase type 4	Xu et al. (2012)
uppD (gumB)	*Atu1238*	Polysaccharide export protein, membrane protein	Xu et al. (2012)
uppE (exoP)	*Atu1239*	Polysaccharide transport protein, membrane protein	Xu et al. (2012)
uppF	*Atu1240*	Possible acetyltransferase	Xu et al. (2012)
pssA	*Atu0102*	Homologue of *uppE* and *exoP*, membrane protein	Xu et al. (2012)

2 Cyclic β-(1, 2)-Glucan

2.1 Structure and Biosynthesis

Cyclic-β-(1, 2)-glucans are polymers of β-(1→2) linked D-glucose molecules. They are found in many members of the Rhizobiacea (Breedveld and Miller 1994). In agrobacteria, the polymers contain between 17 and 25 glucose residues. They may contain substitutions which vary with species; in agrobacteria, they are principally 1-phosphoglycerol linked by a phosphodiester bond to C-6 of glucose (Fig. 1b). Substitutions are added to the neutral cyclic glucan after it is synthesized. The cyclic glucans are predominately located in the periplasmic space, where they may reach concentrations as high as 15 mM. In stationary phase or at elevated temperatures, varying amounts may be secreted to the medium (Breedveld and Miller 1994).

Table 2 Genes Involved in the Regulation of Exopolysaccharide Synthesis in *A. tumefaciens* C58

Gene name	Gene number	Exopolysaccharides regulated	Phenotype of mutant	Other characteristics	References
celG	*Atu8186*	Cellulose	Cellulose overproduction		Matthysse et al. (2005)
celI	*Atu3105*	Cellulose	Cellulose overproduction	Transcriptional regulator	Matthysse et al. (2005)
chvG	*Atu0033*	Succinoglycan	Decreased succinoglycan, increased biofilm formation	Sensor of acidic pH	Cheng and Walker (1998), Heckel et al. (2014), Tomlinson et al. (2010), Wu et al. (2016)
chvI	*Atu0034*	Succinolgycan	Decreased succinoglycan, increased biofilm formation	Response regulator, acidic pH	Cheng and Walker (1998, Heckel et al. (2014), Tomlinson et al. (2010), Wu et al. 2016)
claR	*Atu1631*	UPP and cellulose	Decreased biofilm formation	CheY-like protein	Feirer et al. (2017)
crdR	*Atu0361*	Curdlan	Decreased curdlan production	Transcriptional regulator	Stasinopoulos et al. (1999)
dcgA	*Atu1257*	UPP and cellulose	Decreased UPP and cellulose, decreased biofilm	EAL/GGDEF protein	Xu et al. (2013)
dcgB	*Atu1691*	UPP and cellulose	Decreased UPP and cellulose, decreased biofilm	EAL/GGDEF protein	Xu et al. (2013)
dcgC	*Atu2179*	Cellulose	Triple mutant *degA*, *B*, and *C*: lacks cellulose	EAL/GGDEF protein	Xu et al. (2013)
exoC (*pgm*, *pscA*)	*Atu4074*	Cyclic β-1,2 glucan, succinoglycan	Decreased cyclic β-1,2 glucan, decreased succinoglycan	Phosphoglucose mutase	Marks et al. (1987)
dcpA	*Atu3495*	UPP and cellulose	Increased biofilm, increased c-di-GMP	EAL/GGDEF protein	Feirer et al. (2015)

(continued)

Table 2 (continued)

Gene name	Gene number	Exopolysaccharides regulated	Phenotype of mutant	Other characteristics	References
exoR	*Atu1715*	Succinoglycan	Decreased biofilm, nonmotile, increased succinoglycan, increased ChvI/ChvG activity	Regulator of acid-sensing two-component system (ChvG/ChvI)	Tomlinson et al. (2010)
exoX[b]	*Atu4061*	Succinoglycan	Increased succinoglycan	Post-transcriptional regulator	Kamoun et al. (1989)
exsB (*psdA*)	*Atu3346*	Succinoglycan	Increased succinoglycan	Post-transcriptional regulator	Kamoun et al. (1989)
ntrB[a]	*Atu1445*	Curdlan	Decreased curdlan production, cannot grow on nitrate	Two component system	Stanisich and Stone (2009), Yu et al. (2011a)
ntrC[a]	*Atu1446*	Curdlan	Decreased curdlan production, cannot grow on nitrate	Two-component system	Stanisich and Stone (2009), Yu et al. (2011a)
ntrX[a]	*Atu1448*	Curdlan	Essential gene	Maybe a positive regulator of curdlan, two-component system	Ruffing and Chen (2012), Stanisich and Stone (2009)
ntrY[a]	*Atu1447*	Curdlan	Decreased curdlan	Two-component system	Ruffing and Chen (2012), Stanisich and Stone (2009)
odc	*Atu3196*	Cellulose	Increased cellulose, increases cyclic-di-GMP	Ornithine decarboxylase, required for spermidine synthesis	Wang et al. (2016)
phoB	*Atu0425*	UPP	Essential gene	Low phosphate required for UPP synthesis	Danhorn et al. (2004)
pleD (*celR*)	*Atu1297*	UPP, cellulose	Decreased UPP and cellulose, decreased biofilm		Barnhart et al. (2013)
ppx1[a]	*Atu0619*	Curdlan	Increased curdlan in log phase	Exopolyphosphatase	Ruffing and Chen (2012)

(continued)

Table 2 (continued)

Gene name	Gene number	Exopolysaccharides regulated	Phenotype of mutant	Other characteristics	References
pruA	*Atu1130*	UPP, cellulose	Increased UPP and cellulose	Pterine synthesis	Feirer et al. (2015)
pruR		UPP, cellulose	Increased UPP and cellulose	Regulates esterase activity of DcpA	Feirer et al. (2015)
relA[a]	*Atu1030*	Curdlan	Decreased curdlan	Required for the stringent response	Ruffing and Chen (2012)
ros	*Atu0916*	Succinoglycan	Decreased succinoglycan	Transcriptional regulator, binds zinc	Brightwell et al. (1995), Hussain and Johnston (1997)
visN	*Atu0524*	UPP, cellulose	Increased UPP and cellulose, decreased succinoglycan, nonmotile, aggregates in liquid	Required for motility	Xu et al. (2013)
visR	*Atu0525*	UPP, cellulose	Increased UPP and cellulose, decreased succinoglycan, nonmotile, aggregates in liquid	Required for motility	Xu et al. (2013)
[a]*unnamed*	*Atu1114*	Curdlan	Decreased curdlan	EAL/GGDEF protein	Wu et al. (2016)

[a]This gene has been identified as playing a role in the regulation of curdlan synthesis in *Agrobacterium species* LTU50 (ATCC 3179). The regulation of curdlan synthesis in *A. tumefaciens* C58 has not been studied

[b]This gene has been identified as playing a role in the regulation of succinoglycan synthesis in *S. meliloti*. Its role in *A. tumefaciens* has not been described

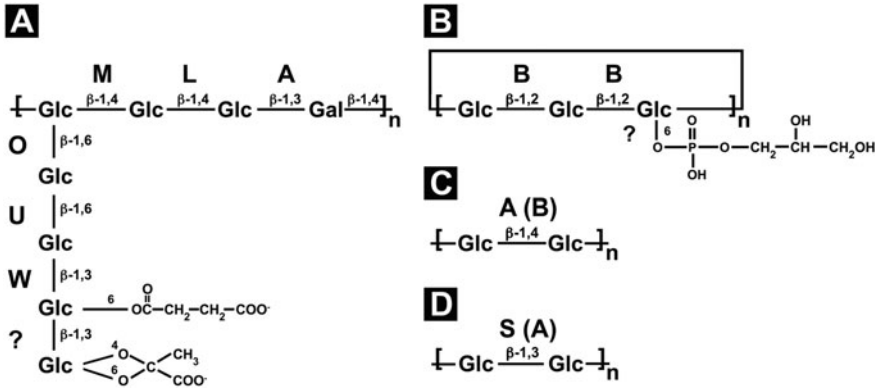


Fig. 1 **a** Structure of the repeating unit of succinoglycan of *A. tumefaciens*. Letters M, L, A, O, U, and W indicate the Exo enzymes which catalyze the formation of the indicated bond. ? indicates that the Exo enzyme which forms this bond has not been identified. The addition of succinic acid is catalyzed by ExoH, and the addition of pyruvate is catalyzed by ExoV. **b** Structure of cyclic-β-(1, 2)-glucan. The letter B indicates ChvB which is responsible for forming the bond between the glucose units and for forming a cyclic molecule. The cyclic ring contains between 17 and 25 glucose units. The number of phosphorylglycerol substitutions per ring is variable but averages about 3. ? indicates that the enzyme which adds the phosphorylglycerol has not been identified. **c** Structure of the repeating unit of cellulose. A (B) indicates that CelA is responsible for the synthesis. CelA requires the presence of an adjacent CelB molecule to function. **d** Structure of the repeating unit of curdlan. S (A) indicates that CrdS is responsible for the synthetic reaction. CrdS requires the presence of an adjacent CrdA molecule to function

The precursor for the cyclic-β-(1, 2)-glucans is UDP-glucose (Castro et al. 1996). Two genes, *chvA* and *chvB*, are required for the biosynthesis of cyclic-β-(1, 2)-glucans in *A. tumefaciens* (Table 1). Both are cell membrane proteins. ChvB is the glycosyl-transferase which synthesizes the glucose chain. It also appears to catalyze the cyclization of the chain. A large portion of the C-terminal region of ChvB is not required for cyclic-β-(1, 2)-glucan synthesis (Zorreguieta et al. 1988). The function of this region is unknown. ChvA transports the cyclic-β-(1, 2)-glucan across the membrane (O'Connell and Handelsman 1989). The *exoC* (aka *pgm* and *pscA*) gene is also required for cyclic-β-(1, 2)-glucan synthesis (Thomashow et al. 1987; Uttaro et al. 1990). It encodes a phosphoglucomutase which catalyzes the reaction α-D-glucose 1-phosphate ↔ α-D-glucose 6-phosphate. This reaction is required for the synthesis of many polysaccharides including cyclic-β-(1, 2)-glucan, succinoglycan, cellulose, curdlan, and lipopolysaccharide and is not specific to the cyclic glucans.

2.2 Regulation

The production of cyclic-β-(1, 2)-glucans is controlled by ionic and osmotic strength. As far as is known, the expression of the *chvA* and *chvB* genes is

constitutive (Ingram-Smith and Miller 1998). Enzymatic activity of ChvB is inhibited in vitro by as little as 10 mM NaCl. Cyclic-β-(1, 2)-glucans are made when the cell is in low-ionic strength medium. Neutral solutes such as glycine betaine or trehalose added to the medium at 1 M concentration stimulate cyclic-β-(1, 2)-glucan synthesis and reverse the inhibition by ions (Ingram-Smith and Miller 1998). ChvA and the C-terminal portion of ChvB are not required for these effects. High osmolarity of the medium due to solutes which are not osmoprotectants, such as 0.5 M sucrose or mannitol, reduces the net accumulation of cyclic-β-(1, 2)-glucans but does not affect the activity of the enzyme. An enzyme capable of degrading β-(1, 2)-glucans has been identified in *Chitinophaga aversicola* (Abe et al. 2017). *A. tumefaciens* has a homologue to this gene (*Atu3054*) encoding a protein of unidentified function which might play a role in regulating the amount of cyclic-β-(1, 2)-glucans present in the cell.

Bacteria grown in low-phosphate media do not have phosphoglycerol substituents on their cyclic-β-(1, 2)-glucans. Under these conditions the cyclic-β-(1, 2)-glucans are neutral. The amount of the cyclic-β-(1, 2)-glucans produced is unchanged. In agrobacteria, the neutral glucans seem to function in the same manner as the negatively charged polymers (Breedveld et al. 1995).

2.3 Function

Cyclic-β-(1, 2)-glucans are needed by the bacteria to allow normal growth in low ionic strength media. Mutants which lack ChvB show slower growth in medium consisting of 0.1% yeast extract (Ingram-Smith and Miller 1998). The addition of 0.02% NaCl restored the growth of the mutants to that of the wild-type parent. ChvB mutants also show a number of other phenotypes which are puzzling. They are nonmotile, unable to transfer the plasmid pAGK84 by conjugation, fail to bind to plant surfaces, and are avirulent (Puvanesarajah et al. 1985). These multiple phenotypes suggest that the lack of cyclic-β-(1, 2)-glucans causes changes in the surface of the bacteria. Most of these changes can be at least partially reversed at high osmolarity (Ingram-Smith and Miller 1998). Lack of binding to plant surfaces and virulence can also be reversed by incubating the bacteria with plants at a low temperature, 16 °C (Bash and Matthysse 2002). When first isolated, ChvB mutants were described as unable to bind to plant cells and it was proposed that cyclic-β-(1, 2)-glucans mediate binding to plant surfaces (Douglas et al. 1982). However, cyclic-β-(1, 2)-glucans added to the medium have no effect on binding of either wild-type or mutant bacteria (Swart et al. 1994). The effects on binding and virulence of mutations in genes required for cyclic-glucan production were also observed to vary with the plant host (Hawes and Pueppke 1989). Thus, the major role of these glucose oligomers appears to be in allowing cells to function in low-ionic strength medium.

3 Succinoglycan

3.1 Structure and Biosynthesis

Succinoglycan is a branched chain polysaccharide made up of repeating units of eight sugars. The backbone contains one galactose and three glucose residues with a side chain of four glucose molecules. The basic structure is shown in Fig. 1a. There is generally a succinic acid and a pyruvate substitution on the side chain as shown. In many rhizobia, there is also an acetate substitution on the main chain, but this is usually absent in agrobacteria (Chouly et al. 1995).

The pathway for the biosynthesis of succinoglycan was studied in *S. meliloti* (Reuber and Walker 1993). Some of the genes and reactions are similar or identical in *A. tumefaciens* (Cangelosi et al. 1987; Wu et al. 2016). The pathway described here is that determined for *S. meliloti* except as otherwise noted (Table 1). The precursor for succinoglycan is UDP-glucose. This is formed from glucose-1-phosphate by the action of ExoC (phosphoglucomutase, aka PscA) and ExoN (UTP-glucose-1-phosphate uridyltransferase) (Marks et al. 1987). ExoB converts the UDP-glucose into UDP-galactose. The polysaccharide chain is initiated by transferring galactose from UDP to a lipid carrier, probably bactoprenol. This reaction is carried out by ExoY and requires ExoF. Successive glucose residues are added by glycosyltransferases, all of which belong to the GT2 family except ExoL. These include, in order of reactions they catalyze in the pathway, ExoA, ExoL, ExoM, ExoO, ExoU, and ExoW. The enzyme which adds the last glucose is undetermined. ExoH adds a succinic acid to the C6 of the seventh glucose. ExoV adds pyruvate to C4 and C6 of the terminal glucose on the side chain (Glucksmann et al. 1993). In *S. meliloti,* there is an acetate group on the third sugar residue added by ExoZ. This acetate is generally not present in *A. tumefaciens* (Glucksmann et al. 1993). Most of the enzymes involved in these reactions after the addition of galactose to the lipid carrier are membrane associated. Once the octomeric unit of succinoglycan is made, it is transported out of the cell and polymerized into long chains. These reactions require ExoP, ExoQ, and ExoT. The mechanism of transport and polymerization has not been established (Reuber and Walker 1993). In *S. meliloti* both long-chain succinoglycan and shorter chain succinoglycan are made. The shorter chain molecules appear to be the polymer involved in the symbiotic interaction with the plant leading to the formation of nitrogen-fixing nodules (Cheng and Walker 1998). The shorter chains may be formed by the action of ExoK (York and Walker 1998). There is no evidence for the formation of short chains in *A. tumefaciens*, but there is a homologue of the *exoK* gene located in the operon containing many of the genes for the synthesis of succinoglycan.

It should be noted that other genes homologous to the genes identified as required for the synthesis of succinoglycan are located at various locations in the C58 genome. They include Atu2373 (*exoM*), *Atu2372* (*exoU*), *Atu2375* (*exoT*),

Atu3009 (*exoZ*), *Atu3550* (*exoT*), and *Atu3556* (*exoP*). These genes are presumably involved in the synthesis of other as yet identified exopolysaccharides, but are sometimes named using the succinoglycan biosynthesis gene names.

3.2 Regulation

It is clear that the production of succinoglycan is regulated in *A. tumefaciens* C58. Several genes have been identified which are capable of affecting the succinoglycan synthesis, but no clear regulatory pathway or link between these genes has been identified (Table 2 and Fig. 2). Two genes, *exsB* (aka *psdA*) and *exoX*, have been identified as negative regulators of succinoglycan production. Both of the genes appear to act post-transcriptionally (Kamoun et al. 1989). Another gene, *ros*, has been identified as a negative regulator of the *virC* and *virD* operons and the *ipt* oncogene. Ros is a zinc-finger DNA-binding protein (Chou et al. 1998). Mutations in *ros* result in failure to produce succinoglycan in C58. Ros requires iron in the medium to function (Hussain and Johnston 1997). Paradoxically, low iron levels are reported to increase the synthesis of several Exo proteins (Gonzalez et al. 1996). When the amino-terminal end of the Ros protein was removed, the resulting strains carrying both the mutant and wild-type genes were unable to make succinoglycan (Brightwell et al. 1995). In *A. radiobacter*, *ros* is required for the transcription of *exoY* and thus for succinoglycan formation (Tiburtius et al. 1996). Presumably, the same mechanism accounts for the effect of *ros* on succinoglycan production in *A. tumefaciens* C58. There is no information as to whether Ros interacts directly with the *exoY* promoter or the effect is mediated by the effect of Ros on another

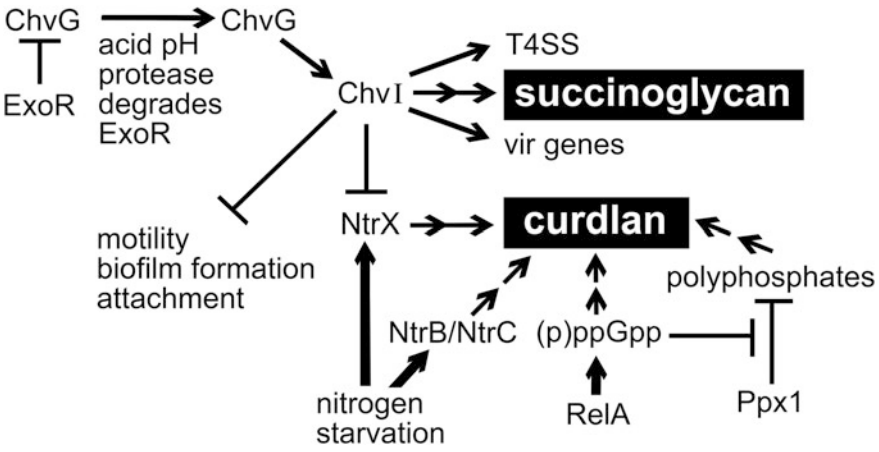


Fig. 2 The regulation of the synthesis of succinoglycan and curdlan and the interactions between the control of the synthesis of these exopolysaccharides. See Sects. 3.2 and 5.2 for an explanation of the genes and interactions involved

gene. The fact that there is no apparent Ros bonding sequence (*ros* box) upstream of *exoY* suggests the later alternative.

The ExoR/ChvG/ChvI system which regulates the response to acidic pH also regulates succinoglycan synthesis (Cheng and Walker 1998; Heckel et al. 2014; Tomlinson et al. 2010; Wu et al. 2012). ChvG/ChvI is a two-component system. ChvG is a sensor located in the cell membrane. At neutral pH, it is blocked from autophosphorylation by the binding of the periplasmic protein ExoR. At acidic pH, ExoR is degraded by a protease and ChvG is able to autophosphorylate and then phosphorylate ChvI. ChvI then activates transcription of type VI secretion, succinoglycan biosynthesis, and virulence genes (Tomlinson et al. 2010; Wu et al. 2016). Phosphorylated ChvI represses motility, biofilm formation, and bacterial attachment. Phosphorylated ChvI may also repress *ntrX* in *A. tumefaciens* (Heckel et al. 2014). NtrX is a negative regulator of succinoglycan biosynthesis in *S. meliloti* (Wang et al. 2013). NtrX is a positive regulator of curdlan biosynthesis in agrobacteria (Ruffing and Chen 2012). Activation of succinoglycan biosynthesis and repression of *ntrX* by phosphorylated ChvI would explain the observation that curdlan and succinoglycan generally appear to be regulated in opposite directions so that when one is made the other is not (Fig. 2).

3.3 Function

Succinoglycan is not required for virulence or for biofilm formation by agrobacteria. Colonies of agrobacteria that lack succinoglycan appear rough rather than mucoid. It is possible that succinoglycan helps the bacteria to retain water and protects them from various environmental stresses. Succinoglycan-minus mutants are more sensitive to acid than are the wild-type parent bacteria (Halder et al. 2017). Other possible functions of succinoglycan have not been explored. Because the bacteria often synthesize succinoglycan in quantities amounting to more than 20% of the total cell dry weight, it must be assumed that this polysaccharide is required in relatively large amounts for normal growth and/or survival in nature (Wu et al. 2016).

4 Cellulose

4.1 Structure and Biosynthesis

Cellulose is a liner polymer of β-(1→4)-linked D-glucose molecules (Fig. 1c). The polymers are generally quite long, containing hundreds to thousands of glucose units. Even quite short chain lengths of cellulose (>6 glucose units) are insoluble in water. Hydrogen bonds between adjacent chains cause the close packing of the molecules, resulting in the formation of microfibrils. When present as microfibrils, cellulose is resistant to degradation by strong base.

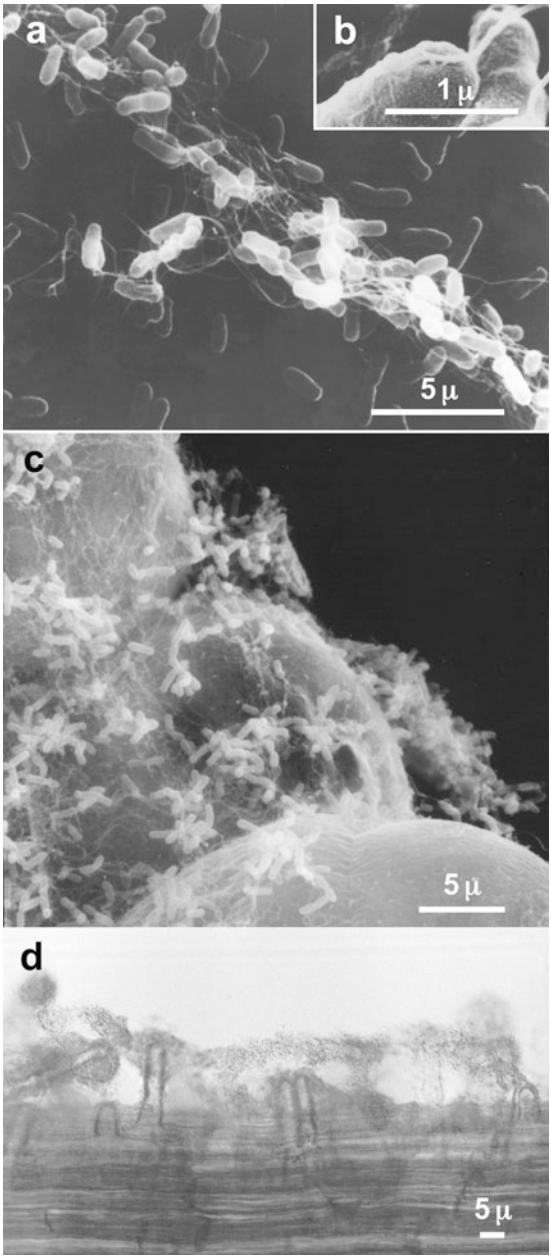


Fig. 3 TEM photomicrographs showing cellulose production by *A. tumefaciens*. **a** Cellulose production by a group of bacteria in liquid medium. **b** Enlargement of bacteria from liquid medium showing cellulose fibrils projecting from the surface. **c** *A. tumefaciens* bound to the surface of carrot suspension cells. The cellulose fibrils holding bacterial aggregates are visible. **d** A cellulose-overproducing mutant of *A. tumefaciens* (*celG* mutant) bound to the surface of a tomato root. Note the large clusters and streamers of bacteria held together by cellulose fibrils. Figures **a**, **b**, and **c** are reprinted from Matthysse et al. (1981). Figure **d** is reprinted from Matthysse et al. (2005)

Cellulose is synthesized from UDP-glucose at the plasma membrane by the addition of glucose to the nonreducing end of the chain and transported out of the cell (Slabaugh et al. 2014). Bacteria produce cellulose in two geometrical patterns. Some groups of bacteria such as gluconacetobacter (komagataeibacter) and *Pseudomonas fluorescens* produce cellulose from an ordered array of complexes oriented in a line along the long axis of the cell (Brown et al. 1976; Spiers et al. 2003). This cellulose forms a sheet as it is exported from one side of the bacterium. Its synthesis results in a layer which the bacteria use as a raft to float on the surface of a liquid and form a pellicle. Other groups of bacteria such as agrobacteria and rhizobia form thin cellulose fibrils which extend out from the sides of the bacterium in all directions (Fig. 3; Matthysse et al. 1981). These bacteria generally do not form pellicles.

Cellulose is synthesized from a UDP-glucose precursor by a plasma membrane protein CelA (called BcsA in some bacteria) (Matthysse et al. 1995; Wong et al. 1990). This protein is highly conserved in those bacteria which synthesize cellulose. It is nearly identical in most of the rhizobiacea (greater than 95% identical amino acid sequence). It is 40% identical with the cellulose synthase (BcsA) from *Rhodobacter spheroides,* whose crystal structure with bound substrates has been determined and a detailed mechanism of action of the synthase has been proposed (Morgan et al. 2016). The conserved motifs HAKAG, TED, and FFCSG which form the pocket in which the donor UDP-glucose is located are all conserved between the *R. spheroides* and *A. tumefaciens* protein sequences with the exception of the replacement of the T by S in the TED site. The transmembrane sequence QxxRW, its position relative to the catalytic D, and the terminal glucose interacting site DxD are also conserved. Thus, it seems likely that the agrobacterium CelA uses the same catalytic mechanism elegantly described for the cellulose synthase of *R. spheroides*.

Two other proteins are generally required for cellulose synthesis in bacteria: CelB is a membrane protein and forms a complex with CelA and CelC, which has homology with cellulases; Table 1 and Fig. 1c). CelB is closely associated with CelA in the cell membrane. In *E. coli* K12, the deletion of *bcsB* prevents cell growth (Baba et al. 2006). However, a strain carrying a deletion of both *bcsA* and *bcsB* is able to grow (Matthysse, unpublished observation). It seems likely that the requirement for BcsB in the presence of BcsA reflects the interaction of the proteins with each other and with the cell membrane. In some strains of gluconacetobacter, *bcsA* and *bcsB* are fused into a single gene and encode a single polypeptide product (Wong et al. 1990). The cellulose synthase complex (CelA and CelB) spans the inner (and possibly the outer) membrane. Transport of the growing cellulose chain is directly across the inner membrane through the channel in the protein which begins with the QxxRW motif in CelA (Morgan et al. 2016). In *R. spheroides, E. coli*, and *P. fluorescens* transport across the outer membrane is probably mediated by BcsC or AlgK (Keiski et al. 2010;Whitney et al. 2011). Agrobacteria lack a homologue to either of these proteins. It is possible that BcsC and AlgK are involved in the orientation of the cellulose molecules to form sheets of cellulose.

The agrobacteria which form strands of cellulose may use a different protein to enable cellulose to cross the outer membrane.

The role of CelC in cellulose synthesis is unclear. The protein has homology to cellulases. Its structure predicts a soluble protein (Matthysse et al. 1995). A homologue of CelC is required for cellulose synthesis in all bacteria in which its role has been examined (Romling 2002). A protein with homology to cellulases has also been implicated in cellulose synthesis in higher plants (Molhoj et al. 2002).

In *A. tumefaciens* two additional genes are required for cellulose synthesis, *celD* and *celE* (Matthysse et al. 1995). These genes are located in an operon immediately adjacent to the operon containing *celABCG*. Their role is unknown. BLAST homologies suggest that they are related to acetylases and to proteins which hydrolyze acetate groups from acetylated proteins. Although some bacterial celluloses appear to be acetylated [e.g., *P. fluorescens* (Spiers et al. 2003)], there is no evidence to suggest that *Agrobacterium* cellulose is acetylated. Possibly one or more of the proteins making up the cellulose synthase complex is regulated by acetylation or *celD* and *celE* may function in some other unidentified process required for cellulose synthesis or for its regulation.

4.2 Regulation

The major regulator of cellulose synthesis is cyclic-di-GMP (Amikam and Benziman 1989). CelA contains a pilZ sequence which is a cyclic-di-GMP binding site. In the absence of cyclic-di-GMP, a salt bridge blocks access of the substrate to the active site of the enzyme. In the presence of cyclic-di-GMP, the salt bridge is disrupted and the enzyme is active (Morgan et al. 2014). Cyclic-di-GMP is synthesized from GTP by diguanylate cyclases and is degraded by esterases. The synthetic enzymes contain the amino acid sequence GGDEF: the esterases contain either an EAL or a HD-GYP sequence (Ausmees et al. 2001). In some cases, the same protein contains both the cyclase and the esterase motifs. *A. tumefaciens* C58 contains 16 predicted proteins which contain GGDEF motifs and 13 predicted proteins which contain both GGDEF and EAL motifs. There is also one predicted protein which contains an EAL motif and one which contains a HD-GYP motif. Thus, the regulation of levels of cyclic-di-GMP is complex. Current evidence suggests that the activity of CelA is controlled not by the general level of cyclic-di-GMP in the cell but instead by specific cyclases and esterases (Table 2 and Fig. 4). The diguanylate cyclase DcgA, encoded by *Atu1257*, appears to have the largest role in controlling the amount of cellulose made (Xu et al. 2013). However, *dcgA* mutants still make a detectable amount of cellulose. Mutations in three genes, *dcgA, dcgB*, and *dcgC*, are required to abolish cellulose synthesis completely. Mutations in the diguanylate cyclase gene *Atu1297* (aka *celR* or *pleD*) also result in decreased cellulose production (Barnhart et al. 2013). There appears to be some cross talk between systems involving cyclic-di-GMP, as overexpression of other diguanylate cyclase genes such as *Atu1060* or *Atu1297* results in increased cellulose production. Deletion of *Atu1060*

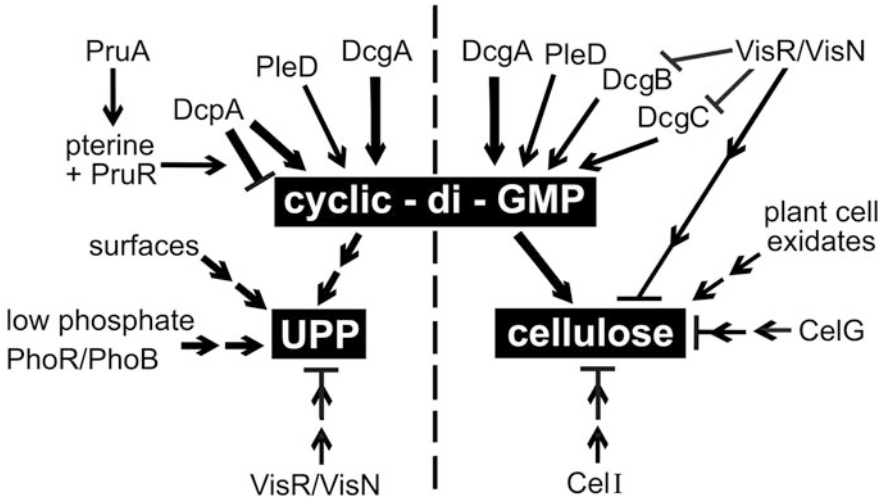


Fig. 4 The regulation of the synthesis of cellulose and the UPP and the central role of cyclic-di-GMP. See Sects. 4.2 and 6.2 for an explanation of the genes and interactions involved

does not affect cellulose production, suggesting that it does not generally contribute to the pool of cyclic-di-GMP involved in the regulation of cellulose synthesis or that its activity can be replaced by the activity of other diguanylate cyclases (Barnhart et al. 2013). Bacteria with mutations in ornithine decarboxylase (*odc*, *Atu3196*) produce increased amounts of cyclic-di-GMP and cellulose. Expression of a cyclic-di-GMP phosphodiesterase in the mutant bacteria reversed this effect (Wang et al. 2016). A pterin-responsive system involving DcpA regulates both cellulose and UPP synthesis and is described in Sect. 6.2.

Two transcriptional regulators, VisR and VisN, which are required for motility, negatively regulate cellulose and UPP synthesis in *A. tumefaciens* (Xu et al. 2013). Deletion of these genes results in an increase in cellulose synthesis. The regulation of motility by these genes is mediated by a transcriptional regulator Rem. However, the regulation of cellulose synthesis does not require *rem* and VisR/VisN have no effect on the transcription of *dcgA,* although they do regulate the transcription of *dcgB* and *dcgC*. The mechanism by which VisR/VisN regulate cellulose production remains to be elucidated (Xu et al. 2013).

Cellulose synthesis also appears to be negatively regulated by two genes, *celG* and *celI*, about which there is little information. *CelG* is the last gene in the *celABCG* operon and is not indicated in some annotations of the C58 genome. Mutations in *celG* result in overproduction of cellulose (Matthysse et al. 2005). Normal levels of cellulose production can be restored in the *celG* mutant by the presence of an intact *celG* gene cloned behind the *lac* promoter on a plasmid. CelI, a putative DNA-binding protein, has homology to the arsenical resistance operon repressor subfamily and contains putative Zn^{+2} binding sites (Matthysse et al. 2005). These proteins typically dissociate from DNA in the presence of metal ions.

Nothing is known about whether CelI regulates transcription of the cellulose synthase operon directly or, as is the case for the other better characterized regulators discussed above, it exerts its effects through changes in the level of cyclic-di-GMP.

Unlike curdlan, there is little evidence for regulation of cellulose synthesis at the level of transcription of the genes required.

Cellulose synthesis may be regulated by signals from the plant. Plant extracts such as soytone or the presence of plant tissue culture cells increase the amount of cellulose made. The nature of the substance to which the bacteria respond is not known. It is low molecular weight and heat stable (Matthysse et al. 1981; Matthysse 1994). Acetosyringone, the inducer of the *vir* genes in *A. tumefaciens*, has no effect on cellulose synthesis (Matthysse, unpublished observation). The mechanism by which a plant-derived small molecule(s) regulates cellulose synthesis is unknown.

4.3 Function

Cellulose production by agrobacteria results in the formation of loose clusters of bacteria held together by cellulose fibrils (Fig. 3). These clusters may form on the surface of plant cells or on inanimate objects or in liquid suspensions. In general, cellulose does not mediate the initial attachment to a surface but after attachment, the elaboration of cellulose fibrils increases the size of the bacterial aggregates on the surface and the strength of the binding to surfaces which contain cellulose, such as plant cell walls or filter paper. Cellulose-minus mutants of *A. tumefaciens* C58 showed reduced colonization of tomato and *Arabidopsis thaliana* roots (Matthysse and McMahan 1998). The bacteria were still able to colonize the roots, but the number of tightly and irreversibly bound bacteria after 10 days incubation was reduced by 10^3 for tomato and 10^1 for *A. thaliana*. Cellulose also participates in bacterial binding to leaf wound sites. The binding of wild-type and cellulose-minus bacteria to wound sites on the surface of *Bryophyllum* leaves was compared. The cellulose-minus bacteria could be removed by water washing, whereas the wild-type bacteria remained bound to the leaves (Sykes and Matthysse 1986). In liquid cultures, the addition of plant extracts causes increased cellulose production which results in the formation of bacterial flocs (Matthysse et al. 1981).

5 Curdlan

5.1 Structure and Biosynthesis

Like cellulose, curdlan is a linear polymer of β-linked D-glucose molecules (Fig. 1d). Curdlan contains β(1→3)-linked D-glucose rather than the 1→4 linkages found in cellulose. This difference changes the interaction of adjacent chains so that

curdlan chains do not form microfibrils as cellulose does. As with cellulose, curdlan polymers are generally quite long, containing hundreds to thousands of glucose units. Curdlan is insoluble in water but does dissolve in basic solutions. Curdlan forms gels in water and is used commercially as a gelling agent in food and as an additive to produce super-workable concrete. In gels, curdlan often exists as a triple helix of adjacent chains (McIntosh et al. 2005).

Most of the research on curdlan production by agrobacteria has been carried out using the ATCC strain 31749. This is an industrial strain which produces large amounts of curdlan. When the bacteria are grown on 4% glucose with limiting nitrogen, more than 85% of the glucose in the medium is converted into curdlan (McIntosh et al. 2005). Three genes are required for curdlan biosynthesis: *crdA, crdS*, and *crdC* (Table 1; Stasinopoulos et al. 1999). These genes are found in an operon which is similar to the cellulose synthesis operon. Curdlan is synthesized from a UDP-glucose precursor at the cell membrane. CrdS, the curdlan synthase, is similar in structure to CelA, the cellulose synthase. Both proteins are located in the cell membrane (Karnezis et al. 2003). Both CelA and CrdS have a conserved QxxRW transmembrane channel and a possible catalytic D located at the same distance from the channel. The donor-binding site HAKAG and SED, and the DxD which interacts with the terminal glucose in cellulose are also conserved in CrdS. However, the FFCSA site which is part of the membrane channel in CelA is replaced by AFCVGTS in CrdS. These conserved motifs suggest that the catalytic mechanism of the two synthases may be similar. In addition to CrdS, CrdA is also required for curdlan synthesis (Stasinopoulos et al. 1999). The protein is associated with the cell membrane. It has no conserved motifs and its function is unknown. CrdC is not absolutely required for curdlan synthesis. Small amounts of the polysaccharide can be made in its absence but they are not secreted into the medium. The synthesis of large amounts of curdlan requires CrdC (McIntosh et al. 2005). CrdC has a signal sequence and is predicted to be on the outer side of the cell membrane, probably in the periplasmic space. It is possible that CrdC may function in the export of curdlan across the outer membrane. The *crdASC* operon is conserved in *A. tumefaciens* C58, which is thus presumed to have a mechanism for curdlan synthesis similar to that of ATCC 3179.

5.2 Regulation

Studies of the regulation of curdlan synthesis have been carried out in the strain ATCC 3179 (Table 2 and Fig. 2). This strain has an unknown history of mutagenesis and selection for increased curdlan production. Thus, it is unclear how many of the conclusions regarding regulation of curdlan synthesis in this strain apply to wild-type strains such as C58 which have not been selected for overproduction of curdlan. Curdlan and succinoglycan have been suggested to be regulated coordinately in opposite directions. Thus, ATCC 3179 makes curdlan but not

succinoglycan and C58 makes succinoglycan and very little curdlan (McIntosh et al. 2005; Stanisich and Stone 2009).

Unlike cellulose, a major control of curdlan synthesis appears to be the expression of the *crdASC* operon. When cells in log phase growth are compared with stationary phase cells which are starved for nitrogen, the expression of the *crd* operon is increased by approximately 100-fold (Ruffing and Chen 2012). The CrdR transcriptional regulator is required for the synthesis of curdlan, but there are no data as to whether it acts directly on the *crd* operon (Stasinopoulos et al. 1999).

Nitrogen starvation is also a major regulator of curdlan synthesis (McIntosh et al. 2005). NtrB (sensor kinase) and NtrC (response regulator) constitute a two-component system which responds to nitrogen starvation. NtrC mutants show decreased synthesis of curdlan (Stanisich and Stone 2009; Yu et al. 2011a, b). In many systems, the response regulator NtrC, once phosphorylated, acts on transcription by binding to RpoN. However, *rpoN* is not required for curdlan synthesis in ATCC 3179. Thus, the action of NtrC must involve some other unidentified protein (Ruffing and Chen 2012; Yu et al. 2011a). There is another gene in the *ntrBC* operon at the start of the operon, *nifR*, whose function is unknown. However, deletion of this gene resulted in lower but still significant levels of curdlan production (30% of wild-type levels). The reduction may reflect a role for *nifR* in regulating curdlan synthesis or simply be an effect of alteration of the operon structure upstream of the required genes *ntrBC* (Ruffing and Chen 2012; Yu et al. 2011a). There is a second two-component system in *Agrobacterium*, *ntrYX*, which also responds to nitrogen starvation. NtrY is a membrane protein which acts as a sensor for nitrogen. Deletion mutants of *ntrY* cannot grow on nitrate as a nitrogen source and do not make curdlan. The response regulator *ntrX* is an essential gene as mutations in it are lethal (Ruffing and Chen 2012; Stanisich and Stone 2009). NtrX is a predicted transcriptional regulator, but there is no information on how it regulates curdlan synthesis. Both NtrX and NtrC may regulate transcription of the *crd* operon or they may act on other genes which then regulate its transcription. One gene whose transcription is activated by phosphorylated NtrC in *E. coli* is *relA*, the gene involved in the stringent response. Mutation of *relA* in ATCC 3179 eliminates curdlan production and decreases the transcription of the *crd* operon more than 50-fold (Ruffing and Chen 2012). It thus seems probable that NtrC (and possibly NtrX) may control curdlan production by their action on the transcription of *relA*. RelA, which is required for the stringent response, produces (p)ppGpp. PppGpp is involved in regulating many processes during stress or stationary phase growth. There is no one binding site for (p)ppGpp, so computer predictions of which proteins might be the intermediaries between (p)ppGpp and curdlan synthesis are not easy (Srivatsan and Wang 2008).

When nitrogen is limiting in liquid growth medium, high phosphate concentrations inhibit curdlan production (Kim et al. 2000). The mechanism of this effect is unknown. During stationary phase, agrobacteria accumulate polyphosphate. During logarithmic growth polyphosphate is hydrolyzed by Ppx1, an exopolyphosphatase (Ruffing and Chen 2012). The activity of this enzyme in *E. coli* is inhibited by (p)ppGpp, the product of the RelA protein (Kuroda et al. 1997).

During stationary phase polyphosphates accumulate. Mutation of *ppx1* causes an early accumulation of polyphosphates and an early onset of curdlan synthesis during logarithmic growth, suggesting that the presence of polyphosphates influences curdlan synthesis by some unknown mechanism (Ruffing and Chen 2012).

Although cyclic-di-GMP plays a major role in the regulation of cellulose and UPP synthesis, there is only one report of an effect of cyclic-di-GMP on curdlan synthesis. Unlike cellulose synthase, curdlan synthase does not contain a PilZ domain and does not bind cyclic-di-GMP. The mutation of one gene encoding a protein containing a GGDEF motif (*Atu1114*) resulted in decreased curdlan synthesis (Wu et al. 2016). This gene has not been further characterized and so the mechanism of this effect is not known.

The regulation of curdlan synthesis in the industrial strain of *Agrobacterium* ATCC 3179 appears to involve RelA and (p)ppGpp which are responsible, at least in part, for the restriction of curdlan synthesis to stationary phase cells and the stimulation of curdlan synthesis by nitrogen starvation. Whether these same regulatory mechanisms are operational in *A. tumefaciens* C58 remains to be determined.

5.3 Function

Curdlan is deposited extracellularly and surrounds the bacterial cells, forming a capsule during growth in liquid or on solid surfaces. In liquid, curdlan production is accompanied by the formation of large bacterial flocs in the medium (Fig. 5). When *A. tumefaciens* ATCC 3179 is grown for a prolonged time (more than 1 week) on solid agar containing high levels of glucose, curdlan and enmeshed cells form a layer which covers the bacteria on the surface of the medium. This layer is sufficiently self-adhesive that it can be stripped off the medium leaving the surface

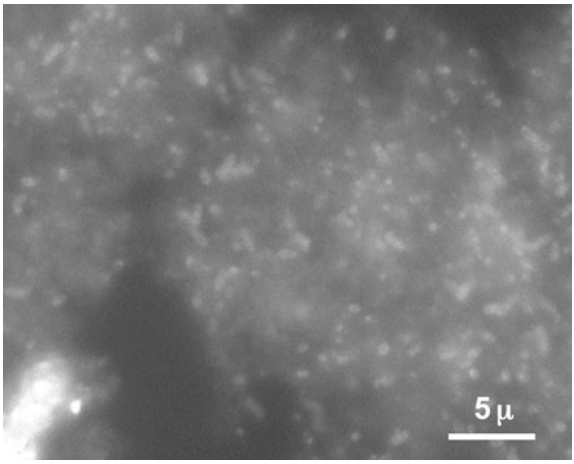


Fig. 5 Fluorescence photomicrograph of a floc of *Agrobacterium* strain LTU50 (a chlorampenicol-resistant derivative of ATCC 3179) grown in liquid culture with limiting nitrogen and stained with aniline blue fluorochrome to show the presence of curdlan

bacteria behind (McIntosh et al. 2005). The curdlan layer protects the bacteria from phagocytosis by protozoa. It may also protect against heat (55 °C) and desiccation (Stanisich and Stone 2009). *A. tumefaciens* C58 has not been observed to produce significant amounts of curdlan, and so the function of this exopolysaccharide for this bacterial strain is unknown. There is no observable phenotype of a *crdS* mutation in *A. tumefaciens* C58 (Matthysse, unpublished observation). However, the regulation of curdlan production by RelA suggests that it may play a role in protection against predation and various environmental stresses such as starvation, high and low temperatures, and desiccation.

6 The Unipolar Polysaccharide (UPP)

6.1 Structure and Biosynthesis

Agrobacterium tumefaciens produces a polysaccharide at the pole of the bacterial cell opposite the circumpolar flagellae referred to as the unipolar polysaccharide (UPP) (Xu et al. 2012, 2013). The structure of this polysaccharide is unknown but its reaction with wheat germ agglutinin and *Dolichos bifloris* lectins suggests that it contains both N-acetlyglucosamine and N-acetylgalactosamine. A similar unipolar polysaccharide made by *Rhizobium leguminosarum* is composed of 55% mannose and 40% glucose with small amounts of rhamnose and galactose (Laus et al. 2006). The genes required for the biosynthesis of the UPP (*uppABCDEF*) are located in two or three operons (Table 1). These genes have homologues in *R. leguminosarum*. They include an *E. coli* WzyC homologue which is predicted to attach a sugar to a lipid carrier, possibly undecaprenol (*Atu1235*), two glycosyltransferases (*Atu1236* and *Atu1237*), a homologue of GumB which may be involved in polysaccharide transport across the membrane (*Atu1238*), a homologue of ExoP also involved in transport and possibly in chain-length determination (*Atu1239*), and an acetyltransferase (*Atu1240*). Other genes may be required for the synthesis as not all steps in the synthesis and export of the UPP can easily be accounted for with the genes in these operons. In particular, a flippase is missing. Mutations in any of these genes, including *uppE*, block the synthesis of the UPP during normal growth. Under conditions in which phosphate is limiting, Atu0102 (presumably usually involved in the synthesis of a different polysaccharide) can substitute for UppE (Xu et al. 2012).

6.2 Regulation

The UPP is only made when cells come into contact with a surface. Planktonic cells rarely make detectable amounts of the UPP as judged by staining of cells with

fluorescent wheat germ agglutinin. However, when the bacteria come into contact with a surface they begin production of the UPP promptly. The production of the UPP is accompanied by a loss of motility and attachment to the surface via the UPP (Danhorn and Fuqua 2007; Li et al. 2012).

As with cellulose synthesis, the major regulator of UPP synthesis is cyclic-di-GMP (Table 2 and Fig. 4; Xu et al. 2013). However, unlike the regulation of cellulose synthesis where cyclic-di-GMP binds directly to cellulose synthase (CelA) and activates the enzyme, none of the known enzymes required for the synthesis of the UPP has an obvious cyclic-di-GMP binding site. Thus, the mechanism by which cyclic-di-GMP controls UPP synthesis is likely to be more complex than that of the control of cellulose synthesis. For both cellulose and UPP, the diguanylate cyclase/phosphodiesterase DcgA appears to be an important regulator of cyclic-di-GMP levels (Xu et al. 2013).

A second enzyme DcpA, which can act as either a cyclase to synthesize cyclic-di-GMP or as a phosphodiesterase to hydrolyze cyclic-di-GMP, plays a key role in the control of synthesis of the UPP. Pterdine reductase, PruA, regulates the activity of DcpA. In its presence DcpA acts as an esterase; in the absence of PruA, DcgA acts as a cyclase. A gene located immediately upstream of *dcpA*, *pruR* which encodes a putative pterin-binding protein, is required for the esterase activity of DcpA. Thus, it seems likely that PruA is required for the synthesis of the pterine which interacts with PruR (or a regulator of PruR) to activate the esterase activity of DcpA, reduce levels of cyclic-di-GMP, and reduce the synthesis of the UPP and of cellulose (Feirer et al. 2015). A CheY-like protein, ClaR, negatively regulates the synthesis of the UPP and cellulose, in part through the pterin pathway. However, it also has an effect in mutants for the DcpA-pterin-dependent pathway, suggesting that regulation by this protein involves more than one pathway (Feirer et al. 2017).

The transcriptional regulators VisR and VisN, which are required for motility, negatively regulate cellulose, and UPP synthesis in *A. tumefaciens*. Deletion of these genes results in an increase in UPP synthesis. The regulation of motility by these genes is mediated by the transcriptional regulator Rem. The regulation of UPP and cellulose synthesis by VisR/VisN does not require the transcriptional regulator Rem, although Rem is required for the regulation of motility (Xu et al. 2013). The mechanism by which VisR/VisN regulate UPP and cellulose production remains to be elucidated.

PleD (aka CelR) also regulates both UPP and cellulose synthesis. It is a diguanylate cyclase and has a cheY domain, which means that it could be regulated by phosphorylation. Mutations in *pleD* result in decreased biofilm formation, and decreased UPP and cellulose (Barnhart et al. 2013; Xu et al. 2013). These mutations also decrease virulence for unknown reasons. Mutations in a cyclic-di-GMP phosphodiesterase (*dcpA*, *Atu3495*) increase biofilm formation and presumably increase the levels of the UPP and cellulose due to increased cyclic-di-GMP (Feirer et al. 2015; Xu et al. 2013).

The synthesis of the UPP is also regulated by environmental conditions, particularly phosphate levels. Low external phosphate is sensed by the two-component system PhoR and PhoB. PhoR is a sensor kinase which responds to low phosphate and phosphorylates, the response regulator PhoB. PhoB controls the synthesis of the UPP so that it is made only under conditions of low phosphate (Danhorn et al. 2004; Xu et al. 2012).

6.3 Function

The UPP and cellulose function coordinately in the formation of biofilms on various animate and inanimate surfaces. The UPP is only made after cells come into contact with a surface (Li et al. 2012; Matthysse et al. 1981). This is in contrast to cellulose, which can be made by stationary-phase planktonic cells if a plant inducer is present. The UPP functions to attach the bacteria by one pole to surfaces as varied as plant roots, soil particles, glass, or plastic surfaces such as cover slips, and nylon thread (Fig. 6; Matthysse 2014). Thus, the nature of the surface is not a major factor in UPP-mediated binding. No active signaling between the surface and the bacteria is needed. Once the bacteria are bound to a surface via the UPP, then cellulose is made. Cellulose increases the size of the bacterial aggregates formed and for some surfaces increases the strength of the binding (Matthysse 1983; Matthysse et al. 2005).

In *A. tumefaciens* as in many other bacteria, cyclic-di-GMP appears to be a key regulator controlling the shift from planktonic to biofilm growth. High levels of cyclic-di-GMP are associated with biofilm formation. The UPP appears to be essential for initial binding and biofilm formation on most surfaces. It mediates the initial attachment of the bacteria under conditions of low phosphate. These attached bacteria can be closely packed on the surface, and thus readily form a biofilm. Other mechanisms of attachment result in sparsely bound bacteria and thus are less likely to result in biofilm formation.

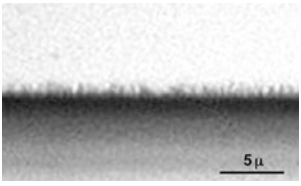


Fig. 6 Photomicrograph showing *A. tumefaciens* C58 bound to a nylon thread by the UPP. Note the end-on attachment of the bacteria

7 Summary and Conclusions

7.1 Structure and Biosynthesis

Three of the exopolysaccharides discussed in this chapter are homopolymers: cyclic-β-(1, 2)-glucan [glucose-β-(1→2) glucose], cellulose [glucose-β-(1→4)-glucose], and curdlan [glucose-β-(1→3)-glucose]. Cyclic-β-(1, 2)-glucans are low molecular weight polymers which contain approximately 17–25 sugar moieties (Breedveld and Miller 1994). They are water soluble. The precursor for all three of these glucose polymers is UDP-glucose. The synthesis of the β-(1, 2)-glucans is carried out by the membrane protein ChvB. Cyclization of the β-(1, 2)-glucan is apparently also carried out by the same protein (Zorreguieta et al. 1985). ChvA is believed to be involved in the transport of the polymer across the membrane (Cangelosi et al. 1989). Cellulose and curdlan are very-long-chain polymers generally containing more than 1000 sugar residues. They are not water soluble. The synthesis and export of these homopolymeric polysaccharides in both cases is carried out by a single protein (CelA or CrdS) located in the cytoplasmic membrane. An accessory membrane protein is generally required (CelB or CrdA) (Matthysse et al. 1995; Stasinopoulos et al. 1999). There is an additional gene in *A. tumefaciens* (*rcdA*, *Atu5090*), which has homology to CelA and CrdS. This gene is wide-spread in the rhizobiaceae and presumably is involved in the synthesis and export of an unidentified homopolymer. The conserved sites in CelA and CrdS are only partially conserved in this protein. The motifs which form the UDP-glucose binding pocket in CelA and CrdS (HAKAG, TED, and FFCSG) are replaced by GSKAG, TED, and FCCGT in RcdA. The QxxRW transmembrane sequence and the DxD sequence, which interacts with the terminal glucose of the growing chain are conserved. The nature of the exopolysaccharide made by RcdA is unknown.

The other two exopolysaccharides discussed in this chapter, succinoglycan and the UPP, are heteropolymers composed of repeating units. In the case of succinoglycan, the repeat contains eight sugar residues (Evans et al. 2000). The composition of the subunit of the UPP is not known except that it includes *N*-acetylglucosamine and *N*-acetylgalactosamine (Heindl et al. 2014). In the case of succinoglycan, the repeat unit is synthesized from UDP-glucose and UDP-galactose by membrane-associated proteins. The initial sugar, galactose, is linked to a lipid, presumably bactoprenol. The finished subunit is flipped to the outside of the membrane, polymerized into a polysaccharide, and exported outside the cell (Reuber and Walker 1993). The synthesis of a heteropolymer requires more genes for synthesis and export of the polymer that are required for the homopolymers, which generally seem to require only two or three genes for their synthesis.

7.2 Regulation

Cyclic-β-(1, 2)-glucans appear to protect the cells from low ionic strength and are made in response to growth in dilute media. They are not needed in high ionic strength and are not made under these conditions. This regulation appears to be a result of a direct effect of salt concentration on the activity of the synthetic enzyme ChvB (Ingram-Smith and Miller 1998). There is little other information on the regulation of the synthesis of this exopolysaccharide.

Succinoglycan appears to be regulated in an opposite fashion to curdlan, cellulose, and the UPP. The major well-characterized regulator of succinoglycan production is the ExoR/ChvG/ChvI system, which is a two-component system which responds to acidic pH (Heckel et al. 2014; Tomlinson et al. 2010; Wu et al. 2012). Under acidic conditions, the response regulator ChvI activates transcription of *vir* genes and succinoglycan biosynthesis. ChvI represses motility, biofilm formation, and bacterial attachment. Presumably, the later phenotypes are a result of an effect of ChvI on UPP and cellulose synthesis either directly or via an effect on cyclic-di-GMP, which is required for both the synthesis of both exopolysaccharides. ChvI may also repress *ntrX*. Because NtrX is a positive regulator of curdlan biosynthesis, this would explain the observation that in general agrobacteria do not make curdlan and succinoglycan at the same time (Ruffing and Chen 2012).

Curdlan appears to be regulated by nitrogen starvation and other stresses which activate the stringent response. RelA, the mediator of the stringent response, is required for curdlan synthesis (Ruffing and Chen 2012). The accumulation of polyphosphates, which normally is associated with the stationary phase, also stimulates the formation of curdlan. The enzyme responsible for the breakdown of polyphosphates, Ppx1, is inhibited by (p)ppGpp, the product of the RelA protein (Ruffing et al. 2011). How polyphosphates stimulate curdlan synthesis is unknown.

Cellulose and the UPP appear to be coordinately regulated under many conditions. Cyclic-di-GMP is the major positive regulator identified for the synthesis of these exopolysaccharides. This compound regulates cellulose synthesis directly as it binds to the cellulose synthase enzyme, CelA, and causes a change in configuration which exposes the active site and allows cellulose synthesis to proceed (Morgan et al. 2014). The mechanism by which cyclic-di-GMP regulates the synthesis of the UPP is not known. Several other regulators of cellulose and/or UPP synthesis may act through their effect on levels of cyclic-di-GMP. These include the VisR/VisN transcriptional regulators, PleD (CelR), and PruA/PruR (Barnhart et al. 2013; Xu et al. 2013). Some conditions affect the synthesis of only one of these exopolysaccharides. The UPP is made only after cells come into contact with a surface (Li et al. 2012). Cellulose can be made by planktonic cells (Matthysse 1983). Low phosphate also increases the synthesis of the UPP but has no effect on cellulose synthesis (Xu et al. 2012). Cellulose synthesis is increased by a low molecular weight compound(s) released by the plant (Matthysse 1994). UPP synthesis is not known to be affected by plant extracts. Thus, cellulose and the UPP are generally, but not always, made in response to the same environmental conditions.

7.3 Function

With the exception of cyclic-β-(1, 2)-glucans, none of these exopolysaccharides is required for virulence on those hosts which have been examined (tomato, Bryophyllum, tobacco, and *A. thaliana*). The requirement for the cyclic-β-(1, 2)-glucans appears to be due to the general effects of their absence on the structure of the cell surface and not to a particular role in the pathogenesis of *A. tumefaciens*. Mutants which cannot make cyclic-β-(1, 2)-glucans are virulent under specific conditions in which cyclic-β-(1, 2)-glucans are no longer required for normal bacterial growth (Bash and Matthysse 2002; Hawes and Pueppke 1987; Swart et al. 1994).

The major role of these exopolysaccharides is in the interaction of the bacteria with their environment. Some of them are known to protect cells from various environmental stresses and dangers. Cyclic-β-(1, 2)-glucans protect the cells against low ionic strength (Miller et al. 1986). Curdlan protects the cells against heat, desiccation, and phagocytosis by protozoa (McIntosh et al. 2005). Others aid in the formation of biofilms on surfaces. These biofilms may be protective and may also play a role in keeping the bacteria in a desirable location. The UPP and curdlan can initiate biofilms on surfaces under different conditions. The UPP is made when the environment is low in phosphate and the cell comes into contact with a surface (Danhorn et al. 2004; Li et al. 2012). Curdlan is made when the cells are starved for nitrogen or exposed to other stresses, which activate the stringent response (Ruffing and Chen 2012; Yu et al. 2011b). UPP-based biofilms are easily permeated by various solutes and can be removed from surfaces by vigorous vortexing or sonication. Curdlan biofilms are dense and not readily penetrated by solutes. The cells cannot be removed easily, and if attempts to remove them are made, they stick together in flocs. Curdlan can be produced in liquid as well as on solid surfaces, resulting in the formation of planktonic flocs. The UPP is the only one of these exopolysaccharides, whose production is known to be stimulated by the presence of a surface. Production of the UPP results in bacterial binding to that surface. Cellulose is often produced as a second stage in the formation of a biofilm which was initiated by the UPP (Fig. 7). Cellulose causes the formation of aggregates of cells loosely attached to each other (Matthysse 1983). In addition, cellulose can directly mediate attachment to surfaces containing cellulose such as roots or filter paper. This is a firm attachment and the bacteria cannot be removed from the surface by vortexing, although sonication using a bath sonicator will remove them. Cellulose is the only one of these exopolysaccharides whose production is known to be increased in the presence of the plant host.

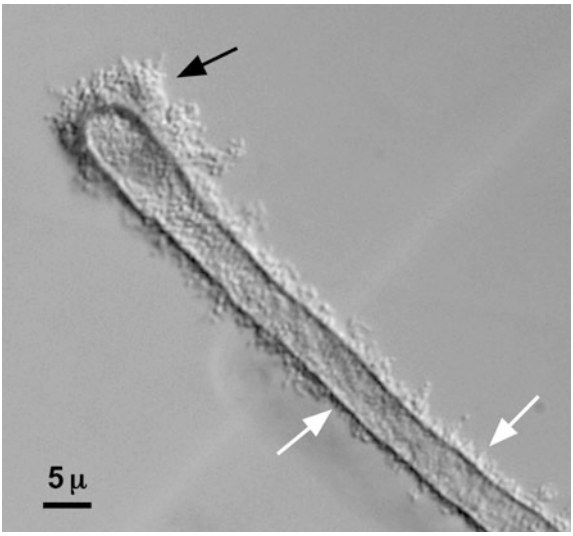


Fig. 7 Photomicrograph showing *A. tumefaciens* C58 bound to a tomato root hair. White arrows show initially attached bacteria bound end-on by the UPP. The black arrow shows the cluster of bacteria formed after the initially bound bacteria elaborate cellulose fibrils

References

Abe K, Nakajima M, Yamashita T et al (2017) Biochemical and structural analyses of a bacterial endo-beta-1,2-glucanase reveal a new glycoside hydrolase family. J Biol Chem 292:7487–7506

Amikam D, Benziman M (1989) Cyclic diguanylic acid and cellulose synthesis in *Agrobacterium tumefaciens*. J Bacteriol 171:6649–6655

Ausmees N, Mayer R, Weinhouse H et al (2001) Genetic data indicate that proteins containing the GGDEF domain possess diguanylate cyclase activity. FEMS Microbiol Lett 204:163–167

Baba T, Ara T, Hasegawa M et al (2006) Construction of *Escherichia coli* K–12 in-frame, single-gene knockout mutants: the Keio collection. Mol Syst Biol 2:2006

Barnhart DM, Su S, Baccaro BE et al (2013) CelR, an ortholog of the diguanylate cyclase PleD of Caulobacter, regulates cellulose synthesis in *Agrobacterium tumefaciens*. Appl Environ Microbiol 79:7188–7202

Bash R, Matthysse AG (2002) Attachment to roots and virulence of a *chvB* mutant of *Agrobacterium tumefaciens* are temperature sensitive. Mol Plant Microbe Interact 15:160–163

Breedveld MW, Benesi AJ, Marco ML et al (1995) Effect of phosphate limitation on synthesis of periplasmic cyclic (beta)-(1,2)-glucans. Appl Environ Microbiol 61:1045–1053

Breedveld MW, Miller KJ (1994) Cyclic beta-glucans of members of the family Rhizobiaceae. Microbiol Rev 58:145–161

Brightwell G, Hussain H, Tiburtius A et al (1995) Pleiotropic effects of regulatory *ros* mutants of *Agrobacterium radiobacter* and their interaction with Fe and glucose. Mol Plant-Microbe Interact 8:747–754

Brown RM Jr, Willison JHM, Richardson CL (1976) Cellulose biosynthesis in *Acetobacter xylinum*: visualization of the site of synthesis and direct measurement of the in vivo process. Proc Natl Acad Sci USA 73:4565–4569

Cangelosi GA, Hung L, Puvanesarajah V et al (1987) Common loci for *Agrobacterium tumefaciens* and *Rhizobium meliloti* exopolysaccharide synthesis and their roles in plant interactions. J Bacteriol 169:2086–2091

Cangelosi GA, Martinetti G, Leigh JA et al (1989) Role for *Agrobacterium tumefaciens* ChvA protein in export of beta-1,2-glucan. J Bacteriol 171:1609–1615

Castro OA, Zorreguieta A, Ielmini V et al (1996) Cyclic beta-(1,2)-glucan synthesis in Rhizobiaceae: roles of the 319-kilodalton protein intermediate. J Bacteriol 178:6043–6048

Cheng HP, Walker GC (1998) Succinoglycan production by *Rhizobium meliloti* is regulated through the ExoS-ChvI two-component regulatory system. J Bacteriol 180:20–26

Chou AY, Archdeacon J, Kado CI (1998) *Agrobacterium* transcriptional regulator Ros is a prokaryotic zinc finger protein that regulates the plant oncogene *ipt*. Proc Natl Acad Sci USA 95:5293–5298

Chouly C, Colquhoun IJ, Jodelet A et al (1995) NMR studies of succinoglycan repeating-unit octasaccharides from *Rhizobium meliloti* and *Agrobacterium radiobacte*r. Int J Biol Macromol 17:357–363

Danhorn T, Fuqua C (2007) Biofilm formation by plant-associated bacteria. Annu Rev Microbiol 61:401–422

Danhorn T, Hentzer M, Givskov M et al (2004) Phosphorus limitation enhances biofilm formation of the plant pathogen *Agrobacterium tumefaciens* through the PhoR-PhoB regulatory system. J Bacteriol 186:4492–4501

Douglas CJ, Halperin W, Nester EW (1982) *Agrobacterium tumefaciens* mutants affected in attachment to plant cells. J Bacteriol 152:1265–1275

Evans LR, Linker A, Impallomeni G (2000) Structure of succinoglycan from an infectious strain of *Agrobacterium radiobacter*. Int J Biol Macromol 27:319–326

Feirer N, Kim D, Xu J, Fernandez et al (2017) The *Agrobacterium tumefaciens* CheY-like protein ClaR regulates biofilm formation. Microbiol 163:1680–1691

Feirer N, Xu J, Allen KD et al (2015) A pterin-dependent signaling pathway regulates a dual-function diguanylate cyclase-phosphodiesterase controlling surface attachment in *Agrobacterium tumefaciens*. MBio 6:e00156

Glucksmann MA, Reuber TL, Walker GC (1993) Genes needed for the modification, polymerization, export, and processing of succinoglycan by *Rhizobium meliloti*: a model for succinoglycan biosynthesis. J Bacteriol 175:7045–7055

Gonzalez JE, York GM, Walker GC (1996) *Rhizobium meliloti* exopolysaccharides: synthesis and symbiotic function. Gene 179:141–146

Halder U, Banerjee A, Bandopadhyay R (2017) Structural and functional properties, biosynthesis, and patenting trends of bacterial succinoglycan: a review. Indian J Microbiol 57:278–284

Hawes MC, Pueppke SG (1987) Correlation between binding of *Agrobacterium-tumefaciens* by root cap cells and susceptibility of plants to crown gall. Plant Cell Rep 6:287–290

Hawes MC, Pueppke SG (1989) Variation in binding and virulence of *Agrobacterium tumefaciens* chromosomal virulence (Chv) mutant bacteria on different plant-species. Plant Physiol 91:113–118

Heckel BC, Tomlinson AD, Morton ER et al (2014) *Agrobacterium tumefaciens exoR* controls acid response genes and impacts exopolysaccharide synthesis, horizontal gene transfer, and virulence gene expression. J Bacteriol 196:3221–3233

Heindl JE, Wang Y, Heckel BC et al (2014) Mechanisms and regulation of surface interactions and biofilm formation in *Agrobacterium*. Front Plant Sci 5:176

Hou CT, Ahlgren JA, Brown W et al (1996) Production of an extracellular polysaccharide by *Agrobacterium* sp DS3 NRRL B-14297 isolated from soil. J Ind Microbiol 16:129–133

Hussain H, Johnston AW (1997) Iron-dependent transcription of the regulatory gene *ros* of *Agrobacterium radiobacter*. Mol Plant Microbe Interact 10:1087–1093

Ingram-Smith C, Miller KJ (1998) Effects of ionic and osmotic strength on the glucosyltransferase of *Rhizobium meliloti* responsible for cyclic beta-(1,2)-glucan biosynthesis. Appl Environ Microbiol 64:1290–1297

Kamoun S, Cooley MB, Rogowsky PM et al (1989) Two chromosomal loci involved in production of exopolysaccharide in *Agrobacterium tumefaciens*. J Bacteriol 171:1755–1759

Karnezis T, Epa VC, Stone BA et al (2003) Topological characterization of an inner membrane (1→3)-{beta}-D-glucan (curdlan) synthase from *Agrobacterium* sp. strain ATCC31749. Glycobiology 13:693–706

Keiski CL, Harwich M, Jain S, Neculai AM et al (2010) AlgK is a TPR-containing protein and the periplasmic component of a novel exopolysaccharide secretin. Structure 18:265–273

Kim MK, Lee IY, Kim KT et al (2000) Residual phosphate concentration under nitrogen-limiting conditions regulates curdlan production in *Agrobacterium* species. J Ind Microbiol Biotechnol 25:180–183

Kuroda A, Murphy H, Cashel M et al (1997) Guanosine tetra- and pentaphosphate promote accumulation of inorganic polyphosphate in *Escherichia coli*. J BiolChem 27:21240–21243

Laus MC, Logman TJ, Lamers GE et al (2006) A novel polar surface polysaccharide from *Rhizobium leguminosarum* binds host plant lectin. Mol Microbiol 59:1704–1713

Li G, Brown PJ, Tang JX et al (2012) Surface contact stimulates the just-in-time deployment of bacterial adhesins. Mol Microbiol 83:41–51

Liu Y, Gu Q, Ofosu FK et al (2016) Production, structural characterization and gel forming property of a new exopolysaccharide produced by *Agrobacterium* HX1126 using glycerol or d-mannitol as substrate. Carbohydr Polym 136:917–922

Marks JR, Lynch TJ, Karlinsey JE et al (1987) *Agrobacterium tumefaciens* virulence locus *pscA* is related to the *Rhizobium meliloti exoC* locus. J Bacteriol 169:5835–5837

Matthysse AG (1983) Role of bacterial cellulose fibrils in *Agrobacterium tumefaciens* infection. J Bacteriol 15:906–915

Matthysse AG (1994) Conditioned medium promotes the attachment of *Agrobacterium tumefaciens* strain NT1 to carrot cells. Protoplasma 183:131–136

Matthysse AG (2014) Attachment of *Agrobacterium* to plant surfaces. Front Plant Sci 5:252

Matthysse AG, Holmes KV, Gurlitz RH (1981) Elaboration of cellulose fibrils by *Agrobacterium tumefaciens* during attachment to carrot cells. J Bacteriol 145:583–595

Matthysse AG, Marry M, Krall L et al (2005) The effect of cellulose overproduction on binding and biofilm formation on roots by *Agrobacterium tumefaciens*. Mol Plant-Microbe Interact 18:1002–1010

Matthysse AG, McMahan S (1998) Root colonization by *Agrobacterium tumefaciens* is reduced in *cel, attB, attD*, and *attR* mutants. Appl Environ Microbiol 64:2341–2345

Matthysse AG, White S, Lightfoot R (1995) Genes required for cellulose synthesis in *Agrobacterium tumefaciens*. J Bacteriol 177:1069–1075

McIntosh M, Stone BA, Stanisich VA (2005) Curdlan and other bacterial (1→3)-beta-D-glucans. Appl Microbiol Biotechnol 68:163–173

Miller KJ, Kennedy EP, Reinhold VN (1986) Osmotic adaptation by Gram-negative bacteria: possible role for periplasmic oligosaccharides. Science 231:48–51

Molhoj M, Pagant S, Hofte H (2002) Towards understanding the role of membrane-bound endo-{beta}-1,4-glucanases in cellulose biosynthesis. Plant Cell Physiol 43:1399–1406

Morgan JL, McNamara JT, Fischer M et al (2016) Observing cellulose biosynthesis and membrane translocation *in crystallo*. Nature 531:329–334

Morgan JL, McNamara JT, Zimmer J (2014) Mechanism of activation of bacterial cellulose synthase by cyclic di-GMP. Nat Struct Mol Biol 21:489–496

O'Connell KP, Handelsman J (1989) *chvA* locus may be involved in export of neutral cyclic beta-1,2-linked D-glucan from *Agrobacterium tumefaciens*. Mol Plant-Microbe Interact 2:11–16

O'Neill MA, Robison PD, Chou KJ et al (1992) Evidence that the acidic polysaccharide secreted by *Agrobacterium radiobacter* (ATCC 53271) has a seventeen glycosyl-residue repeating unit. Carbohydr Res 226:131–154

Puvanesarajah V, Schell FM, Stacey G et al (1985) Role for 2-linked-beta-D-glucan in the virulence of *Agrobacterium tumefaciens*. J Bacteriol 164:102–106

Reuber TL, Walker GC (1993) Biosynthesis of succinoglycan, a symbiotically important exopolysaccharide of *Rhizobium meliloti*. Cell 74:269–280

Reuhs BL, Kim JS, Matthysse AG (1997) Attachment of *Agrobacterium tumefaciens* to carrot cells and *Arabidopsis* wound sites is correlated with the presence of a cell-associated, acidic polysaccharide. J Bacteriol 179:5372–5379

Romling U (2002) Molecular biology of cellulose production in bacteria. Res Microbiol 153:205–212

Ruffing AM, Castro-Melchor M, Hu WS et al (2011) Genome sequence of the curdlan-producing *Agrobacterium* sp. strain ATCC 31749. J Bacteriol 193:4294–4295

Ruffing AM, Chen RR (2012) Transcriptome profiling of a curdlan-producing *Agrobacterium* reveals conserved regulatory mechanisms of exopolysaccharide biosynthesis. Microb Cell Fact 11:17

Slabaugh E, Davis JK, Haigler CH et al (2014) Cellulose synthases: new insights from crystallography and modeling. Trends Plant Sci 19:99–106

Spiers AJ, Bohannon J, Gehrig SM et al (2003) Biofilm formation at the air-liquid interface by the *Pseudomonas fluorescens* SBW25 wrinkly spreader requires an acetylated form of cellulose. Mol Microbiol 50:15–27

Srivatsan A, Wang JD (2008) Control of bacterial transcription, translation and replication by (p) ppGpp. Curr Opin Microbiol 11:100–105

Stanisich VA, Stone BA (2009) Enzymology and molecular genetics of biosynthetic enzymes for (1,3)-beta-glucans: prokaryotes. In: Bacic A, Fincher GB, Stone BA (eds) Chemistry, biochemistry, and biology of (1-3)-beta-glucans and related polysaccharides. Elsevier, Amsterdam, pp 201–232

Stasinopoulos SJ, Fisher PR, Stone BA et al (1999) Detection of two loci involved in (1→3)-beta-glucan (curdlan) biosynthesis by *Agrobacterium* sp. ATCC31749, and comparative sequence analysis of the putative curdlan synthase gene. Glycob 9:31–41

Swart S, Lugtenberg B, Smit G et al (1994) Rhicadhesin-mediated attachment and virulence of an *Agrobacterium tumefaciens chvB* mutant can be restored by growth in a highly osmotic medium. J Bacteriol 176:3816–3819

Sykes LC, Matthysse AG (1986) Time required for tumor induction by *Agrobacterium tumefaciens*. Appl Environ Microbiol 52:597–598

Thomashow MF, Karlinsey JE, Marks JR et al (1987) Identification of a new virulence locus in *Agrobacterium tumefaciens* that affects polysaccharide composition and plant cell attachment. J Bacteriol 169:3209–3216

Tiburtius A, de Luca NG, Hussain H et al (1996) Expression of the *exoY* gene, required for exopolysaccharide synthesis in *Agrobacterium*, is activated by the regulatory *ros* gene. Microbiol 142:2621–2629

Tomlinson AD, Ramey-Hartung B, Day TW et al (2010) *Agrobacterium tumefaciens* ExoR represses succinoglycan biosynthesis and is required for biofilm formation and motility. Microbiol 156:2670–2681

Uttaro AD, Cangelosi GA, Geremia RA et al (1990) Biochemical characterization of avirulent *exoC* mutants of *Agrobacterium tumefaciens*. J Bacteriol 172:1640–1646

Wang D, Xue H, Wang Y et al (2013) The *Sinorhizobium meliloti ntrX* gene is involved in succinoglycan production, motility, and symbiotic nodulation on alfalfa. Appl Environ Microbiol 79:7150–7159

Wang Y, Kim SH, Natarajan R et al (2016) Spermidine Inversely Influences Surface Interactions and Planktonic Growth in *Agrobacterium tumefaciens*. J Bacteriol 198:2682–2691

Weiner R, Langille S, Quintero E (1995) Structure, function and immunochemistry of bacterial exopolysaccharides. J Ind Microbiol 15:339–346

Whitney JC, Hay ID, Li C et al (2011) Structural basis for alginate secretion across the bacterial outer membrane. Proc Natl Acad Sci USA 108:13083–13088

Wong HC, Fear AL, Calhoon RD et al (1990) Genetic organization of the cellulose synthase operon in *Acetobacter xylinum*. Proc Natl Acad Sci USA 87:8130–8134

Wu CF, Lin JS, Shaw GC et al (2012) Acid-induced type VI secretion system is regulated by ExoR-ChvG/ChvI signaling cascade in *Agrobacterium tumefaciens*. PLoS Pathog 8:e1002938

Wu D, Li A, Ma F et al (2016) Genetic control and regulatory mechanisms of succinoglycan and curdlan biosynthesis in genus *Agrobacterium*. Appl Microbiol Biotechnol 100:6183–6192

Xu J, Kim J, Danhorn T et al (2012) Phosphorus limitation increases attachment in *Agrobacterium tumefaciens* and reveals a conditional functional redundancy in adhesin biosynthesis. Res Microbiol 163:674–684

Xu J, Kim J, Koestler BJ et al (2013) Genetic analysis of *Agrobacterium tumefaciens* unipolar polysaccharide production reveals complex integrated control of the motile-to-sessile switch. Mol Microbiol 89:929–948

York GM, Walker GC (1998) The *Rhizobium meliloti* ExoK and ExsH glycanases specifically depolymerize nascent succinoglycan chains. Proc Natl Acad Sci USA 95:4912–4917

Yu LJ, Wu JR, Zheng ZY et al (2011a) Changes of curdlan biosynthesis and nitrogenous compounds utilization characterized in *ntrC* mutant of *Agrobacterium* sp. ATCC 31749. Curr Microbiol 63:60–67

Yu LJ, Wu JR, Zheng ZZ et al (2011b) Changes in gene transcription and protein expression involved in the response of *Agrobacterium* sp. ATCC 31749 to nitrogen availability during curdlan production. Prikl Biokhim Mikrobiol 47:537–543

Zorreguieta A, Geremia RA, Cavaignac S et al (1988) Identification of the product of an *Agrobacterium tumefaciens* chromosomal virulence gene. Mol Plant-Microbe Interact 1:121–127

Zorreguieta A, Ugalde RA, Leloir LF (1985) An intermediate in cyclic beta 1-2 glucan biosynthesis. Biochem Biophys Res Commun 126:352–357

Function and Regulation of *Agrobacterium tumefaciens* Cell Surface Structures that Promote Attachment

Melene A. Thompson, Maureen C. Onyeziri and Clay Fuqua

Contents

1 Introduction 144
2 Flagella 146
2.1 Flagellum Structure, Function, Biogenesis, and Regulation 146
2.2 Role of Flagellar Motility in Attachment 151
3 Pili 152
3.1 Ctp Pili 153
3.2 Conjugative Pili 155
4 Exopolysaccharides 156
4.1 Cellulose 157
4.2 Unipolar Polysaccharide (UPP) 161
4.3 Curdlan 167
5 Protein Adhesins 169
6 Regulation of Attachment 170
6.1 Transcriptional Regulators that Impact Attachment 170
6.2 Cyclic Diguanylate Monophosphate-Dependent Regulation of Surface Attachment Functions 172
6.3 Surface-Contact-Responsive Control of Attachment Functions 176
7 Conclusions 177
References 178

Abstract *Agrobacterium tumefaciens* attaches stably to plant host tissues and abiotic surfaces. During pathogenesis, physical attachment to the site of infection is a prerequisite to infection and horizontal gene transfer to the plant. Virulent and avirulent strains may also attach to plant tissue in more benign plant associations, and as with other soil microbes, to soil surfaces in the terrestrial environment. Although most *A. tumefaciens* virulence functions are encoded on the

M. A. Thompson · M. C. Onyeziri · C. Fuqua (✉)
Department of Biology, Indiana University, Bloomington, IN 47405, USA
e-mail: cfuqua@indiana.edu

Current Topics in Microbiology and Immunology (2018) 418:143–184
DOI 10.1007/82_2018_96

Published Online: 12 July 2018

tumor-inducing plasmid, genes that direct general surface attachment are chromosomally encoded, and thus this process is not obligatorily tied to virulence, but is a more fundamental capacity. Several different cellular structures are known or suspected to contribute to the attachment process. The flagella influence surface attachment primarily via their propulsive activity, but control of their rotation during the transition to the attached state may be quite complex. *A. tumefaciens* produces several pili, including the Tad-type Ctp pili, and several plasmid-borne conjugal pili encoded by the Ti and At plasmids, as well as the so-called T-pilus, involved in interkingdom horizontal gene transfer. The Ctp pili promote reversible interactions with surfaces, whereas the conjugal and T-pili drive horizontal gene transfer (HGT) interactions with other cells and tissues. The T-pilus is likely to contribute to physical association with plant tissues during DNA transfer to plants. *A. tumefaciens* can synthesize a variety of polysaccharides including cellulose, curdlan (β-1,3 glucan), β-1,2 glucan (cyclic and linear), succinoglycan, and a localized polysaccharide(s) that is confined to a single cellular pole and is called the unipolar polysaccharide (UPP). Lipopolysaccharides are also in the outer leaflet of the outer membrane. Cellulose and curdlan production can influence attachment under certain conditions. The UPP is required for stable attachment under a range of conditions and on abiotic and biotic surfaces. Other factors that have been reported to play a role in attachment include the elusive protein called rhicadhesin. The process of surface attachment is under extensive regulatory control and can be modulated by environmental conditions, as well as by direct responses to surface contact. Complex transcriptional and post-transcriptional control circuitry underlies much of the production and deployment of these attachment functions.

Keywords Attachment · Cell surface structures · Biofilms · Regulation

1 Introduction

A wide diversity of bacteria interact with surfaces in their environments, often forming multicellular assemblies known as biofilms. Recent years have seen an explosion of research on biofilms and surface attachment mechanisms, reflecting an appreciation of how ubiquitous these processes are and the extent to which they can influence bacterial physiology (Visick et al. 2016). Adherent bacteria exhibit dramatically different bioactivities than they do in the unattached state and, most notably, biofilm formation can markedly increase tolerance toward antibiotics. For pathogenic bacteria, association with host surfaces is often the first step toward infection, and antibiotic-resistant biofilms formed by pathogens have become a major clinical problem in human medicine (Hoiby 2017). Biofilms on non-host surfaces can also act as disease reservoirs and are conducive to horizontal gene transfer of antibiotic resistance determinants and virulence factors (Madsen et al. 2012). Structures on the bacterial cell surface, including a variety of filamentous and globular protein adhesins, and exopolysaccharides mediate the interactions with

surfaces that lead to stable colonization. Multiple types of surface fibers including pili, fimbriae, and β-amyloid filaments contribute to attachment. These fibers include flagella that drive motility for many bacteria, but in some cases can act as adhesins as well as propulsive structures. Several different types of secretion systems, such as Type III (T3S), Type IV (T4S), and Type VI (T6S) systems, can also influence surface interactions. Certain bacterial taxa produce large surface proteins with multi-repeat domains that can function as adhesins (Hinsa et al. 2003). A variety of polysaccharides, including outer membrane lipopolysaccharides, have been implicated in stable surface attachment and biofilm formation (Branda et al. 2005). The production of specific cell surface adhesive structures is often under elaborate regulatory control.

Similar to human and animal pathogens, many pathogens of plants must productively colonize host surfaces to cause disease. Multicellular aggregates also play roles for pathogenic bacteria in foliar, vascular, and root environments (Danhorn and Fuqua 2007). Pathogenic *Agrobacterium* species must physically associate with surfaces to drive interkingdom gene transfer to plants and other aspects of *Agrobacterium*-induced disease. For crown gall disease, the details of the physical interactions between *A. tumefaciens* and plant cells which lead to T4S-mediated introduction of transferred DNA (T-DNA) into the plant cell cytoplasm remain poorly understood. It is clear that *A. tumefaciens* is an effective colonizer of plant surfaces during pathogenic and non-pathogenic interactions, as well as associating with abiotic surfaces in the soil environment. How this general surface attachment progresses, or switches to the physical association leading to T-DNA transfer, is still being actively studied. A large cluster of so-called Attachment (*Att*) genes were purported to be required for *A. tumefaciens* association with plant tissues, and to be necessary for virulence (Matthysse et al. 2000). Subsequently, the *Att* genes were shown to be encoded on the pAt plasmid, and this plasmid was demonstrated to be dispensable for virulence (Nair et al. 2003). More recent work has suggested the possibility that mutations in the *Att* cluster may have a dominant inhibitory effect on attachment (Matthysse et al. 2008). Thus, despite their historical identification, the *Att* genes are currently not thought to play a direct role in attachment to plants or other surfaces. Rather, there must be other cell surface structures that function in this capacity. Flagella, several different forms of pili, and multiple complex polysaccharides are produced by *A. tumefaciens*, and several of these are now known to promote general attachment to surfaces. The virulence (Vir) proteins involved in interkingdom gene transfer to plants may also contribute to attachment on host tissues. As with mammalian pathogens, deployment of these and other cell surface attributes in *A. tumefaciens* can be elaborately regulated by transcriptional and post-transcriptional mechanisms. In this review, we describe the current understanding of *A. tumefaciens* cell surface structures that contribute to surface attachment mechanisms, including those for host tissues and abiotic materials, the molecular composition and biosynthesis of these structures, and the recognized systems that control their activity.

2 Flagella

Flagella play an important role in attachment in addition to their more general function in enabling diverse bacteria to propel themselves through their environment. Flagellar propulsion can enable bacteria to move toward conditions that are favorable such as high nutrients, and avoid conditions that inhibit growth or damage cells. Much of what is known about flagella structure and assembly derives from studies of the peritrichous flagella of *Escherichia coli* and *Salmonella enterica* (Chevance and Hughes 2008; MccNab 1996). However, outside of these model systems, there is a significant variety of flagellar organization, number, and composition (Schuhmacher et al. 2015). An example of this is the type strain *A. tumefaciens* C58, which extrudes four to six flagella that are each ~10–12 nm diameter filaments (Chesnokova et al. 1997; Shaw et al. 1991). As with several other members of the *Rhizobiaceae*, the *A. tumefaciens* flagellar filament is composed of multiple flagellin proteins and exhibits a complex ultrastructure (Götz et al. 1982). The primary flagellin FlaA is strictly required for motility. The other three flagellins (FlaB, FlaC, and FlaD) play more ancillary roles but are proposed to be important for flagellar filament structural integrity (Deakin et al. 1999).

2.1 Flagellum Structure, Function, Biogenesis, and Regulation

The structure and activity of the bacterial flagellum is recognized as one of the molecular marvels of the natural world (prompting some individuals to conclude that their structure is evidence for divine intervention; Pallen and Matzke 2006). The biogenesis of these remarkable rotary nanomachines is a prime example of an ordered molecular assembly process. Flagella are assembled from the inside out, with their basal bodies comprised of a series of ring structures (Chevance and Hughes 2008). In gram-negative bacteria, the C-ring is assembled near the cytoplasmic face of the inner membrane (Fig. 1). Associated with the C-ring is the MS-ring that forms within the cytoplasmic membrane, and it houses a T3S system which exports specific flagellar components through the center of this ring. The next components to assemble are the P-ring (embedded in the peptidoglycan) and the L-ring (embedded in the outer membrane), containing proteins secreted into the periplasm via the general secretion system. All of these rings are made of multiple copies of the same proteins, and at least some of the structures are thought to be dynamic in the numbers of monomers that form the ring (Branch et al. 2014; Lele et al. 2012). The motor and stator complex is embedded in the inner membrane and has a large domain that associates with the C-ring (Fig. 1). Motor proteins assemble around the C-ring complex. Other proteins, including the switch proteins that can alter flagellar rotation, can associate with the inner face of the C-ring (Chevance and Hughes 2008). The rod structure connects these rings through the bacterial

envelope with the proteins that make up the flagellar filament (hook subunits, assembly chaperones, linkers, and flagellins). These proteins are exported through the lumen of the rod, via the T3S system, and are added to the growing filament at the distal end. The hook is connected to the rod, and approximately 130 copies of the hook protein assemble, directed by the hook chaperone, to form this flexible universal joint (Fig. 1). Additional linker proteins and an assembly chaperone facilitate sequential addition of the flagellin subunits to the hook and subsequent extension of the helical filament building from the distal end, comprised of as many as 30,000 flagellin monomers (although this number varies significantly with bacterial taxon and average flagellar length; Blair 2003). Rotation of the flagellum is driven by proton translocation, and in *E. coli* and *Salmonella* spp., it is estimated that ~550 protons are translocated per single rotation of the flagellum. In the enteric model systems, the rotation of the flagellum is reversible, with counter-clockwise (CCW) rotation driving straight swimming, and clockwise (CW) rotation generating cellular tumbles that reorient the cell. The ratio of swimming to tumbling

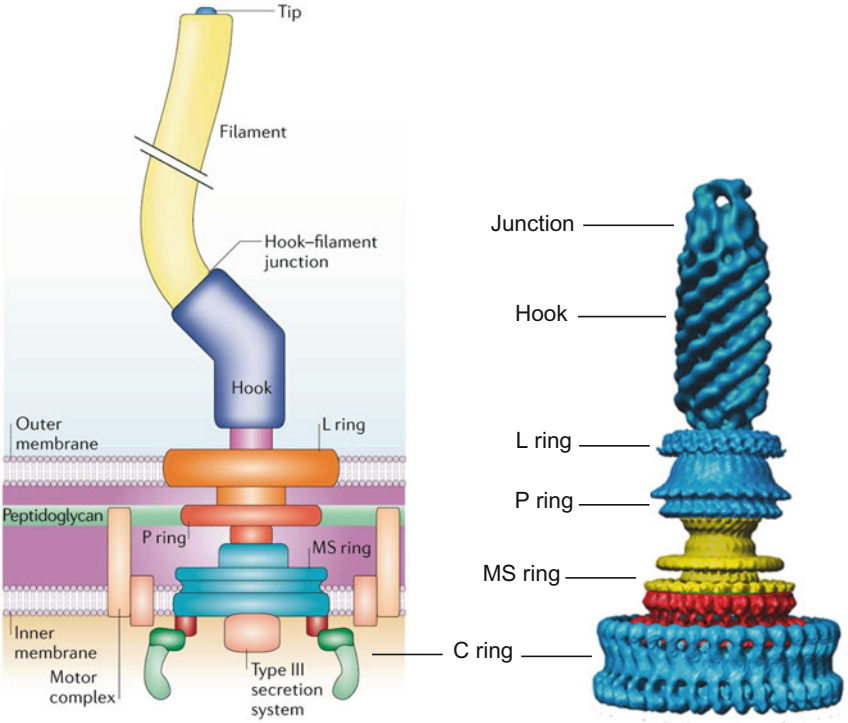


Fig. 1 General structure of the bacterial flagellum. Diagrammatic representation of a bacterial flagellum structure based on flagella from *Salmonella enterica* serovar *Typhimurium* (*Salmonella typhimurium*), alongside a flagellum structure determined by cryoelectron microscopy of the flagellum from *Treponema primitia*. Combined figure adapted with permissions from Nat Rev Microbiol (Pallen and Matzke 2006) and Curr. Biol. (DeRosier 2006)

is under the control of the chemotaxis system, which enables directed motility (Wadhams and Armitage 2004).

The general properties of the *A. tumefaciens* flagellum are consistent with those from the enteric model systems, but with several important differences. Among the *Rhizobiaceae,* the best-studied model is *Sinorhizobium meliloti*, and this system shares several distinguishing features with *A. tumefaciens*. For example, each of these systems has complex flagella composed of four different flagellins (Deakin et al. 1999; Götz et al. 1982). Rotation of these complex flagella is consistently in the CW direction, and in contrast to *E. coli* and *Salmonella,* this is not reversed to generate tumbles. Rather, the current model is that the rate of CW flagellar rotation is modulated by the chemotaxis system, and that asynchronous rotation of multiple flagella causes tumbling (Sourjik and Schmitt 1996).

The arrangement of the flagellar filaments varies among even closely related bacteria (Schuhmacher et al. 2015). In *A. tumefaciens,* the flagella are organized into a polar tuft of 4-6 filaments (Chesnokova et al. 1997), whereas *S. meliloti* has a peritrichous organization (Götz et al. 1982). There is significant variation for flagellar placement among the rhizobia that has been recognized for many years (Leifson and Erdman 1958). *A. tumefaciens* demonstrates a swimming pattern with long, straight runs (Mohari et al. 2015). The bacterium has strong positive chemotactic responses to the sugars sucrose, glucose, and fructose, with slightly weaker responses to a variety of other sugars, as well as responses to the amino acids valine and arginine (Ashby et al. 1988).

The many proteins that contribute to assembly of flagella and motility are typically encoded in large operons or gene clusters. In *A. tumefaciens* C58, the genes encoding the structural components of the flagella, assembly factors, and several motility regulators are in one large gene cluster ($\sim$36 kbp, Atu0541-Atu0585; Table 1) on the circular chromosome (Deakin et al. 1997a, b, 1999). Additionally, the core chemotaxis genes reside in a single gene cluster (>13.5 kbp, Atu0514-0526) that also includes several likely flagellar genes and the two master motility regulators *visN* and *visR.* The *Che* cluster is located close to the *Fla* cluster, separated only by $\sim$14 kbp. Together, there are more than 50 genes in these two clusters (Table 1) and most if not all appear to be dedicated to motility, although only a subset of these have been experimentally validated for *A. tumefaciens*. There are a few scattered chemotaxis gene homologues throughout the genome (predominantly methyl-dependent chemotaxis protein homologues), but the majority of motility and chemotaxis functions are in these two clusters (Liu and Ochman 2007).

Expression of flagellar genes is often under complex, stepwise morphogenetic control which ensures that each component of the flagellum is produced at the appropriate step of biogenesis (Chevance and Hughes 2008). Based on the enteric model systems, these genes are often divided into three classes, initiating with the master regulators of motility (Class I), the genes encoding the T3S system through which the flagellum is assembled, the hook–basal body complex (Class II), and those proteins that make up the flagellar filament (Class III). Flagellar gene expression in several rhizobia, including *A. tumefaciens*, is controlled through a complex regulatory hierarchy the details of which remain to be fully defined.

Table 1 Core chemotactic and flagellar gene clusters

Gene number	Gene[1] name	Gene length—bp (Protein length—aa)[1]	Predicted function[1]
Atu0514		1707 (568)	Methyl-accepting chemotaxis protein
Atu0515	cheX	300 (99)	Chemotaxis protein
Atu0516	cheY1	366 (121)	Chemotaxis receiver protein
Atu0517	cheA	2214 (737)	Chemotaxis histidine kinase
Atu0518	cheR	909 (302)	Chemotaxis methyltransferase
Atu0519	cheB	1056 (351)	Chemotaxis methylesterase
Atu0520	cheY2	390 (129)	Chemotaxis receiver protein
Atu0521	cheD	546 (181)	Methyl-accepting chemotaxis protein
Atu0522		184 (127)	Undefined function
Atu0523	fliF	1701 (566)	Flagellar M-ring protein
Atu0524	visN	681 (226)	LuxR-type transcriptional regulator
Atu0525	visR	756 (251)	LuxR-type transcriptional regulator
Atu0526	mclA	1848 (615)	Methyl-accepting chemotaxis protein
Atu0542	fla	942 (313)	Flagellin
Atu0543	flaB	963 (320)	Flagellin
Atu0544		258 (85)	Undefined function
Atu0545	flaA	921 (306)	Flagellin
Atu0546	fliP	738 (245)	Flagellar export protein
Atu0547	fliL	501 (166)	Flagellar protein
Atu0548	flgH	720 (239)	Flagellar L-ring precursor
Atu0549		537 (178)	Undefined function
Atu0550	flgI	1122 (373)	Flagellar P-ring precursor
Atu0551	flgA	489 (162)	Flagellar P-ring protein
Atu0552	flgI	789 (262)	Flagellar rod protein
Atu0553	fliE	339 (112)	Flagellar hook–basal body protein
Atu0554	flgC	420 (139)	Flagellar body-rod protein
Atu0555	flgB	393 (130)	Flagellar basal body-rod protein
Atu0556		405 (134)	Undefined function
Atu0557	fliI	1422 (473)	Flagellum-specific ATPase
Atu0558	flgF	735 (244)	Flagellar basal body-rod protein
Atu0559		636 (211)	Undefined function
Atu0560	motA	873 (290)	Flagellar motor protein
Atu0561	fliM	960 (319)	Flagellar motor switch protein
Atu0562	fliN	540 (179)	Flagellar motor switch protein
Atu0563	fliG	1044 (347)	Flagellar motor switch protein
Atu0564	flhB	1083 (360)	Flagellar export protein
Atu0565		438 (145)	Undefined function
Atu0566		213 (70)	Undefined function
Atu0567	flaD	1293 (430)	Flagellin

(continued)

Table 1 (continued)

Gene number	Gene[1] name	Gene length—bp (Protein length—aa)[1]	Predicted function[1]
Atu0568		657 (218)	Undefined function
Atu0569	motB	1302 (433)	Flagellar motor protein
Atu0570	motC	1281 (426)	Chemotaxis protein
Atu0571	motD	1350 (449)	Chemotaxis protein
Atu0572		444 (147)	Undefined function
Atu0573	rem	672 (223)	OmpR-type transcriptional regulator
Atu0574	flgE	1278 (425)	Flagellar hook protein
Atu0575	flgK	1479 (492)	Flagellar hook-associated protein
Atu0576	flgL	1104 (367)	Flagellar hook-associated protein
Atu0577	flaF	345 (114)	Flagellar biosynthesis regulatory protein
Atu0578	flbT	450 (149)	Flagellar biosynthesis regulatory protein
Atu0579	flgD	474 (157)	Hook formation protein
Atu0580	fliQ	267 (88)	Flagellar export protein
Atu0581	flhA	2088 (695)	Flagellar export protein
Atu0582	fliR	723 (240)	Flagellar export protein
Atu0583		420 (139)	Undefined function
Atu0584		528 (175)	Undefined function
Atu0585		369 (122)	Undefined function
Atu8132		534 (177)	Undefined function

[1]Gene identity, length, and predicted function annotations are based on KEGG Gene Ontology (Kyoto University Bioinformatics Center) and Osterman et al. 2015

Class IA is composed of the LuxR-type transcription factors VisN and VisR (Vital in swimming), both of which are required for flagellar motility (Sourjik et al. 2000). VisN and VisR activate expression of the class IB gene *rem* (regulator of exponential growth motility) (Xu et al. 2013). Rem is an OmpR-type response regulator (Rotter et al. 2006) with no recognized cognate histidine kinase, and lacking the canonical Asp residue at which most response regulators are phosphorylated. Rem activates expression of the Class II flagellar genes which include the components of the flagellar hook and basal body as well as the motor, and it is also required for expression of Class III genes, including the flagellin genes and several chemotaxis genes (Sourjik et al. 2000; Zatakia et al. 2018). In *Brucella melitensis*, flagellin synthesis is inversely regulated by the flagellin activator FlbT and the repressor FlaF (Ferooz et al. 2011). Indeed, this level of control may be broadly conserved throughout the Alphaproteobacteria, as a similar pathway has been delineated in *C. crescentus* (Mangan et al. 1999). *A. tumefaciens* has homologues of FlaF and FlbT, although their roles are largely unexplored. Many flagellar assembly pathways include the activity of a specialized sigma factor dedicated to the transcription of subsets of the flagellar genes (e.g., σ^{28} in *E. coli* and σ^{D} in *B. subtilis*), with promoter sequences that are quite distinct from those for σ^{70} promoters (Aldridge and Hughes 2002). The motility sigma factor is often controlled through an

anti-sigma factor (Hughes and Mathee 1998). Promoters regulated by Rem in *S. meliloti*, and by extension in *A. tumefaciens*, clearly have a non-σ^{70} architecture (Rotter et al. 2006), suggesting the presence of an alternate sigma factor, but there is no such annotated sigma factor encoded in the motility gene cluster. Thus far, this presumptive sigma factor has not been identified for any of the rhizobia.

Numerous factors outside of the hierarchy of flagellar regulators affect *A. tumefaciens* motility through various mechanisms. The periplasmic succinoglycan-regulatory protein ExoR is required for motility through its effects on flagellar gene expression (Tomlinson et al. 2010); the broader role of ExoR will be discussed later in this review (Sect. 6.1.1). Flagella synthesis as well as flagellin expression and motility are elevated when *A. tumefaciens* cells are grown in the absence of light, but the mechanism for this is unclear (Oberpichler et al. 2008). Flagellar motility is also affected by the cell cycle regulators *divK*, *pdhS1*, and *pleC*, mutants in which form branched cells and have altered flagellar placement (Kim et al. 2013). A *pdhS2* mutant does not manifest aberrant cell shape, and these cells produce flagella, but they do not swim. Given the asymmetric cell division mechanism of *A. tumefaciens* (Brown et al. 2012), it is not surprising that flagellar biogenesis would be integrated with control of the cell cycle.

2.2 Role of Flagellar Motility in Attachment

Flagella play a role in *A. tumefaciens* attachment to model surfaces. Aflagellate mutants deleted for the hook protein FlgE, and *motA* deletion mutants with unpowered flagella are both highly deficient in surface attachment and biofilm formation under static conditions (Merritt et al. 2007). Given the requirement for active flagellar rotation, it was concluded that swimming motility drives the frequency or productivity of surface contact. Interestingly, an aflagellate hook mutant formed biofilms more robustly and rapidly than did wild-type cells in a flow cell, suggesting that the flow regime promoted high frequency surface contact and that perhaps the lack of motility limited emigration from the surface (Merritt et al. 2007). A straight swimming *cheA* mutant manifested only a modest attachment defect in static culture, but quantitative analysis of flow cell biofilms revealed a different three-dimensional structure. Motility in these straight swimming *cheA* mutants is compromised in motility agar dispersal assays, but spontaneous suppressor mutants can be readily isolated (Mohari et al. 2015). These mutants regain tumbling activity, and hence migration through motility agar because of structural deformations in the flagellum. These changes, however, result in a dramatic loss of attachment, revealing how proper coordination of motility is important during surface colonization. Further evidence for the connection between motility and attachment of *A. tumefaciens* came from a screen for regulators of attachment in which null mutations of the Class IA master regulators *visN* and *visR* resulted in increased attachment, even though they abolished motility (Xu et al. 2013). The increased attachment of these mutants results from a complex regulatory pathway as detailed below (Sect. 6.2).

3 Pili

Many bacteria promote attachment to surfaces with proteinaceous surface appendages known as pili (or sometimes called fimbriae). Although many pili are considered static appendages, they are all actively extruded, and in many cases actively retracted. These filaments can also drive twitching motility in some bacteria (Mattick 2002). A subset of pili are involved in plasmid conjugation and are referred to as sex pili. *A. tumefaciens* encodes production of several different types of pili. The chromosomes of agrobacteria often carry a cluster of genes (Fig. 2)

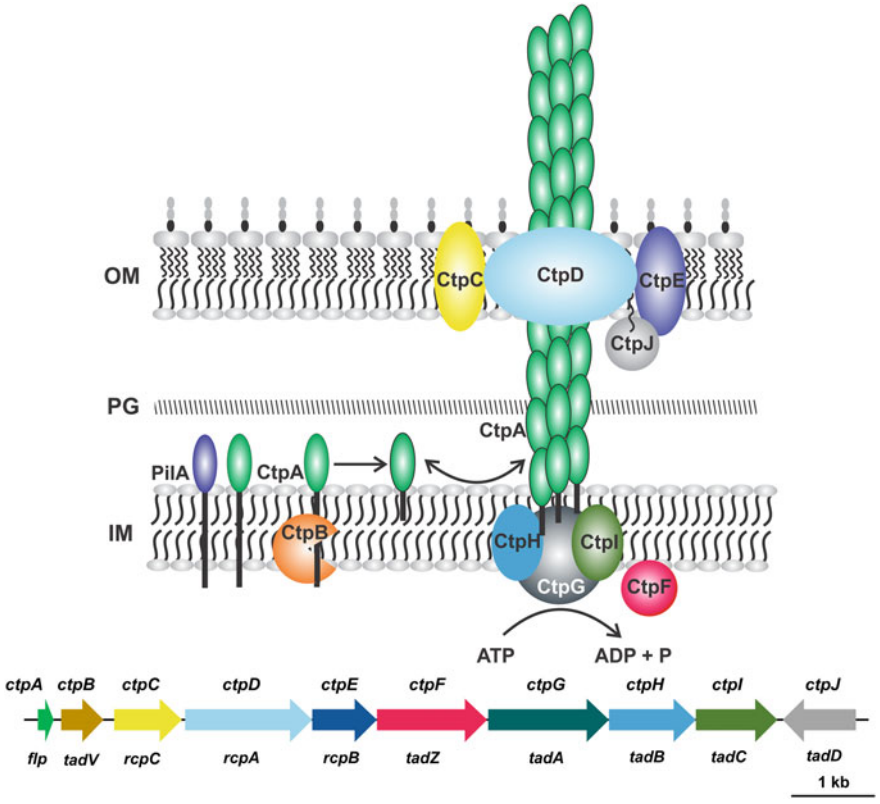


Fig. 2 Genetic basis of Ctp pili and model for assembly. Predicted localization and assembly mechanism for Ctp proteins based on the general model for Tad pilus assembly (Tomich et al. 2007). The CtpA pilin is processed by CtpB cleavage and incorporated into the emerging pilus. CtpG hydrolyzes ATP to drive pilus assembly. Protein names indicated in the figure. Gene map indicates the *ctp* gene names above the gene arrows, and gene names in the generalized Tad-type pilus system are provided below. Gene colors match protein colors in the diagram. OM, outer membrane; PG, peptidoglycan; IM, inner membrane

annotated as the *Ctp* cluster (Cpa-type pilus; named after homologues in *Caulobacter crescentus*; Skerker and Shapiro 2000), required to form Type IV pili of the Tad (Tight adherence) subclass, also known as common pili (Wang et al. 2014). Two distinct types of conjugative pili are encoded by the Ti plasmid: the Trb conjugal pili, required to conjugatively transfer the entire plasmid to recipient bacteria (Cook et al. 1997), and the T-pilus, required for T-DNA transfer to plants (Fullner et al. 1996). The At plasmid also encodes its own set of conjugative pili, called the AvhB system (Chen et al. 2002).

3.1 *Ctp Pili*

Electron microscopy of *A. tumefaciens* C58 reveals thin filaments (~3 nm in diameter) interspersed across the cell body (Lai et al. 2000; Wang et al. 2014). The *A. tumefaciens* C58 genome sequence contains a cluster of nine genes (Atu0224-0216), annotated *ctpABCDEFGHI*, that encode Type IV pilus assembly homologues (Fig. 2). Although initially defined as Type IVb pili, a recent study has re-classified the *Ctp* pili genes as Type IVc pilus assembly genes, conserved in diverse bacteria (Ellison et al. 2017). The *A. tumefaciens Ctp* locus is homologous and syntenous among different members of the *Rhizobiaceae*, as well as in more diverse Alphaproteobacteria such as *C. crescentus* (Skerker and Shapiro 2000). Individual non-polar, in-frame deletions of most of the genes in the *A. tumefaciens* C58 *Ctp* cluster result in loss of piliation as evaluated by transmission electron microscopy, and these mutants are significantly inhibited for attachment and biofilm formation (Wang et al. 2014). Transcriptional fusion analysis suggests the existence of two promoters, one upstream of *ctpA* (*ctpABCD*) and a second upstream of *ctpE* (*ctpEFGHI)* (Fig. 2, Wang et al. 2014).

The *ctpA* gene encodes a small pilin homologue (64aa), often called Flp (fimbrial low molecular-weight protein) pilin. The protein contains the hydrophobic "Flp-motif" that includes the conserved pilin processing site (G/XXXXEY) (Kachlany et al. 2001). A second Flp pilin homologue, annotated *pilA*, is encoded elsewhere in the C58 genome (Atu3514), and if ectopically expressed can complement *ctpA* mutants for pilus assembly (Wang et al. 2014). Immediately downstream of *ctpA* is *ctpB*, a prepilin peptidase homologue, required to process pilin during assembly. The *ctpC*, *ctpD,* and *ctpE* genes encode proteins that are homologous to the RcpC, RcpB, and RcpA proteins, respectively, from *Aggregatibacter actinomycetemcomitans* which putatively encode an outer membrane complex centered around the CtpD secretin (Tomich et al. 2007). The CtpF protein is a homologue of CpaE from *C. crescentus* and TadZ from *A. actinomycetemcomitans*, proteins that have MinD/ParE homology and are predicted localization factors. Interestingly, *ctpF* is the only *A. tumefaciens* C58 gene in the *Ctp* cluster for which deletion does not cause loss of piliation (Wang et al. 2014). CtpG is an ATPase likely localized to the cytoplasm, and based on similarity to TadA from *A. actinomycetemcomitans*, is involved in powering Ctp pilus

biogenesis. Finally, the CtpH and CtpI proteins are so-called platform proteins associated with the cytoplasmic membrane upon which the pilus is assembled, similar to TadB/C, and are likely to have arisen via a gene duplication event (Tomich et al. 2007). In many systems, a TadD homologue with a tetratricopeptide repeat (TPR) motif is located downstream and convergent to the pilus gene cluster. This gene is Atu0215 in *A. tumefaciens* C58, and here we tentatively designate it as *ctpJ* (Fig. 2). For the Tad system, this homologue is speculated to be a pilotin, with a lipid linkage to the outer membrane. However, the CtpJ product has no secretion signal, and its role in Ctp pilus function is not known.

In the current model (Fig. 2), pools of the CtpA pilin (or alternatively PilA) associate with the cytoplasmic membrane. Pilus biogenesis requires the CtpB prepilin peptidase to cleave the CtpA monomers, and it is this processed form that interacts with the assembly machine, incorporating into the growing pilus at the base. Pilus assembly is powered by the cytoplasmic ATPase (CtpG), driving conformational changes that promote interactions of the pilin subunits on the platform proteins at the cytoplasmic face. The CtpC, CtpD, and CtpE proteins interact with the emerging pilus at the outer membrane, with the pilus spanning the membrane through the CtpD secretin. Recent work has also shown that Type IVc pili can retract, and the released pilin proteins can reassociate with the cytoplasmic membrane (Ellison et al. 2017). Many of the specific aspects of this overall model remain to be evaluated experimentally in *A. tumefaciens* and other similar systems.

Mutagenesis of the *ctpA*, *ctpB*, and *ctpG* genes leads to significant deficiencies in biofilm formation for *A. tumefaciens*, and this was correlated with problems in reversible attachment (Wang et al. 2014). Deletion of the entire *Ctp* gene cluster causes a similar loss of attachment. Individual mutations in any of the other genes in the *Ctp* cluster (*ctpC*, *ctpD*, *ctpE*, *ctpF*, *ctpH*, and *ctpI*), with the exception of the *ctpF* mutant, abolish pilus biogenesis, but surprisingly all of these mutants are stimulated rather than diminished for attachment. Disruption of *ctpA* in these hyperadherent *Ctp* mutants abolished their stimulated attachment. Likewise, mutations that prevent production of the unipolar polysaccharide (UPP) adhesin (see Sect. 4.2) completely prevent bacterial attachment (Wang et al. 2014). Although the reason for this hyperattachment in certain *Ctp* mutants remains unclear, it is possible that genetic disruption of these pilus functions, and perhaps accumulation of the CtpA pilin in the cytoplasmic membrane, causes feedback regulation in the cell to activate a pilus-independent, but UPP-dependent attachment mechanism. Similar signaling in response to pilin levels in the inner membrane has been reported for T4 Pa pili in *Pseudomonas aeruginosa* (Kilmury and Burrows 2016).

As with other Alphaproteobacteria that produce unipolar polysaccharide adhesins which promote attachment, *A. tumefaciens* transitions from reversible to stable polar attachment via just-in-time production of the adhesive material (see Sect. 6.3). Production of the UPP is strictly surface-contact dependent (Li et al. 2012). For *C. crescentus*, recent studies showed that its Type IVc Cpa pilus is involved in triggering surface-contact-dependent production of its polar adhesive, known as the holdfast (Ellison et al. 2017). The Cpa pili are localized to the same cellular pole

from which the holdfast will be produced, and resistance to pilus depolymerization stimulates holdfast production. It is hypothesized that physical association with surfaces during reversible interactions inhibits Cpa pilus depolymerization, thereby stimulating holdfast production. In *A. tumefaciens,* the Ctp pilus is required for reversible attachment, but electron microscopy does not thus far support a polar localization for these pili (Lai et al. 2000; Wang et al. 2014). However, it is certainly possible that a similar pilus-dependent surface-stimulation of UPP production plays a role in the transition to stable attachment for *A. tumefaciens*.

3.2 Conjugative Pili

Conjugative plasmids have been reported to promote biofilm formation in diverse bacteria (Ghigo 2001). This is, however, not attributed to the conjugative pili encoded by the plasmids, but rather, to other plasmid-encoded functions. These plasmids often drive their conjugation via cellular interaction with recipient bacteria mediated through their filamentous conjugative pili. For many years, it was proposed that conjugative pili simply act in promoting the physical association of donor and recipient cells, and that DNA transfer occurred via mating pair formation and an ill-defined conduit between cells (Lessl and Lanka 1994). However, more recent work has resurrected the notion that the single-stranded DNA delivered by conjugative plasmids transits through the lumen of the conjugal pilus (Babic et al. 2008; Costa et al. 2016). Conjugative pili promote interactions between bacterial cells, and it is certainly conceivable that these surface structures could also play a role in surface attachment. Virulent *A. tumefaciens* produce two distinct conjugative pili encoded by the Ti plasmid (Vir and Tra/Trb) and often also a third conjugative pilus type encoded by the At plasmid (AvhB). These conjugative pili are all considered components of their respective Type IV secretion (T4S) systems. The pili that function in T4S plasmid conjugation systems should not be confused with the Ctp Type IVc pili (Sect. 3.1) or other Type IV pili. The Ti plasmid encodes the so-called T-pilus as a component of the machinery for T-DNA transfer to plant cells (Fullner et al. 1996). Also encoded by the Ti plasmid is the Trb conjugative pilus, required for horizontal transfer of the plasmid to other agrobacteria (Cook et al. 1997).

T-pilus production is strictly regulated along with the *vir* genes in response to plant wound conditions (Fullner et al. 1996). Encoded within the *virB* operon that also specifies the other components of the T4S system, the VirB2 protein is the pilin, and VirB3 is the prepilin peptidase (Lai et al. 2002). The VirB5 protein localizes to the terminus of the T-pilus and is proposed to mediate interactions between the T-pilus and the target plant cell (Aly and Baron 2007). Whether this interaction can be considered a component of the surface attachment process, or rather some other aspect of the intimate associations that lead to T-DNA transfer, is unclear. UPP-dependent attachment drives polar interactions with biotic and abiotic surfaces. Multiple studies from different groups, examining protein localization using VirB T4SS protein fusions with autofluorescent proteins, have suggested a

unipolar localization for the T4SS secretion complex, including recognized components of the T-pilus (Judd et al. 2005). However, more recent work using immunolocalization has reported localization of VirB proteins in lateral arrays along the length of the cell, and this work has invoked a model in which T-DNA transfer occurs via these longitudinally associated cells (Aguilar et al. 2011). It remains unclear how to accommodate these apparently conflicting observations. It is possible that the polar attachment to surfaces (both plants and abiotic surfaces) and the attachment to plant tissues that leads to T-DNA transfer are mechanistically distinct processes. Another possible model is that there is a temporal progression from polar attachment via the UPP to deployment of the VirB T4SS machinery. This transition could include a reorientation to a longitudinal association between the bacteria and the plant cells, or simply transfer of T-DNA at the site of polar associations. The polar or longitudinal association mechanisms are not mutually exclusive. Understanding the extent to which surface attachment is integrated with T-DNA transfer remains an area of active study.

The Ti plasmid conjugative pilus is encoded by the *traI-trbABCDEFGHI* operon and is under the strict control of conjugal opines and, through these, TraR-dependent quorum sensing (Fuqua and Winans 1994; Piper et al. 1993). Thus, the Ti plasmid conjugative pilus is only produced by cells that are exposed to specific opines, and have reached a high population density. Given the strict conditionality of Trb pilus production, this structure cannot play a general role in surface interactions, but it remains uncertain whether under the appropriate conditions they might do so. In contrast, the At plasmid conjugative transfer system characterized for *A. tumefaciens* C58 is expressed constitutively in laboratory culture, and pAtC58 can be conjugatively transferred to recipients at a significant rate (Chen et al. 2002). However, there have been no reports indicating a role for these presumptive surface structures in surface attachment outside of interactions between bacterial cells. The At conjugative pili have never been visualized microscopically.

4 Exopolysaccharides

The process by which *A. tumefaciens* transitions from a planktonic, swimming cell to a sessile, surface-attached cell is determined, among other factors, by the production and subsequent extrusion of exopolysaccharides (EPS) from the cell. So far, it has been shown that *A. tumefaciens* C58 produces at least five different forms of EPS: cellulose, unipolar polysaccharide (UPP) adhesin, succinoglycan, cyclic-β-1,2-glucan, and curdlan (Berne et al. 2015; Li et al. 2012; Schmid et al. 2015; Xu et al. 2013), although it is possible that other yet-to-be-identified EPS species are produced. Not all the known EPS types, however, play roles in surface attachment. Succinoglycan, for instance, is necessary for symbiosis in *S. meliloti* (Reuber and Walker 1993) but is not required for attachment or biofilm formation in *A. tumefaciens* (Matthysse 2014; Tomlinson et al. 2010). Cyclic-β-1,2-glucan is a periplasmic polysaccharide believed to play a role in osmoregulation, and the

inability of *A. tumefaciens* to synthesize this polysaccharide results in pleiotropic effects, including increased sensitivity to osmotic stress, reduced motility, and reduced attachment. The attachment-deficient phenotype of a cyclic-β-1,2-glucan null mutant is likely indirect (Matthysse 2014), and no direct evidence implicating cyclic-β-1,2-glucan in attachment has been reported. This section will focus on the exopolysaccharides implicated in *A. tumefaciens* attachment.

4.1 Cellulose

Cellulose is an abundant crystalline polymer that is commonly found in the plant kingdom but is also produced by a broad range of bacterial species. It is common in the genera of Proteobacteria including *Komagataeibacter* (formerly *Gluconoacetobacter* and *Acetobacter*), *Azotobacter, Aerobacter, Escherichia*, *Salmonella, Rhizobium*, and *Agrobacterium* (Arioli et al. 1998; Shoda and Sugano 2005). Cellulose, a polymer of glucose joined by β (1→4) glycosidic bonds, forms fibers where the individual cellulose chains are arranged in parallel structures held together by hydrogen bonds and van der Waals forces. These fibers are water-insoluble and mechanically strong (Römling 2002). The physical properties of cellulose are illustrated by its presence in plant cell walls as large bundles of microfibrils and other higher-order structures, where it functions to determine plant cell shape and protect plant cells from osmotic stress and other environmental damage. In bacteria, however, cellulose is rarely contained in the cell wall or plasma membrane but rather is secreted outside of the cell as thin microfibril ribbons, the size of which is estimated at one-hundredth that of plant cellulose (Shoda and Sugano 2005; Williamson et al. 2002). Cellulose synthesis was first identified in *Agrobacterium tumefaciens* as thin fibrils responsible for floc formation during log-phase growth (Deinema and Zevenhuizen 1971), and then later reported to be involved in functions such as attachment and plant infection (Matthysse 1981, 1983, 1987).

4.1.1 Genetic Basis and Biosynthesis

Cellulose biosynthesis genes have been characterized in many different bacteria and are often encoded in conserved gene clusters. Core functionalities are found in all such systems, but there are other genes that are specific to certain subgroups. All genes involved in cellulose biosynthesis have recently been classified as the bacterial cellulose synthesis (*bcs*) genes (Römling and Galperin 2015). Historically, *A. tumefaciens* cellulose biosynthesis genes have been designated as *cel* genes, and for the purposes of this review, we will maintain this nomenclature but also provide the corresponding *bcs* designation (Table 2 and Fig. 3). In *A. tumefaciens* C58, cellulose synthesis is largely directed by the products of seven genes found in two presumptive operons convergent to one another on the linear chromosome (Fig. 3). The first operon is composed of five genes, *celHABCG*, and is convergent with the

Table 2 *A. tumefaciens* cellulose biosynthesis genes

Gene	Number	bp (aa)	BCS name[1]	Enzymatic function[1]	Predicted localization[1]
celH	Atu8187	900 (300)	*bcsN*	Periplasmic, single TM domain	Periplasmic
celA	Atu3309	2190 (789)	*bcsA*	Cellulose synthase, subunit A	Transmembrane (7 TMs)
celB	Atu3308	2484 (828)	*bcsB*	Cellulose synthase, subunit B	Periplasmic (SS + 1 TM)
celC	Atu3307	1056 (352)	*bcsZ*	Endo-β-1,4-glucanase, (cellulase), periplasmic	Periplasmic (SS)
celG	Atu3306	2343 (781)	*bcsK*	Tetratricopeptide (TPR) motif peptidoglycan interaction	Periplasmic (SS)
celE	Atu3305	1155 (385)	*bcsL*	Acetyltransferase (TPR)	Cytoplasmic
celD	Atu3304	1647 (549)	*bcsM*	Aminohydrolase (deacetylase?)	Cytoplasmic

[1]Based on Römling and Galperin 2015

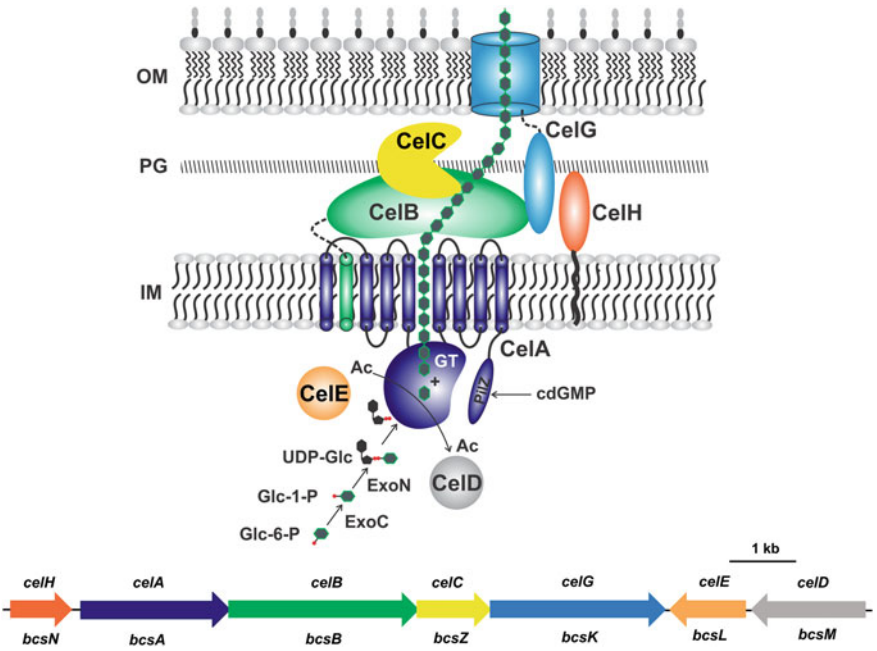


Fig. 3 Cellulose: Genetic basis and biosynthesis model. Predicted localization and mechanism of biosynthesis for cellulose in *A. tumefaciens* based on the generalized model (Römling and Galperin 2015). Cellulose strand depicted as linked green hexagons. Cel protein names indicated in the figure. For CelA, GT is the predicted glycosyl transferase domain and PilZ is the cdGMP-binding domain. Black squiggle on CelH is a predicted lipid linkage. Gene colors match protein colors in the diagram; Cel names are above with corresponding Bcs nomenclature below. OM, outer membrane; PG, peptidoglycan; IM, inner membrane, Ac, acetyl groups; Glc-6-P, glucose-6-phosphate; Glc-1-P, glucose-1-phosphate; UDP-Glc, uridyl diphosphate glucose

celDE genes (Fig. 3). The *celABC* genes have homologues in most *bcs* systems, but *celH*, *celG, celD*, and *celE* are found in a restricted subgroup designated Type IIIa (Römling and Galperin 2015).

A tentative molecular model for cellulose biosynthesis by *A. tumefaciens* can be formulated based on both limited experimental data from *A. tumefaciens* and several extensively studied systems, most notably *Komagataeibacter xylinus*. The proteins involved in cellulose synthesis are believed to form multi-protein complexes, with around 50 such complexes visible in a row along the longitudinal axis of the bacterial rod as observed by cryoEM in *K. xylinus* (formerly *Acetobacter xylinum*) (Kimura et al. 2001). Precursors for cellulose synthesis derive from the glycolytic intermediate glucose-6-phosphate (G6P), which is isomerized to glucose-1-phosphate (G1P) by a phosphoglucomutase likely to be ExoC in *A. tumefaciens* (Fig. 3). G1P is then conjugated to the nucleotide uridyl diphosphate (UDP) by a UTP–glucose-1-phosphate uridylyltransferase to form UDP-glucose (UDP-Glc), and it is likely that *A. tumefaciens* ExoN, or a paralogue, drives this reaction. CelA is predicted to be the complex cellulose synthase that utilizes UDP-Glc and adds each Glc residue to the growing cellulose molecule via a β-1,4 linkage. CelA is homologous to the cellulose synthase catalytic subunits of *Rhizobium leguminosarum bv trifolii,* and *Sinorhizobium meliloti*, and corresponds more broadly to the BcsA component of other cellulose biosynthesis systems (Römling and Galperin 2015). As such, CelA has eight transmembrane domains, a large cytoplasmic loop that comprises a glycosyl transferase (GT) domain which functions to add glucose residues to the growing glucan chain, and a C-terminal PilZ domain which regulates the enzyme (Morgan et al. 2013) (see Sect. 4.1.2). Although earlier models suggested a lipid linkage for the nascent cellulose polymer (Matthysse et al. 1995a), more recent structural work with the Type IIIa BcsA protein of *Rhodobacter sphaeroides* suggests that there is no lipid linkage and that the cellulose chain is synthesized by addition of one glucose subunit at a time to the interior end of the molecule (Morgan et al. 2013). The growing polymer is extruded into the periplasm through a channel via a ratcheting motion within the BcsA glycosyl transferase domain. CelB (BcsB) is often considered as a second non-catalytic subunit of cellulose synthase, and it is associated with CelA via a single transmembrane domain (Fig. 3). The majority of CelB is predicted to be periplasmic and plays a role in navigation of the emerging cellulose chain through the periplasm. The cellulose strand then transits the periplasm and crosses the outer membrane via a β-barrel-type protein channel; in many Bcs systems, this function is performed by the BcsC component.

In Type IIIa systems such as *A. tumefaciens,* there are no BcsC homologues and it is now thought that CelG (BcsK) plays an analogous role, as an outer membrane secretin (Römling and Galperin 2015). As with the BcsK component in other Cel systems, CelG has a large tetratricopeptide repeat (TPR) domain, considered to drive protein–protein interactions, which extends into the periplasm presumably in contact with other periplasmic Cel proteins and perhaps peptidoglycan (Fig. 3). Although the outer membrane channel for cellulose would presumably be critical for cellulose biosynthesis, one study reported that a *celG* transposon mutant

exhibited elevated cellulose production (Matthysse et al. 2005). The mechanistic basis for the surprising phenotype of this mutant, given the predicted central role for CelG, warrants further investigation. The CelC protein (BcsZ) belongs to a family of glycoside hydrolases that cleave glycosidic bonds between carbohydrates. Transposon insertions in *celC* block cellulose synthesis in *A. tumefaciens,* and in vitro cellulose synthesis experiments with *cel* mutants are consistent with a role for the *celC* product in the hydrolysis of glucose oligomers (Matthysse et al. 1995a, b). Immediately upstream of *celA* is a gene we designate as *celH* (*bcsN*), encoding a predicted periplasmic protein with a single transmembrane domain, found only in Type IIIa systems. The function of CelH remains poorly defined, but it is required for cellulose synthesis in *A. tumefaciens* (Kim and Fuqua, unpublished). Encoded in a second cellulose synthesis operon convergent to *celHABCG* are *CelE* and *CelD*, again only found in Type IIIa Bcs systems. Both CelE and CelD proteins are cytoplasmic (Matthysse et al. 1995a), and sequence similarity suggests that they may encode an acetyltransferase and a possible deacetylase, perhaps controlling the level of acetylation for cellulose precursors.

4.1.2 Regulation of Cellulose Synthesis

To date, the best understood form of cellulose regulation is allosteric, via the cdGMP second messenger and its interaction with the CelA C-terminal cytoplasmic domain, which contains a PilZ motif found in many cdGMP-responsive proteins. High intracellular levels of cdGMP stimulate cellulose production (Barnhart et al. 2013; Xu et al. 2013). The cdGMP signal molecule is now recognized as a nearly ubiquitous second messenger that controls various important cellular processes such as cell cycle progression, cell division, motility, attachment, and virulence in a wide range of bacteria (Jenal et al. 2017). In fact, cellulose synthase activity in *A. tumefaciens* was one of the first systems shown to be regulated by cdGMP (Amikam and Benziman 1989). Intracellular cdGMP levels are regulated by the activities of diguanylate cyclases (DGCs) which synthesize cdGMP from two molecules of GTP, and phosphodiesterases (PDEs) which typically break down cdGMP into pGpG molecules (Jenal et al. 2017; Römling et al. 2013) (see Sect. 6.2). A pair of recent studies have suggested that a particular DGC in *A. tumefaciens*, Atu1297 that these investigators named CelR (also called PleD), has a strong effect on *A. tumefaciens* cellulose biosynthesis (Barnhart et al. 2013, 2014). Several other *A. tumefaciens* DGCs and PDEs can also regulate cellulose biosynthesis (Feirer et al. 2015; Xu et al. 2013). Expression of the *A. tumefaciens cel* genes is hypothesized to be regulated by a MarR-type protein designated CelI that was identified in a genetic screen. A transposon mutant in *celI* overproduced cellulose, but no direct measurements were performed on *cel* gene expression (Matthysse et al. 2005).

4.1.3 Role in Attachment

An important early step in *A. tumefaciens* pathogenesis is attachment to a host plant surface. Cellulose synthesis has been reported to play a central role in stable plant attachment, although it is now clear that this cellulose-mediated attachment is not sufficient for virulence (Matthysse 1983, 1987). One early study showed that, although mutants unable to synthesize cellulose were easily washed off from the site of inoculation and failed to form large aggregates compared to the parent strain, they could still attach to carrot cells and were virulent (Matthysse 1983). Furthermore, washing the inoculation site drastically affected the ability of these mutants to form tumors compared to the parent strain, suggesting that cellulose may be important for properly anchoring *A. tumefaciens* to plant cells thereby facilitating tumorigenesis. On abiotic surfaces such as a glass or a plastic coverslip, however, cellulose-deficient mutants are fully proficient for attachment and are not easily dislodged following washing (Xu et al. 2012). This result suggests a differential role for cellulose synthesis depending on the nature of the surface being colonized. Overproduction of cellulose can, however, increase biofilm formation by *A. tumefaciens* even on abiotic surfaces (Wang et al. 2014; Xu et al. 2013). It seems reasonable to speculate that the cellulose fibers present in plant cells may form hydrogen bonds with the cellulose fibers synthesized by *A. tumefaciens* upon contact with the plant surface, and this interaction may anchor *A. tumefaciens* tightly to its host, thereby facilitating infection (Matthysse 2014).

4.2 Unipolar Polysaccharide (UPP)

Many Gram-negative bacteria can either attach to surfaces via one pole or laterally (Meadows 1971). For the well-studied pseudomonads, it is thought that the transition from reversible to irreversible attachment occurs as a switch from polar to lateral attachment (Sauer et al. 2002). However, for many bacteria, particularly those in the Alphaproteobacteria group, cells attach stably by their poles and do not readily transition to a lateral state (Li et al. 2012). Polar attachment has been observed for several Alphaproteobacteria and is often facilitated by a secreted polysaccharide-containing adhesin. The best characterized of these is the so-called holdfast localized to the end of the cellular appendage known as the stalk in the prosthecate bacterium *C. crescentus* (Bodenmiller et al. 2004; Fiebig et al. 2014; Toh et al. 2008), and also found in other members of the Order *Caulobacterales* (Berne et al. 2015; Fritts et al. 2017). Polar polysaccharides are now known to be prevalent in the order *Rhizobiales* but are best characterized in *Rhizobium leguminosarum* and *A. tumefaciens*. This structure is localized to a single, consistent pole (the old pole of the daughter after cell septation) and is therefore designated as the unipolar polysaccharide (UPP) in *Agrobacterium tumefaciens* (Gu et al. 2011; Laus et al. 2006; Tomlinson and Fuqua 2009).

4.2.1 Composition

UPP-type polysaccharides are often chemically distinct among members of the *Rhizobiales*, yet they share similar genetic, biosynthetic, and functional characteristics as highlighted in subsequent subsections. The chemical composition of most alphaproteobacterial UPP-type polysaccharides is not clear, partly due to insolubility as well as limitations in the quantity synthesized by the bacteria (Berne et al. 2015; Tomlinson et al. 2010). However, the UPP-type polysaccharide of *R. leguminosarum*, which functions in host legume association through binding by a host-specific lectin, is composed mostly of mannose and glucose and has thus been designated a glucomannan (Laus et al. 2006). Even with the intensely studied *C. crescentus* holdfast, the only information on polysaccharide chemistry is that it includes *N*-acetylglucosamine (GlcNAc), through its recognition by the GlcNAc-specific lectin Wheat Germ Agglutinin (WGA) (Berne et al. 2015). The *A. tumefaciens* UPP is not only recognized by the WGA lectin, but also by the *Dolichos biflora* (DBA) lectin which specifically binds *N*-acetylgalactosamine (Heindl et al. 2014; Tomlinson et al. 2010; Xu et al. 2012). Notwithstanding, it is probable that other sugars are present in the UPP. *Rhodopseudomonas palustris* also produces a UPP recognized by WGA (Fritts et al. 2017). Chemical analyses, including mass spectrometry and NMR of extracted UPP-type polysaccharides, should help to determine the chemical composition of these exopolysaccharide adhesins.

4.2.2 Genetic Basis and Biosynthesis

In *C. crescentus,* holdfast production requires proteins encoded by *h*old*f*ast *s*ynthesis (*hfs*) genes in a genomic cluster comprising *hfsEFGHCBAD* (Berne et al. 2015; Toh et al. 2008). Several homologues and analogues of these proteins have been identified in *A. tumefaciens.* Notably, *uppE and uppC* of *A. tumefaciens* have sequence similarity to *hfsE* and *hfsD*, respectively (Fritts et al. 2017). The *uppE* and *uppC* genes are contained in a cluster on the *A. tumefaciens* circular genome with four other genes designated *uppA, uppB, uppD,* and *uppF* spanning loci Atu1235-1240 (Fig. 4; Fritts et al. 2017; Xu et al. 2012). Deletion of the entire *uppA-F* gene cluster, as well as individual genes such as *uppE*, abolishes UPP production in *A. tumefaciens.* This effect can be observed as a lack of binding to surfaces, and hence a lack of biofilm production (Xu et al. 2012, 2013). Consistent with its role in attachment, mutations which cause UPP overproduction greatly enhance attachment and biofilm formation (Feirer et al. 2015, 2017; Wang et al. 2014; Xu et al. 2013).

The occurrence of a *upp*-type gene cluster in members of the *Rhizobiales* appears to be common among plant-associated microbes, as several plant symbionts such as *R. leguminosarum, B. japonicum*, and *S. meliloti* possess highly conserved *upp* clusters (Fritts et al. 2017). In *R. leguminosarum,* these genes are designated *gms* because mutation of one of these genes, *gmsA* (homologous to *uppE*),

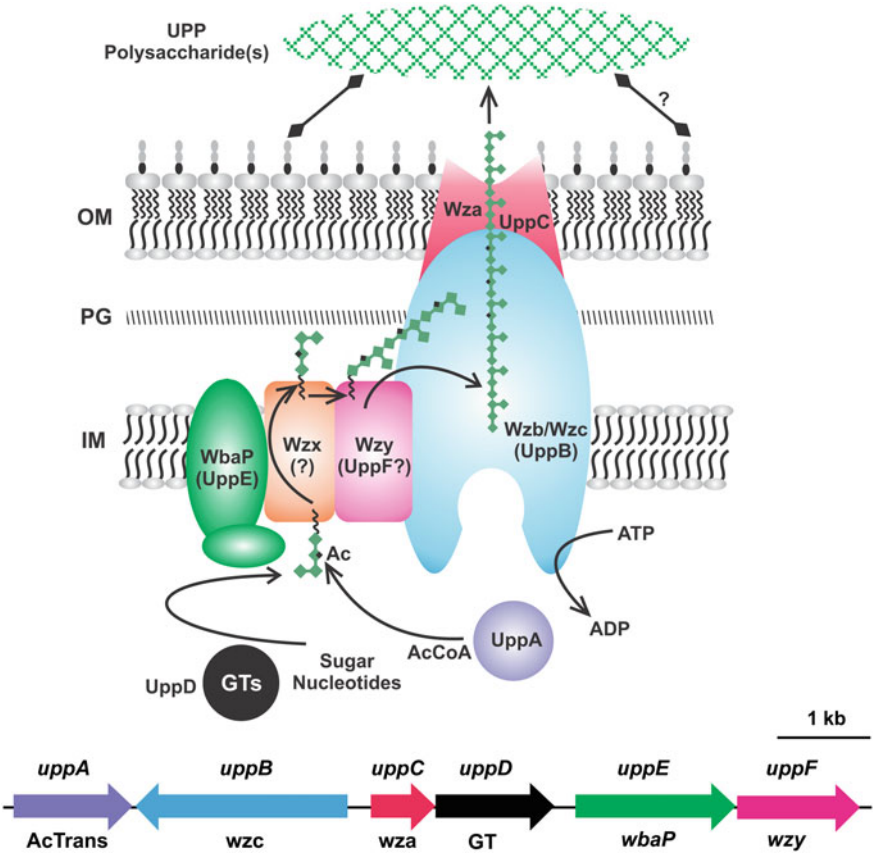


Fig. 4 Wzx/Wzy polysaccharide pathway: Genetic basis and biosynthesis model for UPP. Predicted localization and mechanism of biosynthesis for UPP in *A. tumefaciens* based on Wzx–Wzy generalized model (Cuthbertson et al. 2009). Branched green filament of diamonds is meant to depict the polysaccharide strand. The black square on some residues is putative acetylation. The Wzb/c protein may hydrolyze ATP to power polymerization. Upp protein names and the general Wzx–Wzy components are indicated in the figure. Black squiggle on the polysaccharide subunit is the undecaprenyl phosphate. Gene colors match protein colors in the diagram; upp gene names above and corresponding Wzx–Wzy components below (AcTrans—acetyltransferase; GT glycosyl transferase). Assembled UPP structure is represented by green cross-hatched oval outside cell, and the diamond headed lines are putative linkages to the cell pole surface. OM, outer membrane; PG, peptidoglycan; IM, inner membrane

abolishes synthesis of the unipolar glucomannan (Laus et al. 2006). In addition, free-living members of the *Rhizobiales* such as *R. palustris* possess a *upp* cluster closely similar to that of *A. tumefaciens* (Fig. 4). Certain alphaproteobacterial mammalian pathogens also have *upp*-type genes, including *Ochrobactrum anthropi* ATCC 49188 with a conserved and syntenous set of all six *upp* genes, and *Brucella melitensis* and *B. suis* that have a cluster homologous to three of the six *upp* genes.

Conserved *upp* genes appear to be lacking in some obligate intracellular pathogens of the *Rhizobiales*, suggesting a loss of the *upp* cluster following adaptation to the intracellular lifestyle. Although they are required to synthesize a polar polysaccharide and are hypothesized to share a common ancestor with the *hfs* genes of *C. crescentus*, it is noteworthy that the *upp*-type genes are distinct and broadly distributed among the *Rhizobiales* (Fritts et al. 2017).

Apart from the genes that make up the core *hfs* cluster in *C. crescentus*, other genes, present elsewhere in the genome, can also play important roles in holdfast biosynthesis. The *pssY* and *pssZ* genes are paralogues of *hfsE,* and single deletion mutants for each gene synthesized holdfasts comparable to those of wild-type. However, a triple deletion mutant is completely deficient for attachment and holdfast synthesis (Toh et al. 2008). Furthermore, *hfsC* deletion results in a holdfast synthesis defect when combined with a deletion of its paralogue, *hfsI,* but not when each gene is singly deleted. In *A. tumefaciens*, the *uppE* gene is required for UPP production and attachment under phosphorus (P_i)-replete conditions, but under P_i limitation *uppE* is functionally redundant with Atu0102, an orthologue of *pssY* (Xu et al. 2012). Disruption of both genes is required to prevent UPP production and attachment in P_i limitation. It is possible that additional core *upp* genes exist elsewhere in the *A. tumefaciens* genome. Preliminary evidence reveals that there are other genes outside of the *A. tumefaciens uppABCDEF* cluster required for UPP production (Natarajan et al., in preparation).

Bacterial polysaccharides are synthesized via one of three main pathways, largely differentiated by membrane topology, characteristic components, and the type of polysaccharides synthesized. These pathways are the Wzx/Wzy-dependent pathway, the ABC-transporter-dependent pathway, and the synthase-dependent pathway. Both the ABC-transporter- and synthase-dependent pathways synthesize mainly homopolymeric exopolysaccharides with single-sugar repeating units (with some exceptions such as certain LPS pathways), and use a mechanism that largely involves complete polysaccharide synthesis in the cytoplasm before export out of the cell, (Mi et al. 2017). In contrast, the Wzx/Wzy-dependent pathway more commonly synthesizes heteropolymers with repeating units of three to six sugars. These repeating units are polymerized in the periplasm before final export out of the cell (Islam and Lam 2013, 2014; Schmid et al. 2015).

The holdfast synthesis genes of *C. crescentus*, as well as the *upp*-type genes from *A. tumefaciens and R. palustris*, are homologous to a Wzx/Wzy-dependent polysaccharide biosynthesis pathway (Toh et al. 2008; Fritts et al. 2017; Heindl et al. 2014). The Wzx/Wzy-dependent pathway is named after the integral, inner-membrane oligosaccharide-transferase protein identified in several gram-positive and gram-negative bacteria responsible for O-antigen assembly, encoded by the *wzy* gene (Islam and Lam 2013, 2014; Kalynych et al. 2014). This pathway consists of cytoplasmic glycosyltransferases (GTs) which add sugar nucleotide precursors to a polyisoprenoid lipid carrier molecule, undecaprenyl phosphate (UndP), which is embedded within the cytoplasmic leaflet of the inner membrane (Fig. 4). The first sugar is added to UndP by the initiating GT (WbaP), a polyisoprenylphosphate hexose-1 phosphate (PHPT), to form a pyrophosphate

linkage with the carrier (UndPP) (Cuthbertson et al. 2009). For holdfast synthesis in *C. crescentus,* this initiating PHPT activity is performed by the GT HfsE (encoded within the *hfs* gene cluster), PssY, or PssZ (Toh et al. 2008). In *A. tumefaciens* the PHPT is UppE, with the Atu0102 PssY homologue functioning in the same capacity under P_i limitation (Fig. 4; Xu et al. 2012). Following addition of the first sugar moiety, subsequent sugars are added by additional GTs until a complete repeating polysaccharide unit is formed. This molecule is then flipped from the cytoplasmic face of the inner membrane to its periplasmic face via a Wzx translocase, also called a flippase (Fig. 4). Flippases of this type are integral membrane proteins, usually with twelve transmembrane domains. Wzx flippases belong to the polysaccharide transport (PST) family of the larger multidrug/oligosaccharidyl-lipid/polysaccharide (MOP) flippase superfamily (Hong and Reeves 2014). A flippase homologue is encoded within the core *hfs* gene cluster of *C. crescentus* (HfsF), but such a homologue is absent from the UPP-type gene cluster of the *Rhizobiales* (Fritts et al. 2017; Toh et al. 2008). Despite the vast array of O-antigen and other polysaccharides synthesized through this pathway, several studies on flippases indicate a strong specificity for their substrates, which may explain the high sequence variability among Wzx proteins in different organisms and even within a single organism (Islam and Lam 2013, 2014; Wang et al. 2012). Polymerization of the flipped repeating units occurs in the periplasm via the Wzy protein in conjunction with the chain-length determining Wzc protein (sometimes in conjunction with a Wzb-type protein), also called the polysaccharide co-polymerase or PCP (Fig. 4). In *C. crescentus* there are two Wzy-type proteins, HfsC and HfsI, that are functionally redundant for holdfast synthesis (Hardy et al. 2018). The PCP function is thought to be split into two proteins, HfsA (Wzc) and HfsB (Wzb). For the UPP-type systems in the *Rhizobiales,* there are no clear Wzy-type proteins within *upp* cluster, although UppF is annotated as an O-antigen ligase, which could plausibly drive a similar polymerization reaction. The UppA protein is a PCP homologue, but mutations in this gene result in only modest effects on UPP production in *A. tumefaciens* (Natarajan et al., in prep). There are multiple unassigned Wzx-type, Wzy-type, and Wzz-type proteins encoded in the *A. tumefaciens* genome that may contribute to UPP biosynthesis. The polymerizing polysaccharide strand is then exported across the outer membrane via a secretin, a β-barrel protein designated Wza or OPX for outer membrane polysaccharide export (Cuthbertson et al. 2009; Islam and Lam 2014). In *C. crescentus,* this protein is annotated HfsD, and in the UPP-type systems this is annotated UppC. Mutations in these genes result in abolishment of polar polysaccharide synthesis (Berne et al. 2015; Natarajan et al. in prep).

Stable association with the pole of the cell is one of the properties that distinguishes polar polysaccharides from capsular or secreted exopolysaccharides. In *C. crescentus*, the holdfast attachment (*hfa*) genes, arranged in a three gene operon, *hfaABD,* encode proteins that function in anchoring the holdfast to the *C. crescentus* stalk (Cole et al. 2003; Hardy et al. 2010). Independent mutants in several *hfa* genes shed their holdfasts from cells and fail to attach to abiotic surfaces. HfaA is a β-amyloid protein similar to curlin from *E. coli* (Blanco et al. 2012). HfaA export requires the outer membrane lipoprotein HfaB and the outer membrane

protein HfaD to which it is thought to anchor. HfaA forms a structure analogous to a β-amyloid filament that tethers the holdfast to the stalk tip (Hardy et al. 2010). Whereas *hfs* and *hfa* genes are well conserved within the *Caulobacterales,* no *hfa* homologues have been identified in the *A. tumefaciens* genome or in the genomes of most other *Rhizobiales* (Fritts et al. 2017). There are several uncharacterized proteins that are predicted to adopt a β-amyloid fold, but their role, if any, is unknown. It is also plausible that the UPP is anchored to another outer membrane structure such as LPS, but it is not clear how this would be specifically retained at the pole.

4.2.3 Regulation of UPP Synthesis

Thus far, no differential gene expression for the *uppABCDEF* cluster has been observed, even in mutants or under conditions which strongly promote or inhibit UPP production (Xu et al. 2013). Therefore, most of the regulatory effects observed are hypothesized to be post-transcriptional, and likely through allosteric control of biosynthesis.

Holdfast synthesis is intricately linked to cell cycle progression in the dimorphic life cycle of *C. crescentus*, where the holdfast is elaborated at the end of the stalk during cell development (Toh et al. 2008). Interactions with the surface can accelerate production of the holdfast (Li et al. 2012). In contrast to *C. crescentus,* UPP production in *A. tumefaciens* is only observed upon surface or cell–cell contact (Li et al. 2012; Xu et al. 2013). Following attachment, UPP is visible very rapidly after surface contact. Surface-contact-dependent stimulation of polar polysaccharide secretion has also been observed for other Alphaproteobacteria such as *Asticcacaulis biprosthecum* (Li et al. 2012). The mechanism of surface-contact recognition in *A. tumefaciens* is an area of active study.

Among the widespread roles of cdGMP in many bacteria is the regulation of polysaccharide production such as cellulose (Sect. 4.1.2) and UPP (Jenal et al. 2017; Xu et al. 2013). Ectopic expression of the *A. tumefaciens* DGC PleD in *A. tumefaciens* resulted not only in increased cell–cell aggregation and biofilm formation, but also a decoupling of UPP synthesis from surface contact (Xu et al. 2013). These observations suggest that elevated intracellular cdGMP levels bypass the surface-contact-dependent requirement for UPP production, and is evidence for the role of cdGMP in UPP regulation. The *A. tumefaciens* genome encodes over 30 DGCs that regulate intracellular cdGMP levels, and several of these have been directly implicated in control of UPP production (Heindl et al. 2014) (see Sect. 6.2). A number of mutants identified for dysregulated UPP production have been isolated, and thus far all of these appear to connect in some way to cdGMP pools (Feirer et al. 2015; Feirer et al. 2017; Xu et al. 2013). Several environmental conditions which modulate UPP-dependent attachment include P_i limitation (increasing attachment) and limitation for the divalent cations Fe^{2+} and Mn^{2+} (decreased attachment) (Danhorn et al. 2004; Heindl et al. 2016; Xu et al. 2012). Interestingly, UPP production and cellulose synthesis are often, but not always,

co-regulated by specific cdGMP-dependent pathways (Wang et al. 2016; Xu et al. 2013). The integration of control for these two important polysaccharides during surface colonization is a subject of active investigation.

4.2.4 Role in Attachment

Biofilm formation on biotic and abiotic surfaces in *A. tumefaciens* follows a stepwise trend beginning with reversible attachment during which the organism must overcome repulsive electrostatic and hydrodynamic forces near the surface, followed by a more permanent attachment largely mediated by UPP (Heindl et al. 2014). In fact, surface association stimulates UPP production very rapidly, and its deployment is coincident with stable surface attachment (Li et al. 2012). Mutations of key *upp* genes such as *uppE* and *uppC*, as well as deletion of the *uppABCDEF* cluster, abolishes binding to abiotic surfaces (Xu et al. 2012) and results in loose aggregate formation and weak binding to plant tissues (Natarajan et al., manuscript in preparation), suggesting its importance in permanent, irreversible attachment. In *C. crescentus* and its close relatives, studies have shown that the cohesive and adhesive properties of the holdfast are determined by the degree and pattern of acetylation of the polysaccharide, regulated by the putative deacetylase HfsH (Wan et al. 2013). A putative acetyltransferase encoding gene, *uppA*, may play a similar role in *A. tumefaciens*, but this is yet to be rigorously tested. Mutations in genes that block UPP production retain virulence (Natarajan et al. in preparation; Gelvin et al., unpublished), as evaluated by assays that directly inoculate wounded plant tissues, suggesting that UPP-mediated attachment per se is not required for T-DNA transfer to plants. This is consistent with the chromosomal location of the *upp* genes (Xu et al. 2012), their activity in avirulent, Ti-plasmidless agrobacteria, and their conservation among diverse *Rhizobiales* (Fritts et al. 2017).

4.3 Curdlan

Many species and strains of *Agrobacterium* synthesize the β-1,3 glucan called curdlan, and several species are the industrial source for this polysaccharide, which is used as a cement additive. There is indirect evidence that curdlan may have the capacity to function in attachment.

4.3.1 Composition

Curdlan is a neutral, water-insoluble exopolysaccharide that is one of three structural classes of (1→3)-β-glycosidic linkages of glucose. These classes are the linear (1→3)-β-D-glucans, the side-chain-branched (1→3, 1→2)-β-glucans, and the cyclic (1→3, 1→6)-β-glucans. Curdlan is mostly linear, but may have a few intra- and

inter-chain (1→6)-linkages (McIntosh et al. 2005; Zhan et al. 2012). A single curdlan molecule may have as many as 12,000 glucose units (McIntosh et al. 2005), displaying an average molecular weight of 5.3×10^4–2.0×10^6 Da when dissolved in alkaline solutions such as 0.3 N NaOH. Curdlan was first identified in *Alcaligenes faecalis* var. *myxogenes* (now *Agrobacterium* sp. ATCC31749) and has since been identified in other *Agrobacterium* and some *Rhizobium* species (Nakanishi et al. 1976).

4.3.2 Genetic Basis and Biosynthesis

The molecular genetics of curdlan biosynthesis has been studied predominantly in *Agrobacterium* sp. ATCC31749 (Zhan et al. 2012). Isolation of curdlan-deficient transposon mutants of *Agrobacterium* sp. ATCC31749 resulted in the identification of curdlan synthesis genes. These genes include *crdA, crdS, crdC*, and *crdR* (Karnezis et al. 2003; Ruffing et al. 2011; Stasinopoulos et al. 1999), which are necessary for curdlan synthesis, as well as pss_{AG} (Karnezis et al. 2002), which enhances curdlan production. The *crdASC* genes form a cluster and are likely to be co-transcribed, whereas *crdR* and pss_{AG} occur elsewhere in the genome (McIntosh et al. 2005). The *crdS* gene encodes a 540 amino acid integral membrane protein of the HasA family of β-glycosyl transferases, with seven transmembrane helices and a large intracellular hydrophilic region (Karnezis et al. 2003; Stasinopoulos et al. 1999). CrdS has high sequence similarity to BcsA, the bacterial cellulose synthase, and thus is almost certainly the curdlan synthase (Stasinopoulos et al. 1999). However, CrdS lacks the C-terminal PilZ cdGMP-binding domain present in CelA, nor does it have any other known c-di-GMP-binding domains (Barnhart et al. 2013). In contrast to *crdS, crdA* and *crdC* have no known counterparts in the sequence databases. The *crdA* gene is predicted to encode a membrane-anchored protein that may play a role in transfer of the polymer across the cytoplasmic membrane, whereas *crdC* is predicted to encode a periplasmic protein which may function in passage of the polymer through the periplasm. However, *crdC* mutants still make curdlan, whereas *crdA* and *crdS* mutants do not (McIntosh et al. 2005). CrdA, CrdS, and CrdC may form a multi-protein complex spanning the inner membrane and the periplasm (Karnezis et al. 2003). The pss_{AG} gene encodes a membrane protein that plays a role in phosphatidylserine production and functions to enhance curdlan production (Karnezis et al. 2002). Phosphatidylserine is a membrane phospholipid as well as a precursor for phosphatidylethanolamine, suggesting a role for phospholipid composition in increased curdlan biosynthesis. No outer membrane secretin has been identified to function in extracellular export of curdlan. It is noteworthy that although *A. tumefaciens* C58 has the complete set of genes for its synthesis, curdlan is not detectably produced under standard growth conditions (Matthysse 2014).

The model for the biosynthesis of curdlan is thought to resemble that of cellulose, with UDP-Glc, derived from G6P, as the precursor for the polysaccharide synthesis (Fig. 3; Stasinopoulos et al. 1999). Polymerization presumably occurs

within the GT catalytic domain on the large cytoplasmic loop of CrdS followed by extrusion of the polysaccharide from the cell. The details of the curdlan export mechanism remain to be defined.

4.3.3 Regulation of Curdlan Synthesis

Curdlan biosynthesis is transcriptionally activated by the helix–turn–helix protein CrdR (Yu et al. 2015). Transcriptome profiling of *Agrobacterium* ATCC31749 to investigate regulation of curdlan synthesis showed a 100-fold increase of *crdASC* expression under nitrogen limiting conditions (Ruffing and Chen 2012). This regulation occurs through the sensor kinase NtrB and its response regulator NtrC, global regulators of nitrogen metabolism (McIntosh et al. 2005). Deletion of the sigma factor *rpoN*, often associated with nitrogen limitation responses, led to elevated levels of curdlan synthesis (Ruffing and Chen 2012). Transcriptional profiling under nitrogen-limited conditions in *Agrobacterium* sp. ATCC31749 also revealed increased expression of a pair of DGC proteins, predicted to impact cdGMP pools (Ruffing and Chen 2012). Mutation of one of these DGCs (AGRO_3967) resulted in markedly decreased curdlan production (Ruffing and Chen 2012). The CrdS protein lacks a PilZ domain that would impart cdGMP-responsiveness, but there are other possible inputs of cdGMP into curdlan biosynthesis (Barnhart et al. 2013).

4.3.4 Role in Attachment

Presently, curdlan has not directly been shown to play a role in the attachment of cells to biotic or abiotic surfaces, although it seems a likely candidate. There is no published evidence that mutations which impact curdlan production in *A. tumefaciens* C58 affect virulence or root colonization. The curdlan over-producing strain ATCC31749 forms a fragile, easily detachable blanket-like structure when incubated with tomato roots or on agar plates containing Aniline Blue (Matthysse 2014).

5 Protein Adhesins

Many bacteria externalize specialized adhesin proteins, distinct from pili, that promote surface interactions. These include the large adhesin proteins (Laps) from *Pseudomonas* (Hinsa et al. 2003), the filamentous hemagglutinin (FHA) proteins originally identified in *Bordatella* species (Locht et al. 1993), and β-amyloid curli fibers from uropathogenic *E. coli* (Barnhart and Chapman 2006; Kai-Larsen et al. 2010). Proteinaceous adhesins have also been implicated in attachment of rhizobia, including *A. tumefaciens*, to root tissues. Early studies on the attachment of rhizobia

to pea root hairs under conditions with high calcium ion concentrations identified a Ca^{2+}-binding protein defined as rhicadhesin, and addition of a crude preparation of this protein to live rhizobia diminished their attachment (Smit et al. 1989). Similar activities were characterized in extracts of *A. tumefaciens* (Dardanelli et al. 2003). Cell surface proteins defined as the Rap (*Rhizobium*-adhering proteins) adhesins were identified from a subset of rhizobia (Ausmees et al. 2001; Russo et al. 2006) and may be related to the earlier defined rhicadhesin activity. The predicted genes for these proteins are not conserved within the *A. tumefaciens* genome. It remains unclear what encodes the rhicadhesin activity initially detected in *A. tumefaciens* and how this putative activity relates to attachment. *A. tumefaciens* C58 does have a small cluster of genes that encode proteins that are predicted to adopt β-amyloid structures due to their similarity with the CsgA protein of *E. coli* (Barnhart and Chapman 2006), although their role in attachment, if any, remains unexplored.

6 Regulation of Attachment

The transition from a motile state to a sessile growth mode is a crucially important process for many bacteria, and for *A. tumefaciens* this is the case under a variety of contexts. As such, the attachment process is under complex regulatory control at multiple levels. The individual surface features are themselves often under complex transcriptional control networks. Some of the relevant expression control pathways represent global transcription regulatory networks, such as the response to decreasing pH through the RGI pathway (Heckel et al. 2014), or phosphorus (P_i)-limitation via the PhoR-PhoB system (Danhorn et al. 2004). Other control circuits are more specific to a given process, such as the VisR-VisN-Rem regulation of flagellar and chemotaxis gene expression, target functions that can also impact attachment processes (Xu et al. 2013). Beyond simply identifying transcriptional control systems that influence attachment, the understanding of how these regulatory circuits are integrated leading up to and during intimate interactions with surfaces remains rudimentary. In addition to the variety of different transcriptional pathways, post-transcriptional mechanisms, particularly allosteric control of polysaccharide biosynthesis, exert a major influence on attachment.

6.1 Transcriptional Regulators that Impact Attachment

Genetic analysis in *A. tumefaciens* C58 has identified several different transcriptional regulators, mutations in which cause significant changes in attachment proficiency or biofilm formation.

6.1.1 ExoR-ChvG-ChvI (RGI) Pathway

Transposon mutations in the *exoR* regulatory gene result in profound loss of attachment. As described above (Sect. 2.1), ExoR is not itself a transcription factor, but rather a periplasmic regulator of the ChvG sensor kinase, which in turn controls the downstream response regulator ChvI, altogether comprising the RGI pathway (Heckel et al. 2014). Under neutral conditions ExoR maintains ChvG in the inactive form. At low pH ExoR is proteolytically cleaved, resulting in active ChvG, and leading to phosphorylation of ChvI (Wu et al. 2012). Mutations in *exoR* lead to high level phosphorylation of ChvI, and differential expression of hundreds of *A. tumefaciens* genes (Heckel et al. 2014). Null mutants for *exoR* are non-motile, hypermucoid, and unable to attach to surfaces. Strikingly, the attachment deficiency is not due to the loss of motility or hypermucoidy (Tomlinson et al. 2010). Among the genes under *exoR* control are a number that may potentially impact attachment, including genes directing motility and chemotaxis (decreased), succinoglycan biosynthesis (increased), a Type VI secretion system (T6SS) (increased), plasmid conjugation (decreased), cdGMP synthesis that would affect dcGMP (decreased), and Tad-type pili (decreased). However, no single class of genes is responsible for the severe attachment deficiency of the *exoR* mutant, and this phenotype may be due to the combinatorial effects on these genes. It is also unclear which target genes ChvI regulates directly, with the exception of the T6SS genes, a promoter to which ChvI has been shown to bind in vitro (Wu et al. 2012).

6.1.2 Additional Environmentally Responsive Pathways

Ectopic expression of the response regulator PhoB simulates the Pi limitation response and increases biofilm formation in *A. tumefaciens* C58, consistent with the observation that limiting P_i has the same effect (Danhorn et al. 2004). PhoR and PhoB comprise a P_i-responsive two-component system that has been intensely studied in several different systems, including the rhizobia (McDermott 2000). The PhoB-regulated target genes that lead to increased attachment of *A. tumefaciens* in limiting Pi are not yet known, but it is clear that P_i limitation increases the number of cells which attach at their pole via the UPP adhesin (Xu et al. 2012). It is possible that this stimulatory effect is mediated through cdGMP pools.

A genetic screen for attachment deficiencies in *A. tumefaciens* C58 led to the identification of a regulatory gene that was designated *sinR* (surface interaction Regulator), mutations in which resulted in less dense and thinner biofilms than wild-type on abiotic surfaces (Ramey et al. 2004). *A. tumefaciens* SinR is a transcription factor of the CRP-FNR superfamily and is one of four FNR-type regulators encoded by the *A. tumefaciens* C58 genome. The canonical FNR regulator characterized in *E. coli* has an N-terminal domain associated with an iron–sulfur (4Fe-4S) cluster via a conserved set of cysteine residues, that is responsive to oxygen levels, and a C-terminal DNA-binding domain (Körner et al. 2003). FNR is activated under limiting oxygen, and its regulon is composed of many different

genes that are differentially expressed under these conditions. As with several other subgroups of regulators in the FNR family, SinR does not have the conserved Cys residues and as such is not thought to contain an iron–sulfur cluster. Transcription of the *sinR* gene is strongly activated under oxygen limitation and upregulated in mature biofilms, and this is dependent on FnrN, another *A. tumefaciens* FNR-type protein. In contrast to SinR, FnrN has the conserved Cys residues (Ramey et al. 2004). Ectopic expression of *sinR* increases the density of *A. tumefaciens* biofilms relative to those of wild-type. It is hypothesized that as surface-adherent biomass increases, oxygen limitation is sensed through FnrN, which stimulates *sinR* expression, thereby controlling late stages of biofilm formation. The SinR-regulated functions that impart these effects on biofilm formation have not been identified.

Analysis of the genome sequence of plant pathogen *Xylella fastidiosa* identified a transcriptional regulator eventually designated BigR (Biofilm growth-associated Repressor) within a cluster of genes predicted to be involved in sulfur metabolism, including hydrogen sulfide detoxification (Barbosa and Benedetti 2007; Guimaraes et al. 2011). The regulator and most of the presumptive operon are conserved in *A. tumefaciens* C58 (Atu3465-68). Based on its phenotypes in *X. fastidiosa,* the *bigR* homologue (Atu3466) of *A. tumefaciens* was disrupted, and the mutant had elevated expression of this operon and increased attachment on abiotic and biotic surfaces (Barbosa and Benedetti 2007). BigR is a member of the ArsR/SmtB family of repressors and has a pair of redox-reactive Cys residues. Oxidation of these Cys residues and their resulting disulfide linkage cause changes in the protein that lead to derepression of the promoter upstream of Atu3465. This promoter is also active in *A. tumefaciens* biofilms, and it is speculated that BigR is oxidized by unknown intermediates of sulfur metabolism under the oxygen limitation in the biofilm. Expression of the *bigR* operon leads to the detoxification of hydrogen sulfide and other growth inhibitory sulfur metabolites (Guimaraes et al. 2011). It is not known whether BigR regulates any additional genes outside of its own operon.

6.2 *Cyclic Diguanylate Monophosphate-Dependent Regulation of Surface Attachment Functions*

Multiple lines of evidence have indicated that cdGMP levels play a profound role during the transition of *A. tumefaciens* from the free-living to the surface-attached state. Early studies by Benzimann and colleagues suggested that *A. tumefaciens* cellulose synthase is cdGMP-regulated (Amikam and Benziman 1989), and more recent work has shown that both cellulose and UPP production are controlled in parallel to regulate the motile-to-sessile switch (Xu et al. 2013). Over the last two decades, and particularly emerging in the last ten years, the pervasive role for cdGMP in modulating bacterial physiology and behavior has gained recognition (Wolfe and Visick 2010). Diguanylate cyclase enzymes (DGCs) synthesize cdGMP from two GTP molecules via a catalytic site that contains the GGDEF motif

(sometimes GGEEF). Specific cdGMP phosphodiesterases (PDEs) typically hydrolyze the signal into pGpG, which then rapidly degrades to GMP. PDEs predominantly have an EAL catalytic motif, but a small fraction of cdGMP PDEs have an HD-GYP catalytic motif (Jenal et al. 2017). There are additional residues outside the catalytic site that define the DGC and PDE domains. It is common to identify single proteins with both DGC (GGDEF) and PDE (EAL/HD-GYP) domains. For a significant subset of these dual domain proteins, one of the two domains is degenerate and non-catalytically active. For a small fraction of these proteins, both domains are active and can catalyze synthesis of cdGMP, as well as its turnover. The bias in the ratio between DGC and PDE activity can be under regulatory control, comprising an environmentally responsive cdGMP module. In fact, many DGC, PDE, and DGC-PDE proteins have additional domains that impart environmentally responsive control of their catalytic activities, and it is common that these proteins contain transmembrane domains. Certain bacteria can have as few as a single DGC protein (e.g., *Yersinia pestis*), and others may have as many as 60 different proteins (e.g., *Vibrio* species) with DGC domains, with or without PDE domains. The integration of cdGMP signaling by multiple synthesis and degradation proteins in a single bacterial cell is a topic garnering significant attention, and current models suggest a complex combination of (i) conditional expression of specific DGC, PDE, and DGC-PDE genes; (ii) differential allosteric control of the proteins under varying environmental conditions; (iii) rapid fluxes of the signal across the cell by localized synthesis and degradation; and (iv) in some cases, compartmentalization. As might be expected with such a pervasive regulator, there are also multiple mechanisms by which cdGMP can control output functions in the cell. At least seven structurally distinct cdGMP-binding motifs have been identified. The first identified were so-called PilZ domains, but there are now five other documented protein-binding motifs as well as a cdGMP-responsive riboswitch (Jenal et al. 2017). Several of these may function simultaneously in the same bacterium.

A. tumefaciens was the second bacterial species for which cdGMP control was reported for cellulose biosynthesis (Amikam and Benziman 1989), after its initial discovery in the bacterium now called *K. xylinus* (Ross et al. 1987). It is now clear that cdGMP regulates the cellulose synthase enzyme CelA in *A. tumefaciens*, very likely through allosteric interactions with its conserved PilZ domain. In addition, cdGMP exerts profound control over UPP biosynthesis. Mutations that increase cdGMP levels stimulate UPP production, and those that decrease cdGMP can diminish UPP-mediated attachment (Xu et al. 2013). None of the proteins known to be required for UPP biosynthesis have a recognizable PilZ domain or cdGMP-binding motif (Xu et al. 2012). Production of cdGMP by *A. tumefaciens* depends on a suite of proteins with DGC domains. The *A. tumefaciens* C58 genome encodes 29 proteins with a conserved GGDEF (or GGEEF) catalytic motif (Fig. 5a), and 13 of these also contain an EAL motif (Fig. 5b). Two proteins are presumptive solo PDEs, one with an EAL motif, and a second with an HD-GYP motif, in both cases not associated with a DGC domain. There is experimental evidence in *A. tumefaciens* implicating four of the DGCs (Atu1257, Atu1297, Atu1691, and Atu2179) and one DGC-PDE protein (Atu3495) in the control of

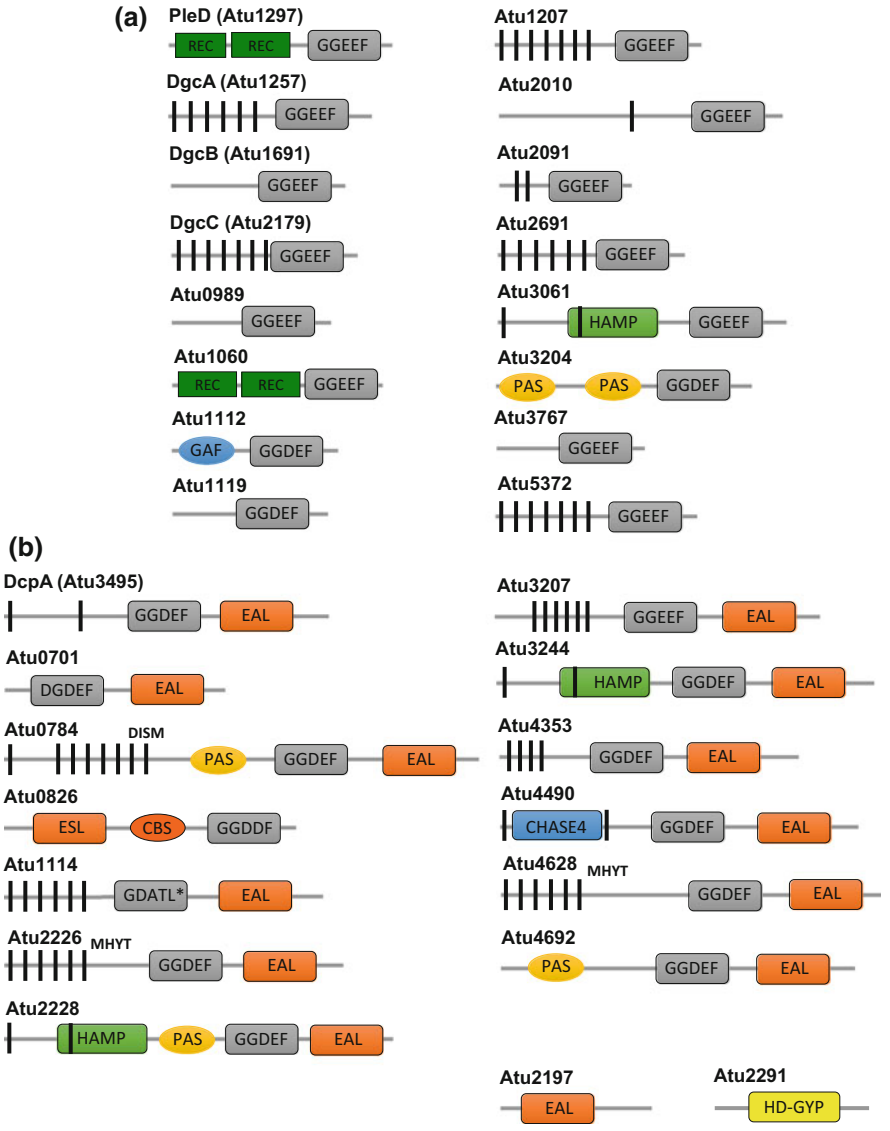


Fig. 5 Diguanylate cyclases and phosphodiesterases of *A. tumefaciens*. **a** DGC proteins, **b** DGC-PDE proteins and PDE/HD-GYP only proteins. Atu gene numbers and if assigned, genetic names are provided for each protein. Predicted transmembrane domains are indicated by vertical lines and any recognizable functional domains are indicated. Protein domains. REC—two-component-type receiver domain; GAF—cGMP-specific phosphodiesterases, Adenylyl cyclases and FhlA; HAMP—Histidine kinases, Adenyl cyclases, Methyl-accepting proteins, and Phosphatases; PAS—Per, ARNT, Sim; CBS—Cystathionine-Beta-Synthase; CHASE4, predicted ligand-binding module; MHYT, 6-transmembrane domain; DISM, 7-transmembrane domain, Diverse Intracellular Signaling Modules

polysaccharide production and attachment. Additionally, a gene from *Agrobacterium* sp. ATCC 31749 (AGRO_3967) homologous to the DGC homologue Atu1119 (Fig. 5A) is co-regulated with β-1,3 glucan (curdlan) production, and its mutation diminishes production of this polysaccharide (Ruffing and Chen 2012). A recent characterization of cdGMP-related proteins in the related bacterium *S. meliloti* provides a useful comparison for *A. tumefaciens*. *S. meliloti* has six presumptive DGC proteins, 12 DGC-PDEs, and two PDEs (Schäper et al. 2016). A subset of 13 of these proteins have likely orthologues in *A. tumefaciens*, whereas seven are not found in *A. tumefaciens*. Conversely, 16 of the *A. tumefaciens* cdGMP-related proteins are not found in *S. meliloti*. Several of the *S. meliloti* proteins have roles in attachment and biofilm formation, but none are as striking in their phenotypes as several examples from *A. tumefaciens*.

The Atu1297 gene encoding a DGC homologue (Fig. 5a) has been designated *pleD* on the basis of its predicted amino acid sequence and its position immediately downstream of the cell cycle regulator *divK*, similar to the well-studied *pleD* DGC from *C. crescentus* (Paul et al. 2004; Sommer and Newton 1989). As with its homologue, the N-terminus of the *A. tumefaciens* PleD includes a dual two-component response regulator domain upstream of the DGC domain (Fig. 5a). Expression of this protein in *E. coli* and *A. tumefaciens* results in strong overproduction of cdGMP, dependent upon the GGDEF catalytic motif (Xu et al. 2013). In *C. crescentus*, the phosphorylation state of PleD is regulated by the sensor kinases DivJ and PleC of the complex CtrA phosphorelay, and it plays a prominent role in cell cycle control. The genetic linkage of *pleD* to *divK* in *A. tumefaciens* suggested that it might also function in cell cycle control, as *divK* mutants manifest cell division defects (Kim et al. 2013). However, null mutations in *pleD* have minimal impacts on cell division, and only marginally impact motility and attachment (Kim et al. 2013). In contrast, ectopic expression of PleD in *A. tumefaciens* strongly stimulates cellulose and UPP production concomitant with elevated cdGMP levels (Xu et al. 2013). Direct measurements in a *pleD* mutant suggest that it is diminished for cellulose production, but this does not manifest itself as strong effects on attachment, even exhibiting mild stimulation of attachment to plants (Barnhart et al. 2013). Ectopic expression of another DGC homologue, Atu1060, which has similar domain structure to PleD (Fig. 5a), also increased polysaccharide synthesis, but mutation of Atu1060 had no significant impact (Barnhart et al. 2013). It seems likely that elevated PleD levels (and similarly Atu1060) impact attachment processes by increasing cytoplasmic pools of cdGMP, but its role under native conditions is not clear.

In genetic studies on UPP production, mutations in the motility master regulators *visN* and *visR* resulted in increased UPP and cellulose production, enhancing attachment to abiotic surfaces and bypassing surface-contact dependence (Xu et al. 2013). Genetic suppressor analysis revealed that mutation of a specific DGC homologue, designated DgcA (Atu1257; Fig. 5a), prevented elevated polysaccharide production in a *visR* mutant. Furthermore, transcriptome analysis identified two additional DGC genes, *dgcB* (Atu1691) and *dgcC* (Atu2179), that were elevated in expression in a *visR* mutant. DgcA and DgcB (Fig. 5a) were capable of increasing

cdGMP levels when expressed in *E. coli* and in *A. tumefaciens* (dependent on their GGDEF catalytic motifs), whereas DgcC was not. Directed mutation of *dgcA* blocked the elevated attachment of the *visR* mutant, whereas mutation of *dgcB* had a more modest effect. In the wild-type background, mutation of *dgcB* significantly diminishes attachment and a *dgcA* mutation has a less pronounced effect. A *dgcAdgcB* double mutant abolished all detectable attachment. In contrast, mutations in *dgcC* had no effect on attachment, irrespective of the genetic background tested. Intriguingly, DgcB is orthologous to DgcB of *C. crescentus,* which was recently reported to be responsible for stimulation of polar holdfast production via interactions with specific biosynthetic proteins in response to the drag on flagellar rotation experienced during surface interactions (Hug et al. 2017; see Sect. 6.3).

In the same genetic screen that identified the role for VisNR in regulating attachment, a mutant was isolated in a dual DGC-PDE, Atu3495 (Fig. 5b), that strongly upregulated UPP and cellulose production, again bypassing surface-contact dependence (Xu et al. 2013). This protein is now designated DcpA (diguanylate cyclase/phosphodiesterase A), as it exhibits both cdGMP synthesis and degradation activities (Feirer et al. 2015). Mutations in *dcpA* cause elevated levels of cdGMP synthesis in laboratory-grown *A. tumefaciens*, indicating that the PDE activity is dominant under these conditions. However, a complex regulatory pathway has been discovered that dictates the bias between DGC and PDE activity for DcpA, and hence its control of attachment processes. Surprisingly, this pathway involves low molecular-weight metabolites called pterins, compounds virtually ubiquitous in living systems and related to folic acid and its derivatives (Feirer and Fuqua 2017). Production of a specific type of pterin derivative, known as a monapterin, is required to maintain DcpA in its PDE dominant state, and this control requires a protein encoded in the same operon as *dcpA*, named PruR (Pterin-responsive Upp Regulator), which contains a presumptive pterin-binding motif (Feirer et al. 2015). It is clear that this pathway plays a prominent role in regulation of attachment through modulation of the relevant cdGMP pools. Several observations have implicated the DcpA system as an important hub for environmentally responsive attachment control (Feirer et al. 2017; Heindl et al. 2016; Wang et al. 2016). The mechanism by which the monapterin cofactors are involved, and the reason for this complex regulation, are the subject of current studies.

6.3 *Surface-Contact-Responsive Control of Attachment Functions*

In wild-type *A. tumefaciens,* production of the UPP is strictly surface-contact-dependent (Li et al. 2012). Elevation of cellular cdGMP levels, either through DGC expression or loss of PDE activity, leads to the uncoupling of this requisite surface-contact stimulation (Feirer et al. 2015; Xu et al. 2013). This finding

suggests that the response to surface contact may function through fluxes in cdGMP levels, perhaps establishing transient gradients within the cell. Recent work in *C. crescentus* has revealed a role for cdGMP in surface-contact stimulation of holdfast production, both in response to inhibition of Cpa pilus retraction and the inhibition of flagellar rotation (Hug et al. 2017; Ellison et al. 2017). It is hypothesized that the drag on flagellar rotation is transduced through an increase in the activity of the DgcB DGC, and that cdGMP allosterically regulates the glycosyl transferase HfsJ, required for holdfast biosynthesis. The *C. crescentus* DgcB is orthologous to the *A. tumefaciens* DgcB ($DgcB_{At}$), mutations in which lead to dramatic reductions in surface attachment in otherwise wild-type *A. tumefaciens*. It is intriguing to speculate that $DgcB_{At}$ functions similarly to stimulate deployment of the UPP, thereby initiating stable attachment.

7 Conclusions

From the breadth of this review, it should be clear that there are multiple extracellular structures which can contribute to the attachment of *A. tumefaciens* cells to surfaces. The molecular details of these structures and their surface assembly or export continues to be a fertile area of study. Furthermore, the deployment of these specific structures is highly regulated, both at the level of gene expression and through allosteric control mechanisms. It is likely that additional attachment factors remain to be identified for *A. tumefaciens* and incorporated into the current models. Most importantly, the coordination of these different factors to drive productive surface attachment in benign associations, as well as disease progression, remains poorly understood. It is clear that *A. tumefaciens* is well adapted for transitioning to a sessile state on a variety of surfaces, and that factors such as pili, flagella, the UPP, and cellulose can play important roles in this general attachment process. However, the relationship of these general attachment mechanisms to the interactions with plant tissues that eventually lead to T-DNA transfer is still uncertain. Standard tumorigenesis assays that often involve inoculation with large numbers of cells that are either strongly induced for virulence or applied to tissues that are extensively wounded prior to inoculation do not accurately reflect the natural infection process. In these assays, the general attachment functions seem to be dispensable and only mutants which directly disrupt the T-DNA transfer process or severely compromise bacterial metabolism manifest deficiencies (Heindl et al. 2016; Kim et al. 2013). Qualitative evaluation of plant tissue attachment has revealed large-scale trends and deficiencies (Matthysse et al. 2005; Ramey et al. 2004; Tomlinson et al. 2010), but these are difficult to relate to virulence. Experiments which require more natural modes of infection or that measure attachment directly are likely required to evaluate effectively the role for general attachment functions in transformation.

The mechanisms that drive general surface attachment are certainly fundamental cellular attributes, as the genes required are encoded on the chromosomes, whereas

the majority of primary virulence functions are carried on the curable (unstable) and horizontally transmitted Ti plasmid (Goodner et al. 2001; Wood et al. 2001). *A. tumefaciens* derivatives that lack the Ti plasmid are avirulent but are still fully competent for surface attachment. Although it is logical to posit that the fundamental attachment process leading to stable polar association of cells with surfaces is functionally connected with the attachment processes that culminate in T-DNA transfer, this remains to be validated convincingly. One model would be a temporal process in which the general attachment functions promote stable, proximal association to potential infection sites on the plant, among other more benign interactions. Under the appropriate environmental signals, this attachment may be followed by induction of the *Vir* genes and related functions that drive pathogenesis and disease progression, including direct engagement with plant surface structures (Zhu et al. 2003). An alternative model is that the general attachment processes promote benign interactions with plants and other surfaces, but that the attachment that leads to T-DNA transfer is completely distinct. Whether one of these models is correct, or perhaps an amalgamation of the two, remains an area of active study.

Acknowledgements Research on surface attachment functions for *A. tumefaciens* in the Fuqua lab is supported by the National Institutes of Health GM120337. We thank Justin Eagan for his comments on the manuscript.

References

Aguilar J, Cameron TA, Zupan J, et al (2011) Membrane and core periplasmic *Agrobacterium tumefaciens* virulence Type IV secretion system components localize to multiple sites around the bacterial perimeter during lateral attachment to plant cells. MBio 2:e00218-11

Aldridge P, Hughes KT (2002) Regulation of flagellar assembly. Curr Opin Microbiol 5:160–165

Aly KA, Baron C (2007) The VirB5 protein localizes to the T-pilus tips in *Agrobacterium tumefaciens*. Microbiology 153:3766–3775

Amikam D, Benziman M (1989) Cyclic diguanylic acid and cellulose synthesis in *Agrobacterium tumefaciens*. J Bacteriol 171:6649–6655

Arioli T, Peng L, Betzner AS et al (1998) Molecular analysis of cellulose biosynthesis in *Arabidopsis*. Science 279:717–720

Ashby AM, Watson MD, Loake GJ et al (1988) Ti plasmid-specified chemotaxis of *Agrobacterium tumefaciens* C58C1 toward vir-inducing phenolic compounds and soluble factors from monocotyledonous and dicotyledonous plants. J Bacteriol 170:4181–4187

Ausmees N, Jacobsson K, Lindberg M (2001) A unipolarly located, cell-surface-associated agglutinin, RapA, belongs to a family of *Rhizobium*-adhering proteins (Rap) in *Rhizobium leguminosarum* bv. *trifolii*. Microbiology 147:549–559

Babić A, Lindner AB, Vulić M et al (2008) Direct visualization of horizontal gene transfer. Science 319:1533–1536

Barbosa RL, Benedetti CE (2007) BigR, a transcriptional repressor from plant associated bacteria, regulates an operon implicated in biofilm growth. J Bacteriol 189:6185–6194

Barnhart MM, Chapman MR (2006) Curdlan biogenesis and function. Annu Rev Microbiol 60:131–147

Barnhart DM, Su S, Baccaro BE et al (2013) CelR, an ortholog of the diguanylate cyclase PleD of *Caulobacter*, regulates cellulose synthesis in *Agrobacterium tumefaciens*. Appl Environ Microbiol 79:7188–7202

Barnhart DM, Su S, Farrand SK (2014) A signaling pathway involving the diguanylate cyclase CelR and the response regulator DivK controls cellulose synthesis in *Agrobacterium tumefaciens*. J Bacteriol 196:1257–1274

Berne C, Ducret A, Hardy GG et al (2015) Adhesins involved in attachment to abiotic surfaces by Gram-negative bacteria. Microbiol Spectr 3

Blair DF (2003) Flagellar movement driven by proton translocation. FEBS Lett 545:86–95

Blanco LP, Evans ML, Smith DR et al (2012) Diversity, biogenesis and function of microbial amyloids. Trends Microbiol 20:66–73

Bodenmiller D, Toh E, Brun YV (2004) Development of surface adhesion in *Caulobacter crescentus*. J Bacteriol 186:1438–1447

Branch RW, Sayegh MN, Shen C, Nathan VSJ, Berg HC (2014) Adaptive remodeling by FliN in the bacterial rotary motor. J Mol Biol 426:3314–3324

Branda SS, Vik Å, Friedman L et al (2005) Biofilms: the matrix revisited. Trends Microbiol 13:20–26

Brown PJ, de Pedro MA, Kysela DT et al (2012) Polar growth in the Alphaproteobacterial order Rhizobiales. Proc Natl Acad Sci USA 109:1697–1701

Chen L, Chen Y, Wood DW et al (2002) A new type IV secretion system promotes conjugal transfer in *Agrobacterium tumefaciens*. J Bacteriol 184:4838–4845

Chesnokova O, Coutinho JB, Khan IH et al (1997) Characterization of flagella genes of *Agrobacterium tumefaciens*, and the effect of a bald strain on virulence. Mol Microbiol 23:579–590

Chevance FF, Hughes KT (2008) Coordinating assembly of a bacterial macromolecular machine. Nat Rev Microbiol 6:455

Cole JL, Hardy GG, Bodenmiller D et al (2003) The HfaB and HfaD adhesion proteins of *Caulobacter crescentus* are localized in the stalk. Mol Microbiol 49:1671–1683

Cook DM, Li PL, Ruchaud F et al (1997) Ti plasmid conjugation is independent of vir: reconstitution of the tra functions from pTiC58 as a binary system. J Bacteriol 179:1291–1297

Costa TR, Ilangovan A, Ukleja M et al (2016) Structure of the bacterial sex F pilus reveals an assembly of a stoichiometric protein-phospholipid complex. Cell 166:1436–1444

Cuthbertson L, Mainprize IL, Naismith JH et al (2009) Pivotal roles of the outer membrane polysaccharide export and polysaccharide copolymerase protein families in export of extracellular polysaccharides in gram-negative bacteria. Microbiol Mol Biol R 73:155–177

Danhorn T, Fuqua C (2007) Biofilm formation by plant-associated bacteria. Annu Rev Microbiol 61:401–422

Danhorn T, Hentzer M, Givskov M et al (2004) Phosphorus limitation enhances biofilm formation of the plant pathogen *Agrobacterium tumefaciens* through the PhoR-PhoB regulatory system. J Bacteriol 186:4492–4501

Dardanelli M, Angelini J, Fabra A (2003) A calcium-dependent bacterial surface protein is involved in the attachment of rhizobia to peanut roots. Can J Microbiol 49:399–405

Deakin WJ, Parker VE, Wright EL et al (1999) *Agrobacterium tumefaciens* possesses a fourth flagellin gene located in a large gene cluster concerned with flagellar structure, assembly and motility. Microbiology 145:1397–1407

Deakin WJ, Furniss CS, Parker VE et al (1997a) Isolation and characterisation of a linked cluster of genes from *Agrobacterium tumefaciens* encoding proteins involved in flagellar basal-body structure. Gene 189:135–137

Deakin WJ, Sanderson JL, Goswami T et al (1997b) The *Agrobacterium tumefaciens* motor gene, *motA*, is in a linked cluster with the flagellar switch protein genes, *fliG*, *fliM* and *fliN*. Gene 189:139–141

Deinema MH, Zevenhuizen LPTM (1971) Formation of cellulose fibrils by gram-negative bacteria and their role in bacterial flocculation. Arch Microbiol 78:42–57

DeRosier D (2006) Bacterial flagellum: visualizing the complete machine in situ. Curr Biol 16: R928–R930

Ellison CK, Kan J, Dillard RS et al (2017) Obstruction of pilus retraction stimulates bacterial surface sensing. Science 358:535–538

Feirer N, Fuqua C (2017) Pterin function in bacteria. Pteridines 28:23–36

Feirer N, Kim D, Xu J et al (2017) The *Agrobacterium tumefaciens* CheY-like protein ClaR regulates biofilm formation. Microbiology 163:1680–1691

Feirer N, Xu J, Allen KD et al (2015) A pterin-dependent signaling pathway regulates a dual-function diguanylate cyclase-phosphodiesterase controlling surface attachment in *Agrobacterium tumefaciens*. MBio 6:e00156-15

Ferooz J, Lemaire J, Letesson JJ (2011) Role of FlbT in flagellin production in *Brucella melitensis*. Microbiology 157:1253–1262

Fiebig A, Herrou J, Fumeaux C et al (2014) A cell cycle and nutritional checkpoint controlling bacterial surface adhesion. PLoS Genet 10:e1004101

Fritts RK, LaSarre B, Stoner AM et al (2017) A Rhizobiales-specific unipolar polysaccharide adhesin contributes to *Rhodopseudomonas palustris* biofilm formation across diverse photoheterotrophic conditions. Appl Environ Microbiol 83:e03035–16

Fullner KJ, Lara JC, Nester EW (1996) Pilus assembly by *Agrobacterium* T-DNA transfer genes. Science 273:1107-1109

Fuqua WC, Winans SC (1994) A LuxR-LuxI type regulatory system activates *Agrobacterium* Ti plasmid conjugal transfer in the presence of a plant tumor metabolite. J Bacteriol 176:2796–2806

Ghigo JM (2001) Natural conjugative plasmids induce bacterial biofilm development. Nature 412:442

Goodner B, Hinkle G, Gattung S et al (2001) Genome sequence of the plant pathogen and biotechnology agent *Agrobacterium tumefaciens* C58. Science 294:2323–2328

Götz R, Limmer N, Ober K et al (1982) Motility and chemotaxis in two strains of *Rhizobium* with complex flagella. Microbiology 128:789–798

Gu X, Lee SG, Bar-Peled M (2011) Biosynthesis of UDP-xylose and UDP-arabinose in *Sinorhizobium meliloti* 1021: first characterization of a bacterial UDP-xylose synthase, and UDP-xylose 4-epimerase. Microbiology 157:260–269

Guimarães BG, Barbosa RL, Soprano AS et al (2011) Plant pathogenic bacteria utilize biofilm growth-associated repressor (BigR), a novel winged-helix redox switch, to control hydrogen sulfide detoxification under hypoxia. J Biol Chem 286:26148–26157

Hardy GG, Allen RC, Toh E et al (2010) A localized multimeric anchor attaches the *Caulobacter* holdfast to the cell pole. Mol Microbiol 76:409–427

Hardy GG, Toh E, Berne C et al (2018) Mutations in sugar-nucleotide synthesis genes restore holdfast polysaccharide anchoring to *Caulobacter crescentus* holdfast anchor mutants. J Bacteriol 200:e00597–17

Heckel BC, Tomlinson AD, Morton ER et al (2014) *Agrobacterium tumefaciens exoR* controls acid response genes and impacts exopolysaccharide synthesis, horizontal gene transfer, and virulence gene expression. J Bacteriol 196:3221–3233

Heindl JE, Hibbing ME, Xu J et al (2016) Discrete responses to limitation for iron and manganese in *Agrobacterium tumefaciens*: influence on attachment and biofilm formation. J Bacteriol 198:816–829

Heindl JE, Wang Y, Heckel BC et al (2014) Mechanisms and regulation of surface interactions and biofilm formation in *Agrobacterium*. Front Plant Sci 5:176

Hinsa SM, Espinosa-Urgel M, Ramos JL et al (2003) Transition from reversible to irreversible attachment during biofilm formation by *Pseudomonas fluorescens* WCS365 requires an ABC transporter and a large secreted protein. Mol Microbiol 49:905–918

Høiby N (2017) A short history of microbial biofilms and biofilm infections. APMIS 125:272–275

Hong Y, Reeves PR (2014) Diversity of O-antigen repeat unit structures can account for the substantial sequence variation of Wzx translocases. J Bacteriol 196:1713–1722

Hug I, Deshpande S, Sprecher KS et al (2017) Second messenger–mediated tactile response by a bacterial rotary motor. Science 358:531–534
Hughes KT, Mathee K (1998) The anti-sigma factors. Annu Rev Microbiol 52:231–286
Islam ST, Lam JS (2013) Wzx flippase-mediated membrane translocation of sugar polymer precursors in bacteria. Environ Microbiol 15:1001–1015
Islam ST, Lam JS (2014) Synthesis of bacterial polysaccharides via the Wzx/Wzy-dependent pathway. Can J Microbiol 60:697–716
Jenal U, Reinders A, Lori C (2017) Cyclic di-GMP: second messenger extraordinaire. Nat Rev Microbiol 15:271
Judd PK, Kumar RB, Das A (2005) The type IV secretion apparatus protein VirB6 of *Agrobacterium tumefaciens* localizes to a cell pole. Mol Microbiol 55:115–124
Kachlany SC, Planet PJ, DeSalle R et al (2001) flp-1, the first representative of a new pilin gene subfamily, is required for non-specific adherence of Actinobacillus actinomycetemcomitans. Mol Microbiol 40:542–554
Kai-Larsen Y, Lüthje P, Chromek M et al (2010) Uropathogenic *Escherichia coli* modulates immune responses and its curli fimbriae interact with the antimicrobial peptide LL-37. PLoS Pathog 6:e1001010
Kalynych S, Morona R, Cygler M (2014) Progress in understanding the assembly process of bacterial *O*-antigen. FEMS Microbiol Rev 38:1048–1065
Karnezis T, Epa VC, Stone BA et al (2003) Topological characterization of an inner membrane (1→3)-β-D-glucan (curdlan) synthase from *Agrobacterium* sp. strain ATCC31749. Glycobiology 13:693–706
Karnezis T, Fisher HC, Neumann GM et al (2002) Cloning and characterization of the phosphatidylserine synthase gene of *Agrobacterium* sp. strain ATCC 31749 and effect of its inactivation on production of high-molecular-mass (1→3)-β-D-glucan (curdlan). J Bacteriol 184:4114–4123
Kilmury SL, Burrows LL (2016) Type IV pilins regulate their own expression via direct intramembrane interactions with the sensor kinase PilS. PNAS 113:6017–6022
Kim J, Heindl JE, Fuqua C (2013) Coordination of division and development influences complex multicellular behavior in *Agrobacterium tumefaciens*. PLoS ONE 8:e56682
Kimura S, Chen HP, Saxena IM et al (2001) Localization of c-di-GMP-binding protein with the linear terminal complexes of *Acetobacter xylinum*. J Bacteriol 183:5668–5674
Körner H, Sofia HJ, Zumft WG (2003) Phylogeny of the bacterial superfamily of Crp-Fnr transcription regulators: exploiting the metabolic spectrum by controlling alternative gene programs. FEMS Microbiol Rev 27:559–592
Lai EM, Chesnokova O, Banta LM et al (2000) Genetic and environmental factors affecting T-pilin export and T-pilus biogenesis in relation to flagellation of *Agrobacterium tumefaciens*. J Bacteriol 182:3705–3716
Lai EM, Eisenbrandt R, Kalkum M et al (2002) Biogenesis of T pili in *Agrobacterium tumefaciens* requires precise VirB2 propilin cleavage and cyclization. J Bacteriol 184:327–330
Laus MC, Logman TJ, Lamers GE et al (2006) A novel polar surface polysaccharide from *Rhizobium leguminosarum* binds host plant lectin. Mol Microbiol 59:1704–1713
Leifson E, Erdman LW (1958) Flagellar characteristics of *Rhizobium* species. A van Leeuw J Microb 24:97–110
Lele PP, Branch RW, Nathan VJS, Berg HC (2012) Mechanism for remodeling of the bacterial flagellar switch. Proc Natl Acad Sci USA 109:20018–20022
Lessl M, Lanka E (1994) Common mechanisms in bacterial conjugation and Ti-mediated T-DNA transfer to plant cells. Cell 77:321–324
Li G, Brown PJ, Tang JX et al (2012) Surface contact stimulates the just-in-time deployment of bacterial adhesins. Mol Microbiol 83:41–51
Liu R, Ochman H (2007) Origins of flagellar gene operons and secondary flagellar systems. J Bacteriol 189:7098–7104
Locht C, Berlin P, Menozzi FD et al (1993) The filamentous haemagglutinin, a multifaceted adhesin produced by virulent *Bordetella* spp. Mol Microbiol 9:653–660

Madsen JS, Burmølle M, Hansen LH et al (2012) The interconnection between biofilm formation and horizontal gene transfer. FEMS Immunol Med Mic 65:183–195
Mangan EK, Malakooti J, Caballero A et al (1999) FlbT couples flagellum assembly to gene expression in *Caulobacter crescentus*. J Bacteriol 181:6160–6170
Matthysse AG (1983) Role of bacterial cellulose fibrils in *Agrobacterium tumefaciens* infection. J Bacteriol 154:906–915
Matthysse AG (1987) Characterization of nonattaching mutants of *Agrobacterium tumefaciens*. J Bacteriol 169:313–323
Matthysse AG (2014) Attachment of *Agrobacterium* to plant surfaces. Front Plant Sci 5:252
Matthysse AG, Holmes KV, Gurlitz RH (1981) Elaboration of cellulose fibrils by *Agrobacterium tumefaciens* during attachment to carrot cells. J Bacteriol 145:583–595
Matthysse AG, Jaeckel P, Jeter C (2008) *attG* and *attC* mutations of *Agrobacterium tumefaciens* are dominant negative mutations that block attachment and virulence. Can J Microbiol 54:241–247
Matthysse AG, Marry M, Krall L et al (2005) The effect of cellulose overproduction on binding and biofilm formation on roots by *Agrobacterium tumefaciens*. Mol Plant Microbe Interac 18:1002–1010
Matthysse AG, Thomas DL, White AR (1995a) Mechanism of cellulose synthesis in *Agrobacterium tumefaciens*. J Bacteriol 177:1076–1081
Matthysse AG, White S, Lightfoot R (1995b) Genes required for cellulose synthesis in *Agrobacterium tumefaciens*. J Bacteriol 177:1069–1075
Matthysse AG, Yarnall H, Boles SB et al (2000) A region of the *Agrobacterium tumefaciens* chromosome containing genes required for virulence and attachment to host cells. Biophys Biochi. Acta 1490:208–212
Mattick JS (2002) Type IV pili and twitching motility. Annu Rev Microbiol 56:289–314
McDermott TR (2000) Phosphorus assimilation and regulation in the rhizobia. In: Triplett EW (ed) Prokaryotic nitrogen fixation: a model system for the analysis of a biological process. Horizon Scientific Press, Norfolk, United Kingdom, pp 529–548
McIntosh M, Stone BA, Stanisich VA (2005) Curdlan and other bacterial (1→3)-β-D-glucans. Appl Microbiol Biotech 68:163–173
Macnab RM (1996) Flagella and motility. In: Neidhardt FC (ed) *Escherichia coli* and *Salmonella*: cellular and molecular biology. ASM Press, Washington DC, pp 123–145
Meadows PS (1971) The attachment of bacteria to solid surfaces. Arch Microbiol 75:374–381
Merritt PM, Danhorn T, Fuqua C (2007) Motility and chemotaxis in *Agrobacterium tumefaciens* surface attachment and biofilm formation. J Bacteriol 189:8005–8014
Mi W, Li Y, Yoon SH et al (2017) Structural basis of MsbA-mediated lipopolysaccharide transport. Nature 549:233
Mohari B, Licata NA, Kysela DT et al (2015) Novel pseudotaxis mechanisms improve migration of straight-swimming bacterial mutants through a porous environment. MBio 6:e00005-15
Morgan JL, Strumillo J, Zimmer J (2013) Crystallographic snapshot of cellulose synthesis and membrane translocation. Nature 493:181
Nair GR, Liu Z, Binns AN (2003) Reexamining the role of the accessory plasmid pAtC58 in the virulence of *Agrobacterium tumefaciens* strain C58. Plant Physiol 133:989–999
Nakanishi I, Kimura K, Suzuki T et al (1976) Demonstration of curdlan-type polysaccharide and some other β-1, 3-glucan in microorganisms with aniline blue. J Gen Appl Microbiol 22:1–11
Oberpichler I, Rosen R, Rasouly A et al (2008) Light affects motility and infectivity of *Agrobacterium tumefaciens*. Environ Microbiol 10:2020–2029
Osterman IA, Dikhtyar YY, Bogdanov AA et al (2015) Regulation of flagellar gene expression in bacteria. Biochemistry (Moscow) 80:1447–1456
Pallen MJ, Matzke NJ (2006) from the origin of species to the origin of bacterial flagella. Nat Rev Microbiol 4:784
Paul R, Weiser S, Amiot NC et al (2004) Cell cycle-dependent dynamic localization of a bacterial response regulator with a novel di-guanylate cyclase output domain. Gene Dev 18:715–727

Piper KR, Beck von Bodman S, Farrand SK (1993) Conjugation factor of *Agrobacterium tumefaciens* regulates Ti plasmid transfer by autoinduction. Nature 362:448–450

Ramey BE, Matthysse AG, Fuqua C (2004) The FNR-type transcriptional regulator SinR controls maturation of Agrobacterium tumefaciens biofilms. Mol Microbiol 52:1495–1511

Reuber TL, Walker GC (1993) Biosynthesis of succinoglycan, a symbiotically important exopolysaccharide of *Rhizobium meliloti*. Cell 74:269–280

Römling U (2002) Molecular biology of cellulose production in bacteria. Res Microbiol 153:205–212

Römling U, Galperin MY, Gomelsky M (2013) Cyclic di-GMP: the first 25 years of a universal bacterial second messenger. Microbiol Mol Biol Rev 77:1–52

Römling U, Galperin MY (2015) Bacterial cellulose biosynthesis: diversity of operons, subunits, products, and functions. Trends Microbiol 23:545–557

Ross P, Weinhouse H, Aloni Y, Michaeli D, Weinbergerohana P, Mayer R, Braun S, Devroom E, Vandermarel GA, Vanboom JH, Benziman M (1987) Regulation of cellulose synthase in *Acetobacter xylinum* by cyclic diguanylic acid. Nature 325:279–281

Rotter C, Mühlbacher S, Salamon D et al (2006) Rem, a new transcriptional activator of motility and chemotaxis in *Sinorhizobium meliloti*. J Bacteriol 188:6932–6942

Ruffing AM, Castro-Melchor M, Hu WS et al (2011) Genome sequence of the curdlan-producing *Agrobacterium* sp. strain ATCC 31749. J Bacteriol 193:4294–4295

Ruffing AM, Chen RR (2012) Transcriptome profiling of a curdlan-producing *Agrobacterium* reveals conserved regulatory mechanisms of exopolysaccharide biosynthesis. Microb Cell Fact 11:17

Russo DM, Williams A, Edwards A et al (2006) Proteins exported via the PrsD-PrsE type I secretion system and the acidic exopolysaccharide are involved in biofilm formation by *Rhizobium leguminosarum*. J Bacteriol 188:4474–4486

Sauer K, Camper AK, Ehrlich GD et al (2002) *Pseudomonas aeruginosa* displays multiple phenotypes during development as a biofilm. J Bacteriol 184:1140–1154

Schäper S, Krol E, Skotnicka D et al (2016) Cyclic di-GMP regulates multiple cellular functions in the symbiotic alphaproteobacterium *Sinorhizobium meliloti*. J Bacteriol 198:521–535

Schmid J, Sieber V, Rehm B (2015) Bacterial exopolysaccharides: biosynthesis pathways and engineering strategies. Front Microbiol 6:496

Schuhmacher JS, Thormann KM, Bange G (2015) How bacteria maintain location and number of flagella? FEMS Microbiol Rev 39:812–822

Shaw CH, Loake GJ, Brown AP et al (1991) Isolation and characterization of behavioural mutants and genes of *Agrobacterium tumefaciens*. Microbiology 137:1939–1953

Shoda M, Sugano Y (2005) Recent advances in bacterial cellulose production. Biotechnol Bioproc E 10:1

Skerker JM, Shapiro L (2000) Identification and cell cycle control of a novel pilus system in *Caulobacter crescentus*. EMBO J 19:3223–3234

Smit G, Logman TJ, Boerrigter ME et al (1989) Purification and partial characterization of the *Rhizobium leguminosarum* biovar viciae Ca2+-dependent adhesin, which mediates the first step in attachment of cells of the family Rhizobiaceae to plant root hair tips. J Bacteriol 171:4054–4062

Sommer JM, Newton A (1989) Turning off flagellum rotation requires the pleiotropic gene *pleD*: *pleA*, *pleC*, and *pleD* define two morphogenic pathways in *Caulobacter crescentus*. J Bacteriol 171:392–401

Sourjik V, Muschler P, Scharf B, Schmitt R (2000) VisN and VisR are global regulators of chemotaxis, flagellar, and motility genes in *Sinorhizobium* (*Rhizobium*) *meliloti*. J Bacteriol 182:782–788

Sourjik V, Schmitt R (1996) Different roles of CheY1 and CheY2 in the chemotaxis of *Rhizobium meliloti*. Mol Microbiol 22:427–436

Stasinopoulos SJ, Fisher PR, Stone BA et al (1999) Detection of two loci involved in (1→3)-β-glucan (curdlan) biosynthesis by *Agrobacterium* sp. ATCC31749, and comparative sequence analysis of the putative curdlan synthase gene. Glycobiol 9:31–41

Toh E, Kurtz HD, Brun YV (2008) Characterization of the *Caulobacter crescentus* holdfast polysaccharide biosynthesis pathway reveals significant redundancy in the initiating glycosyltransferase and polymerase steps. J Bacteriol 190:7219–7231

Tomich M, Planet PJ, Figurski DH (2007) The tad locus: postcards from the widespread colonization island. Nat Rev Microbiol 5:363

Tomlinson AD, Fuqua C (2009) Mechanisms and regulation of polar surface attachment in *Agrobacterium tumefaciens*. Curr Opin Microbiol 12:708–714

Tomlinson AD, Ramey-Hartung B, Day TW et al (2010) *Agrobacterium tumefaciens* ExoR represses succinoglycan biosynthesis and is required for biofilm formation and motility. Microbiology 156:2670–2681

Visick KL, Schembri MA, Yildiz F et al (2016) Biofilms 2015: multidisciplinary approaches shed light into microbial life on surfaces. J Bacteriol 198:2553–2563

Wadhams GH, Armitage JP (2004) Making sense of it all: bacterial chemotaxis. Nat Rev Mol Cell Biol 5:1024–1037

Wan Z, Brown PJ, Elliott EN, Brun YV (2013) The adhesive and cohesive properties of a bacterial polysaccharide adhesin are modulated by a deacetylase. Mol Microbiol 88:486–500

Wang Y, Haitjema CH, Fuqua C (2014) The Ctp type IVb pilus locus of *Agrobacterium tumefaciens* directs formation of the common pili and contributes to reversible surface attachment. J Bacteriol 196:2979–2988

Wang Y, Kim SH, Natarajan R et al (2016) Spermidine inversely influences surface interactions and planktonic growth in *Agrobacterium tumefaciens*. J Bacteriol 198:2682–2691

Wang X, Yang F, von Bodman SB (2012) The genetic and structural basis of two distinct terminal side branch residues in stewartan and amylovoran exopolysaccharides and their potential role in host adaptation. Mol Microbiol 83:195–207

Williamson RE, Burn JE, Hocart CH (2002) Towards the mechanism of cellulose synthesis. Trends Plant Sci 7:461–467

Wolfe AJ, Visick KL (eds) (2010) The second messenger cyclic di-GMP. ASM Press, USA

Wood DW, Setubal JC, Kaul R et al (2001) The genome of the natural genetic engineer *Agrobacterium tumefaciens* C58. Science 294:2317–2323

Wu CF, Lin JS, Shaw GC et al (2012) Acid-induced type VI secretion system is regulated by ExoR-ChvG/ChvI signaling cascade in *Agrobacterium tumefaciens*. PLoS Pathog 8:e1002938

Xu J, Kim J, Danhorn T et al (2012) Phosphorus limitation increases attachment in *Agrobacterium tumefaciens* and reveals a conditional functional redundancy in adhesin biosynthesis. Res Microbiol 163:674–684

Xu J, Kim J, Koestler BJ et al (2013) Genetic analysis of *Agrobacterium tumefaciens* unipolar polysaccharide production reveals complex integrated control of the motile-to-sessile switch. Mol Microbiol 89:929–948

Yu X, Zhang C, Yang L et al (2015) CrdR function in a curdlan-producing *Agrobacterium* sp. ATCC31749 strain. BMC Microbiol 15:25

Zatakia HM, Arapov TD, Meier VM et al (2018) Cellular stoichiometry of methyl-accepting chemotaxis proteins in *Sinorhizobium meliloti*. J Bacteriol 200:e00614–e00617

Zhan XB, Lin CC, Zhang HT (2012) Recent advances in curdlan biosynthesis, biotechnological production, and applications. Appl Microbiol Biotechnol 93:525–531

Zhu Y, Nam J, Humara JM et al (2003) Identification of *Arabidopsis rat* mutants. Plant Physiol 132:494–505

Coping with High Temperature: A Unique Regulation in *A. tumefaciens*

Dvora Biran, Or Rotem, Ran Rosen and Eliora Z. Ron

Contents

1 Introduction 185
2 The Heat-Shock Response 188
3 Conclusions 191
References 191

Abstract Elevation of temperature is a frequent and considerable stress for mesophilic bacteria. Therefore, several molecular mechanisms have evolved to cope with high temperature. We have been studying the response of *Agrobacterium tumefaciens* to temperature stress, focusing on two aspects: the heat-shock response and the temperature-dependent regulation of methionine biosynthesis. The results indicate that the molecular mechanisms involved in *A. tumefaciens* control of growth at high temperature are unique and we are still missing important information essential for understanding how these bacteria cope with temperature stress.

1 Introduction

Most bacteria are exposed to frequent changes in the environment, especially changes in temperature. Exposure to elevated temperatures results in major physiological changes, including decreased activity of enzymes due to structural changes of the proteins. Therefore, immediate and precise adaptation to temperature changes is essential. The reaction to elevated temperatures involves two different processes—one is a quick response to the change, a rescue response aimed at repairing damaged processes and preventing further injury. The second process

D. Biran · O. Rotem · R. Rosen · E. Z. Ron (✉)
School of Molecular Cell Biology and Biotechnology, Faculty of Life Sciences, Tel Aviv University, 69978 Tel Aviv, Israel
e-mail: eliora@post.tau.ac.il

Current Topics in Microbiology and Immunology (2018) 418:185–194
DOI 10.1007/82_2018_119

Published Online: 5 September 2018

involves adaptation to the new temperature by regulating the biochemical and physiological processes for maintaining balanced growth at the new temperature. In the two processes, the level of response to the elevated temperature is proportional to the severity of the change.

The adaptation to the change in temperature is regulated by a global regulatory network that controls the simultaneous expression of a large number of stress-related genes—the heat-shock response (Arsene et al. 2000; Craig 1985; Ghazaei 2017; Guan et al. 2017; Lindquist 1986; Mathew and Morimoto 1998; Mathew et al. 2000; Ron 2009; Schumann 2016; Segal and Ron 1998; Yura and Nakahigashi 1999).

The first attempts to study the temperature-dependent regulatory networks were performed in *Escherichia coli* using O'Farrell two-dimensional gels to compare proteomes in bacteria exposed to elevated temperatures. These experiments resulted in the identification of a large group of proteins, the heat-shock proteins, which are induced following the increase in temperature (Neidhardt et al. 1981; O'Farrell 1975). These data indicated the existence of global regulatory systems that control the expression of a large regulon, the heat-shock regulon. This regulated heat-shock response protects against the increase in temperature and provides thermo-tolerance as well as tolerance to additional stress conditions (Inbar and Ron 1993; Ramsay 1988).

The heat-shock response is the first discovered global regulatory system (Craig 1985), and is found in all living cells examined: bacteria, yeast, insects (*Drosophila melanogaster*) (Michaud et al. 1997), worms (*Caenorhabditis elegans*) (Rose and Rankin 2001), and mammals (Christians et al. 2002; Li et al. 2002; Srivastava 2002). The heat-shock response is characterized by the induction of a large set of proteins (heat-shock proteins; HSPs) as a result of a rapid increase in the environmental temperature. The proteins involved in the response to elevated temperature include mostly components of protein quality control, such as chaperons (e.g., GroEL, GroES, DnaK, and DnaJ) and ATP-dependent proteases (e.g., ClpP, Lon (La) and HslVU). Although many of the HSPs are highly conserved, their regulatory elements differ considerably between organisms, as well as the conditions under which the response systems are activated. Clearly, mesophiles respond to much lower temperatures than the thermophiles.

The complementary process, maintenance of balanced growth as a function of temperature, was studied in *Enterobacteriaceae* (gamma proteobacteria) and in *Agrobacterium tumefaciens* (alpha proteobacterium). These findings indicate that the maintenance of balanced growth at elevated temperatures is regulated by the availability of methionine (Biran et al. 1995; Gur et al. 2002; Ron 1975; Ron et al. 1990; Ron and Davis 1971; Ron and Shani 1971; Rotem et al. 2013).

Here, we discuss the two response mechanisms to elevated temperature, maintaining balanced growth and the heat-shock response, in *A. tumefaciens*.

Growth at elevated temperature is limited by the availability of methionine

When grown in minimal medium (Chilton et al. 1974), growth of *A. tumefaciens* is severely inhibited at 39 °C (Fig. 1a). However, the inhibition is relieved by the

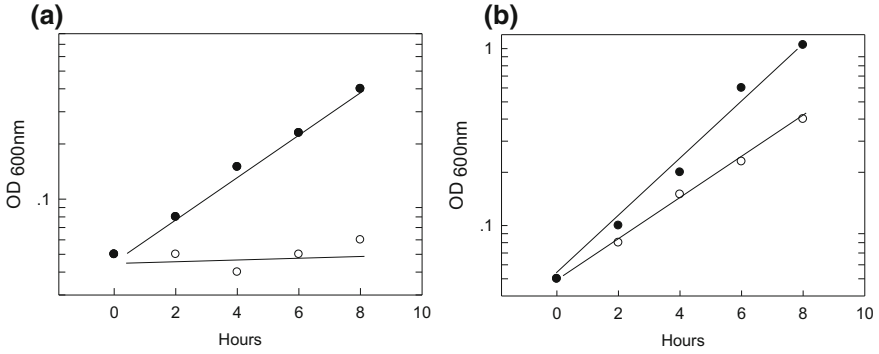


Fig. 1 Growth of *A. tumefaciens* at elevated temperatures. Bacteria were grown in minimal medium (Chilton et al. 1974) **a** or LB broth (Difco) **b** at 30 °C (filled circles) or 39 °C (empty circles). Growth was monitored by turbidity at OD 600 nm

addition of nutrients found in LB medium (Difco) (Fig. 1b). These results are similar to results obtained with bacteria belonging to the *Enterobacteriaceae,* where growth at elevated temperatures is limited by the availability of methionine (Ron 1975). *E. coli*, for example, depends on exogenous methionine for growth above 43 °C because the activity of the first enzyme in the methionine biosynthetic pathway, homoserine trans-succinylase, is extremely temperature-sensitive (Gur et al. 2002).

The first step in methionine biosynthesis is the activation of homoserine. In *E. coli*, this activation uses succinyl-coA. However, in *A. tumefaciens*, the first enzyme for methionine synthesis is not a trans-succinylase, but a trans-acetylase (acetyl transferase), using acetyl-CoA for the activation of homoserine (Rotem et al. 2013). *A. tumefaciens* MetA has severely reduced activity at 39 °C, a temperature at which it undergoes proteolysis. Proteolysis of the first enzyme in the methionine biosynthetic pathway was also shown in *E. coli* (Biran et al. 2000; Gur et al. 2011; Katz et al. 2009). It therefore appears that *A. tumefaciens* maintains balanced growth at elevated temperatures by regulating the availability and activity of the MetA protein, homoserine acetyl transferase.

The similarity of this regulatory mechanism between *A. tumefaciens* and the *Enterobacteriaceae* is not trivial, as they are phylogenetically quite distant (α and γ proteobacteria, respectively) and the enzymes have different substrates and temperature-dependent activity. It was also unexpected that a biosynthetic enzyme is not stable but has a short half-life. These findings can therefore be explained by assuming that regulation of growth rate as a function of elevated temperature is a critical necessity for mesophilic bacteria, and that such regulation is best performed via the availability of methionine. Methionine is a key biochemical compound because, in addition to its structural role as a component of proteins, it is involved in the initiation of protein synthesis and in many other biochemical processes such as the biosynthesis of S-adenosyl-methionine, purines, pyrimidines, fatty acids, and polyamines.

2 The Heat-Shock Response

The heat-shock response involves the induction of a large group of proteins, the heat-shock proteins, by an increase in temperature. The heat-shock response is a universal regulatory network involving highly conserved heat-shock proteins. These include the chaperones Hsp60 (GroEL) and Hsp70 (DnaK), and ATP-dependent proteases. Although the heat-shock proteins are highly conserved, the regulatory mechanisms responsible for the induction of these proteins are variable. The best studied regulation of the heat-shock response is in *E. coli* where the heat-shock genes are transcribed by a special sigma factor, σ32, from specific promoters (Fig. 2) (Erickson et al. 1987; Grossman et al. 1984; Taylor et al. 1984; Yura et al. 1990; Zhou et al. 1988). The transcriptional activator σ32 is unstable at low temperatures, as it is degraded by a specific protease FtsH (HflB), but it is stabilized as the temperature increases (Herman et al. 1995; Shenhar et al. 2009; Tomoyasu et al. 1995). This mechanism ensures the selective transcription of heat-shock promoters only at elevated temperatures. However, the induced transcription of the heat-shock operon is transient, and after an initial burst, its rate eventually returns to a basal level typical for the new temperature. This decrease in transcription of heat-shock genes is explained by the "titration model" which assumes that the level of HSPs is regulated in correlation with the amount of their substrates, unfolded proteins (Guisbert et al. 2004; Straus et al. 1990; Tomoyasu et al. 1998).

In *Bacillus subtilis*, there is no heat-shock-specific sigma factor and the heat-shock genes use regular σ70 promoters. However, upstream of the promoter of heat-shock genes there is an inverted repeat, CIRCE (Conserved Inverted Repeat Control Element), which binds the repressor HrcA. At elevated temperatures, the

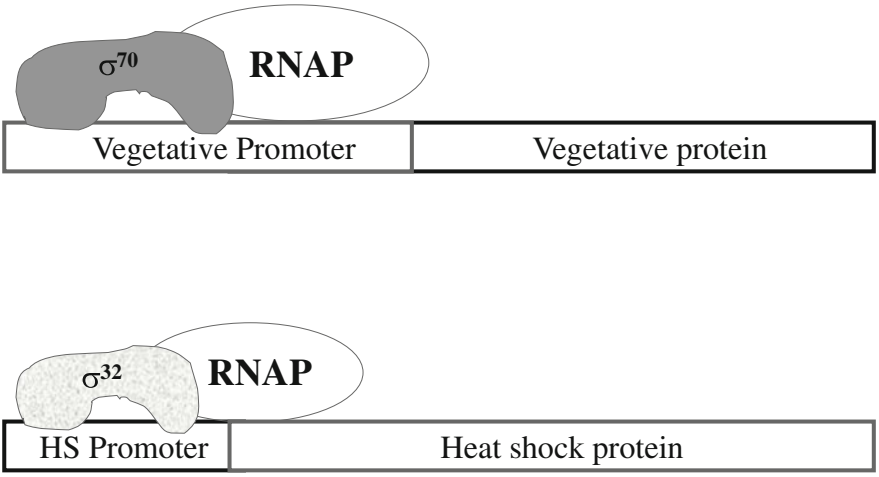


Fig. 2 The *E. coli* Heat-Shock Response. RNA polymerase (RNAP) binds with σ70 to promoters of vegetative genes (top) and with σ32 to promoters of heat-shock genes (bottom)

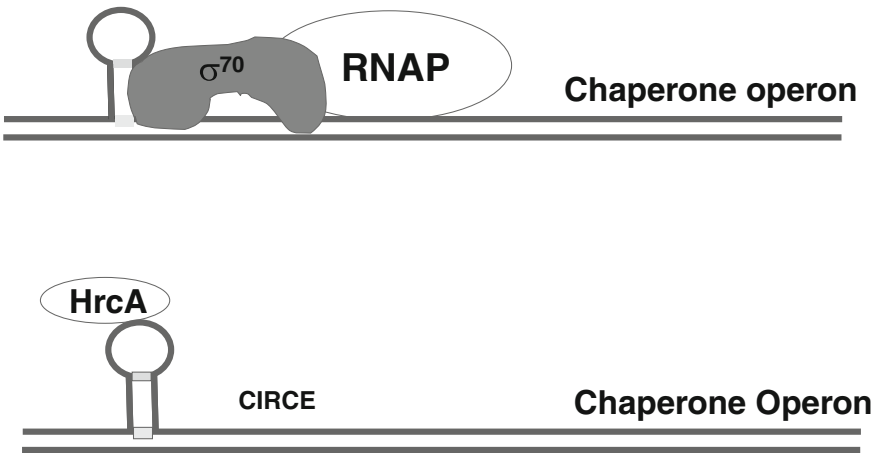


Fig. 3 The *B. subtilis* Heat-Shock Response.RNA polymerase (RNAP) binds with σ70 to promoters of heat-shock chaperone genes at elevated temperatures (top) but not at low temperatures when transcription is inhibited by a CIRCE element repressed by an HcrA repressor (bottom)

repressor is released, thus enabling transcription (Fig. 3) (Hecker et al. 1996; Schumann 2003).

Proteomic studies of *A. tumefaciens* (Rosen and Ron 2011) indicated that there are at least 40 heat-shock-induced proteins, 16 of which are also induced by other stress conditions. These include the set of heat-shock proteins which are conserved in bacteria, such as the Lon and Clp proteases and the DnaK and GroESL chaperones (Boshoff et al. 2008; Rosen et al. 2001, 2002; Rosen and Ron 2002). At least two of the heat-shock proteins, Lon and HspL, are important for virulence (Hwang et al. 2015; Su et al. 2006; Tsai et al. 2009, 2010, 2012). The regulation of the heat-shock response in *A. tumefaciens* is interesting, as there is a combination of the two regulatory systems (Segal and Ron 1996b). The *groESL* operon, encoding Hsp60 and Hsp10, is regulated by the CIRCE inverted repeat and the HrcA protein (Segal and Ron 1993, 1995b, 1996a). However, there is also a heat-shock-specific sigma factor which recognizes a specific heat-shock promoter upstream of heat-shock genes such as *dnaK* (Fig. 4) (Nakahigashi et al. 1998, 1999; Segal and Ron 1995a). These promoters are not similar to the heat-shock promoters of *E. coli*, in contrast to the σ70 promoters which are highly conserved (Fig. 5) (Segal and Ron 1995a). It is interesting to note that *metA* expression is regulated by a σ32 promoter and an S-adenosyl-methionine riboswitch and is poorly transcribed in mutants deleted for σ32 (Rotem et al. 2013). This transcription control directly links the heat-shock response and the growth balancing process under elevated temperature, probably facilitating propagation only after physiological adaptation has been attained.

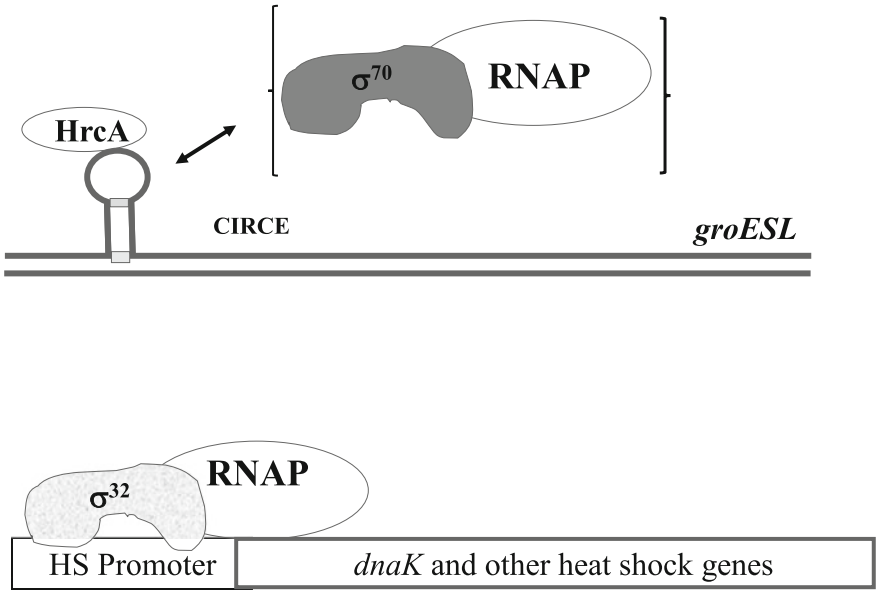


Fig. 4 Heat-Shock Response in *A. tumefaciens*. Transcription of the *groESL* chaperon is regulated by a CIRCE-HcrA element (top) while transcription of other heat-shock genes, such as *dnaK*, is regulated by RNAP bound to σ32 (bottom)

Consensus promoter σ70	TTGACA	‹17›	TATAAT
A. tumefaciens σ32 promoter	CTTG	‹17/18›	CYTAT-T—G
E. coli σ32 promoter	TCTC-CCCTTGAA	‹13/14›	CCCCAT-AT

Fig. 5 Consensus promoters. The σ70 promoter DNA sequence is conserved in bacteria (*E. coli, B. subtilis*, and *A. tumefaciens*) whereas the heat-shock promoters of *E. coli* and *A. tumefaciens* are different

Thus, we already know the basic facts, but there are still numerous unanswered questions:

1. HrcA-CIRCE regulation involves release of the repressor as the temperature increases. However, the heat-shock response is gradual and gets stronger with an increase in temperature. The release of a repressor is an on-off event. So how can the strength of the response depend on the increase of temperature? What are the quantitative aspects of HrcA-CIRCE concentrations as a function of temperature? How is the cellular concentration of HrcA regulated?
2. In *E. coli* σ32 has a stronger affinity for the RNA polymerase than does σ70. However, σ32 is an unstable protein which is stabilized at high temperatures. This observation explains how heat-shock genes are transcribed preferentially upon temperature elevation. However, the σ32 of *A. tumefaciens* is stable at low temperatures (Nakahigashi et al. 1995, 1999), so how can we explain the fact that it transcribes the heat-shock genes only when the temperature is elevated?

Are there additional regulatory elements involved, such as temperature-dependent anti-sigma factors?

3. Is there an evolutionary advantage to having two independent heat-shock control systems, the HcrA-CIRCE and the σ32 (Nakahigashi et al. 1999)?
4. Last but not least: a large fraction of the heat-shock proteins are still induced at high temperatures in mutants deleted for σ32 (Rosen et al. 2002). This finding means that there are additional factors that regulate the heat-shock response in *A. tumefaciens*. One can think of additional sigma factors, yet to be discovered. Alternatively, there may exist posttranscriptional mechanisms which stabilize the transcripts of heat-shock genes. For example, the transcript of *yedU* (*hchA*) in *E. coli* is stabilized at elevated temperatures, and there are RNA chaperones which stabilize transcripts of stress genes during the stress (Mujacic and Baneyx 2006; Rasouly et al. 2007).

3 Conclusions

The heat-shock response is a regulatory network which is critical for the maintenance of life at elevated temperatures. Coping with elevated temperatures involves regulatory mechanisms that control the growth rate as a function of temperature. In many Gram-negative bacteria, including *A. tumefaciens*, regulation of the growth rate is achieved by regulating the availability of methionine via temperature-dependent changes in the first biosynthetic enzyme. In addition to maintenance of regulated growth rate as a function of temperature, there is also a response to the change, the heat-shock response. This response involves the induction of many chaperones and proteases which consist of protein quality control and assure the availability of functional proteins during the change in temperature. *A. tumefaciens* has evolved a complicated control system for the heat-shock response, which includes repressor binding to conserved stem-loop structures upstream of heat-shock genes, as well as a specialized sigma factor which activates specific promoters of heat-shock genes. On top of these two regulatory systems, there are likely additional control systems which are still undiscovered, that are responsible for the temperature-dependent induction of about 20 genes, which are not controlled by either σ32 or the CIRCE-HrcA apparatus.

References

Arsene F, Tomoyasu T, Bukau B (2000) The heat shock response of *Escherichia coli*. Int J Food Microbiol 55:3–9

Biran D, Brot N, Weissbach H, Ron EZ (1995) Heat shock-dependent transcriptional activation of the *metA* gene of *Escherichia coli*. J Bacteriol 177:1374–1379

Biran D, Gur E, Gollan L, Ron EZ (2000) Control of methionine biosynthesis in *Escherichia coli* by proteolysis. Mol Microbiol 37:1436–1443

Boshoff A, Stephens LL, Blatch GL (2008) The *Agrobacterium tumefaciens* DnaK: ATPase cycle, oligomeric state and chaperone properties. Int J Biochem Cell B 40:804–812
Chilton MD, Currier TC, Farrand SK, Bendich AJ, Gordon MP, Nester EW (1974) *Agrobacterium-tumefaciens* DNA and Ps8 bacteriophage DNA not detected in crown gall tumors. Proc Natl Acad Sci USA 71:3672–3676
Christians ES, Yan LJ, Benjamin IJ (2002) Heat shock factor 1 and heat shock proteins: critical partners in protection against acute cell injury. Crit Care Med 30:S43–S50
Craig EA (1985) The heat shock response. CRC Crit Rev Biochem 18:239–280
Erickson JW, Vaughn V, Walter WA, Neidhardt FC, Gross CA (1987) Regulation of the promoters and transcripts of *rpoH*, the *Escherichia coli* heat shock regulatory gene. Genes Develop 1:419–432
Ghazaei C (2017) Role and mechanism of the Hsp70 molecular chaperone machines in bacterial pathogens. J Med Microbiol 66:259–265
Grossman AD, Erickson JW, Gross CA (1984) The *htpR* gene product of *E. coli* is a sigma factor for heat-shock promoters. Cell 38:383–390
Guan N, Li J, Shin HD, Du G, Chen J, Liu L (2017) Microbial response to environmental stresses: from fundamental mechanisms to practical applications. Appl Microbiol Biotechnol 101:3991–4008
Guisbert E, Herman C, Lu CZ, Gross CA (2004) A chaperone network controls the heat shock response in *E. coli*. Genes Develop 18:2812–2821
Gur E, Biran D, Gazit E, Ron EZ (2002) *In vivo* aggregation of a single enzyme limits growth of *Escherichia coli* at elevated temperatures. Mol Microbiol 46:1391–1397
Gur E, Biran D, Ron EZ (2011) Regulated proteolysis in Gram-negative bacteria–how and when? Nature Rev Microbiol 9:839–848
Hecker M, Schumann W, Volker U (1996) Heat-shock and general stress response in *Bacillus subtilis*. Mol Microbiol 19:417–428
Herman C, Thevenet D, D'Ari R, Bouloc P (1995) Degradation of sigma 32, the heat shock regulator in *Escherichia coli*, is governed by HflB. Proc Natl Acad Sci USA 92:3516–3520
Hwang HH, Liu YT, Huang SC, Tung CY, Huang FC, Tsai YL, Cheng TF, Lai EM (2015) Overexpression of the HspL promotes *Agrobacterium tumefaciens* virulence in *Arabidopsis* under heat shock conditions. Phytopathol 105:160–168
Inbar O, Ron EZ (1993) Induction of cadmium tolerance in *Escherichia coli* K-12. FEMS Microbiol Lett 113:197–200
Katz C, Rasouly A, Gur E, Shenhar Y, Biran D, Ron EZ (2009) Temperature-dependent proteolysis as a control element in *Escherichia coli* metabolism. Res Microbiol 160:684–686
Li Z, Menoret A, Srivastava P (2002) Roles of heat-shock proteins in antigen presentation and cross-presentation. Curr Opin Immunol 14:45–51
Lindquist S (1986) The heat-shock response. Ann Rev. Biochem 55:1151–1191
Mathew A, Morimoto RI (1998) Role of the heat-shock response in the life and death of proteins. Ann NY Acad Sci 851:99–111
Mathew A, Shi Y, Jolly C, Morimoto RI (2000) Analysis of the mammalian heat-shock response. Inducible gene expression and heat-shock factor activity. Meth Mol Biol 99:217–255
Michaud S, Marin R, Tanguay RM (1997) Regulation of heat shock gene induction and expression during *Drosophila* development. Cell Mol Life Sci 53:104–113
Mujacic M, Baneyx F (2006) Regulation of *Escherichia coli hchA*, a stress-inducible gene encoding molecular chaperone Hsp31. Mol Microbiol 60:1576–1589
Nakahigashi K, Yanagi H, Yura T (1995) Isolation and sequence analysis of *rpoH* genes encoding sigma 32 homologs from gram negative bacteria: conserved mRNA and protein segments for heat shock regulation. Nucl Acids Res 23:4383–4390
Nakahigashi K, Yanagi H, Yura T (1998) Regulatory conservation and divergence of sigma32 homologs from Gram-negative bacteria: *Serratia marcescens, Proteus mirabilis, Pseudomonas aeruginosa*, and *Agrobacterium tumefaciens*. J Bacteriol 180:2402–2408
Nakahigashi K, Ron EZ, Yanagi H, Yura T (1999) Differential and independent roles of a sigma (32) homolog (RpoH) and an HrcA repressor in the heat shock response of *Agrobacterium tumefaciens*. J Bacteriol 181:7509–7515

Nakahigashi K, Yanagi H, Yura T (2001) DnaK chaperone-mediated control of activity of a sigma (32) homolog (RpoH) plays a major role in the heat shock response of *Agrobacterium tumefaciens*. J Bacteriol 183:5302–5310

Neidhardt FC, Phillips TA, VanBogelen RA, Smith MW, Georgalis Y, Subramanian AR (1981) Identity of the B56.5 protein, the A-protein, and the *groE* gene product of *Escherichia coli*. J Bacteriol 145:513–520

O'Farrell PH (1975) High resolution two-dimensional electrophoresis of proteins. J Biol Chem 250:4007–4021

Ramsay N (1988) A mutant in a major heat shock protein of *Escherichia coli* continues to show inducible thermotolerance. Mol Gen Genet 211:332–334

Rasouly A, Shenhar Y, Ron EZ (2007) Thermoregulation of *Escherichia coli hchA* transcript stability. J Bacteriol 189:5779–5781

Ron EZ (1975) Growth rate of *Enterobacteriaceae* at elevated temperatures: limitation by methionine. J Bacteriol 124:243–246

Ron EZ (2009) An update on the bacterial stress response. Res Microbiol 160:243–244

Ron EZ, Davis BD (1971) Growth rate of *Escherichia coli* at elevated temperatures: limitation by methionine. J Bacteriol 107:391–396

Ron EZ, Shani M (1971) Growth rate of *Escherichia coli* at elevated temperatures: reversible inhibition of homoserine trans-succinylase. J Bacteriol 107:397–400

Ron EZ, Alajem S, Biran D, Grossman N (1990) Adaptation of *Escherichia coli* to elevated temperatures: the *metA* gene product is a heat shock protein. Antonie Van Leeuwenhoek 58:169–174

Rose JK, Rankin CH (2001) Analyses of habituation in *Caenorhabditis elegans*. Learn Mem 8:63–69

Rosen R, Ron EZ (2002) Proteome analysis in the study of the bacterial heat-shock response. Mass Spectrom Rev 21:244–265

Rosen R, Ron EZ (2011) Proteomics of a plant pathogen: *Agrobacterium tumefaciens*. Proteomics 11:3134–3142

Rosen R, Buttner K, Schmid R, Hecker M, Ron EZ (2001) Stress-induced proteins of *Agrobacterium tumefaciens*. FEMS Microbiol Ecol 35:277–285

Rosen R, Buttner K, Becher D, Nakahigashi K, Yura T, Hecker M, Ron EZ (2002) Heat shock proteome of *Agrobacterium tumefaciens*: evidence for new control systems. J Bacteriol 184:1772–1778

Rotem O, Biran D, Ron EZ (2013) Methionine biosynthesis in *Agrobacterium tumefaciens*: study of the first enzyme. Res Microbiol 164:12–16

Schumann W (2003) The *Bacillus subtilis* heat shock stimulon. Cell Stress Chaperones 8:207–217

Schumann W (2016) Regulation of bacterial heat shock stimulons. Cell Stress Chaperones 21:959–968

Segal G, Ron EZ (1993) Heat shock transcription of the *groESL* operon of *Agrobacterium tumefaciens* may involve a hairpin-loop structure. J Bacteriol 175:3083–3088

Segal G, Ron EZ (1995a) The *dnaKJ* operon of *Agrobacterium tumefaciens*: transcriptional analysis and evidence for a new heat shock promoter. J Bacteriol 177:5952–5958

Segal G, Ron EZ (1995b) The *groESL* operon of *Agrobacterium tumefaciens*: evidence for heat shock-dependent mRNA cleavage. J Bacteriol 177:750–757

Segal G, Ron EZ (1996a) Heat shock activation of the *groESL* operon of *Agrobacterium tumefaciens* and the regulatory roles of the inverted repeat. J Bacteriol 178:3634–3640

Segal G, Ron EZ (1996b) Regulation and organization of the *groE* and *dnaK* operons in Eubacteria. FEMS Microbiol Lett 138:1–10

Segal G, Ron EZ (1998) Regulation of heat-shock response in bacteria. Ann NY Acad Sci 851:147–151

Shenhar Y, Rasouly A, Biran D, Ron EZ (2009) Adaptation of *Escherichia coli* to elevated temperatures involves a change in stability of heat shock gene transcripts. Environ Microbiol 11:2989–2997

Srivastava P (2002) Roles of heat-shock proteins in innate and adaptive immunity. Nature Rev Immunol 2:185–194

Straus D, Walter W, Gross CA (1990) DnaK, DnaJ, and GrpE heat shock proteins negatively regulate heat shock gene expression by controlling the synthesis and stability of sigma 32. Genes Developt 4:2202–2209

Su SC, Stephens BB, Alexandre G, Farrand SK (2006) Lon protease of the alpha-proteobacterium *Agrobacterium tumefaciens* is required for normal growth, cellular morphology and full virulence. Microbiol-Sgm 152:1197–1207

Taylor WE, Straus DB, Grossman AD, Burton ZF, Gross CA, Burgess RR (1984) Transcription from a heat-inducible promoter causes heat shock regulation of the sigma subunit of *E. coli* RNA polymerase. Cell 38:371–381

Tomoyasu T, Gamer J, Bukau B, Kanemori M, Mori H, Rutman AJ, Oppenheim AB, Yura T, Yamanaka K, Niki H et al (1995) *Escherichia coli* FtsH is a membrane-bound, ATP-dependent protease which degrades the heat-shock transcription factor sigma 32. EMBO J 14:2551–2560

Tomoyasu T, Ogura T, Tatsuta T, Bukau B (1998) Levels of DnaK and DnaJ provide tight control of heat shock gene expression and protein repair in *Escherichia coli*. Mol Microbiol 30:567–581

Tsai YL, Wang MH, Gao C, Kluesener S, Baron C, Narberhaus F, Lai EM (2009) Small heat-shock protein HspL is induced by VirB protein(s) and promotes VirB/D4-mediated DNA transfer in *Agrobacterium tumefaciens*. Microbiol-Sgm 155:3270–3280

Tsai YL, Chiang YR, Narberhaus F, Baron C, Lai EM (2010) The Small heat-shock protein HspL is a VirB8 chaperone promoting Type IV secretion-mediated DNA transfer. J Biol Chem 285:19757–19766

Tsai YL, Chiang YR, Wu CF, Narberhaus F, Lai EM (2012) One out of Four: HspL but no other small heat shock protein of *Agrobacterium tumefaciens* acts as efficient virulence-promoting VirB8 chaperone. PLoS ONE 7:e49685

Yura T, Nakahigashi K (1999) Regulation of the heat-shock response. Curr Opin Microbiol 2:153–158

Yura T, Kawasaki Y, Kusukawa N, Nagai H, Wada C, Yano R (1990) Roles and regulation of the heat shock sigma factor sigma 32 in *Escherichia coli*. Antonie Van Leeuwenhoek 58:187–190

Zhou YN, Kusukawa N, Erickson JW, Gross CA, Yura T (1988) Isolation and characterization of *Escherichia coli* mutants that lack the heat shock sigma factor sigma 32. J Bacteriol 170:3640–3649

Small Noncoding RNAs in *Agrobacterium tumefaciens*

Keunsub Lee and Kan Wang

Contents

1 Introduction 196
1.1 Discovery of ncRNAs 196
1.2 Classification and Mode of Action 196
1.3 Genome-Wide Identification of ncRNAs 199
2 Identification of Small ncRNAs in *Agrobacterium* 199
3 Functions of *Agrobacterium* ncRNAs 202
3.1 Thi-Box Riboswitch 203
3.2 RepE 203
3.3 AbcR1 204
3.4 RNA1111 205
3.5 Atr35C 206
4 Challenges 208
5 Conclusions 209
References 210

Abstract During the last decade, small noncoding RNAs (ncRNAs) have emerged as essential post-transcriptional regulators in bacteria. Nearly all important physiological and stress responses are modulated by ncRNA regulators, such as riboswitches, *trans*-acting small RNAs (sRNAs), and *cis*-antisense RNAs. Recently, three RNA-seq studies identified a total of 1534 candidate ncRNAs from *Agrobacterium tumefaciens*, a pathogen and biotechnology tool for plants. Only a few ncRNAs have been functionally characterized in *A. tumefaciens*, and some of them appear to be involved in virulence. AbcR1 regulates multiple ABC transporters and modulates uptake of a quorum-sensing inhibitor produced by plants. RNA1111, a Ti plasmid-encoded sRNA, might regulate the dispersal of the Ti

K. Lee · K. Wang (✉)
Crop Bioengineering Center, Iowa State University,
Ames, IA, USA
e-mail: kanwang@iastate.edu

K. Lee · K. Wang
Department of Agronomy, Iowa State University, Ames, IA, USA

Current Topics in Microbiology and Immunology (2018) 418:195–213
DOI 10.1007/82_2018_84

Published Online: 20 March 2018

plasmid and virulence. In addition, a chromosomally encoded sRNA Atr35C is induced by the vir gene regulator VirG and its expression is affected by iron, manganese, and hydrogen peroxide, suggesting a possible role in oxidative stress responses and *Agrobacterium*–plant interactions. Progress in ncRNA functional analysis is slow, likely resulting from innate challenges, such as poor sequence conservation and imperfect base-pairing between sRNAs and mRNAs, which make computational target predictions inefficient. Advances in single-cell-based RNA-seq and proteomics approaches would provide valuable tools to reveal regulatory networks involving ncRNA regulators.

1 Introduction

1.1 Discovery of ncRNAs

Regulatory noncoding RNAs (ncRNAs) have emerged as important regulators of physiological responses in bacteria to survive in ever-changing environments. RNA-mediated response regulation is more advantageous for bacteria than is regulation by proteins (e.g., transcription factors) because it requires less time and energy for synthesis (transcription only vs. transcription and translation) and the responses can be rapidly reversed when needed thanks to a short ncRNA turnover time. Numerous RNA molecules have been discovered that modulate most biological processes and stress responses via various mechanisms. The first studied bacterial small ncRNAs were exosome-encoded antisense RNAs that block plasmid replication (Stougaard et al. 1981; Tomizawa et al. 1981) and inhibit transposon movement (Simons and Kleckner 1983). Although these findings precede the discovery of microRNAs (miRNAs) and small interfering RNAs (siRNAs), the importance of bacterial ncRNAs as regulators had not been much appreciated until the early 2000s when genome-wide identification of chromosomally encoded ncRNAs from *E. coli* and other bacteria were reported (reviewed in Livny and Waldor 2007). Since then, tens to hundreds of candidate ncRNAs have been identified from diverse bacterial species including plant-associated bacteria (reviewed in Becker et al. 2014; Harfouche et al. 2015).

1.2 Classification and Mode of Action

Ribosomal RNAs (rRNAs) and transfer RNAs (tRNAs) are well-characterized ncRNAs regulating protein translation but they are not discussed in this review. We will focus on riboswitches, *trans*-acting small RNAs (sRNAs), and *cis*-antisense RNAs (asRNAs) in this chapter (Waters and Storz 2009). Each group of RNAs uses a variety of mechanisms to modulate physiological and stress responses. Below, we review how these ncRNAs exert regulatory effects in the model bacterial species.

1.2.1 Riboswitches

Riboswitches are part of untranslated regions (UTRs) of mRNAs and affect cognate gene expression at the transcriptional or post-transcriptional levels. Riboswitches are evolutionarily conserved among distantly related bacteria and their functional counterparts are also found in archaea, plants, and fungi. Bacterial riboswitches were first discovered in 2002 as sensors of intracellular small molecules (Mironov et al. 2002; Nahvi et al. 2002; Winkler et al. 2002). Three independent studies reported that part of the mRNA binds to vitamin B derivatives and affects downstream gene expression via transcription attenuation or translation inhibition. Since then, many different riboswitches have been identified and functionally characterized in various bacteria (Winkler and Breaker 2005; Serganov and Nudler 2013). Most riboswitches have two distinct parts: the ligand-binding aptamer domain and the expression platform domain. The aptamer is the sensor region which binds to a specific ligand or metabolite, and the expression platform domain is the response region which adopts alternative structures to affect gene expression. Most well-characterized riboswitches are metabolite sensors and are located in the 5′ UTR of mRNAs encoding enzymes responsible for the biosynthesis the metabolites. Under normal conditions, the ribosome binding site (RBS) is open and accessible to the translation machinery, resulting in the production of functional proteins. Under high metabolite conditions, however, a metabolite binds to the aptamer domain, leading to a conformational change. This change can result in transcriptional attenuation by forming a terminator or translational inhibition by masking the RBS. This type of negative feedback loop prevents overproduction of a specific metabolite, ensuring balanced resource utilization.

Riboswitches are highly selective and many different types of riboswitches have been discovered (reviewed in Serganov and Nudler 2013). A wide range of ligands can be sensed by riboswitches: fluoride anions, metals, purines and their derivatives, cofactors, and amino acids. Recent studies showed that these *cis*-acting regulatory elements can also affect ncRNA expression and modulate RNA–protein interactions (reviewed in Mellin and Cossart 2015). In addition, some riboswitches can act as catalytic enzymes (Tinsley et al. 2007) or as *trans*-acting sRNAs (Loh et al. 2009), suggesting that bacteria have unexpectedly complex regulatory networks involving different classes of ncRNAs.

1.2.2 Trans-encoded sRNAs

sRNAs are 50–500 nt in size and are often encoded in intergenic regions. This group of ncRNAs represents the most well-known RNA regulators and are involved in many physiological and stress responses (Waters and Storz 2009; Gottesman and Storz 2011). sRNAs appear to evolve rapidly, because primary sequence conservation is very limited among closely related bacterial species (Gottesman and Storz 2011). Many sRNAs modulate gene expression via imperfect base-pairing with their target mRNAs, which are transcribed from distinct genomic locations (Waters

and Storz 2009). Another feature of sRNAs is their requirement for the global RNA chaperone Hfq for optimal sRNA–target interactions (De Lay et al. 2013). Hfq-binding affects RNA secondary structure and might facilitate sRNA–mRNA interactions, presumably by binding to both molecules (Gottesman and Storz 2011). Interestingly, however, Hfq is not required for sRNA functions in some bacteria, such as *Staphylococcus aureus* and *Bacillus subtilis*. A recent study by Smirnov et al. (2016) discovered another major sRNA-binding protein, ProQ, which forms stable complexes with small RNAs. It is possible that evolutionarily distant bacteria might have adopted different RNA-binding proteins to facilitate sRNA–mRNA interactions. The exact mechanisms of how sRNAs select and interact with their target mRNAs are still largely unknown.

Another group of sRNAs modulate RNA-binding proteins by sequestering them or directly affecting enzymatic activity (reviewed in Storz et al. 2011). For example, 6S RNA sequesters the house-keeping RNA polymerase (Wassarman and Storz 2000) and the CsrB family sRNAs negatively regulate the activity of CsrA (carbon storage regulator), the regulator of secondary metabolism, by sequestering multiple subunits (Romeo 1998). Many protein-binding sRNAs contain multiple protein-binding sequences, and direct competition by mimicry is the underlying mechanism of these sRNAs (Storz et al. 2011). The presence of multiple protein-binding sRNAs and RNA-binding proteins suggests that bacteria utilize complex sRNA-protein pairs to fine-tune the regulatory networks.

1.2.3 Cis-antisense RNAs

asRNAs are transcribed from the complementary strand of a target gene, and thus exert effects via base-pairing with perfect complementarity (Waters and Storz 2009). The most well-studied asRNAs are encoded on mobile elements, such as plasmids, transposons, and bacteriophages, and maintain proper copy numbers via various mechanisms (Waters and Storz 2009). Common mechanisms are to inhibit plasmid replication by blocking replication primer formation and to inhibit the translation of transposases and protein toxins encoded by these mobile elements (Brantl 2007; Wagner and Simons 1994).

There are an increasing number of asRNAs discovered from diverse bacteria (Georg and Hess 2011; Thomason et al. 2015). However, it is not clear how many of these are actually regulatory asRNAs. Because a low level of pervasive transcription occurs throughout the entire genome (reviewed in Wade and Grainger 2014; Lloréns-Rico et al. 2016), systematic approaches need to be developed to distinguish regulatory asRNAs from these pervasive antisense transcripts. Adding another level of complexity, recent studies discovered that genes and operons encoding proteins performing opposing functions can modulate the expression of genes encoded on the opposite strand (reviewed in Sesto et al. 2013). The total number of asRNAs inventories in bacterial genomes will likely increase in the near future. However, it remains challenging to study how these asRNAs exert regulatory effects, if any, to counter ever-changing environmental stresses.

1.3 Genome-Wide Identification of ncRNAs

Early genome-wide ncRNA identification studies utilized whole genome sequencing data and searched for conserved bacterial promoter and *rho*-independent terminator sequences in the conserved intergenic regions of *E. coli* (reviewed in Livny and Waldor 2007; Livny et al. 2008). The predicted ncRNAs were experimentally validated to prove the effectiveness of the computational predictions. However, the bias toward intergenic regions and low throughput validation procedures limited the thorough discovery of inventories of regulatory RNAs in a genome. High-resolution tiling arrays were successfully used to identify 20–50 sRNAs (Landt et al. 2008; Toledo-Arana et al. 2009) before whole transcriptome sequencing (RNA-seq) became a powerful tool to discover many hundreds of candidate ncRNAs from various bacterial species (Sharma and Vogel 2009). In this chapter, we review recent advances in regulatory ncRNA research in the "natural genetic engineer" *A. tumefaciens* C58 and discuss current challenges and remaining questions.

2 Identification of Small ncRNAs in *Agrobacterium*

Three RNA-seq studies have identified numerous small ncRNAs from *A. tumefaciens* strain C58 thus far (Table 1; Wilms et al. 2012a; Lee et al. 2013; Dequivre et al. 2015). Each study employed some unique approaches, and therefore provided nonredundant identification of novel candidate ncRNAs.

The first genome-wide identification study was done by Wilms et al. (2012a), who used a Roche FLX platform and identified 228 novel ncRNAs. Briefly, total RNA was extracted from *A. tumefaciens* C58 cultures grown under two different conditions: AB minimal medium in the presence (+Vir) or absence (−Vir) of the *vir* gene inducer acetosyringone (AS). To identify transcription start sites (TSS), each total RNA sample was treated or not treated with the Terminator™ 5′-Phosphate-Dependent Exonuclease (TEX), which selectively degrades transcripts containing a 5′-mono-phosphate, to enrich primary transcripts which contain a 5′-tri-phosphate. Four cDNA libraries were constructed and sequenced using a Roche FLX sequencer. The resulting 422,204 cDNA sequences were compared to the C58 reference genome and 348,998 sequences longer than 18 nt were mapped. Sequences mapping to intergenic regions or complementary to protein-coding genes were manually analyzed for ncRNA discovery. Putative ncRNAs were identified if there were a minimum of five cDNA reads in at least one of the four cDNA libraries. A total of 228 candidate ncRNAs were identified from all four replicons: 129 on the circular chromosome, 59 on the linear chromosome, 20 on the pAt plasmid, and 20 on the Ti plasmid. The list also included widely conserved ncRNAs, such as 6S RNA, SRP RNA 4.5S, RNase P, and tmRNA, as well as the previously published *Agrobacterium* ncRNAs RepE, AbcR1, and AbcR2.

Table 1 Comparison of three *Agrobacterium* RNA-seq studies

RNA-seq studies	Wilms et al. (2012a)	Lee et al. (2013)	Dequivre et al. (2015)
Agrobacterium strain	*Agrobacterium tumefaciens* C58	*Agrobacterium tumefaciens* C58	*Agrobacterium tumefaciens* C58
Growth conditions	AB minimal medium with or without AS	YEP complex medium (log and stationary phases); AB minimal medium with or without AS	YPG rich medium (log and stationary phases); AB minimal medium (log and stationary phases)
Total RNA extraction method	Hot acid phenol method (Aiba et al. 1981)	RNeasy Protect Bacteria mini kit	Frozen acid phenol method (Maes and Messens 1992)
Sample treatment	Terminator™ 5′-phosphate-dependent exonuclease	Terminator™ 5′-phosphate-dependent exonuclease; MICROBExpress™ kit	Size fractionation (25–500 nt); First Strand cDNA synthesis kit: RNaseH and DNase I treatment
NGS platform	Roche FLX	Illumina GAII (PE: 2x 50 bp)	Illumina GAII (1x 36 bp)
Total UMR (×1000)	308	48,302	28,386
Identified ncRNAs	228	475	1108
Experimentally validated ncRNAs	22	36	14
Validation methods	RNA gel blot hybridization	RNA gel blot hybridization and RACE	RACE

Twenty-two ncRNAs were validated by RNA-blot hybridization: 10 from the circular chromosome, six from the linear chromosome, two from the pAt plasmid, and four from the pTi plasmid. Among these, 152 were intergenic sRNAs, whereas 76 were antisense to known protein-coding genes (asRNAs). Several independently validated ncRNAs were differentially expressed under varying growth conditions such as medium, temperature, and pH. One ncRNA encoded by the Ti plasmid, Ti2, was highly induced under Vir gene induction conditions, and its expression was diminished in *virA* and *virG* deletion mutants.

Lee et al. (2013) identified 475 highly expressed candidate ncRNAs under four different growth conditions: YEP logarithmic and stationary phases and AB minimal medium in the presence or absence of AS. Because ribosomal RNAs represent the vast majority of the total cellular RNA (He et al. 2010), two commercial kits were used to deplete rRNAs and tRNAs, the TEX and MICROBExpress™ kits. All four total RNA samples were treated with reagents in the MICROBExpress™ kit, which removed ∼55% of the 16S and 23S rRNAs, followed by TEX treatment

(±TEX). RNA samples were fragmented to ~200–300 nt before cDNA library construction. A total of eight cDNA libraries (four growth conditions; ±TEX) were sequenced using the Illumina GAII platform. A total of 842 million 50-bp reads were obtained and 48.3 million reads were mapped exactly once to the reference genome (Uniquely Mapped Reads, UMRs). These UMRs were used for data analysis. The use of TEX treatment substantially improved the UMR ratio from 7.5 (−TEX) to 12.5% (+TEX), indicating RNA-seq sensitivity was considerably enhanced. The highly expressed ncRNA identification procedure began with calculating the depth of coverage data for each individual nucleotide position on both forward/reverse strands of all four replicons. Candidate ncRNAs were identified in the intergenic regions or complementary sequences of protein-coding genes, where the average depth of coverage of a candidate ncRNA region was at least 10 times greater than those of immediate upstream and downstream regions. A total of 101 sRNAs and 310 asRNAs were identified, as well as 20 5′ UTR leader sequences. Thirty-six ncRNAs were experimentally validated by RNA-blot hybridization and RACE (Rapid Amplification of cDNA Ends)-PCR (Gerhart et al. 2014). Twenty-two ncRNAs were differentially expressed by Vir gene induction: 15 were up-regulated and seven were down-regulated. Fourteen of the 15 AS-induced ncRNAs contain a putative *vir* box, a conserved motif for VirG binding, in the promoter regions. In addition to the identified ncRNAs, a stranded RNA-seq approach also revealed interesting features of *Agrobacterium* transcripts: (i) highly expressed antisense transcripts from the complementary strands of important *vir* genes and operons and (ii) novel transcripts within the known protein-coding genes (e.g., virD4* internal transcript; Lee et al. 2013). It is likely that many putative asRNAs might have been ignored due to the high stringency of the informatics cutoff, i.e., a minimum of 10 times higher expression level compared to adjacent regions, whereas internal transcripts had not been considered for ncRNA identification.

Most recently, Dequivre et al. (2015) conducted another RNA-seq study using size-fractionated RNA samples (25–500 nt). *A. tumefaciens* strain C58 was grown under four different growth conditions: logarithmic/stationary phases in YPG rich medium and in AB minimal medium. tRNAs were depleted using a First Strand cDNA synthesis kit. tRNA-specific primers were used to synthesize the first strand cDNAs, and RNaseH was used to degrade tRNAs in the RNA–DNA duplex, followed by DNase I treatment. A total of 193.1 million reads were obtained and 28.4 million reads were mapped once to the reference genome (UMRs). Genomic regions whose average depth was at least 10 times greater than the adjacent regions were considered transcribed, and candidate ncRNAs were identified only when a transcript was presented in all four libraries. A total of 1108 candidate ncRNAs were evenly distributed among all four replicons: 602 on the circular chromosome, 291 on the linear chromosome, 140 on the pAt plasmid, and 75 on the Ti plasmid. Four hundred and seven were intergenic sRNAs and 262 were asRNAs. Additionally, 402 and 37 were derived from 5′ and 3′ UTRs, respectively. Seventeen candidate ncRNAs were independently validated by RACE-PCR. An intergenic sRNA encoded by the Ti plasmid, RNA1111, was conserved among

Table 2 Summary of the three RNA-seq identifications of ncRNAs in *A. tumefaciens* C58

Dataset	Number of ncRNAs exclusively identified by the combined dataset		
	Wilms et al. (2012a)	Lee et al. (2013)	Dequivre et al. (2015)
Wilms et al. (2012a)	98	28[a]	48
Lee et al. (2013)	28[a]	300	94
Dequivre et al. (2015)	48	94	912
All three datasets[b]	54	54	54
Total	228	475	1108

[a]A sRNA identified by Lee et al. (2013) corresponds to two tandemly encoded sRNAs identified by Wilms et al. (2012a)

[b]Fifty four ncRNAs (3.5%) were identified by all three studies, while 170 (11.1%) and 1310 (85.4%) were identified by two and one studies, respectively

other Ti plasmids, and the deletion mutant exhibited reduced tumorigenicity in tomato, suggesting the involvement of this sRNA in bacterial virulence.

These three genome-wide RNA-seq studies identified a total of 1534 candidate ncRNAs from *A. tumefaciens* C58. As summarized in Table 2, 54 (3.5%) candidate ncRNAs were identified by all three studies, whereas 170 (11.1%) and 1310 (85.4%) ncRNAs were identified by two and one studies, respectively. The numbers presented in Table 2 are slightly different from those presented in the Venn Diagram of Dequivre et al. (2015) due to two small differences. First, one sRNA identified by Lee et al. (2013) corresponds to two tandemly encoded sRNAs identified by Wilms et al. (2012a); thus the total number of ncRNAs identified by Wilms et al. (2012a) was 228, not 227. The second was a simple calculation error, as the total number of ncRNAs is 1534 (=98 + 300 + 912 + 28 + 48 + 94 + 54; Table 2) not 1560 (Dequivre et al. 2015). As each RNA-seq investigation employed unique approaches, the collective efforts led to a thorough inventory of *Agrobacterium* ncRNAs. Functional analyses of these ncRNAs, however, have not been comprehensive; only several conserved ncRNAs have been characterized in *Agrobacterium* and other closely related species.

3 Functions of *Agrobacterium* ncRNAs

Although over a thousand candidate ncRNAs have been discovered from *A. tumefaciens*, the regulatory functions of all but a few remain unknown. Only a handful of ncRNAs have been functionally characterized thus far: *repE* (Chai and Winans 2005), *AbcR1* (Wilms et al. 2011), a TPP riboswitch (Lee et al. 2013), and *RNA1111* (Dequivre et al. 2015). Here we describe how these ncRNAs have been discovered and how they exert regulatory functions via various mechanisms. We also report how a chromosomally encoded and AS-induced sRNA, Atr35C (Lee et al. 2013), is expressed under iron deficiency and oxidative stress conditions.

3.1 Thi-Box Riboswitch

Thiamine, also known as vitamin B1, is an essential coenzyme for carbohydrate and branched-chain amino acid metabolism in all living cells. Maintaining a proper level of thiamine is critical and a highly conserved RNA structure called the Thi-box riboswitch or TPP (thiamine pyrophosphate) riboswitch regulates the biosynthesis and transport of thiamine in bacteria, archaea, and eukaryotes (Serganov and Nudler 2013; RF00059 in Rfam database). The Thi-box riboswitch binds to TPP to cause RNA structural changes which can lead to transcriptional attenuation or translational inhibition (Serganov and Nudler 2013). Three TPP riboswitches have been identified in *A. tumefaciens* C58, two on the circular and one on the linear chromosome (Rfam database: http://rfam.xfam.org/search?q=Agrobacterium%20fabrum%20AND%20rna_type:%22riboswitch%22%20and%20TPP%20AND%20alignment_type:%22full%22). All three Thi-box riboswitches are located in the 5′ UTR of operons encoding putative thiamine biosynthesis enzymes or transporters. Lee et al. (2013) demonstrated, using Northern Blot analysis, that a Thi-box riboswitch (C1_2541934R) located in the 5′ UTR of the thiamine biosynthesis operon *thiCOGG* indeed regulates gene expression via transcriptional attenuation. The *thiCOGG* mRNA was detected when *A. tumefaciens* was grown in minimal medium lacking thiamine, but only the ∼110 nt riboswitch accumulated when grown in nutrient-rich medium-containing thiamine, suggesting that the *thiCOGG* promoter has constitutive activity and a transcriptional attenuator is formed to block transcription of the full-length mRNA of the thiamine biosynthesis genes (Lee et al. 2013). Thi-box riboswitch-mediated transcriptional attenuation was also observed in the nitrogen-fixing bacterium *R. etli* (Miranda-Ríos et al. 2001).

In addition to the Thi-box riboswitches, the *A. tumefaciens* C58 genome was predicted to encode six Cobalamin (vitamin B12), two SAM (S-Adenosyl Methionine), one Flavin mononucleotide (FMN; vitamin B2), and one glycine riboswitches (http://rfam.xfam.org/search?q=Agrobacterium%20fabrum%20AND%20rna_type:%22riboswitch%22), but their functional roles have not yet been confirmed.

3.2 RepE

The first characterized sRNA in *Agrobacterium* was *RepE*, a sRNA that regulates the replication of an octopine-type tumor-inducing Ti plasmid (Chai and Winans 2005). *RepE* is encoded in the intergenic region of the *repABC* operon, whose products are responsible for the replication of the Ti plasmid (Chai and Winans 2005). RepABC-type replication is widespread among plasmids found in alpha-proteobacteria, especially in Rhizobiales (Palmer et al. 2000). All known

repABC operons include at least three genes: *repA*, *repB*, and *repC* (reviewed in Cevallos et al. 2008). RepA and RepB are involved in plasmid partitioning and segregation, whereas RepC is responsible for initiation of the DNA synthesis. Chai and Winans (2005) demonstrated that *RepE* is ~54 nt in size and suppresses the replication of a mini-Ti plasmid when expressed in *trans* (Chai and Winans 2005). In addition, mutations introduced at the promoter region resulted in downregulation of *RepE*, which subsequently increased plasmid copy number, further suggesting that *RepE* is a negative regulator of Ti plasmid replication. *RepE* likely form duplexes with *repC* mRNAs resulting in transcriptional attenuation (Brantl et al. 2002). Because the *repE*-encoded intergenic region is highly conserved in other repABC-type plasmids, it is likely that *repE*-mediated transcriptional attenuation is an important mechanism to maintain plasmid copy numbers in Rhizobiales.

3.3 *AbcR1*

The *AbcR1* (ABC regulator) was the first studied chromosomally encoded sRNA in α-proteobacteria (Wilms et al. 2011). *AbcR1* was discovered by a computational search (Wilms et al. 2011) in the conserved intergenic region between *atu2186* and *atu2187*, in tandem with a homologous sRNA *AbcR2*. Both *AbcR1* and *AbcR2* are well conserved in α-proteobacteria and belong to the α*r15* sRNA family (del Val et al. 2012). *AbcR1*/*AbcR2* orthologues have been identified in other α-proteobacteria: *Sinorhizobium meliloti* (*SmrC15*/*SmrC16*; del Val et al. 2007), *Rhizobium etli* (*ReC58*/*ReC59*; Vercruysse et al. 2010), and *Brucella abortus* 2308 (*AbcR1*/*AbcR2*; Caswell et al. 2012). Hfq is likely required for *AbcR1*-meditated negative regulation of atleast for some target genes because their expression levels were elevated in both *hfq* and *abcR1* knockout mutants (Wilms et al. 2012b).

AbcR1 regulons have been identified by one- and two-dimensional PAGE analysis (Wilms et al. 2011; Overlöper et al. 2014) or computational predictions using the CopraRNA algorithm (Wright et al. 2013). *AbcR1* regulates at least 16 mRNAs including several periplasmic substrate-binding proteins required for sugar and amino acid ABC transporters (Wilms et al. 2011, 2012b; Overlöper et al. 2014): AtpH, AttC, Atu0857, Atu1879, Atu2422, Atu3114, Atu4046, Atu4259, Atu4431, Atu4577, Atu4678, ChvE, DppA, FrcB, and NocT. Among these, several target genes are involved in *A. tumefaciens* virulence. Atu2422 encodes a periplasmic protein which is responsible for uptake of the plant defense molecule γ-amino butyric acid (GABA) (Chevrot et al. 2006). GABA can suppress the quorum-sensing signal within *A. tumefaciens*, thus attenuating bacterial virulence (Chevrot et al. 2006). ChvE is a sugar-binding protein that senses host-released sugars and directly interacts with the VirA/VirG two-component system to induce *vir* gene expression (He et al. 2009; Hu et al. 2013). AttC and NocT are responsible for the uptake of spermidine/putrescine and nopaline, respectively (Matthysse et al. 1996).

AbcR1 possesses two separate target-binding regions (Overlöper et al. 2014) and each region binds to a set of target mRNAs either near the ribosomal binding site (RBS) to block translation and accelerate target mRNA turnover, or the coding DNA sequence (CDS) to cause transcriptional attenuation (Wilms et al. 2011, 2012b; Overlöper et al. 2014). Further studies may greatly expand the AbcR1 regulon because a large number of proteins differentially expressed by an *avcR1* deletion have yet to be validated (Overlöper et al. 2014).

Recent studies showed that regulation by *AbcR1*/*AbcR2* orthologues has become diversified in α-proteobacteria. In *A. tumefaciens*, *AbcR1*/*AbcR2* have near identical promoter sequences and are highly expressed in late stationary phase, but only *AbcR1* has regulatory functions (Wilms et al. 2011). Conversely, *AbcR1*/*AbcR2* orthologues in the human pathogen *B. abortus* 2308 have some redundant functions, because only the *abcR1*/*abcR2* double knockout mutant exhibited reduced survival in cultured murine macrophages (Caswell et al. 2012). *B. abortus AbcR1*/*AbcR2* have multiple unique and shared target genes (Caswell et al. 2012). In the nitrogen-fixing bacterium *S. meliloti* Rm1021, *AbcR1*/*AbcR2* orthologues (SmrC15/SmrC16) are divergently expressed: *AbcR1* was expressed in actively growing cells but was not detected in stationary phase, whereas *AbcR2* was highly expressed in the stationary phase and under various stress conditions (Torres-Quesada et al. 2014). Together, these data suggest that *AbcR1*/*AbcR2* orthologues may have evolved rapidly in α-proteobacteria, but it is not yet known whether AbcR1 regulons in the plant pathogenic *A. tumefaciens* are also evolutionarily conserved in the human pathogen *B. abortus* or the nitrogen-fixing symbionts *S. meliloti* and *R. etli*.

3.4 RNA1111

RNA1111 is a recently identified sRNA from the intergenic region between *atu6186* (*virE3*) and *Atu6188* (*virE0*) on the complementary strand (Dequivre et al. 2015). *RNA1111* is ∼173 nt in length and highly conserved among the nopaline-type Ti plasmids. Although *RNA1111* was located within the *vir* gene region, its expression level was not affected by *vir* gene induction conditions (Dequivre et al. 2015). Interestingly, however, an *rna1111* deletion mutant exhibited reduced virulence on tomato plants: an *rna1111* mutant strain harboring an empty expression vector produced an average of two tumors per plant, whereas the wild-type and *rna1111* mutant harboring the complementation construct produced 20 and 9.5 tumors per plant, respectively. The complementation construct alone does not restore a full level of virulence, presumably because the deleted *rna1111* gene region contains the *vir* box of *virE0*.

Because *RNA1111* may be involved in *A. tumefaciens* virulence, the next step was to identify the regulatory targets of this sRNA. Three sRNA target search programs (RNApredator, sTarPicker, and IntaRNA) were utilized to identify a total of eight putative target genes, which were predicted by all three programs. Six candidate target genes were encoded on the pTiC58 plasmid, including three

virulence-related genes (*6b*, *virC2,* and *virD3*), two conjugal transfer genes (*traA* and *trbD*), and a gene encoding a hypothetical protein (*atu6072*). Interestingly, Möller et al. (2014) found that *virC2*, *virD3,* and *traA* mRNAs were enriched by Hfq tagged by 3xFlag. Further studies are needed to determine if *RNA1111* interacts with Hfq to regulate its putative target genes.

Quantitative reverse transcription PCR (qRT-PCR) analyses showed that *trbD* RNA was not detectable, and *virC2*/*virD3* did not show altered expression in the *rna1111* mutant compared to the wild-type strain. Three genes, *6b*, *traA*, and *atu6072*, however, exhibited significantly lower expression levels in the *rna1111* mutant. Importantly, *6b*, *traA*, and *atu6072* expression levels were not different in the *rna1111* mutant strain harboring the complementation construct from those in the wild-type strain, further suggesting that *RNA1111* might stabilize these target mRNAs or protect them from degradation. Together, these results suggest that *RNA1111* might regulate genes involved in *A. tumefaciens*–plant interactions as well as in the dispersal of the Ti plasmid.

3.5 Atr35C

In our previous RNA-seq study, we identified 475 candidate ncRNAs from *A. tumefaciens* C58 (Lee et al. 2013). Fifteen of these were up-regulated by the *vir* gene inducer acetosyringone, and among these was a chromosomally encoded sRNA, C2_132595F (=*Atr35C*), which belongs to the *αr35* sRNA family (http://rfam.xfam.org/family/ar35). *Atr35C* is encoded in the intergenic region between *atu3124* (hypothetical protein) and *atu3126* (hypothetical protein) on the linear chromosome. The first *αr35* RNA family member, *Smr35B*, was identified from the symbiotic bacterium *S. meliloti* 1021 by computational prediction and experimental validation (del Val et al. 2007). A comparative genomics approach suggested that this sRNA family is conserved among certain members of the order Rhizobiales, which include both symbiotic (e.g., *R. etli and R. leguminosarum*) and pathogenic species (e.g., *A. tumefaciens* and *Ochrobactrum anthropi*; del Val et al. 2012).

The expression of *αr35* RNA was first reported in *S. meliloti* 1021 (del Val et al. 2007) and interestingly, it was induced by luteolin, the plant flavone that induces nodulation genes, suggesting a possible role during host–bacterial interactions. Similarly, our previous RNA-seq study found that *Atr35C* is induced by the *vir* gene inducer AS (Lee et al. 2013). qRT-PCR analysis confirmed that *Atr35C* is indeed induced by AS (Fig. 1a). To verify further that *Atr35C* is regulated by VirG, a *virG* mutant was generated as previously described (Lee et al. 2013), and *A. tumefaciens* strains were grown in induction medium (IM) containing (+Vir) or lacking AS (−Vir). qRT-PCR analysis showed that *Atr35C* expression was 21-fold lower in the *virG* mutant than in the wild-type strain in the presence of AS, and *Atr35C* expression was 15-fold higher in the presence of AS in the wild-type strain (Fig. 1a). By comparison, in our previous RNA-seq study, the *Atr35C* level was 6.1-fold higher in the presence of AS in the wild-type strain C58 (Lee et al. 2013,

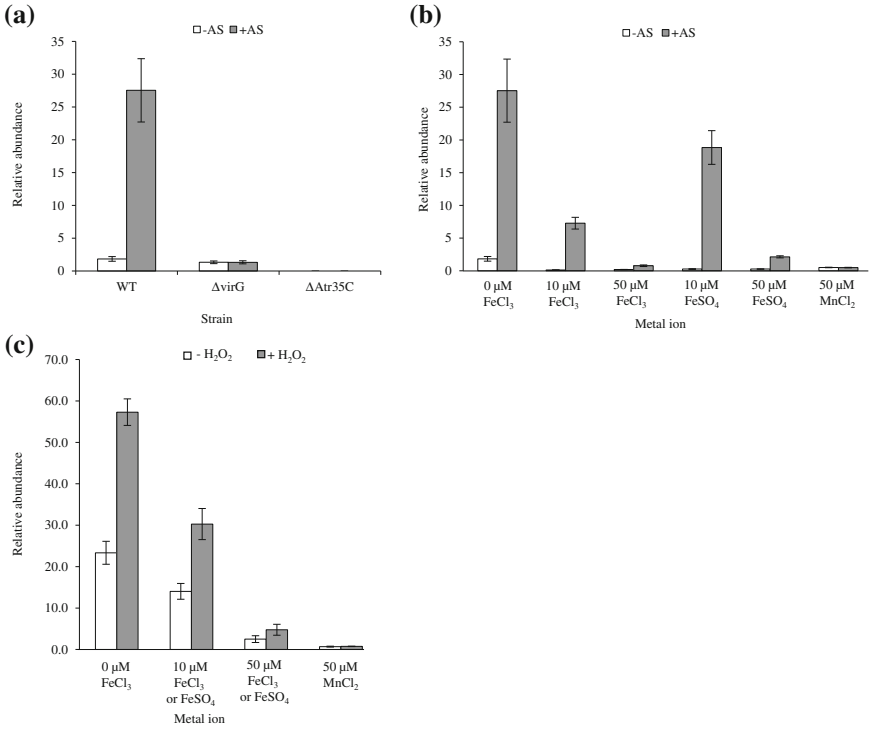


Fig. 1 *Atr35C* transcript levels were estimated by RT-qPCR using the $2^{-\Delta\Delta CT}$ method (Livak and Schmittgen 2001) as described previously (Lee et al. 2013). **a** *Atr35C* transcript level was not induced by AS in the *virG* mutant, suggesting VirG-dependent expression. **b** *Atr35C* expression was negatively correlated with the concentration of iron and manganese ions. **c** *Atr35C* transcript level was further enhanced by 9 mM hydrogen peroxide (H_2O_2). Error bars represent standard errors

Table S4A. C2_132595F). Together, these results strongly suggest that the chromosomally encoded sRNA *Atr35C* is regulated by VirG.

Because *Atr35C* appears to be part of the VirG regulon, we examined if an *atr35C* mutant has altered virulence. However, neither transient GUS expression using an *Arabidopsis* seedling assay (Wu et al. 2014) nor tumorigenicity assay using tobacco leaf disks (Clemente 2006) showed significant differences between the *atr35C* mutant and wild-type C58 strains.

In search of environmental/stress stimuli that trigger *Atr35C* expression, several transition metals were added to the IM, and qRT-PCR assays were used to monitor *Atr35C* transcript levels. When added to a concentration of 100 μM, $FeCl_3$ and $MnCl_2$ greatly reduced *Atr35C* expression levels in the presence of AS (Fig. 1b), whereas $CuSO_4$ and $ZnSO_4$ did not have a significant impact. Because typical IM contains 10 μM $FeSO_4$ (Gelvin 2006), we tested if there were a dosage effect of iron and manganese. Addition of 10 μM ferric ($FeCl_3$) and ferrous ($FeSO_4$) irons to

IM resulted in a mild reduction of *Atr35C* expression levels by about 3.8- (27.5 vs. 7.3) and 1.5-fold (27.5 vs. 18.9), respectively. A higher iron concentration further reduced *Atr35C* transcript levels: addition of 50 μM $FeCl_3$ and $FeSO_4$ reduced *Atr35C* transcript levels by 35.3- (27.5 vs. 0.8) and 12.9-fold (27.5 vs. 2.1), respectively. Thus, *A. tumefaciens* is more responsive to ferric than to ferrous iron (3.8- vs. 1.5-fold changes at 10 μM; 35.3- vs. 12.9-fold change at 50 μM). Addition of 50 μM $MnCl_2$ reduced *Atr35C* transcript levels by 55-fold (27.5 vs. 0.5). These results strongly suggest that *Atr35C* might be involved in iron and manganese homeostasis.

Interestingly, iron and manganese play important roles in oxidative stress responses and virulence in *A. tumefaciens* (Saenkham et al. 2008; Kitphati et al. 2007). We therefore tested if hydrogen peroxide (H_2O_2), a primary defense molecule of plants (Wojtaszek 1997; Dan et al. 2015), affects *Atr35C* expression. Wild-type *A. tumefaciens* C58 was grown in the presence of AS for 24 h and 9 mM H_2O_2 was added to the culture and further incubated for 30 min. *Atr35C* transcript levels increased by ∼twofold in the presence of 9 mM H_2O_2 compared to the control with 0, 10, or 50 μM $FeCl_3$ or $FeSO_4$ (Fig. 1c). However, *Atr35C* transcript levels were not affected by 9 mM H_2O_2 in the presence of 50 μM $MnCl_2$. Taken together, our results suggest that there is cross-talk between the Ti plasmids and the chromosomally encoded sRNA *Atr35C*, which might be involved in oxidative stress responses or iron/manganese homeostasis. Further studies are needed to identify the target genes regulated by *Atr35C* and to elucidate how this sRNA exerts regulatory functions.

4 Challenges

Regulatory ncRNAs are versatile and provide bacteria many adaptive advantages in rapidly changing environments. As mentioned above, however, the biological functions of most ncRNAs remain largely unknown; only a small number of ncRNAs have been functionally characterized in *A. tumefaciens*. This can be attributed to the characteristics of ncRNAs and their interactions with targets: (1) poor sequence conservation in homologous ncRNAs, (2) imperfect complementarity in sRNA–mRNA base-pairing, and (3) quantitative changes in target gene expression.

Bacterial ncRNA homologs have a low level of primary sequence conservation among evolutionarily distantly related species. Consequently, most ncRNA homologs are only found among closely related bacteria. This observation significantly limits data mining, which can provide useful information such as conserved domains, putative functions, and interactions with putative targets and transcription factors. For instance, although 1534 candidate ncRNAs have been discovered in *A. tumefaciens* C58 thus far, only 45 families are found in the Rfam database (http://rfam.xfam.org/search?q=UP000000813%20AND%20entry_type:%22Family%22). Among these are 5S rRNA (RF00001), RNase P RNA (RF00010), SRP RNA

(RF00169), and 6S RNA (RF00013). Other than these broadly conserved ncRNAs, α-proteobacterial ncRNA families, such as αr7 (RF02342), αr9 (RF02343), αr14 (RF02344), αr15 (RF02345), αr35 (RF02346), and αr45 (RF02347) still lack known functions. Extended searches for conserved secondary structures and adjacent protein-coding genes have proven useful to facilitate homologous ncRNA discovery (Barrick et al. 2005), but it still remains challenging to identify functional analogs among distantly related bacteria.

In contrast to eukaryotic miRNAs and siRNAs that base-pair with target mRNAs with near-perfect complementarity (Brodersen et al. 2008), bacterial ncRNAs, especially sRNAs, interact with target mRNAs via base-pairing with a less perfect complementarity and with gaps (Storz et al. 2011). This finding poses a difficult challenge to identify sRNA targets using existing computational algorithms (Pain et al. 2015). Even validated sRNA-mRNA target pairs are not predicted as top candidates (Pain et al. 2015), which strongly suggests that there are unknown crucial factors determining sRNA-mRNA specificities, or that current computational algorithms need further optimization. Many RNA-seq-based approaches have recently been developed to identify sRNA targets (reviewed in Saliba et al. 2017), but these approaches are costly and time consuming for extensive optimization and data analyses. Undoubtedly, additional experimentally validated ncRNA–target interactions will improve ncRNA target prediction algorithms in the future, but it is important to expand the search algorithms to include protein databases because some ncRNAs directly interact with protein targets.

As post-transcriptional regulators, some ncRNAs do not dramatically alter target gene transcript levels, whereas others only affect target mRNA translation without altering mRNA stability (Storz et al. 2011). In addition, as demonstrated by Levine et al. (2007), ncRNA-mediated gene regulation is largely affected by the rate of transcription of the target genes. Therefore, it is crucial to define the conditions under which a specific ncRNA exerts regulatory effects on target gene expression. Moreover, a high level of heterogeneity exists among the individual cells in bacterial colonies (Martins and Locke 2015), but standard procedures measure only the average levels in a population. In this regard, single-cell-based analyses may provide useful platforms to measure precisely the regulatory effects of ncRNAs on target gene expression. Recent advances in single-cell-based RNA-seq and proteomics look promising to provide more accurate genome-wide pictures of complex regulatory networks, including ncRNA regulators (Shapiro et al. 2013; Martins and Locke 2015).

5 Conclusions

RNA-seq approaches allowed identification of 1534 candidate ncRNAs from *A. tumefaciens* C58 (Table 2). This is, however, only the beginning of the regulatory ncRNA era, and a number of questions remain unanswered. For example, how many ncRNAs are true regulators? Do asRNAs represent important regulators or

mere by-products of transcriptional noise? Which ncRNAs, if any, modulate *Agrobacterium*–plant interactions? Although there are many challenges for ncRNA research, accumulating evidence has solidified the importance of ncRNA regulators for many aspects of biological reactions and stress responses. Technical advances, such as single-cell-based RNA-seq and proteomics, will provide new tools to reveal how ncRNAs specify targets, both RNAs and proteins, and how multiple layers of regulatory networks interact harmoniously with one another to maximize bacterial fitness. This in turn offers an excellent opportunity to improve the efficiency and host-range of *A. tumefaciens*-mediated plant genetic transformation.

Acknowledgements The authors thank Abbagail Johnson, Juan Carlos Martinez-Nicolas, Alan Eggenberger, and Jonah Miller for their assistances. This work was partially supported by the USDA National Institute of Food and Agriculture, Hatch project number #IOW05162, by State of Iowa funds, and by Charoen Pokphand Indonesia.

References

Aiba H, Adhya S, de Crombrugghe B (1981) Evidence for two functional gal promoters in intact *Escherichia coli* cells. J Biol Chem 256:11905–11910

Barrick JE, Sudarsan N, Weinberg Z et al (2005) 6S RNA is a widespread regulator of eubacterial RNA polymerase that resembles an open promoter. RNA 11:774–784

Becker A, Overlöper A, Schlüter JP et al (2014) Riboregulation in plant-associated-proteobacteria. RNA Biol 11:550–562

Brantl S (2007) Regulatory mechanisms employed by *cis*-encoded antisense RNAs. Curr Opin Microbiol 10:102–109

Brantl S, Wagner E, Gerhard H (2002) An antisense RNA-mediated transcriptional attenuation mechanism functions in *Escherichia coli*. J Bacteriol 184:2740–2747

Brodersen P, Sakvarelidze-Achard L, Bruun-Rasmussen M et al (2008) Widespread translational inhibition by plant miRNAs and siRNAs. Science 30:1185–1190

Caswell CC, Gaines JM, Ciborowski P et al (2012) Identification of two small regulatory RNAs linked to virulence in *Brucella abortus* 2308. Mol Microbiol 85:345–360

Cevallos MA, Cervantes-Rivera R, Gutiérrez-Ríos RM (2008) The repABC plasmid family. Plasmid 60:19–37

Chai Y, Winans SC (2005) A small antisense RNA downregulates expression of an essential replicase protein of an *Agrobacterium tumefaciens* Ti plasmid. Mol Microbiol 56:1574–1585

Chevrot R, Rosen R, Haudecoeur E et al (2006) GABA controls the level of quorum-sensing signal in *Agrobacterium tumefaciens*. Proc Natl Acad Sci USA 103:7460–7464

Clemente T (2006) *Nicotiana* (*Nicotiana tobaccum, Nicotiana benthamiana*) In: Wang K (ed) *Agrobacterium* protocols, 2nd edn. Humana Press Inc., New Jersey, pp 143–154

Dan Y, Zhang S, Zhong H et al (2015) Novel compounds that enhance *Agrobacterium*-mediated plant transformation by mitigating oxidative stress. Plant Cell Rep 34:291–309

De Lay N, Schu D, Gottesman S (2013) Bacterial small RNA-based negative regulation: Hfq and its accomplices. J Bio Chem 288:7996–8003

del Val C, Rivas E, Torres-Quesada O et al (2007) Identification of differentially expressed small non-coding RNAs in the legume endosymbiont *Sinorhizobium meliloti* by comparative genomics. Mol Microbiol 66:1080–1091

del Val C, Romero-Zaliz R, Torres-Quesada O et al (2012) A survey of sRNA families in α-proteobacteria. RNA Biol 9:119–129

Dequivre M, Diel B, Villard C et al (2015) Small RNA deep-sequencing analyses reveal a new regulator of virulence in *Agrobacterium fabrum* C58. Mol Plant-Microbe Interact 28:580–589

Gelvin SB (2006) *Agrobacterium* virulence gene induction. In: Wang K (ed) *Agrobacterium* protocols, 2nd edn. Humana Press Inc, New Jersey, pp 77–84

Georg J, Hess WR (2011) *cis*-antisense RNA, another level of gene regulation in bacteria. Microbiol Mol Biol Rev 75:286–300

Gerhart E, Wagner H, Vogel J (2014) Approaches to identify novel non-messenger RNAs in bacteria to investigate their biological functions: functional analysis of identified non-mRNAs. In: Hartmann RK, Bindereif A, Schön A, Westhof E (eds) Handbook of RNA biochemistry, 2nd edn. Wiley-VCH Verlag GmbH & Co., Weinheim, pp 719–786

Gottesman S, Storz G (2011) Bacterial small RNA regulators: versatile roles and rapidly evolving variations. Cold Spring Harb Perspect Bio 3:1–16

Harfouche L, Haichar F, Achouak W (2015) Small regulatory RNAs and fine-tuning of plant-bacteria interactions. New Phytol 206:98–106

He F, Nair GR, Soto CS et al (2009) Molecular basis of ChvE function in sugar binding, sugar utilization, and virulence in *Agrobacterium tumefaciens*. J Bacteriol 191:5802–5813

He S, Wurtzel O, Singh K et al (2010) Validation of two ribosomal RNA removal methods for microbial metatranscriptomics. Nat Methods 7:807–812

Hu X, Zhao J, DeGrado WF et al (2013) *Agrobacterium tumefaciens* recognizes its host environment using ChvE to bind diverse plant sugars as virulence signals. Proc Natl Acad Sci USA 110:678–683

Kitphati W, Ngok-Ngam P, Suwanmaneerat S et al (2007) *Agrobacterium tumefaciens fur* has important physiological roles in iron and manganese homeostasis, the oxidative stress response, and full virulence. Appl Environ Microbiol 73:4760–4768

Landt SG, Abeliuk E, McGrath PT et al (2008) Small non-coding RNAs in *Caulobacter crescentus*. Mol Microbiol 68:600–614

Lee K, Huang X, Yang C et al (2013) A genome-wide survey of highly expressed non-coding RNAs and biological validation of selected candidates in *Agrobacterium tumefaciens*. PloS One 8:e70720

Levine E, Zhang Z, Kuhlman T et al (2007) Quantitative characteristics of gene regulation by small RNA. PLoS Biol 5(9):e229

Livak KJ, Schmittgen TD (2001) Analysis of relative gene expression data using real-time quantitative PCR and the $2^{-\Delta\Delta CT}$ method. Methods 25:402–408

Livny J, Waldor MK (2007) Identification of small RNAs in diverse bacterial species. Curr Opin Microbiol 10:96–101

Livny J, Teonadi H, Livny M et al (2008) High-throughput, kingdom-wide prediction and annotation of bacterial non-coding RNAs. PLoS ONE 3:e3197

Lloréns-Rico V, Cano J, Kamminga T et al (2016) Bacterial antisense RNAs are mainly the product of transcriptional noise. Sci Adv 2(3):e1501363

Loh E, Dussurget O, Gripenland J et al (2009) A trans-acting riboswitch controls expression of the virulence regulator PrfA in Listeria monocytogenes. Cell 139:770–779

Maes M, Messens E (1992) Phenol as grinding material in RNA preparations. Nucleic Acids Res 20:4374

Martins BMC, Locke JCW (2015) Microbial individuality: how single-cell heterogeneity enables population level strategies. Curr Opin Microbiol 24:104–112

Matthysse AG, Yarnall HA, Young N (1996) Requirement for genes with homology to ABC transport systems for attachment and virulence of *Agrobacterium tumefaciens*. J Bacteriol 178:5302–5308

Mellin JR, Cossart P (2015) Unexpected versatility in bacterial riboswitches. Trends Genet 31:150–156

Miranda-Ríos J, Navarro M, Soberón M (2001) A conserved RNA structure (thi box) is involved in regulation of thiamin biosynthetic gene expression in bacteria. Proc Natl Acad Sci USA 98:9736–9741

Mironov AS, Gusarov I, Rafikov R et al (2002) Sensing small molecules by nascent RNA: a mechanism to control transcription in bacteria. Cell 111:747–756
Möller P, Overlöper A, Förstner KU et al (2014) Profound impact of Hfq on nutrient acquisition, metabolism and motility in the plant pathogen *Agrobacterium tumefaciens*. PLoS ONE 9: e110427
Nahvi A, Sudarsan N, Ebert M et al (2002) Genetic control by a metabolite binding mRNA. Chem Biol 9:1043–1049
Overlöper A, Kraus A, Gurski R et al (2014) Two separate modules of the conserved regulatory RNA AbcR1 and address multiple target mRNAs in and outside of the translation initiation region. RNA Biol 11:624–640
Pain A, Ott A, Amine H et al (2015) An assessment of bacterial small RNA target prediction programs. RNA Biol 12:509–513
Palmer KM, Turner SL, Young JPW (2000) Sequence diversity of the plasmid replication gene *repC* in the *Rhizobiaceae*. Plasmid 44:209–219
Romeo T (1998) Global regulation by the small RNA-binding proteins CsrA and the non-coding RNA molecule *CsrB*. Mol Microbiol 29:1321–1330
Saenkham P, Utamapongchai S, Vattanaviboon P et al (2008) *Agrobacterium tumefaciens* iron superoxide dismutases have protective roles against singlet oxygen toxicity generated from illuminated Rose Bengal. FEMS Microbiol Lett 289:97–103
Saliba AE, Santos SC, Vogel J (2017) New RNA-seq approaches for the study of bacterial pathogens. Curr Opin in Microbiol 35:78–87
Serganov A, Nudler E (2013) A decade of riboswitches. Cell 152:17–24
Sesto N, Wurtzel O, Archambaud C et al (2013) The excludon: A new concept in bacterial antisense RNA-mediated gene regulation. Nat Rev Microbiol 11:75–82
Shapiro E, Biezuner T, Linnarsson S (2013) Single-cell sequencing-based technologies will revolutionize whole-organism science. Nat Rev Genet 14:618–630
Sharma CM, Vogel J (2009) Experimental approaches for the discovery and characterization of regulatory small RNAs. Curr Opin Microbiol 12:536–546
Simons RW, Kleckner N (1983) Translational control of IS10 transposition. Cell 34:683–691
Smirnov A, Förstner KU, Holmqvist E et al (2016) Grad-seq guides the discovery of ProQ as a major small RNA-binding protein. Proc Natl Acad Sci USA 113:11591–11596
Storz G, Vogel J, Wassarman KM (2011) Regulation by small RNAs in bacteria: expanding frontiers. Mol Cell 43:880–891
Stougaard P, Molin S, Nordström K (1981) RNAs involved in copy number control and incompatibility of plasmid R1. Proc Natl Acad Sci USA 78:6008–6012
Thomason MK, Bischler T, Eisenbart SK et al (2015) Global transcriptional start site mapping using differential RNA sequencing reveals novel antisense RNAs in *Escherichia coli*. J Bacteriol 197:18–28
Tinsley RA, Furchak JRW, Walter NG (2007) Trans-acting *glmS* catalytic riboswitch: locked and loaded. RNA 13:468–477
Toledo-Arana A, Dussurget O, Nikitas G et al (2009) The *Listeria* transcriptional landscape from saprophytism to virulence. Nature 459:950–956
Tomizawa J, Itoh T, Selzer G et al (1981) Inhibition of ColE1 RNA primer formation by a plasmid-specified small RNA. Proc Natl Acad Sci USA 78:1421–1425
Torres-Quesada O, Reinkensmeier J, Schlűter JP et al (2014) Genome-wide profiling of HFq-binding RNAs uncovers extensive post-transcriptional rewiring of major stress response and symbiotic regulons in *Sinorhizobium meliloti*. RNA Biol 11:563–579
Vercruysse M, Fauvart M, Cloots L et al (2010) Genome-wide detection of predicted non-coding RNAs in *Rhizobium etli* expressed during free-living and host-associated growth using a high resolution tiling array. BMC Genom 11:53
Wade JT, Grainger DC (2014) Pervasive transcription: illuminating the dark matter of bacterial transcriptomes. Nat Rev Microbiol 12:647–653
Wagner EG, Simons RW (1994) Antisense RNA control in bacteria, phages, and plasmids. Annu Rev Microbiol 48:713–742

Wassarman KM, Storz G (2000) 6S RNA regulates *E. coli* RNA polymerase activity. Cell 101:613–623

Waters LS, Storz G (2009) Regulatory RNAs in bacteria. Cell 13:615–628

Wilms I, Voss B, Hess WR et al (2011) Small RNA-mediated control of the *Agrobacterium tumefaciens* GABA binding protein. Mol Microbiol 80:492–506

Wilms I, Overlöper A, Nowrousian M et al (2012a) Deep sequencing uncovers numerous small RNAs on all four replicons of the plant pathogen *Agrobacterium tumefaciens*. RNA Biol 9 (4):446–457

Wilms I, Möller P, Stock AM et al (2012b) Hfq influences multiple transport systems and virulence in the plant pathogen *Agrobacterium tumefaciens*. J Bacteriol 194:5209–5217

Winkler WC, Breaker RR (2005) Regulation of bacterial gene expression by riboswitches. Annu Rev Microbiol 59:487–517

Winkler W, Nahvi A, Breaker RR (2002) Thiamine derivatives bind messenger RNAs directly to regulate bacterial gene expression. Nature 419:952–956

Wojtaszek P (1997) Oxidative burst: an early plant response to pathogen infection. Biochem J 322:681–692

Wright PR, Richter AS, Papenfort K et al (2013) Comparative genomics boosts target prediction for bacterial small RNAs. Proc Natl Acad Sci USA 110:E3487–E3496

Wu HY, Liu KH, Wang YC et al (2014) AGROBEST: an efficient *Agrobacterium*-mediated transient expression method for versatile gene function analyses in *Arabidopsis* seedlings. Plant Methods 10:19

The *Agrobacterium* Type VI Secretion System: A Contractile Nanomachine for Interbacterial Competition

Chih-Feng Wu, Delaney A. Smith, Erh-Min Lai and Jeff H. Chang

Contents

1 Introduction 216
2 Functions of the T6SS 217
3 The Structure of the T6SS 218
4 Structure of the T6SS of *Agrobacterium* 220
5 Loading Effectors onto the T6SS of *A. tumefaciens* C58 221
6 Biochemical Activity of the T6SS-Secreted Effectors of *A. tumefaciens* C58 222
7 Ecological Role of T6SS 223
8 Regulation of the T6SS of *Agrobacterium* 224
9 Summary 226
References 227

Abstract The bacterial type VI secretion system (T6SS) is a contractile nanomachine dedicated to delivering molecules out of bacterial cells. T6SS-encoding loci are in the genome sequences of many Gram-negative bacteria, and T6SS has been implicated in a plethora of roles. In the majority of cases, the T6SSs deliver effector proteins in a contact-dependent manner to antagonize other bacteria. Current models suggest that the effectors are deployed to influence social interactions in microbial communities. In this chapter, we describe the structure, function, and regulation of the T6SS and its effectors. We provide focus on the T6SS of *Agrobacterium tumefaciens*, the causative agent of crown gall disease, and relate the role of the T6SS to the ecology of *A. tumefaciens*.

C.-F. Wu · D. A. Smith · J. H. Chang (✉)
Department of Botany and Plant Pathology, Oregon State University, Corvallis, OR, USA
e-mail: changj@science.oregonstate.edu

E.-M. Lai
Institute of Plant and Microbial Biology, Academia Sinica, Taipei 11529, Taiwan

J. H. Chang
Center for Genome Research and Biocomputing, Oregon State University, Corvallis, OR, USA

Current Topics in Microbiology and Immunology (2018) 418:215–231
DOI 10.1007/82_2018_99

Published Online: 11 July 2018

1 Introduction

There are multiple secretion systems that can be used by Gram-negative bacteria to secrete or translocate proteins into the extracellular milieu or into other cells (Chang et al. 2014). The deployed proteins can increase the fitness of the delivering cell by mediating cooperative behavior with kin, engaging in antagonistic or competitive behavior with genetically distinct microbes, driving evolution by exchanging nucleic acids, exploiting eukaryotic hosts for nutrients, and mediating the mutualistic reciprocation of services with eukaryotic hosts.

We focus on the type VI secretion system (T6SS), a nanomachine that is structurally similar to contractile injection systems of bacteriophages (Basler 2015; Bonemann et al. 2010; Cianfanelli et al. 2016b; Leiman et al. 2009; Zoued et al. 2014). A T6SS-encoding locus designated as *imp* (based on the corresponding mutant being impaired in nitrogen fixation) was first discovered in *Rhizobium leguminosarum* Biovar *trifolii*, a lineage of plant mutualistic bacteria in the *Rhizobiaceae* family (Bladergroen et al. 2003; Roest et al. 1997). Similarly, a T6SS locus was discovered and implicated in pathogenesis of the fish pathogen *Edwardsiella tarda*, but the term type VI secretion system (T6SS) was not coined until 2006 upon discovery of a homologous gene cluster in *Vibrio cholerae* (Pukatzki et al. 2006; Rao et al. 2004). Since that discovery, T6SS-encoding loci have been identified in more than 25% of the sequenced genomes of various taxa of Gram-negative bacteria and there has been significant progress in understanding the mechanisms and ecological roles of T6SSs (Boyer et al. 2009). However, efforts have focused largely on select species, and the roles that T6SS provide to plant-associated bacteria have not been examined to the same depth and breadth as some other taxa of bacteria.

Agrobacteria are motile, Gram-negative alphaproteobacteria. These soil bacteria are attracted to roots. They can irreversibly attach, proliferate, and form community structures, such as a biofilm, on the surface of roots (Barton et al. 2018). Populations with cells carrying oncogenic plasmids can cause disease if plant tissue is wounded. Wounds provide ingression points and may also release phenolic compounds, which when coupled to the acidic environment of the rhizosphere trigger a virulence program that relies on the unusual ability of inter-kingdom gene transfer (Gelvin, 2000). The transferred genes on T-(transferred) DNA lead to the overproduction of plant growth promoting hormones, misregulation of growth, and the synthesis of opines. Because cognate opine transport and catabolism genes are located distal to T-DNA on the oncogenic Ti (tumor inducing) plasmid of the infecting bacteria, the "opine hypothesis" suggests that opines are specific carbon and nitrogen sources that provide the infecting cells a fitness advantage. However, catabolism of opines has been demonstrated in other taxa of bacteria, and 16S amplicon sequencing of crown gall tissues suggests that communities of many species of bacteria co-exist in diseased tissues (Canfield and Moore 1991; Faist et al. 2016). Hence, agrobacteria, whether in the soil or in association with healthy or diseased eukaryotes, are likely in environments that require mechanisms to compete with other microbes for resources.

Agrobacteria are diverse, polyphyletic, and classified on the basis of multiple schemes (Costechareyre et al. 2010; Farrand et al. 2003; Lassalle et al. 2011). We will rely on the traditional classification scheme to discriminate between *Agrobacterium* and *Rhizobium* (Farrand et al. 2003). In addition, we use *Agrobacterium tumefaciens* (biovar 1) to discriminate from *Agrobacterium rhizogenes* and *Agrobacterium radiobacter* (biovar 2). The distinction between biovar 1 and biovar 2 is important because only members of the former encode T6SS (Slater et al. 2009). It is important to note that not all sequenced isolates of *A. tumefaciens* have T6SS-encoding loci. In some cases, entire genomospecies appear to lack T6SS-encoding loci, whereas in other genomospecies T6SS loci are polymorphic in presence/absence (Davis, Wu, Weisberg, and Chang unpublished). Genomospecies are genetically defined clusters within the *A. tumefaciens* group (Costechareyre et al. 2010). Lastly, each of the narrow host-range pathogens, *Agrobacterium vitis*, *Agrobacterium larrymoorei*, and *Agrobacterium rubi*, are predicted to encode T6SS. There have been few, if any, reported studies of the T6SS in the different genospecies or lineages of narrow host-range agrobacteria.

2 Functions of the T6SS

The primary function of T6SSs is to launch protein effectors into physically contacted cells. The T6SS-associated effectors are broadly defined into classes based on how they are associated with the delivery apparatus. Specialized effectors are identified on the basis of being covalently fused to proteins that are core components of the secretion apparatus (Cianfanelli et al. 2016b; Durand et al. 2014). Cargo effectors are not covalently associated and require other mechanisms to associate with the secretion apparatus. Early findings showed that some specialized effectors have domains indicative of being virulence factors that function within eukaryotic cells. VgrG-1, which is the spike of the T6SS (described in the next section), of *Vibrio cholerae* has a carboxy-terminal domain that can crosslink actin and mediate inflammation to promote survival or replication in mice (Ma and Mekalanos 2010; Pukatzki et al. 2007). Similarly, VgrG1 of *Aeromonas hydrophila* harbors an actin ADP-ribosylation domain and can disrupt the actin cytoskeleton and induce host cell toxicity (Suarez et al. 2010). Effectors of *Edwardsiella tarda* and *Francisella tularensis* have also been implicated in virulence of eukaryotes (Chen et al. 2017; Rigard et al. 2016).

A key finding reported that Tse1, a T6SS-delivered effector of *Pseudomonas aeruginosa*, is an amidase that degrades peptidoglycan, the main component of the cell wall of bacteria (Hood et al. 2010; Russell et al. 2011). Several subsequently identified T6SS-associated effectors can degrade the cell membrane or nucleic acids of bacteria, but some effectors with novel targets have also been identified (Ma et al. 2014 Russell et al. 2013). Nonetheless, the prokaryotic targets of most T6SS-associated effectors appear to be general features of bacterial cells and, as a consequence, effectors can cause self-intoxication. To protect against this, effector

genes are typically encoded adjacent to cognate immunity-encoding genes, i.e., effector–immunity pairs (Russell et al. 2014). For example, the gene adjacent to *Tse1* encodes a protein that confers immunity and protects Tse1-encoding bacteria from their own toxic protein (Russell et al. 2011). Loci-encoding effector–immunity pairs are vertically inherited, and immunity proteins can provide protection against T6SS-mediated attacks by sibling cells (Russell et al. 2014). These key discoveries led to a flood of findings that collectively pointed toward T6SSs in influencing bacterial social interactions such as competition, kin recognition, swarming, and biofilm formation (Alteri et al. 2013; Russell et al. 2014).

T6SSs have also been implicated in other activities that can give bacteria an advantage over competitors. T6SS activity influences horizontal gene transfer and can potentially drive the evolution of T6SS-using lineages of bacteria. In *V. cholerae*, genes encoding the T6SS and competence proteins, which are components necessary for the uptake of nucleic acids, are co-expressed in chitin-rich environments. When grown in culture, attacking cells can take up DNA released from lysed cells (Borgeaud et al. 2015). Uptake of DNA by cultured bacteria can also diversify the spectrum of T6SS-encoding effector–immunity pairs and provide a competitive advantage over neighboring cells (Thomas et al. 2017). T6SSs have been implicated in nutrient acquisition and, in these roles, do not necessarily function in a contact-dependent manner. Nucleic acids have nutritional value, and the uptake of DNA could provide access to carbon and nitrogen. Moreover, several findings have shown that metal chelators can be released into the extracellular milieu via T6SS. The metal chelators pyoverdine of *Pseudomonas taiwanensis*, YezP of *Yersinia pseudotuberculosis*, and TseM of *Burkholderia thailandensis,* for example, help bacteria acquire micronutrients such as iron, zinc, and manganese, respectively (Chen et al. 2016; Si et al. 2017; Wang et al. 2015). In addition, TseF secreted by a T6SS of *P. aeruginosa* facilitates iron acquisition by interacting with the iron-binding Pseudomonas quinolone signal (Lin et al. 2017). In all, the T6SS is a versatile nanomachine that is employed for multiple purposes and by various taxa of bacteria with diverse lifestyles.

3 The Structure of the T6SS

Within members of the proteobacteria phylum, T6SSs are associated with 13 proteins, named Type Six Subunits (Tss) A-M, that are considered core to the secretion system (Boyer et al. 2009). Many of these Tss proteins have structural similarity to proteins that form the bacteriophage T4 tail and, as a consequence, the T6SS is predicted to assemble into three major subcomplexes that come together to form a membrane-associated inverted phage tail-like structure (Fig. 1; Basler 2015; Leiman et al. 2009). The cell envelope-spanning complex consists of the integral membrane proteins TssL and TssM, which are located in the inner membrane and associate with the lipoprotein TssJ (Felisberto-Rodrigues et al. 2011; Ma et al. 2009). The lipoprotein connects the inner membrane-associated complex to the

outer membrane. The baseplate complex is comprised of TssA, TssE, TssF, TssG, and TssK (Brunet et al. 2015; Planamente et al. 2016). This complex associates with the membrane complex and functions as a platform for the third subcomplex, which resembles a contractile tail-like structure. This third subcomplex consists of TssB, TssC, Hcp (TssD), and VgrG (TssI). The formerly listed two proteins form an outer sheath, whereas Hcp and VgrG form an inner tube structure and a spike consisting of a trimer of molecules, respectively (Lossi et al. 2013; Mougous et al. 2006; Pukatzki et al. 2007). ClpV (TssH), an AAA-type ATPase, is also core to T6SS (Kapitein et al. 2013; Schlieker et al. 2005). The PAAR protein, which "sharpens" the spike, has recently been suggested to be a core protein on the basis that PAAR-encoding genes are found in all T6SS loci and PAAR proteins are necessary for T6SS function (Burkinshaw et al. 2018; Cianfanelli et al. 2016a). Other proteins associated with the T6SS are lineage-specific and regulate, assemble, or modify the T6SS (Hsu et al. 2009).

Current models suggest that T6SSs deliver effectors by rapidly contracting the cytoplasmic outer sheath, thereby extending the Hcp inner tube outward (Fig. 1; Basler et al. 2012). On the tips of Hcp tubes are VgrG puncturing devices, which are additionally sharpened by the PAAR-containing proteins and decorated with effectors (Shneider et al. 2013). This invokes images of the types of gruesome weapons that were used during the Middle Ages. In addition, effectors can be loaded within the hollow of the tube formed by Hcp (Silverman et al. 2013). The Hcp and VgrG proteins, despite being integral to the functionality of the T6SS, are also released into the extracellular milieu and prey cells and are thus often used as

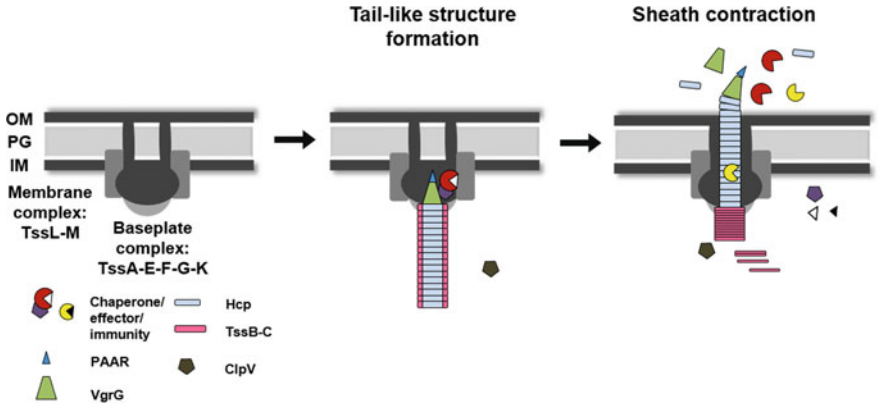


Fig. 1 Model of the T6SS biogenesis pathway. The membrane complex associates with the baseplate complex and spans the inner and outer membranes (left panel). The VgrG spike-PAAR assembles in association with the baseplate complex and the inner Hcp tube and outer TssB-C sheath complex (middle panel). Contraction of the outer sheath propels the Hcp tube outward, delivering and releasing the spike, Hcp, and effectors (right panel). The contracted sheath is subsequently disassembled by the ClpV ATPase. Immunity proteins remain within the cell and protect against attack by cognate effectors and self-intoxication. OM = outer membrane; PG = peptidoglycan; IM = inner membrane

markers for T6SS activity (Vettiger and Basler 2016). After firing, the contracted TssB/TssC tubular polymers are rapidly disassembled and recycled by the ClpV AAA-type ATPases (Kapitein et al. 2013).

Regardless of whether an effector is specialized or cargo, there is specificity between an effector and the Hcp tube or VgrG spike protein (Bondage et al. 2016; Silverman et al. 2013). This specificity is easy to comprehend for the specialized effectors, as they are covalently attached to one of the T6SS-associated components. For the cargo effectors, specificity is hypothesized to be conferred by adaptor proteins that associate a cognate effector to a specific VgrG molecule. These adaptor proteins tend to encode a DUF4123, DUF2169, or DUF1795 domain, which function to load cognate effectors onto the VgrG spike protein and are also reported to stabilize the effectors (Bondage et al. 2016; Cianfanelli et al. 2016a; Ma et al. 2014; Unterweger et al. 2015). It is often the case that the genes encoding the VgrG–PAAR–adaptor–effector–immunity protein partners are genetically linked and co-transcribed.

4 Structure of the T6SS of *Agrobacterium*

The T6SS of *A. tumefaciens* has primarily been examined in one isolate, C58 of genomospecies G8 (also known as *A. fabrum*; Lassalle et al. 2011). The T6SS was discovered in *A. tumefaciens* on the basis of high levels of Hcp that accumulated extracellular to bacteria grown in acidic medium (Wu et al. 2008). The T6SS-encoding locus of *A. tumefaciens* C58 is located on the secondary chromid and consists of two divergently transcribed operons (Fig. 2; Wu et al. 2008). The *imp* operon has 14 genes (*atu4343* to *atu4330*), and a systematic deletion of coding sequences showed that 11 of the 13 genes encode proteins that are core to T6SS of proteobacteria (Lin et al. 2013). These proteins form the main structure of the T6SS, the transmembrane complex, baseplate complex, and outer sheath. The *imp* operon also encodes accessory proteins that are implicated in mediating post-translational regulation of the T6SS, which is described in a subsequent section (Lin et al. 2014, 2018).

The *hcp* operon, which encodes nine genes (*atu4344* to *atu4352*), is the second operon of the T6SS of *A. tumefaciens*. Four of the encoded proteins are core to T6SS of proteobacteria (Lin et al. 2013). These are the Hcp tube, VgrG spike, the ClpV ATPase, and a PAAR protein. The *hcp* operon also encodes an adaptor and two sets of effector–immunity pairs. The genes encoding Tae and its cognate immunity protein are located just downstream of *hcp*. The second effector–immunity gene pair, *tde1* and *tdi1*, is located downstream of *vgrG* (Ma et al. 2014). Isolate C58 has a second *vgrG* locus located on the secondary chromid but distal to the *imp* operon. To distinguish between the two copies of *vgrG*, the one located within the *hcp* operon is designed *vgrG1*, and the other is designed *vgrG2* (Lin et al. 2013). The latter operon includes the effector–immunity gene pair *tde2* and *tdi2* (Ma et al. 2014). The genetic linkage between *vgrGs* and *tde–tdi* pairs is

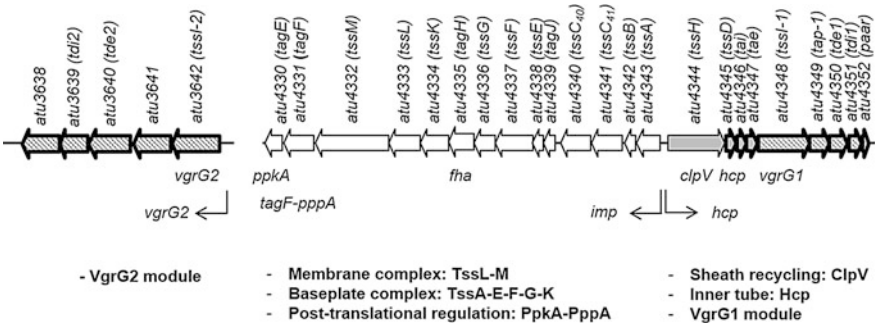


Fig. 2 T6SS-encoding gene clusters of *A. tumefaciens* C58. The *imp* (*atu4343–atu4330*) and *hcp* (*atu4344–atu4352*) gene clusters are divergently expressed operons that encode the main structures of the T6SS. The *vgrG2* (*atu3642–atu3638*) operon encodes a second VgrG module, complete with spike, adaptor, effector and immunity pair, and a hypothetical protein. The Tss or tag names, proposed by Shalom et al. (2007), are indicated above the annotated names, whereas specific names are listed at the bottom of the figure. The *Hcp–Tae* and *vgrG–effector* encoding modules are outlined in bold

common to many T6SS-encoding proteobacteria (Ma et al. 2014). Surveys of the C58 genome sequence failed to reveal additional paralogs of *vgrG*. Moreover, a polymutant deleted of the three effector–immunity gene pairs is as compromised in its ability to antagonize cells as is the *tssL* mutant, which cannot assemble the T6SS (Ma et al. 2014). Collectively, these results suggest that isolate C58 encodes only three T6SS effector–immunity pairs.

5 Loading Effectors onto the T6SS of *A. tumefaciens* C58

Tde1 and Tde2 are cargo effectors that are specific to VgrG1 and VgrG2, respectively. Yet, VgrG1 and VgrG2 are functionally redundant with respect to secretion of Hcp or Tae (Bondage et al. 2016; Lin et al. 2013). The *vrgG1* locus is associated with *tap-1*, a gene that encodes an adaptor with a DUF4123 domain. Evidence suggests that Tap-1 is necessary to stabilize Tde1 and link Tde1 to VgrG1 (Bondage et al. 2016). The *vrgG2* locus is associated with a gene that encodes a DUF2169 domain. Hence, it is hypothesized that the DUF2169 domain fulfills the role of adaptor and binds to the PAAR-like domain of Tde2 to associate the effector to the VgrG2 spike protein (Bondage et al. 2016).

The two VgrG proteins of *A. tumefaciens* C58 are 816 and 754 amino acids long, respectively (Bondage et al. 2016). They share ~92% amino acid identity, but sequence identity is biased toward the amino-termini and VgrG1 has an additional extension on its carboxy-terminal end. Truncation of the VgrG proteins immediately after the most conserved regions abrogates their functionality, as neither Hcp nor effector could be detected in the extracellular fraction of culture-grown bacteria.

Fusions with the C-terminal end of the other VgrG protein regained functionality but with altered specificity. Evidence suggested that VgrG2 with the C-terminus of VgrG1 secreted Tde1, whereas VgrG1 with the C-terminus of VgrG2 secreted Tde2. Bondage et al. (2016) also showed that secretion of Tde1 specifically requires its cognate Tap-1 chaperone, whereas Tde2 requires its cognate DUF2169-containing protein. Hence, the *vgrG*–chaperone–effector–immunity loci of *A. tumefaciens* C58 encode functional units in which the C-termini of the VgrG proteins are necessary for specificity. Given that the spike of the T6SS consists of a trimer of VgrG, it is still unclear whether a single T6SS apparatus consists of all one VgrG isoform and delivers each effector individually, or a 1:2 or 2:1 combination of VrgG + Tde1: VgrG2 + Tde2.

The *Tae* gene is not immediately downstream of a *vgrG* gene, suggesting that its mechanism of secretion is different. Indeed, unlike Tde1 that requires both Hcp and a cognate VgrG, secretion of Tae requires only Hcp and no specific VgrG (Bondage et al. 2016). Second, Tae co-immunoprecipitates with Hcp (Lin et al. 2013). Thus, Tae of *A. tumefaciens* may be delivered in a manner similar to that of Tse2 of *P. aeruginosa*, which is hypothesized to be loaded into the internal pore of the Hcp tube (Silverman et al. 2013).

6 Biochemical Activity of the T6SS-Secreted Effectors of *A. tumefaciens* C58

Queries using the amino acid sequence of Tde1 revealed its membership in a superfamily with a toxin_43 domain (Ma et al. 2014). Members of this superfamily have HxxD, a putative catalytic motif, and are suggested to be nucleases. Indeed, in vitro and heterologous expression of Tde1 in *E. coli* confirmed its ability to degrade DNA but not RNA and showed the necessity of the HxxD motif for activity (Ma et al. 2014). Tde2 also has a C-terminal toxin_43 domain, and heterologous expression of Tde2 in *E. coli* led to the degradation of DNA. In a seminal study, Ma et al. (2014) showed that both of the Tde effectors can provide *A. tumefaciens* a competitive fitness advantage when the bacteria are directly co-inoculated with susceptible prey cells into leaves of the model plant, *Nicotiana benthamiana* (Ma et al. 2014). The advantage was observed when *A. tumefaciens* C58 was paired with susceptible genotypes of *A. tumefaciens* C58 or *P. aeruginosa* as prey.

The function of the Tae effector is less clear. Tae accumulates in the supernatant fluid of bacteria, and deletion of its corresponding gene has no effect on Hcp secretion (Lin et al. 2013). On the basis of homology, Tae is hypothesized to be an amidase. In addition, Tae has two motifs that are conserved in proteins predicted to hydrolyze bonds of peptidoglycan and is homologous to a T6SS-secreted effector of *Serratia marcescens* (English et al. 2012; Lin et al. 2013). Yet, toxicity toward bacteria has not yet been demonstrated and when present as the sole effector, Tae has no significant activity in conferring a competitive *in planta* advantage against

susceptible genotypes of *A. tumefaciens* C58 or *P. aeruginosa* (Ma et al. 2014). Tae could potentially target specific species of prey or could be a "helper" that in specific situations promotes toxicity by facilitating the delivery of Tde1 and Tde2. Helper proteins have been associated with the type III secretion system and are secreted and hypothesized to assist with the translocation of other effector proteins (Chang et al. 2014).

7 Ecological Role of T6SS

In agar plate-based assays, T6SSs provide a quantifiable competitive advantage over susceptible genotypes. The T6SSs are very efficient when *V. cholerae*, *Acinetobacter baumannii*, and *Pseudomonas putida* compete against *E. coli* as a prey (Bernal et al. 2017; Carruthers et al. 2013; Zheng et al. 2011). However, T6SSs and their effectors display different degrees of toxicity (Chatzidaki-Livanis et al. 2016; Ma et al. 2014; Schwarz et al. 2010). Moreover, the agar plate-based assays typically rely on starting with dense cultures in which the predator genotype is present at a greater concentration than are susceptible prey genotypes. For *A. tumefaciens*, a fitness advantage is often observed when it competes at a significantly higher ratio relative to the prey cell (Ma et al. 2014). It is also often the case that T6SS-dependent competition is tested between bacterial genotypes that are mixed to increase the potential for physical interactions between predator and prey cells. There is also an example of some specificity between effector and species of prey. The T6SS-1 of *B. thailandensis* is only efficacious against *Pseudomonas putida*, *Pseudomonas fluorescens*, and *Serratia proteamaculans,* which represent only 3 of 31 tested genotypes of bacteria selected to represent the diversity of proteobacteria and Gram-positive bacteria (Schwarz et al. 2010). Therefore, it is not clear whether the "winner-take-all" model accurately reflects the role of the T6SS in the natural ecology of bacteria.

In natural ecosystems, T6SS-mediated interactions likely do not result in an outright winner. Bacterial populations typically start at low densities, and stochastic processes will lead to the sectoring of the expanding genotypes. This sectoring limits competition to only those cells located at borders of interacting genotypes (Borenstein et al. 2015). Both empirical data and mathematical models suggest that bacterial lineages lacking T6SS-encoding loci can escape attack by T6SSs if the susceptible cells reside within established microbial colonies or if their growth rate exceeds their death rate (Borenstein et al. 2015; McNally et al. 2017; Wong et al. 2016). In addition, natural ecosystems are dynamic and fluctuate in resources and community members. T6SSs could therefore be deployed at certain times to influence specific stages of ecological succession. For example, *Salmonella enterica* serovar Typhimurium can, in a T6SS-dependent manner, target specific gut commensal *Klebsiella oxytoca* and *Klebsiella variicola* cells during invasion and establish infection in the gut of hosts (Sana et al. 2016). T6SS can also defend the habitat and protect against invasion by other genotypes. The plant commensal

Pseudomonas putida KT2440 can, in a T6SS-dependent manner, defend its territory and protect its plant host from invasion by phytopathogens (Bernal et al. 2017).

In native ecosystems, microbial communities may include competing genotypes that also encode T6SS and be competent in attacking and/or encode immunity proteins. Analysis of genome sequences of strains representing *Vibrio* species revealed that the regions downstream to the T6SS-encoding cluster encode diverse repertoires with modules of pairs of effector–immunity-encoding genes (Kirchberger et al. 2017). These regions also have long arrays of orphan immunity genes. The acquisition and maintenance of immunity genes may allow bacteria to withstand T6SS-mediated antagonism and co-exist in close proximity with other genotypes. Alternatively, the immunity proteins may contribute to kin recognition and further promote sectoring of genotypes into patches (Russell et al. 2014).

Therefore, whereas T6SS activity is a demonstrable molecular weapon of culture-grown bacteria, its role in the ecosystem appears to be far more complex. It is possible that antagonism mediated by T6SS could contribute to spatial structuring of microbial communities to promote niche partitioning (McNally et al. 2017). Segregation is critical for limiting the effects of social cheating and enhancing benefits derived from intragroup cooperative behaviors (Driscoll and Pepper 2010; Nadell et al. 2010, 2016; Julou et al. 2013). Whether the T6SS confers upon *A. tumefaciens* such an ability is unknown, as the variation of T6SS-associated effectors and T6SS-associated fitness across agrobacterial taxa and genotypes has yet to be characterized.

8 Regulation of the T6SS of *Agrobacterium*

T6SSs are not ruthlessly and constitutively deployed. T6SS-using bacteria each have unique strategies to regulate expression. Understanding the conditions that regulate T6SS of *A. tumefaciens* will provide significant insights into the ecological roles for this secretion apparatus (Miyata et al. 2013). Acidic conditions (pH = 5.5) are predicted to mimic the rhizosphere environment and have long been recognized as a necessary but insufficient trigger for virulence gene expression in *A. tumefaciens* (Li et al. 2002). Induction of virulence also requires wound-associated phenolics, such as acetosyringone (Stachel et al. 1985). Models predict that the low pH of the rhizosphere results in the proteolysis of the negative regulator ExoR (Wu et al. 2012). As a consequence, the ChvG sensor kinase is derepressed and the ChvI response regulator is activated, correlating with significant whole-transcriptome changes (Heckel et al. 2014; Yuan et al. 2008). Key functions associated with the genes that change expression in either an *exoR*- or acid-dependent manner are motility and chemotaxis (repressed) and social behaviors (induced) such as biofilm formation and T6SS (Heckel et al. 2014; Wu et al. 2012; Yuan et al. 2008). For the latter, upregulation of both the *imp* and *hcp* genes was observed in cells grown at pH 5.5 relative to those grown at pH 7.0. Thus, acidic conditions appear to mimic

the environmental signals that are critical in causing a transition in the lifecycle of *A. tumefaciens*.

The behavior of the T6SS is consistent with the predictions made on the basis of whole-transcriptome studies. Hcp accumulates to low levels in the extracellular milieu when the bacteria are grown in minimal AB-MES medium at a neutral pH but is very abundant when cells are grown at pH 5.5 (Wu et al. 2012). Moreover, in an *exoR* mutant grown at neutral pH, the levels of secreted Hcp are higher compared to those of wild-type *A. tumefaciens* C58. To investigate the role of ChvG/ChvI two-component system in regulating the T6SS, a substitution of aspartic acid 52 with glutamic acid to mimic phosphorylation was introduced into ChvI. This phosphomimetic mutant could bind a 230-bp-long intergenic region located between the divergent T6SS-associated operons. Furthermore, the levels of secreted Hcp were abundant when cells were grown at neutral pH in the *chvI* phosphomimetic mutant, consistent with the hypothesis that ChvI directly regulates T6SS expression (Wu et al. 2012).

However, the regulation of the T6SS in *A. tumefaciens* is more complex than acid-induced transcriptional activation via an ExoR-ChvG/ChvI regulatory cascade. Despite the low activity of the T6SS of *A. tumefaciens* grown in neutral pH, the baseline intracellular levels of proteins encoded by the *hcp* operon are relatively abundant. This result suggests that in neutral pH conditions, the endogenous expression of the *hcp* locus is regulated via a ChvI-independent mechanism. Perhaps the constitutive expression of immunity proteins provides protection against attack. Second, whereas acetosyringone represses the acid-dependent Hcp secretion, it neither abrogates activity nor reduces the levels of the T6SS-associated proteins (Wu et al. 2008, 2012). The mechanism by which repression occurs is unknown. Regardless, because acetosyringone induces the type IV secretion system (T4SS) necessary for delivering the T-DNA into plant cells, the anticorrelated behavior of the secretion systems could imply some interference and a need to downregulate T6SS activity.

In addition to transcriptional regulation, the T6SS of *A. tumefaciens* is further controlled via post-translational regulation (Fig. 3). In some taxa of bacteria, T6SS-encoding gene loci encode a Ser/Thr kinase PpkA and a cognate phosphatase PppA. These two proteins form a threonine phosphorylation pathway (TPP) that post-translationally regulates T6SSs. PpkA and PppA reciprocally modulate the threonine phosphorylation of the protein Fha, also encoded in some T6SS-associated loci. In *Pseudomonas aeruginosa*, phosphorylation of Fha is required for T6SS activity (Mougous et al. 2007).

The T6SS of *A. tumefaciens* is positively regulated by the threonine phosphorylation pathway (Lin et al. 2014). However, the mechanism by which the threonine phosphorylation pathway functions in *A. tumefaciens* is different from that reported for *P. aeruginosa*. The suggested target of *A. tumefaciens* PpkA is TssL, not Fha (Lin et al. 2014). In turn, Fha binds a phosphorylated motif in TssL, which then triggers the assembly and firing of the T6SS. In *A. tumefaciens*, TssL is highly phosphorylated. The T6SS of *A. tumefaciens* is negatively regulated by the TagF repressor, which is a dual-domain protein with an N-terminal TagF domain fused to

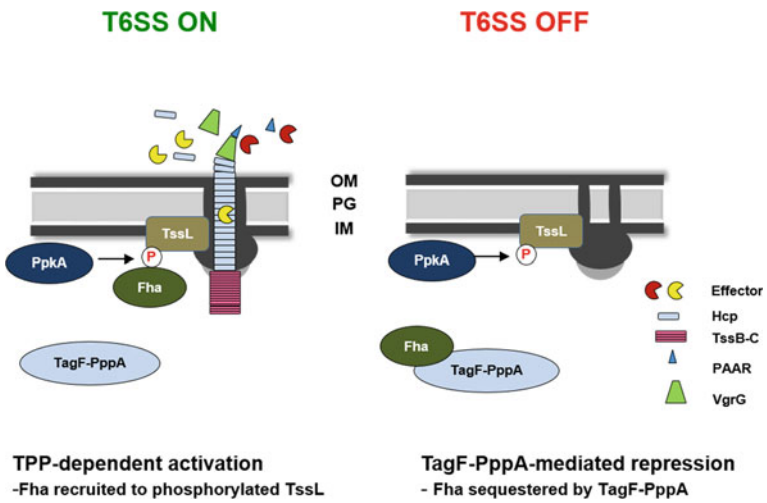


Fig. 3 Model of the post-translational regulation of the T6SS of *A. tumefaciens* C58. The T6SS is hypothesized to be regulated by the threonine phosphorylation pathway (TPP) and the TagR repressor. T66S ON: (1) activation of PpkA leads to the phosphorylation of TssL, (2) Fha binds the phosphorylated TssL, (3) T6SS is assembled, and (4) T6SS fires. T6SS OFF: TagF-PppA binds and sequesters Fha from binding the phosphorylated TssL. OM = outer membrane; PG = peptidoglycan; IM = inner membrane

a C-terminal PppA domain (Lin et al. 2014). Overexpression of just the TagF domain was sufficient to eliminate T6SS-dependent secretion, whereas overexpression of only the PppA domain reduced T6SS activity (Lin et al. 2018). However, in neither case were the levels of the T6SS-associated proteins affected. Furthermore, and unexpectedly, the phosphorylation status of TssL remained unaffected in either strain overexpressing TagF or PppA. Models suggest that an active PpkA leads to phosphorylation of TssL and binding by Fha to trigger the assembly of the T6SS. TagF, on the other hand, interacts with Fha to prevent binding and assembly. The signals that activate PpkA and TagF are not known.

9 Summary

A. tumefaciens is an important pathogen that causes disease with significant costs to agricultural industries. *A. tumefaciens* is also a vital tool for fundamental plant biology and plant biotechnology and, as a model, is studied to inform on T4SS, cell division, attachment, and social behaviors. As described herein, *A. tumefaciens* is emerging as a model for studying the mechanisms of the T6SS. Findings derived from laboratory-based studies of *A. tumefaciens* have formed a solid knowledge base. Based on current data, we posit that the T6SS plays a critical role during the transition from the motile phase to the sessile phase. This key point in the life cycle

of *A. tumefaciens* is controlled by the ExoR-ChvG/ChvI signaling cascade, and genes involved in chemotaxis, motility, and succinoglycan biosynthesis are co-regulated with those associated with the T6SS. This transition is also associated with a change from an individual to a cooperative lifestyle. The T6SS may have critical functions in helping *A. tumefaciens* sector communities into patches by recognizing kin, claiming its territory on the host, defending its territory against competing strains, and/or establishing borders between co-resident and competing genotypes. However, once a wound is perceived, T6SS activity is reduced and *A. tumefaciens* directs its efforts toward entering host tissues, T4SS-mediated virulence, and proliferating within host tissues. Henceforth, T6SS activity may influence the competitiveness of *A. tumefaciens* for access to plant hosts and have consequential impact on the spread and fitness of specific lineages of bacteria. *A. tumefaciens* has multiple desirable features that make it an excellent model to begin testing these hypotheses in ecological settings.

Acknowledgements We thank members of the Chang lab for their stimulating discussions and insightful comments on the manuscript. We thank Dr. Devanand D Bondage for giving us permission to use his illustration. Work in the Chang lab is supported in part by the National Institute of Food and Agriculture, US Department of Agriculture award 2014-51181-22384. Funding for the Lai lab is provided by the Ministry of Science and Technology of Taiwan (MOST 104-2311-B-001-025-MY3). DAS is a 2018 Barry Goldwater Scholar.

References

Alteri CJ, Himpsl SD, Pickens SR, Lindner JR, Zora JS, Miller JE, Arno PD, Straight SW, Mobley HL (2013) Multicellular bacteria deploy the type VI secretion system to preemptively strike neighboring cells. PLoS Pathog 9:e1003608

Barton IS, Fuqua C, Platt TG (2018) Ecological and evolutionary dynamics of a model facultative pathogen: *Agrobacterium* and crown gall disease of plants. Environ Microbiol 20:16–29

Basler VA (2016) Type VI secretion system substrates are transferred and reused among sister cells. Cell 167:99–110

Basler M (2015) Type VI secretion system: secretion by a contractile nanomachine. Philos Trans R Soc Lond B Biol Sci 370:20150021

Basler M, Pilhofer M, Henderson GP, Jensen GJ, Mekalanos JJ (2012) Type VI secretion requires a dynamic contractile phage tail-like structure. Nature 483:182–186

Bernal P, Allsopp LP, Filloux A, Llamas MA (2017) The *Pseudomonas putida* T6SS is a plant warden against phytopathogens. ISME J 11:972–987

Bladergroen MR, Badelt K, Spaink HP (2003) Infection-blocking genes of a symbiotic *Rhizobium leguminosarum* strain that are involved in temperature-dependent protein secretion. Mol Plant-Microbe Interact 16:53–64

Bondage DD, Lin JS, Ma LS, Kuo CH, Lai EM (2016) VgrG C terminus confers the type VI effector transport specificity and is required for binding with PAAR and adaptor-effector complex. Proc Natl Acad Sci U S A 113:E3931–E3940

Bonemann G, Pietrosiuk A, Mogk A (2010) Tubules and donuts: a type VI secretion story. Mol Microbiol 76:815–821

Borenstein DB, Ringel P, Basler M, Wingreen NS (2015) Established microbial colonies can survive Type VI secretion assault. PLoS Comp Bio 11:e1004520

Borgeaud S, Metzger LC, Scrignari T, Blokesch M (2015) The type VI secretion system of *Vibrio cholerae* fosters horizontal gene transfer. Science 347:63–67

Boyer F, Fichant G, Berthod J, Vandenbrouck Y, Attree I (2009) Dissecting the bacterial type VI secretion system by a genome wide *in silico* analysis: what can be learned from available microbial genomic resources? BMC Genom 10:104

Brunet Y, Zoued A, Boyer F, Douzi B, Cascales E (2015) The Type VI secretion TssEFGK-VgrG phage-like baseplate is recruited to the TssJLM membrane complex via multiple contacts and serves as assembly platform for tail tube/sheath polymerization. PLoS Genet 11:e1005545

Burkinshaw BJ, Liang X, Wong M, Le ANH, Lam L, Dong TG (2018) A type VI secretion system effector delivery mechanism dependent on PAAR and a chaperone-co-chaperone complex. Nat Microbiol 3:632–640

Canfield ML, Moore LW (1991) Isolation and characterization of opine-utilizing strains of *Agrobacterium tumefaciens* and fluorescent strains of *Pseudomonas* spp. from rootstocks of Malus. Phytopathol 81:440–443

Carruthers MD, Nicholson PA, Tracy EN, Munson RS Jr (2013) *Acinetobacter baumannii* utilizes a type VI secretion system for bacterial competition. PLoS ONE 8:e59388

Chang JH, Desveaux D, Creason AL (2014) The ABCs and 123 s of bacterial secretion systems in plant pathogenesis. Annu Rev Phytopathol 52:317–345

Chatzidaki-Livanis M, Geva-Zatorsky N, Comstock LE (2016) *Bacteroides fragilis* type VI secretion systems use novel effector and immunity proteins to antagonize human gut *Bacteroidales* species. Proc Natl Acad Sci U S A 113:3627–3632

Chen H, Yang D, Han F, Tan J, Zhang L, Xiao J, Zhang Y, Liu Q (2017) The bacterial T6SS effector EvpP prevents NLRP3 inflammasome activation by inhibiting the Ca2+-dependent MAPK-Jnk pathway. Cell Host Microbe 21:47–58

Chen WJ, Kuo TY, Hsieh FC, Chen PY, Wang CS, Shih YL, Lai YM, Liu JR, Yang YL, Shih MC (2016) Involvement of type VI secretion system in secretion of iron chelator pyoverdine in *Pseudomonas taiwanensis*. Sci Rep 6:32950

Cianfanelli FR, Alcoforado Diniz J, Guo M, De Cesare V, Trost M, Coulthurst SJ (2016a) VgrG and PAAR proteins define distinct versions of a functional type VI secretion system. PLoS Pathog 12:e1005735

Cianfanelli FR, Monlezun L, Coulthurst SJ (2016b) Aim, load, fire: the type VI secretion system, a bacterial nanoweapon. Trends Microbiol 24:51–62

Costechareyre D, Rhouma A, Lavire C, Portier P, Chapulliot D, Bertolla F, Boubaker A, Dessaux Y, Nesme X (2010) Rapid and efficient identification of *Agrobacterium* species by *recA* allele analysis: *Agrobacterium recA* diversity. Microb Ecol 60:862–872

Driscoll WW, Pepper JW (2010) Theory for the evolution of diffusible external goods. Evolution 64:2682–2687

Durand E, Cambillau C, Cascales E, Journet L (2014) VgrG, Tae, Tle, and beyond: the versatile arsenal of Type VI secretion effectors. Trends Microbiol 22:498–507

English G, Trunk K, Rao VA, Srikannathasan V, Hunter WN, Coulthurst SJ (2012) New secreted toxins and immunity proteins encoded within the Type VI secretion system gene cluster of *Serratia marcescens*. Mol Microbiol 86:921–936

Faist H, Keller A, Hentschel U, Deeken R (2016) Grapevine (*Vitis vinifera*) Crown galls host distinct microbiota. Appl Environ Microbiol 82:5542–5552

Farrand SK, Van Berkum PB, Oger P (2003) *Agrobacterium* is a definable genus of the family *Rhizobiaceae*. Int J Syst Evol Microbiol 53:1681–1687

Felisberto-Rodrigues C, Durand E, Aschtgen MS, Blangy S, Ortiz-Lombardia M, Douzi B, Cambillau C, Cascales E (2011) Towards a structural comprehension of bacterial type VI secretion systems: characterization of the TssJ-TssM complex of an *Escherichia coli* pathovar. PLoS Pathog 7:e1002386

Gelvin SB (2000) *Agrobacterium* and plant genes involved in T-DNA transfer and integration. Annu Rev Plant Mol Biol 51:223–256

Heckel BC, Tomlinson AD, Morton ER, Choi JH, Fuqua C (2014) A*grobacterium tumefaciens exoR* controls acid response genes and impacts exopolysaccharide synthesis, horizontal gene transfer, and virulence gene expression. J Bacteriol 196:3221–3233

Hood RD et al (2010) A type VI secretion system of *Pseudomonas aeruginosa* targets a toxin to bacteria. Cell Host Microbe 7:25–37

Hsu F, Schwarz S, Mougous JD (2009) TagR promotes PpkA-catalysed type VI secretion activation in *Pseudomonas aeruginosa*. Mol Microbiol 72:1111–1125

Julou T, Mora T, Guillon L, Croquette V, Schalk IJ, Bensimon D, Desprat N (2013) Cell-cell contacts confine public goods diffusion inside *Pseudomonas aeruginosa* clonal microcolonies. Proc Natl Acad Sci U S A 110:12577–12582

Kapitein N, Bonemann G, Pietrosiuk A, Seyffer F, Hausser I, Locker JK, Mogk A (2013) ClpV recycles VipA/VipB tubules and prevents non-productive tubule formation to ensure efficient type VI protein secretion. Mol Microbiol 87:1013–1028

Kirchberger PC, Unterweger D, Provenzano D, Pukatzki S, Boucher Y (2017) Sequential displacement of Type VI Secretion System effector genes leads to evolution of diverse immunity gene arrays in *Vibrio cholerae*. Sci Rep 7:45133

Lassalle F et al (2011) Genomic species are ecological species as revealed by comparative genomics in *Agrobacterium tumefaciens*. Genome Bio Evol 3:762–781

Leiman PG, Basler M, Ramagopal UA, Bonanno JB, Sauder JM, Pukatzki S, Burley SK, Almo SC, Mekalanos JJ (2009) Type VI secretion apparatus and phage tail-associated protein complexes share a common evolutionary origin. Proc Natl Acad Sci U S A 106:4154–4159

Li L, Jia Y, Hou Q, Charles TC, Nester EW, Pan SQ (2002) A global pH sensor: *Agrobacterium* sensor protein ChvG regulates acid-inducible genes on its two chromosomes and Ti plasmid. Proc Natl Acad Sci U S A 99:12369–12374

Lin J, Zhang W, Cheng J, Yang X, Zhu K, Wang Y, Wei G, Qian PY, Luo Z, Shen X (2017) A *Pseudomonas* T6SS effector recruits PQS-containing outer membrane vesicles for iron acquisition. Nat Commun 8:14888

Lin J, Ma LS, Lai EM (2013) Systematic dissection of the *Agrobacterium* type VI secretion system reveals machinery and secreted components for subcomplex formation. PLoS ONE 8:e67647

Lin JS, Pissaridou P, Wu HH, Tsai MD, Filloux A, Lai EM (2018) TagF-mediated repression of bacterial type VI secretion systems involves a direct interaction with the cytoplasmic protein Fha. J Biol Chem. (In press)

Lin JS, Wu HH, Hsu PH, Ma LS, Pang YY, Tsai MD, Lai EM (2014) Fha interaction with phosphothreonine of TssL activates type VI secretion in *Agrobacterium tumefaciens*. PLoS Pathog 10:e1003991

Lossi NS, Manoli E, Forster A, Dajani R, Pape T, Freemont P, Filloux A (2013) The HsiB1C1 (TssB-TssC) complex of the *Pseudomonas aeruginosa* type VI secretion system forms a bacteriophage tail sheathlike structure. J Biol Chem 288:7536–7548

Ma AT, Mekalano J (2010) In vivo actin cross-linking induced by *Vibrio cholerae* type VI secretion system is associated with intestinal inflammation. Proc Natl Acad Sci U S A 107:4365–4370

Ma LS, Hachani A, Lin JS, Filloux A, Lai EM (2014) *Agrobacterium tumefaciens* deploys a superfamily of type VI secretion DNase effectors as weapons for interbacterial competition in planta. Cell Host Microbe 16:94–104

Ma LS, Lin J, Lai EM (2009) An IcmF family protein, ImpLM, is an integral inner membrane protein interacting with ImpKL, and its walker a motif is required for type VI secretion system-mediated Hcp secretion in *Agrobacterium tumefaciens*. J Bacteriol 191:4316–4329

McNally L, Bernardy E, Thomas J, Kalziqi A, Pentz J, Brown SP, Hammer BK, Yunker PJ, Ratcliff WC (2017) Killing by Type VI secretion drives genetic phase separation and correlates with increased cooperation. Nat Commun 8:14371

Miyata ST, Bachmann V, Pukatzki S (2013) Type VI secretion system regulation as a consequence of evolutionary pressure. J Med Microbiol 62:663–676

Mougous JD et al (2006) A virulence locus of *Pseudomonas aeruginosa* encodes a protein secretion apparatus. Science 312:1526–1530

Mougous JD, Gifford CA, Ramsdell TL, Mekalanos JJ (2007) Threonine phosphorylation post-translationally regulates protein secretion in *Pseudomonas aeruginosa*. Nat Cell Biol 9:797–803

Nadell CD, Foster KR, Xavier JB (2010) Emergence of spatial structure in cell groups and the evolution of cooperation. PLoS Comput Biol 6:e1000716

Nadell CD, Drescher K, Foster KR (2016) Spatial structure, cooperation and competition in biofilms. Nat Rev Microbiol 14:589–600

Planamente S, Salih O, Manoli E, Albesa-Jove D, Freemont PS, Filloux A (2016) TssA forms a gp6-like ring attached to the type VI secretion sheath. EMBO J 35:1613–1627

Pukatzki S, Ma AT, Revel AT, Sturtevant D, Mekalanos JJ (2007) Type VI secretion system translocates a phage tail spike-like protein into target cells where it cross-links actin. Proc Natl Acad Sci U S A 104:15508–15513

Pukatzki S, Ma AT, Sturtevant D, Krastins B, Sarracino D, Nelson WC, Heidelberg JF, Mekalanos JJ (2006) Identification of a conserved bacterial protein secretion system in *Vibrio cholerae* using the *Dictyostelium* host model system. Proc Natl Acad Sci U S A 103:1528–1533

Rao PS, Yamada Y, Tan YP, Leung KY (2004) Use of proteomics to identify novel virulence determinants that are required for *Edwardsiella tarda* pathogenesis. Mol Microbiol 53:573–586

Rigard M et al (2016) *Francisella tularensis* IglG belongs to a novel family of PAAR-like T6SS proteins and harbors a unique N-terminal extension required for virulence. PLoS Pathog 12: e1005821

Roest HP, Mulders IH, Spaink HP, Wijffelman CA, Lugtenberg BJ (1997) A *Rhizobium leguminosarum* biovar *trifolii* locus not localized on the sym plasmid hinders effective nodulation on plants of the pea cross-inoculation group. Mol Plant-Microbe Interact 10:938–941

Russell AB, Hood RD, Bui NK, LeRoux M, Vollmer W, Mougous JD (2011) Type VI secretion delivers bacteriolytic effectors to target cells. Nature 475:343–347

Russell A, LeRoux M, Hathazi K, Agnello DM, Ishikawa T, Wiggins PA, Wai SN, Mougous JD (2013) Diverse type VI secretion phospholipases are functionally plastic antibacterial effectors. Nature 496:508–512

Russell AB, Peterson SB, Mougous JD (2014) Type VI secretion system effectors: poisons with a purpose. Nat Rev Microbiol 12:137–148

Sana TG, Flaugnatti N, Lugo KA, Lam LH, Jacobson A, Baylot V, Durand E, Journet L, Cascales E, Monack DM (2016) *Salmonella Typhimurium* utilizes a T6SS-mediated antibacterial weapon to establish in the host gut. Proc Natl Acad Sci U S A 113:E5044–E5051

Schlieker C, Zentgraf H, Dersch P, Mogk A (2005) ClpV, a unique Hsp100/Clp member of pathogenic proteobacteria. Biol Chem 386:115–1127

Schwarz S, West TE, Boyer F, Chiang WC, Carl MA, Hood RD, Rohmer L, Tolker-Nielsen T, Skerrett SJ, Mougous JD (2010) Burkholderia type VI secretion systems have distinct roles in eukaryotic and bacterial cell interactions. PLoS Pathog 6:e1001068

Shalom G, Shaw JG, Thomas MS (2007) In vivo expression technology identifies a type VI secretion system locus in *Burkholderia pseudomallei* that is induced upon invasion of macrophages. Microbiology 153:2689–2699

Shneider MM, Buth SA, Ho BT, Basler M, Mekalanos JJ, Leiman PG (2013) PAAR-repeat proteins sharpen and diversify the type VI secretion system spike. Nature 500:350–353

Si M, Zhao C, Burkinshaw B, Zhang B, Wei D, Wang Y, Dong TG, Shen X (2017) Manganese scavenging and oxidative stress response mediated by type VI secretion system in *Burkholderia thailandensis*. Proc Natl Acad Sci U S A 114:E2233–e2242

Silverman JM, Agnello DM, Zheng H, Andrews BT, Li M, Catalano CE, Gonen T, Mougous JD (2013) Haemolysin coregulated protein is an exported receptor and chaperone of type VI secretion substrates. Mol Cell 51:584–593

Slater SC et al (2009) Genome sequences of three agrobacterium biovars help elucidate the evolution of multichromosome genomes in bacteria. J Bacteriol 191:2501–2511

Stachel SE, Messens E, Montagu MV, Zambryski P (1985) Identification of the signal molecules produced by wounded plant cells that activate T-DNA transfer in *Agrobacterium tumefaciens*. Nature 318:624–629

Suarez G, Sierra JC, Erova TE, Sha J, Horneman AJ, Chopra AK (2010) A type VI secretion system effector protein, VgrG1, from *Aeromonas hydrophila* that induces host cell toxicity by ADP ribosylation of actin. J Bacteriol 192:155–168

Thomas J, Watve SS, Ratcliff WC, Hammer BK (2017) Horizontal gene transfer of functional Type VI killing genes by natural transformation. mBio 8:e00654–17

Unterweger D, Kostiuk, Otjengerdes R, Wilton A, Diaz-Satizabal L, Pukatzki S (2015) Chimeric adaptor proteins translocate diverse type VI secretion system effectors in *Vibrio cholerae*. EMBO J 34:2198–2210

Wang T et al (2015) Type VI secretion system transports Zn2+ to combat multiple stresses and host immunity. PLoS Pathog 11:e1005020

Wong M, Liang X, Smart M, Tang L, Moore R, Ingalls B, Dong TG (2016) Microbial her protection mediated by antagonistic interaction in polymicrobial communities. Appl Environ Microbiol 82:6881–6888

Wu CF, Lin JS, Shaw GC, Lai EM (2012) Acid-induced type VI secretion system is regulated by ExoR-ChvG/ChvI signaling cascade in *Agrobacterium tumefaciens*. PLoS Pathog 8:e1002938

Wu HY, Chung PC, Shih HW, Wen S, Lai EM (2008) Secretome analysis uncovers an Hcp-family protein secreted via a type VI secretion system in *Agrobacterium tumefaciens*. J Bacteriol 190:2841–2850

Yuan ZC, Liu P, Saenkham P, Kerr K, Nester EW (2008) Transcriptome profiling and functional analysis of *Agrobacterium tumefaciens* reveals a general conserved response to acidic conditions (pH 5.5) and a complex acid-mediated signaling involved in *Agrobacterium*-plant interactions. J Bacteriol 190:494–507

Zheng J, Ho B, Mekalanos JJ (2011) Genetic analysis of anti-amoebae and anti-bacterial activities of the type VI secretion system in *Vibrio cholerae*. PLoS ONE 6:e23876

Zoued A, Brunet YR, Durand E, Aschtgen MS, Logger L, Douzi B, Journet L, Cambillau C, Cascales E (2014) Architecture and assembly of the Type VI secretion system. Biochimi Biophys Acta 1843:1664–1673

The *Agrobacterium* VirB/VirD4 T4SS: Mechanism and Architecture Defined Through In Vivo Mutagenesis and Chimeric Systems

Yang Grace Li and Peter J. Christie

Contents

1 Introduction: A VirB/VirD4 T4SS Primer 234
2 Overview of the VirB/VirD4 Machine Assembly and Architecture 235
3 Defining the Route of DNA Transfer Through the VirB/VirD4 T4SS 237
4 Processing of T-DNA and Effector Protein Substrates and Substrate Recognition Signals for Transfer 239
5 VirD4 Receptors 242
5.1 VirD4's NTD: Spatial Positioning and Interaction with the IMC 243
5.2 VirD4's AAD: Substrate Specificity 244
5.3 VirD4's CTD: Regulator of Substrate Transfer 244
6 The Envelope-Spanning VirB/VirD4 T4SS 245
6.1 Cytoplasmic Entrance: The ATPase Energy Center 245
6.2 The IMC 246
6.3 The OMCC 247
6.4 Signal Activation of the Translocation Channel 248
7 The T-Pilus: Assembly and Function 249
7.1 T-Pilus Subunits 249
7.2 T-Pilus Biogenesis Pathway 251
7.3 Other Attachment Mechanisms 252
8 Conclusions 253
References 254

Y. G. Li · P. J. Christie (✉)
Department of Microbiology and Molecular Genetics, McGovern Medical School, 6431 Fannin St, Houston, TX 77030, USA
e-mail: Peter.J.Christie@uth.tmc.edu

Y. G. Li
e-mail: Yang.G.Li@uth.tmc.edu

Current Topics in Microbiology and Immunology (2018) 418:233–260
DOI 10.1007/82_2018_94

Published Online: 29 May 2018

Abstract The *Agrobacterium tumefaciens* VirB/VirD4 translocation machine is a member of a superfamily of translocators designated as type IV secretion systems (T4SSs) that function in many species of gram-negative and gram-positive bacteria. T4SSs evolved from ancestral conjugation systems for specialized purposes relating to bacterial colonization or infection. *A. tumefaciens* employs the VirB/VirD4 T4SS to deliver oncogenic DNA (T-DNA) and effector proteins to plant cells, causing the tumorous disease called crown gall. This T4SS elaborates both a cell-envelope-spanning channel and an extracellular pilus for establishing target cell contacts. Recent mechanistic and structural studies of the VirB/VirD4 T4SS and related conjugation systems in *Escherichia coli* have defined T4SS architectures, bases for substrate recruitment, the translocation route for DNA substrates, and steps in the pilus biogenesis pathway. In this review, we provide a brief history of *A. tumefaciens* VirB/VirD4 T4SS from its discovery in the 1980s to its current status as a paradigm for the T4SS superfamily. We discuss key advancements in defining VirB/VirD4 T4SS function and structure, and we highlight the power of in vivo mutational analyses and chimeric systems for identifying mechanistic themes and specialized adaptations of this fascinating nanomachine.

1 Introduction: A VirB/VirD4 T4SS Primer

Agrobacterium tumefaciens delivers oncogenic T-DNA and several effector proteins to plant target cells via a type IV secretion system (T4SS) (Alvarez-Martinez and Christie 2009). This nanomachine, elaborated by the 11 VirB proteins (VirB1–VirB11) and the VirD4 subunit, is termed the VirB/VirD4 T4SS. Evidence for the existence of an envelope-spanning T-DNA transfer system, and clues that this system functions by a conjugation-like mechanism, arose through discoveries resulting from sequencing of pTi plasmids nearly 30 years ago (Beijersbergen et al. 1994; Kuldau et al. 1990; Shirasu et al. 1990; Ward et al. 1988). Sequencing of the *virB* locus revealed similarities between the predicted VirB proteins and subunits of conjugation or mating-pair-formation (Mpf) systems. Further sequence analyses revealed similarities between the VirD subunits and the DNA transfer and replication (Dtr) factors involved in processing at origin-of-transfer (*oriT*) sequences associated with conjugative plasmids (Stachel and Zambryski 1986). In quick succession, studies demonstrated nicking activities of VirD1 and VirD2 at T-DNA border repeats in *A. tumefaciens* and *Escherichia coli* and then in vitro using purified proteins (Albright et al. 1987; Filichkin and Gelvin 1993; Stachel et al. 1987; Veluthambi et al. 1988; Yanofsky et al. 1986; Young and Nester 1988). The VirB/VirD4 system also was shown to transfer a conjugative element, the IncQ plasmid RSF1010, which lacks genes for its own T4SS but codes for Dtr factors that both process the plasmid and mobilize its conjugative transfer through various T4SSs. The *A. tumefaciens* system was first shown to deliver RSF1010 to plant cells and subsequently to other agrobacterial cells (Beijersbergen et al. 1992; Buchanan-Wollaston et al. 1987; Fullner 1998; Ward et al. 1991). These findings

established that the VirB system functions as a *bona fide* conjugation system in mediating both interbacterial and interkingdom DNA transfer.

The next major conceptual advance in our understanding of the VirB system in a broader biological context resulted from the discovery in the mid-1990s that the 11 VirB subunits, and homologous Tra subunits encoded by conjugation systems, are also related to subunits of the pertussis toxin (PT) export system of *Bordetella pertussis* (Shirasu and Kado 1993; Winans et al. 1996). This observation expanded our view of the versatility of conjugation-like machines; over evolutionary time, ancestral conjugation machines acquired novel functions enabling interkingdom trafficking of DNA and protein substrates during infection processes. Systems assembled from homologs of VirB and Tra subunits were then grouped as a new secretion superfamily, the type IV secretion systems or T4SSs (Christie 1997; Salmond 1994). With the explosion of sequenced bacterial genomes and mechanistic studies deciphering infection processes during the 1990s and 2000s, new T4SSs were discovered so that now this superfamily includes translocators functioning in most if not all bacterial species (Bhatty et al. 2013). All systems bear at least a few signature subunits enabling their classification, but in fact the T4SSs are a highly compositionally, structurally, and functionally diverse superfamily. This is evidenced at a functional level by the ability of these systems to (i) deliver mobile genetic elements (MGEs) to other bacteria and to eukaryotic cells, (ii) deliver protein substrates (termed 'effectors') to eukaryotic target cells, (iii) inject a protein toxin to bacterial recipients, or (iv) export or import DNA or protein substrates to or from the extracellular milieu (Grohmann et al. 2018). Indeed, the *A. tumefaciens* VirB/VirD4 system itself exhibits striking versatility in its ability to translocate DNA and various protein effectors to plant cells during infection, as well as DNA to other agrobacteria, fungi, and even human cells (Beijersbergen et al. 1992; Bundock et al. 1995; Kunik et al. 2001; Piers et al. 1996). From the early discovery of the *A. tumefaciens* VirB/VirD4 T4SS in the 1980s through today, contributions by many laboratories have helped shape this system as an important paradigm for the T4SS superfamily. In this chapter, we summarize recent progress toward defining the structure and function of the VirB/VirD4 T4SS, with a focus on insights gained through mutational analyses and substitutions of domains, subunits, or entire machine subassemblies with related systems.

2 Overview of the VirB/VirD4 Machine Assembly and Architecture

The *virB* gene arrangements and functions of their products are depicted in Fig. 1a. The *virB* genes comprise an operon inducibly expressed from an upstream *virB* promoter upon sensing by *A. tumefaciens* of plant-derived phenolic compounds, although evidence exists for at least one or two additional internal promoters (Berger and Christie 1994; Stachel and Nester 1986). *virD4* is coexpressed with

virD genes encoding the Dtr processing factors (Stachel and Nester 1986). Studies defining subcellular localizations and membrane topologies of VirB and VirD4 subunits, coupled with definition of subunit interaction networks by two-hybrid and pull-down assays, provided an early view of the architectural arrangement of the VirB/VirD4 T4SS across the *A. tumefaciens* cell envelope (Christie 1997). Importantly, this body of work included the discovery that the outer membrane (OM) lipoprotein VirB7 interacts via a disulfide cross-link with VirB9, and this heterodimer in turn forms a stabilizing interaction with the cell-envelope-spanning subunit VirB10 (Anderson et al. 1996; Baron et al. 1997; Fernandez et al. 1996; Spudich et al. 1996). The VirB7/VirB9/VirB10 complex was originally designated as the 'core' complex and now is termed the outer membrane core complex or OMCC (Christie 1997; Grohmann et al. 2018). Once assembled, the OMCC

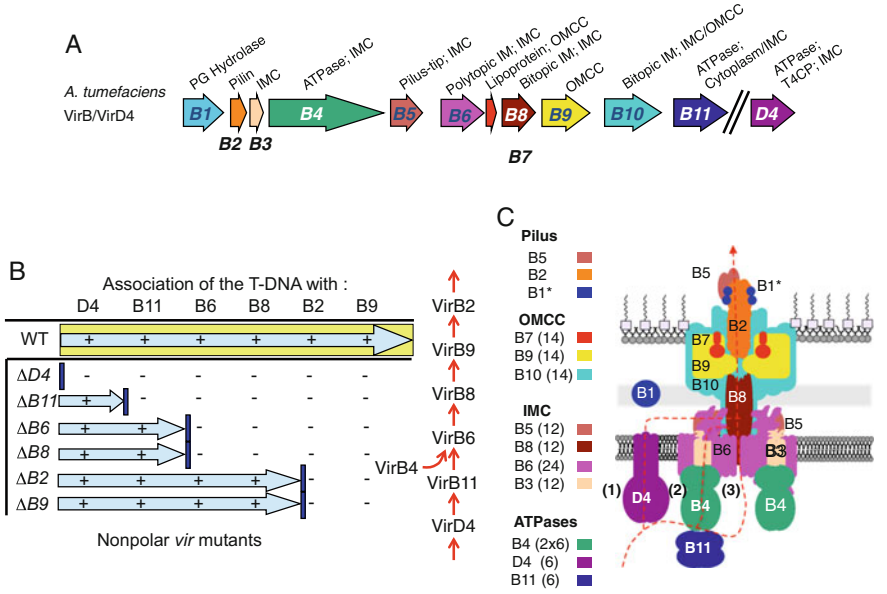


Fig. 1 Composition and architecture of the *A. tumefaciens* VirB/VirD4 T4SS. **a** Order of the *virB* genes in the *virB* operon; *virD4* is encoded by the separate *virD* operon (hashed lines). Functions, localization, or subassembly associations of the encoded products are listed. PG, peptidoglycan; IMC, inner membrane complex; OMCC, outer membrane core complex; T4CP, type IV coupling protein. **b** Summary of transfer DNA immunoprecipitation (TrIP) studies showing that the T-DNA forms formaldehyde (FA)-crosslinkable contacts with the VirD4 and VirB proteins listed at top. The corresponding deletion mutations permit (+) or prevent (-) FA cross-linking of T-DNA with VirB/VirD4 subunits listed, indicative of stage-specific blocks in the transfer pathway denoted at the right. **c** Architecture of the VirB/VirD4 T4SS, with positions of subunits identified. Components of the pilus, OMCC, IMC, and ATPase energy center are listed. The overall architecture and subunit stoichiometries (in parentheses at left) are based on the R388-encoded $VirB_{3-10}$ solved by single-particle, negative-stain electron microscopy (Low et al. 2014). The model shows predicted locations of VirB1, VirB1*, VirB2, VirD4, VirB11, and the T-pilus, which were missing from the $VirB_{3-10}$ substructure

stabilizes other VirB subunits localized in the periplasm or inner membrane (IM) (Beaupre et al. 1997; Fernandez et al. 1996). These early findings led to a proposal that the VirB/VirD4 machine is built by the ordered assembly of the OMCC followed by recruitment and assembly of an inner membrane complex (the IMC) (Christie 1997). The VirB/VirD4 system also is composed of three ATPases (VirB4, VirB11, and VirD4) localized in the cytoplasm or at the cytoplasmic face of the IM. These ATPases are not required for assembly of the OMCC, but rather form part of the IMC either stably or transiently (see below).

In addition to the envelope-spanning translocation channel, the VirB subunits elaborate an extracellular pilus termed the T-pilus (Fullner et al. 1996; Kado 2000). Mutational studies identified interesting classes of mutations that selectively prevent assembly of a functional translocation channel but not the T-pilus (Tra^-, Pil^+) or, conversely, inhibit assembly of the T-pilus production but not the translocation channel (Tra^+, Pil^-). These 'uncoupling' mutations have proven highly informative for assigning VirB and VirD4 building blocks of the channel vs T-pilus. Cells lacking VirD4, for example, are Tra^- but elaborate wild-type T-pili. By contrast, cells lacking the VirB1 lytic transglycosylase efficiently transfer substrates but fail to elaborate T-pili. Further studies identified 'uncoupling' mutations in the VirB11 ATPase, the IMC subunit VirB6, and the OMCC subunits VirB9 and VirB10 (Christie et al. 2005; Jakubowski et al. 2005; Sagulenko et al. 2001; Zhou and Christie 1997). These findings provided strong evidence that the VirB/VirD4 system alternatively assembles as the translocation channel or T-pilus. In an infection setting, it makes sense that upon activation of the *vir* regulon by plant-derived signals, the VirB system would elaborate T-pili for bacterial attachment and colonization of plant tissues. Then, once productive bacterial-plant cell contacts are established, a second signal would stimulate a transition from the T-pilus assembly mode to formation of the translocation channel, in part through recruitment of the VirD4 ATPase for interkingdom transfer of T-DNA and effector proteins (Christie et al. 2005). From a more mechanistic perspective, the ability to isolate Tra^+, Pil^- 'uncoupling' mutations also strongly indicates that substrates are not translocated through the pilus, which is consistent with early genetic and structural findings that *E. coli* conjugation systems mediate the formation of tight donor–target cell junctions as a prerequisite for DNA transfer (see Durrenberger et al. 1991; Samuels et al. 2000).

3 Defining the Route of DNA Transfer Through the VirB/VirD4 T4SS

Although a general architecture of the VirB/VirD4 T4SS was inferred from fractionation and topology studies (See Christie 1997), direct evidence that the VirB subunits form a translocation channel was lacking until a modified chromatin immunoprecipitation assay was employed to identify subunit contacts with the

translocating T-DNA (Fig. 1b) (Cascales and Christie 2004b). In this assay, termed transfer DNA immunoprecipitation (or TrIP), *A. tumefaciens* cells were induced for assembly of the VirB/VirD4 T4SS and treated with formaldehyde (FA) to cross-link channel subunits to the translocating DNA substrate. Following detergent solubilization of the cell envelope, antibodies to the VirB and VirD4 subunits were used to immunoprecipitate the respective machine component and the precipitated material was analyzed for the presence of T-DNA by PCR amplification. In a semiquantitative variation of this assay, termed QTrIP, labeled radionucleotide was added in the logarithmic phase of PCR product amplification. By TrIP, six of the 12 VirB/VirD4 subunits formed FA-crosslinkable contacts with the T-DNA transfer intermediate as well as that of the RSF1010 plasmid substrate. These proteins included the VirD4 and VirB11 ATPases residing at the cytoplasmic face of the IM, the integral IM subunits VirB6 and VirB8, the VirB2 pilin, and the OM-associated VirB9 (Fig. 1b). TrIP studies with various mutant strains and strains producing subsets of VirB and VirD4 proteins confirmed this sequence of DNA–VirB/VirD4 subunit contacts and further identified contributions of the remaining VirB subunits to substrate transfer across the cell envelope (Atmakuri et al. 2004; Jakubowski et al. 2004, 2005).

The DNA substrate first contacts the VirD4 ATPase, which can occur even in the absence of other VirB proteins (Fig. 1b, c) (Cascales and Christie 2004b). Mutations in the nucleotide-binding site of VirD4 did not abrogate DNA binding, indicating that ATP energy does not drive the VirD4–DNA substrate interaction (Atmakuri et al. 2004). In a second step of the transfer pathway, VirD4 delivers the DNA to the VirB11 ATPase (Cascales and Christie 2004b). This transfer reaction requires VirD4 and VirB11, but proceeds independently of other IM-associated subunits. Mutations in the nucleotide sites of both proteins also did not affect DNA binding, further indicating that ATP energy also does not drive DNA substrate transfer from VirD4 to VirB11. Interestingly, in reconstitution experiments, this transfer step proceeds only if the core subunits (VirB7–VirB10) are coproduced, which led to a proposal that VirD4 and/or VirB11 physically interact with one or more components of the core complex for substrate transfer (Atmakuri et al. 2004). In a third reaction, VirD4 and VirB11 coordinate with the third ATPase, VirB4, to deliver the DNA substrate to two integral IM proteins, VirB6 and VirB8 (Cascales and Christie 2004b). This is an ATP energy-driven reaction, as evidenced by a lack of substrate transfer in strains bearing NTP-binding site mutant forms of any of the three ATPases (Atmakuri et al. 2004). The reconstitution experiments further showed that the core complex also is essential for DNA transfer to VirB6 and VirB8, suggesting a coordination of function between the IM- and the OM-associated subassemblies. Finally, VirB6 and VirB8 deliver the DNA substrate to the VirB2 pilin and the core subunit VirB9. Mutational analyses established that the N- and C-terminal regions of polytopic VirB6 contribute to the delivery of DNA to VirB2 and VirB9, which was consistent with evidence that the IM subassembly interacts directly with the OM-associated channel (Jakubowski et al. 2003). Based on results of these TrIP studies, it was postulated that the VirB2 pilin forms a part of the cell-envelope-spanning channel (Jakubowski et al. 2005). However, efforts to

detect substrate interactions with the extended T-pilus were unsuccessful, further arguing against a role for the T-pilus in mediating T-DNA transfer to plant cells (Cascales and Christie 2004b).

Results of the TrIP studies provided the first direct evidence that MGEs are translocated across the cell envelope through a conjugative machine. Furthermore, these studies identified presumptive components of the 'mating' channel as VirD4, VirB11, VirB6, VirB8, VirB2, and VirB9. Other VirB subunits, e.g., VirB3, VirB4, VirB5, VirB7, VirB10, did not cross-link with the substrate, but null mutations blocked formation of specific substrate–channel subunit contacts in the transfer pathway, establishing their importance for specific stages of substrate translocation (Cascales and Christie 2004b). Gratifyingly, the route of DNA transfer defined using the TrIP assay is compatible with results of biochemical studies defining subcellular locations and interactions among the VirB and VirD4 subunits, as well as more recent ultrastructural findings (Fig. 1b, c). With respect to the latter, the *E. coli* conjugative plasmid R388 codes for a T4SS, termed Trw, which is closely related phylogenetically to the VirB/VirD4 T4SS. A large substructure of the Trw T4SS was purified and structurally analyzed by single-particle negative-stain transmission electron microscopy (NS-EM) (Low et al. 2014). The complex was termed the $VirB_{3\text{-}10}$ substructure for the fact it is composed of homologs of the *A. tumefaciens* VirB3 through VirB10 subunits. Noteworthy features of the $VirB_{3\text{-}10}$ substructure include: (i) a barrel-shaped OMCC of a size and architecture similar to that of the OMCCs associated with the pKM101 Tra and *A. tumefaciens* VirB/VirD4 T4SSs (see below), (ii) a thin stalk structure connecting the OMCC to the IMC, and (iii) a highly asymmetric IMC composed of an inner membrane platform and two side-by-side hexamers of the VirB4-like ATPase extending into the cytoplasm. In the most recent structure, a dimer of VirD4-like TrwB was associated with the base of the IMC, sandwiched between the two VirB4 hexamers (Redzej et al. 2017). In view of sequence and compositional similarities, it is reasonable to speculate that the $VirB_{3\text{-}10}$/VirD4 complex is a structural archetype for the VirB/VirD4 system and other closely related T4SSs. By combining results of the TrIP and ultrastructural studies, possible transfer routes for DNA substrates across the VirB/VirD4 T4SS can be envisioned (see Fig. 1b, c & below).

4 Processing of T-DNA and Effector Protein Substrates and Substrate Recognition Signals for Transfer

Several steps of the DNA transfer pathway delineated by the TrIP studies have been analyzed in molecular and structural detail. The initiating step of T-DNA transfer involves its recruitment to the VirB/VirD4 T4SS. In conjugation systems, prevailing models depict the relaxosome as the set of processing factors bound as a preinitiation complex at *oriT* (de la Cruz et al. 2010). This quiescent complex is activated upon binding to the VirD4 substrate receptor, which results in a switch in

the relaxase to a translocation mode (Lang et al. 2011). The active relaxase cleaves the T-strand, the T-strand is unwound from its template strand, and the relaxase pilots the T-strand through the transfer channel (de la Cruz et al. 2010). Among the conjugation systems, therefore, processing of the transferred strand is spatiotemporally linked to translocation through the conjugation machine. For T-DNA, the process is formally similar in the sense that the VirD2 relaxase is guided to the *oriT*-like T-DNA border sequence by the VirD1 accessory factor to catalyze nicking of the T-strand for transfer (Yanofsky et al. 1986; Albright et al. 1987; Stachel et al. 1987). However, there are also important differences in the requirements for formation of the catalytically active relaxosome (see Fig. 2). For example, efficient recruitment of VirD2 to the T-DNA border sequences requires not only VirD1 but also two other Vir proteins, VirC1 and VirC2 (Atmakuri et al. 2007; Lu et al. 2009). These latter proteins bind a sequence termed *overdrive* located immediately adjacent to the right border repeat sequences of octopine-type Ti plasmids (Toro et al. 1989). Binding of the VirD and VirC accessory factors stimulates T-DNA processing and also leads to the generation of many copies ($\sim$30 or more) of free VirD2-T-strand transfer intermediates in a cell (Atmakuri et al. 2007; Veluthambi et al. 1988). VirC1 is a member of the ParA family of ATPases that mediate partitioning of chromosomes and plasmids during cell division. VirC1 localizes at *A. tumefaciens* cell poles to recruit the VirD2–T-strand complex to the cell poles and to interact with VirD4, the substrate receptor for the VirB/D4 T4SS (Atmakuri et al. 2007). *A. tumefaciens* therefore adapted an ancestral Par-like function for the novel purposes of stimulating a conjugative DNA-processing reaction and promoting DNA substrate docking with a cognate T4SS receptor. Both activities potentially mediate transfer of many copies of T-DNA from each bacterium to susceptible plant host cells, presumably for enhanced probability of infection. At a molecular level, this implies that the relaxosomal complex is activated in the absence of binding to VirD4, and also that unwinding of the T-strand from its template is temporally and spatially uncoupled from its engagement with VirD4 and translocation through the VirB/VirD4 channel (Fig. 2).

How the complex of VirC and VirD proteins bound to the T-strand engage with the VirB/VirD4 T4SS is presently not known, but some general features of this interaction have been identified. For example, VirD2 carries a translocation sequence at its C terminus that mediates its translocation, as demonstrated with the Cre recombinase reporter assay for translocation (CRAfT) (Vergunst et al. 2000, 2005). In this assay, full-length or fragments of protein substrates, e.g., VirD2, are fused at their N-termini to Cre recombinase and translocation is monitored to a bacterial or plant target cell engineered to carry *lox* sites arrayed so that Cre-mediated excision yields a reporter activity, e.g., restoration of antibiotic resistance or GFP fluorescence. By use of the CRAfT assay, VirD2's C-terminal translocation signal was shown to consist of a cluster of positively charged Arg residues, leading to a proposal that the relaxase engages with VirD4 at least in part through ionic interactions. The VirD2–T-strand–VirD4 interaction is considerably more complex than this, however, as evidenced by recent work showing that VirD2 additionally carries an internal signal(s) conferring recognition by the VirB/VirD4

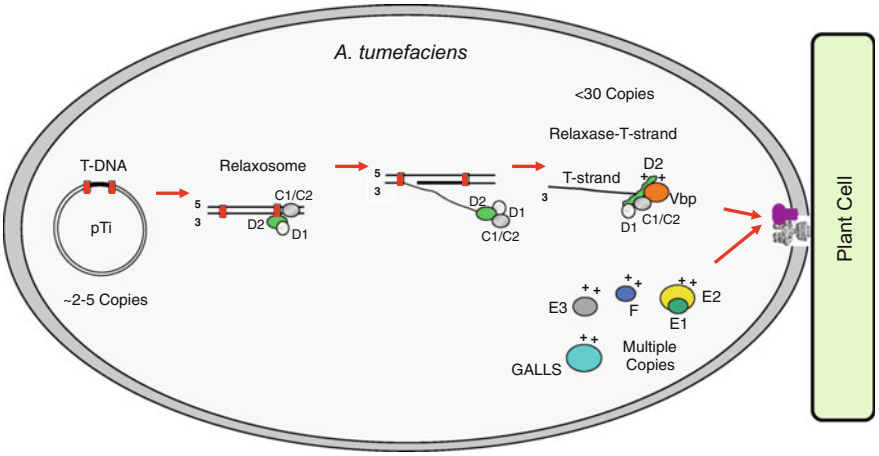


Fig. 2 Substrate processing and recruitment to the VirB/VirD4 T4SS. The VirD2 relaxase and accessory factors (VirD1, VirC1, VirC2) bind the *oriT*-like T-DNA border repeats (red rectangles) of the T-DNA on pTi plasmid, forming the catalytically active relaxosome. VirD2 (green oval) nicks and remains covalently associated with the strand T-DNA destined for transfer (T-strand). In *vir*-induced cells, the transfer intermediate accumulates in as many as 30 copies per cell. VirD2's translocation signal (TS; + charged C terminus), ParA-like VirC1, and VirD2-binding proteins (VBPs) target the transfer intermediate to the VirB/VirD4 T4SS for delivery to the plant cell. Protein effectors are recruited by charged C-terminal TSs by a chaperone-independent (VirE3, VirF, GALLS) or -dependent (VirE2 effector::VirE1 chaperone) mechanism. The VirB/VirD4 T4SS was reported to localize at the *A. tumefaciens* cell poles (Atmakuri et al. 2007; Judd et al. 2005a, b; Kumar and Das 2002), although more recently the VirB subunits were shown to assemble as non-random foci around the cell (Cameron et al. 2012)

machinery (van Kregten et al. 2009). Furthermore, as noted above, the VirC and VirD1 accessory factors form a ternary complex with VirD2 and VirD4, pointing to factors other than just VirD2 for coupling of the VirD2–T-strand substrate with the receptor. Indeed, chromosomal factors termed VirD2-binding proteins (VBPs) also promote recruitment of the VirD2–T-strand complex to VirD4 (Guo et al. 2007; Padavannil et al. 2014). How the VBPs and VirD1/C1/C2 accessory factors coordinate with each other and the VirD2–T-strand transfer intermediate to control reiterative rounds of substrate–VirD4 docking temporally and spatially remains to be defined (Fig. 2).

The VirB/VirD4 T4SS also translocates several effector proteins, including VirD5, VirE2, VirE3, and VirF (from *A. tumefaciens*), and GALLS-CT and GALLS-FL (from *Agrobacterium rhizogenes*) to target cells (Fig. 2) (Hodges et al. 2006; Schrammeijer et al. 2003; Simone et al. 2001; Vergunst et al. 2000). These proteins are translocated independently of the VirD2–T-strand substrate through the VirB/VirD4 T4SS. Where characterized, these effectors carry positively charged C-terminal domains that are required for engagement with VirD4. Although several are translocated independently of a requirement for a secretion chaperone, the single-stranded binding protein VirE2 must bind its cognate chaperone, VirE1, in

A. tumefaciens to prevent self-aggregation and premature binding to the T-DNA substrate prior to translocation (Atmakuri et al. 2003; Vergunst et al. 2003; Zhao et al. 2001).

5 VirD4 Receptors

The VirD4 receptors are alternatively termed type IV coupling proteins or T4CPs because they link substrates to the T4SS (Cabezon et al. 1997). VirD4-like ATPases are associated with nearly all T4SSs, and the presence of a *virD4*-like gene in sequenced bacterial genomes has served as a marker for a possible T4SS gene cluster (Bhatty et al. 2013). A role for VirD4 subunits in substrate reception was first suggested in early genetic studies showing that VirD4 subunits from different T4SSs can sometimes be exchanged (Cabezon et al. 1997). Specifically, the VirD4 homologs TraG and TrwB, encoded respectively by the RP4 and R388 conjugation systems, substituted for each other in mediating transfer of the promiscuous IncQ plasmid RSF1010 through the heterologous R388 and RP4 mating channels. Similarly, *A. tumefaciens* VirD4 functionally substituted for its homolog pTiC58-encoded TraG in supporting transfer of RSF1010 through the pTiC58 channel (Hamilton et al. 2000). These early receptor-swapping studies highlighted two features of the VirD4 subunits. First, these receptors have evolved to recognize specific repertoires of substrates, although promiscuous IncQ plasmids also have evolved mechanisms to bypass certain substrate specificity checkpoints (see below). Second, VirD4 subunits can functionally engage with heterologous translocation channels. This latter finding is in line with recent results of biochemical fractionation and ultrastructural studies indicating that the VirD4 receptors associate only peripherally, and possibly transiently, with the translocation channel (Larrea et al. 2013; Redzej et al. 2017). Further studies defining the 'coupling' activities of VirD4 receptors in substrate recruitment and channel association are described below.

VirD4 subunits are composed of three or four domains that contribute in distinct ways to receptor or 'coupling' functions (Fig. 3) (Alvarez-Martinez and Christie 2009; Gomis-Ruth et al. 2001). Broadly, their N-terminal transmembrane domains (NTDs) are implicated in forming contacts with T4SS channel subunits, and the C-proximal cytosolic moieties mediate substrate docking and energize substrate transfer reactions (Atmakuri et al. 2004; Hormaeche et al. 2004, 2006; Llosa et al. 2003). An X-ray structure of the soluble, ~50-kDa cytoplasmic domain of the TrwB receptor encoded by the conjugative plasmid R388, revealed a globular hexameric assembly in which each subunit is composed of two distinct domains, a nucleotide-binding domain (NBD) and a 7-helix motif called the all-α-domain (AAD) that faces the cytoplasm. The six TrwB protomers assemble to form a globular ring that is ~110 Å in diameter and 90 Å in height, with a ~20-Å-wide channel in the center that constricts to 8 Å at the cytoplasmic pole (Gomis-Ruth et al. 2001). While lacking among homologs associated with well-characterized *E. coli* conjugation systems (e.g., pKM101, R388), many subunits including

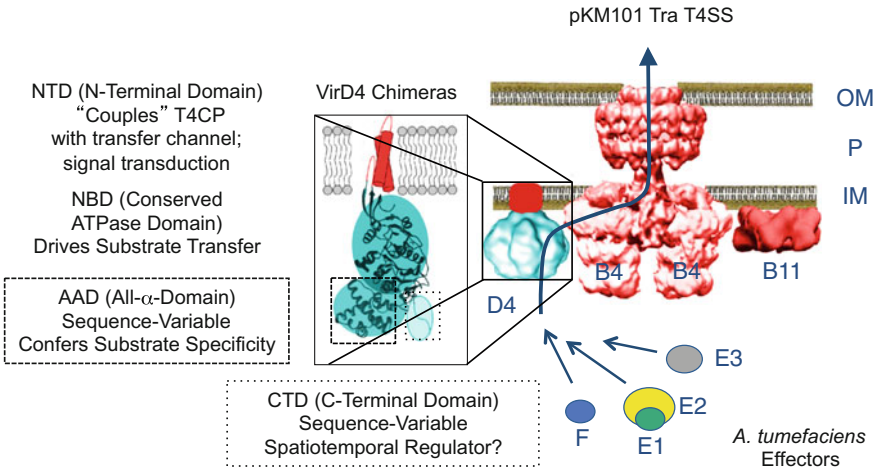


Fig. 3 VirD4 domain composition and functionality of chimeric receptors. Left: VirD4 domains and experimentally derived functions. Right: A VirD4 chimera composed of the N-terminal transmembrane domain (NTD) of the pKM101-encoded TraJ (red) and the cytosolic (NBD/AAD/CTD) moiety of *A. tumefaciens* VirD4 supports transfer of *A. tumefaciens* effector proteins (VirE2, VirE3, VirF) through the 'dedicated' DNA conjugation system encoded by *E. coli* pKM101 (red) (Whitaker et al. 2016)

A. tumefaciens VirD4 also possess sequence-variable C-terminal domains (CTDs) that are typically enriched in acidic residues (Alvarez-Martinez and Christie 2009; Kwak et al. 2017).

5.1 VirD4's NTD: Spatial Positioning and Interaction with the IMC

The N-terminal 87 residues of VirD4 are composed of two membrane-spanning α-helices and an intervening periplasmic loop (Fig. 3) (Das and Xie 1998). Deletion of this NTD does not abolish VirD4's ability to bind the T-strand or VirE2 substrates, but does block substrate transfer to the VirB11 ATPase (Atmakuri et al. 2004; Cascales et al. 2013). These findings suggest that integration of VirD4 into the membrane or NTD-mediated interactions with VirB channel subunits is required for a productive interaction with VirB11. Interestingly, deletion of the entire 30-residue periplasmic loop similarly abolishes substrate transfer to VirB11, whereas smaller deletions of this loop are permissive for this transfer step (Cascales et al. 2013). Early studies showed that VirD4 localizes at *A. tumefaciens* cell poles even in the absence of other VirB channel components. Deletion of the NTD blocks polar localization, establishing the importance of the N-terminal motif for spatial positioning (Atmakuri et al. 2003; Kumar and Das 2002). VirD4's NTD also

interacts with VirB10, which as discussed below is required for transduction of intracellular signals and gating of the distal portion of the VirB channel (see Sect. 6.4).

5.2 VirD4's AAD: Substrate Specificity

The AADs of VirD4 subunits were proposed to play a role in substrate binding, based on structural evidence that the AAD is positioned at the base of the TrwB homohexamer in the crystal structure (Fig. 3) (Gomis-Ruth et al. 2001). These studies further established that the AAD structurally resembles the DNA-binding domain of XerD recombinase, suggestive of a DNA-binding activity (Gomis-Ruth et al. 2001). Mutational studies defined the importance of TrwB's AAD for substrate trafficking, and in the *A. tumefaciens* system a deletion of VirD4's AAD also blocked substrate transfer (de Paz et al. 2010; Whitaker et al. 2015). Further evidence that the VirD4's AAD contributes to substrate recruitment was gained from AAD swapping studies, which showed that VirD4 bearing a heterologous AAD did not support transfer of T-DNA and the VirE2 effector protein to plant cells (Whitaker et al. 2015). Strikingly, however, this VirD4::AAD chimera supported transfer of the mobilizable IncQ plasmid RSF1010 between agrobacterial cells. Furthermore, a purified form of VirD4's AAD bound VirD2 but not VirD1 or heterologous processing factors, and also bound DNA although without sequence or strand specificity. Finally, deletion of VirD2's C-terminal translocation signal did not affect VirD2's relaxase activity but did abolish binding to VirD4 in vitro and correspondingly blocked *A. tumefaciens*-mediated T-DNA transfer to plants (Whitaker et al. 2015). Taken together, these findings support a model in which VirD4's AAD recruits the VirD2 relaxase and covalently associated T-strand through binding of VirD2's C-terminal translocation signal. Interestingly, however, the promiscuous IncQ plasmid RSF1010 retains the ability to bind VirD4 by a mechanism that bypasses the AAD substrate specificity checkpoint (Whitaker et al. 2015).

5.3 VirD4's CTD: Regulator of Substrate Transfer

VirD4 also has a long (104-residue) CTD enriched in acidic residues, which is required for T-strand transfer to plants. Studies exploring the function of VirD4's CTD exploited a discovery that the 'conjugation' system encoded by the *E. coli* plasmid pKM101 can be reconfigured to deliver protein substrates to bacterial recipients by use of engineered chimeric receptors (Fig. 3) (Whitaker et al. 2016). These chimeras consist of the NTD of the VirD4-like TraJ receptor of the pKM101 system joined to the cytosolic moieties of VirD4 homologs associated with effector translocator systems. In the case of the *A. tumefaciens* VirB/VirD4 system, the chimeric receptor is composed of TraJ's NTD fused to VirD4's NBD/AAD/CTD

domains. Intriguingly, this chimeric receptor functionally interacted with the pKM101 translocation channel to support conjugative transfer to *E. coli* recipients of the IncQ plasmid RSF1010, as well as several effector proteins (VirE2, VirF, VirE3) as monitored with the CRAfT assay (Fig. 3). Deletion of VirD4's CTD from the chimeric receptor conferred differential effects on protein trafficking (Whitaker et al. 2016). The ΔCTD variant failed to support VirE2 translocation, but conferred elevated transfer of VirE3 and VirF through the pKM101 channel. Corresponding studies with chimeric receptors composed of receptor domains from VirD4 subunits from *Anaplasma phagocytophilum* and *Wolbachia pipientis* also supported transfer of known or candidate effector proteins from these species through the pKM101 channel. These VirD4 subunits also have long, acidic CTDs, and deletion of the CTDs from the chimeric receptors also attenuated or enhanced transfer frequencies of different protein effectors (Alvarez-Martinez and Christie 2009; Whitaker et al. 2016). Taken together, these findings support the notion that the CTDs of VirD4 receptors have evolved both for expansion of the effector repertoire and as a regulatory domain to exert spatiotemporal control of effector presentation to the T4SS (Fig. 3).

6 The Envelope-Spanning VirB/VirD4 T4SS

6.1 Cytoplasmic Entrance: The ATPase Energy Center

In *A. tumefaciens*, VirD4 physically and functionally interacts with the VirB4 and VirB11 ATPases to drive substrate transfer through the T4SS (Atmakuri et al. 2004; Cascales and Christie 2004b; Pena et al. 2012; Ripoll-Rozada et al. 2013; Savvides et al. 2003). The $VirB_{3–10}$/VirD4 substructure presents a view of the spatial arrangement of VirB4 hexamers and VirD4 dimers, but at this time it is thought that the VirD4 receptors function in vivo as homohexamers. This, coupled with the lack of a VirB11 subunit in the R388-encoded structure, limits our knowledge of how the three ATPases physically and functionally interact to drive early-stage substrate processing reactions. Whereas VirD4 and VirB4 subunits are phylogenetically related and likely adopt similar hexameric structures, VirB11 is by contrast a member of a large family of ATPases termed 'traffic ATPases' that are associated with the types II, III, and IV secretion systems (Kato et al. 2015; Pena and Arechaga 2013; Savvides 2007). VirB11-like ATPases cofractionate with the cytoplasm and IM, suggestive of a dynamic association with the membrane or the VirB/VirD4 T4SS (Rashkova et al. 1997). VirB11 homologs form stable hexameric rings of ∼100–120 Å in diameter (Savvides et al. 2003; Yeo et al. 2000). The N- and C-terminal halves of the six protomers each form rings, giving rise to a double-stacked structure wherein the nucleotide-binding site is at the interface between the two domains. EM studies have shown that the VirB11 hexamers undergo dynamic structural changes upon ATP binding and hydrolysis, although the functional importance of these transitions is not known at this time (Savvides

et al. 2003). VirD4 delivers the recruited DNA substrate to the VirB11 ATPase (Fig. 2). As mentioned earlier, this transfer step does not require ATP hydrolysis by either subunit, suggestive of a direct pass-off from one to the other subunits independently of any energy-driven conformational changes (Atmakuri et al. 2004). Mutational studies have further defined contributions of the VirD4 and VirB11 ATPases to early and late stages of substrate transfer. With regard to the early-stage transfer reactions, VirD4 deleted of its NTD retained the ability to bind the T-DNA substrate, but failed to transfer the substrate to VirB11, suggesting that VirD4 must be docked at the IM to productively engage with VirB11 (Cascales et al. 2013). Also as noted earlier, transfer of T-DNA from VirD4 to VirB11 requires cosynthesis of VirB7, VirB9, and VirB10 (Atmakuri et al. 2004). VirB10 is likely responsible for communicating directly with VirD4 to activate this first substrate transfer step, as suggested by evidence that VirD4 interacts with VirB10 in the *A. tumefaciens* VirB/VirD4 T4SS as well as the related *E. coli* plasmid R388 transfer system (Atmakuri et al. 2004; de Paz et al. 2005; Llosa et al. 2003).

Mutational analyses of VirB11 have been informative in deciphering the role of VirB11 in substrate transfer. For example, one large class of VirB11 mutations imposes blocks in T-DNA transfer across the T4SS at different points by preventing substrate transfer from (i) VirD4 to VirB11, (ii) VirB11 to VirB6/VirB8, or (iii) VirB6/VirB8 to VirB2/VirB9 (Cascales et al. 2013). Mutations conferring these stage-specific blocks are predicted to affect VirB11's interaction with VirD4 in distinct ways or disrupt ATP-binding or hydrolysis activities required for latter-stage transfer reactions (see also Sect. 7). A second broad class of VirB11 mutations was designated as substrate discrimination mutations because they selectively impair transfer of one but not other substrates through the VirB/VirD4 T4SS. Mutations in the ATP-binding pocket block all substrate transfer, but the substrate discrimination mutations map elsewhere and block T-DNA transfer without affecting transfer of the RSF1010 plasmid or VirE2 effector substrates (Cascales et al. 2013; Sagulenko et al. 2001; Zhou and Christie 1997). In a type III secretion system, the VirB11-like subunit InvC dissociates chaperones and unfolds effectors prior to translocation (Akeda and Galan 2005). If VirB11 catalyzes similar reactions, the substrate discrimination mutations might selectively impair an activity required for processing of T-DNA but not the RSF1010 or VirE2 substrates.

6.2 The IMC

Based on the $VirB_{3-10}$ structure, the IMC of the *A. tumefaciens* VirB/VirD4 T4SS is composed of the integral membrane subunits VirB3, VirB6, and VirB8, the N terminus of VirB10, two VirB4 hexamers, and a VirD4 dimer or hexamer (see Fig. 1c) (Low et al. 2014). The IMC is postulated to form a channel across the membrane, but there remain several possible routes by which substrates transit the IM: (1) the lumen of the VirD4 hexamer, (2) the lumen of the VirB4 hexamer, or (3) a channel formed by the VirB6 and VirB8 subunits (Fig. 1c). Favoring the latter pathway, DNA

substrates form close contacts with both VirB6 and VirB8 in the TrIP assay (Cascales and Christie 2004b). Mutations in VirB6 were also shown to permit DNA contacts with VirB6 but block contacts with VirB8, suggesting that the DNA substrate sequentially associates with VirB6 and then VirB8 (Jakubowski et al. 2004). Finally, as mentioned above, the ATP-binding/hydrolysis activities of all three ATPases (VirD4, VirB4, VirB11) are necessary for establishment of the DNA substrate contacts with VirB6 and VirB8, suggesting that ATP energy drives substrate transfer through the IM channel (Atmakuri et al. 2004; Cascales and Christie 2004b).

6.3 The OMCC

Recently, the VirB7/VirB9/VirB10 subassembly was purified from *A. tumefaciens* and shown to adopt a ring-shaped complex resembling equivalent subassemblies of T4SSs encoded by the conjugative plasmids pKM101 and R388 (Fig. 4) (Gordon et al. 2017). Overall, these OMCCs present as large (~1-MDa) double-walled barrels composed of an IM-associated or I layer and an OM-associated or O layer. The I layer is composed of the VirB9 and VirB10 subunits and narrows to form a ring of ~55–85 Å diameter at the IM. The O layer is composed of VirB7, VirB9, and VirB10 subunits that form a main body and narrow cap with a hole of 20–30 Å that is presumed to span the outer membrane. Structural analyses further showed that the OMCC is composed of 14 copies of each of the three subunits (Chandran et al. 2009; Fronzes et al. 2009). The distal end of the OMCC consists of a cap, which is built from 14 copies of an α-helical domain of the VirB10 subunit termed 'the antennae projection' or 'AP.' This cap is thought to span the OM and form a pore through which substrates are conveyed and the T-pilus projects (Banta et al. 2011; Chandran et al. 2009). Interestingly, despite extensive contacts among the VirB7, VirB9, and VirB10 proteins in the OMCC, each of these subunits is tolerant of substitution and small deletion and insertion mutations (Gordon et al. 2017; Jakubowski et al. 2005, 2009). Indeed, certain structural motifs, such as the OM-spanning cap or a short α-helical arm at the base of the OMCC, are completely dispensable for substrate transfer through the T4SS (Gordon et al. 2017). Perhaps the most extreme example of the compositional flexibility of the OMCC was demonstrated through the construction of chimeric T4SSs in *E. coli.* These chimeric T4SSs are composed of the IMC from the pKM101-encoded T4SS joined to OMCCs derived from heterologous T4SSs (Fig. 4). The heterologous OMCCs were from the *E. coli* Trw_{R388} conjugation machine, the *A. tumefaciens* VirB/VirD4 T4SS, and the *Bordetella pertussis* Ptl system, respectively. In their native contexts, these OMCCs mediate transfer of a DNA substrate (Trw_{R388}), a combination of DNA and protein substrates (*A. tumefaciens* VirB/VirD4), and the multi-subunit pertussis toxin (Ptl). Yet, in the context of chimeric T4SSs, the heterologous OMCCs supported transfer of the pKM101 DNA substrate to *E. coli* recipient cells. The functionality of these chimeric T4SSs underscores the highly structural and functional conservation of OMCCs among the T4SSs (Fig. 4) (Gordon et al. 2017).

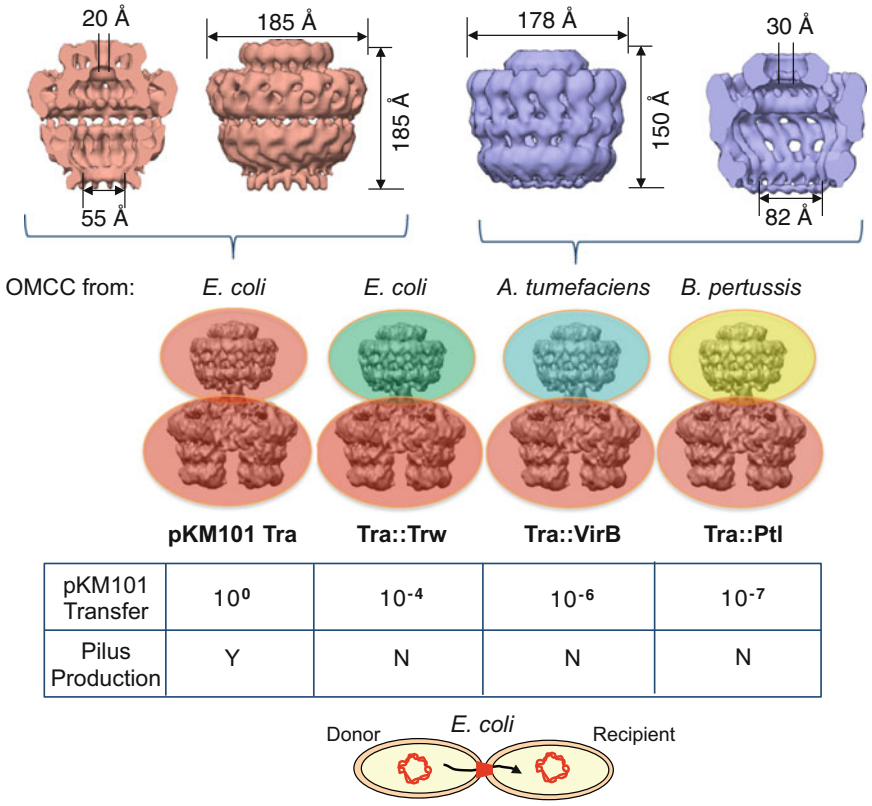


Fig. 4 Structural conservation of T4SS outer membrane subassemblies. Upper: Structures of outer membrane core complexes (OMCCs) associated with the *E. coli* pKM101 Tra and *A. tumefaciens* VirB/VirD4 T4SSs, as determined by single-particle, negative-stain electron microscopy (NS-EM). Dimensions of the overall subassemblies and openings at proximal and distal ends are indicated. Structures are reproduced with permission from (Gordon et al. 2017). Lower: Functionality of chimeric T4SSs composed of the pKM101-encoded Tra IMC joined to heterologous OMCCs from the *E. coli* R388-encoded Trw, *A. tumefaciens* VirB/VirD4, and *B. pertussis* Ptl T4SSs. Transfer frequencies of pKM101 are presented as transconjugants per donor resulting from 2 h filter matings between *E. coli* donors and recipients. Data are from (Gordon et al. 2017)

6.4 *Signal Activation of the Translocation Channel*

The functionality of these chimeric machines is particularly intriguing in view of early findings that *A. tumefaciens* VirB10 undergoes a structural transition that is required for substrate transfer through the distal region of the VirB/VirD4 T4SS (Cascales and Christie 2004a). This structural transition was detected as a change in susceptibility to the *S. griseus* protease upon treatment of spheroplasts and occurs in response to a combination of ATP energy utilization as well as DNA substrate

binding by the VirD4 and VirB11 ATPases (Cascales et al. 2013; Cascales and Christie 2004a). VirB10 thus serves not only as a structural scaffold for the translocation channel, but also to transduce intracellular signals to the distal portion of the channel. Signal activation is thought to induce a conformational change in the channel allowing for substrate passage. Consistent with this mode of action, a mutation in a region of VirB10 located near the OM locks the channel in an activated conformation, resulting in leakage of the VirE2 substrate to the cell surface independently of target cell contact (Banta et al. 2011).

7 The T-Pilus: Assembly and Function

7.1 T-Pilus Subunits

The VirB/VirD4 T4SS and related conjugation systems elaborate conjugative pili in addition to the translocation channel (Fig. 5). The VirB-encoded pilus, termed the T-pilus, is composed of the major pilin subunit VirB2 and a minor subunit VirB5 (Aly and Baron 2007; Kado 2000; Lai and Kado 1998). VirB2 is synthesized as a pro-pilin of ~12.3-kDa in molecular size, which is cleaved of its relatively long (~5-kDa) signal sequence to yield a mature protein of ~7-kDa that inserts into the IM via two hydrophobic domains. The N- and C-termini of the mature protein then undergo a novel cyclization reaction, yielding a covalently joined cyclic peptide (Eisenbrandt et al. 1999). The cyclized pilin monomers form a pool in the IM for recruitment to build the pilus polymer in response to a presently undefined signal. Pilus nucleation is thought to require VirB5, which binds the tip of the T-pilus and likely plays an important role in mediating attachment of *A. tumefaciens* with plant or other agrobacterial target cells (Aly and Baron 2007; Yuan et al. 2005). Early studies established that T-pilus assembly requires most of the same subunits needed to build the translocation channel (Berger and Christie 1994; Fullner et al. 1996). Two important exceptions are VirB1, which is essential for T-pilus production but not a functional translocation channel, and VirD4, which is required for substrate transfer but not T-pilus production (Berger and Christie 1994; Zupan et al. 2007). The dispensability of VirB1 for channel formation is of interest in view of its function as a lytic transglycosylase, which presumably should be necessary for extension of the channel across cell wall. However, the channel might form during periods of active cell wall remodeling or through recruitment of another host-encoded hydrolase, dispensing with the need for a dedicated hydrolase.

For pilus assembly, VirB1 might be required for successful recruitment of VirB2 and/or VirB5 to the T4SS, or it may be necessary for pilus extension beyond the cell surface. Favoring this latter idea, VirB1 undergoes proteolysis to yield separate N-terminal transglycosylase and C-terminal domains (Zupan et al. 2007). The C-terminal domain, designated VirB1*, is required for T-pilus production and, interestingly, is exported across the OM where it associates with the T-pilus.

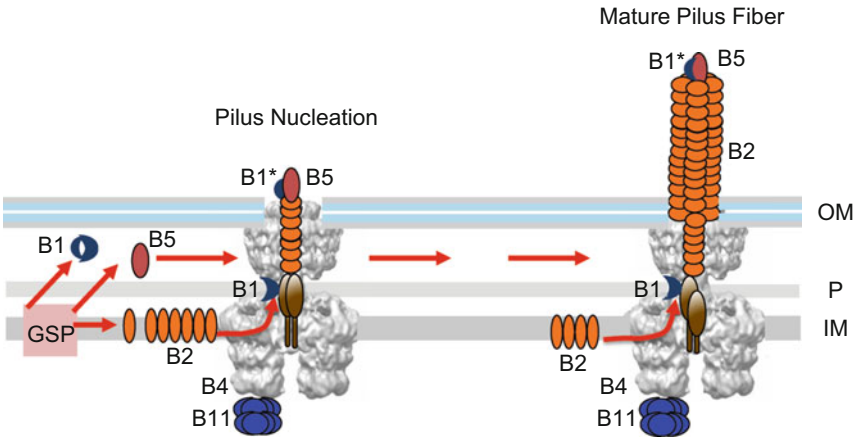


Fig. 5 Postulated T-pilus assembly pathway. Pilus-associated subunits are exported to the periplasm (P) or integrated into the inner membrane (IM) by the general secretory pathway (GSP). The pilus-tip protein VirB5 is recruited to the VirB T4SS where it nucleates pilus assembly from an IM platform composed in part of VirB8 (brown lollipop). VirB1 is proteolytically cleaved in the periplasm, and its N-terminal transglycosylase domain is recruited to the T4SS where it facilitates pilus assembly through localized degradation of the cell wall. The C-terminal domain, VirB1* (B1*), is delivered to the T-pilus tip where it facilitates pilus polymerization and extension. The VirB2 pilin is integrated into the IM, cyclized by covalent linkage of its N- and C-termini, and recruited upon receipt of an unknown signal to the T4SS for pilus assembly. The VirB4 ATPase, with a contribution by the VirB11 ATPase, catalyzes dislocation of the pilin monomers from the IM. The model depicts the IM as the nucleation platform for a thin pilus fiber that extends through the OMCC and across the outer membrane (OM). Once the fiber contacts the distal region of the OMCC, further polymerization leads to extension of a thicker ($\sim$10 nm) helically arrayed filament from the cell surface

VirB1 thus is postulated to contribute in at least two ways to pilus production. Its N-terminal transglycosylase domain provides localized lysis of the cell wall for extension of the pilus through the periplasm, and its C-terminal VirB1* domain promotes T-pilus assembly through protein–protein interactions with T-pilus subunits (Fig. 5) (Zupan et al. 2007). In line with this proposal, certain VirB1 homologs are capable of functionally substituting for VirB1 in supporting T-pilus production in *A. tumefaciens*. For example, homologs associated with T4SSs encoded by *Brucella suis* and the *E. coli* conjugative plasmid pKM101, but not those associated with T4SSs encoded by the *Helicobacter pylori* Cag T4SS and the F plasmid, substitute for *A. tumefaciens* VirB1. Interestingly, the complementing homologs carry C-terminal domains whereas those from *H. pylori* and the F plasmid do not have such domains. These findings underscore the importance of the C-terminal domains of VirB1-like subunits for pilus production (Hoppner et al. 2004).

VirB5 is a minor pilin subunit that associates with the tip of the T-pilus (Fig. 5) (Aly and Baron 2007). VirB5 is implicated in nucleation of the T-pilus from either an inner or outer membrane platform, and also might play a role in binding of plant

cell receptors (Yuan et al. 2005). Indeed, such receptor-binding activity was shown for VirB5-like CagL, which is associated with the *H. pylori* Cag T4SS. CagL binds β-integrin receptors present on the surface of mammalian epithelial cells to promote *H. pylori* binding and infection (Backert et al. 2008) (Barden et al. 2013). In *A. tumefaciens*, a corresponding function is supported by the findings that *A. tumefaciens* cells engineered to overproduce VirB5, or the addition of exogenous VirB5 to an infection site, enhances T-DNA transfer as monitored by a transient expression assay (Lacroix and Citovsky 2011). VirB5 therefore appears to function both in nucleation of the T-pilus and in mediating attachment to plant cell receptors to facilitate establishment of productive *A. tumefaciens*–plant cell mating junctions. Interestingly, the VirB5 homolog, TraC, from the pKM101 T4SS weakly substitutes for VirB5 with respect to T-pilus production, although not for T-DNA transfer (Schmidt-Eisenlohr et al. 1999). It is not surprising that *A. tumefaciens* VirB5 has evolved a specialized attachment motif(s) for binding of plant cells that is missing in homologs associated with dedicated bacterial conjugation systems. VirB5 also binds the *trans*-zeatin enzyme Tzs, which is involved biosynthesis of cytokinin and mediates its localization to the *A. tumefaciens* cell surface (Aly et al. 2008). Cytokinin induces production of At14a, a plant protein that links plant cell walls to the plasma membrane and the cytoskeleton (Sardesai et al. 2013). A recent study further presented evidence that the plant cytoskeleton may be a conduit for translocation of the VirE2 effector through the plant cell (Yang et al. 2017). Thus, VirB5-mediated localization of Tzs to the cell surface could stimulate the localized synthesis of cytokinins, which in turn would promote localized transfer of T-DNA and effector proteins.

7.2 T-Pilus Biogenesis Pathway

Several features of the T-pilus assembly pathway have been defined. In a Cys-accessibility study, the topology of membrane-integrated, cyclic form of VirB2 was mapped by determining the accessibility of engineered Cys residues to membrane-impermeable thiol reactive reagents. These studies established a topology in which the loop formed by cyclization of the N- and C-termini is located in the periplasm and a small loop between the two hydrophobic sequences is in the cytoplasm (Kerr and Christie 2010). Interestingly, Cys residues in the cytoplasmic loop that are inaccessible to thiol modification in the absence of the VirB T4SS became accessible in its presence. Further studies established that coproduction of the VirB4 ATPase and Cys-substituted VirB2 pilins sufficed to promote accessibility of the cytoplasmic Cys residues to thiol modification. VirB4 also was shown to release VirB2 into the periplasm in osmotic shock experiments, and to form an immunoprecipitable complex with VirB2, suggestive of a direct effect of this ATPase on VirB2's membrane topology. Mutation of the Walker A nucleotide triphosphate-binding site abolished all of the observed VirB4-mediated activities. Further, production of VirB11 together with VirB4 affected VirB2's membrane

topology, suggesting a coordination of function between the two ATPases. Together, these data support a model in which VirB4, with an unspecified contribution by VirB11, functions as a dislocase to extract membrane-inserted VirB2 as a prerequisite for T-pilus assembly (Fig. 5) (Kerr and Christie 2010). VirB4 subunits are signatures of all bacterial conjugation machines and thus might function generally as pilin dislocases.

Once pilins are extracted from the membrane, how and where on the cell envelope do they polymerize? These questions are not fully answered, but protein–protein interaction studies have provided some clues. For example, VirB2 interacts with the VirB5 pilus-tip protein and with the IMC subunit VirB8. VirB5 also interacts with VirB8, but additionally with the OMCC subunit VirB10 (Yuan et al. 2005). Based on this interaction network, a VirB2–VirB5 complex might engage with VirB8, which in turn could form a nucleation platform for extension of the pilus through the OMCC to the cell exterior. The chamber of the *A. tumefaciens* OMCC is approximately 100 Å in diameter, sufficiently large to house the T-pilus ($\sim$10 nm in width). However, the OM pore formed by the cap domain is at most only 20–30 Å and clearly not large enough to accommodate the pilus (Chandran et al. 2009; Gordon et al. 2017). Thus, if the T-pilus assembles from an IM platform, pilus extension must induce gross structural changes in the distal region of the OMCC.

In the above model, the OMCC could be considered a passive structural scaffold for the extending T-pilus. However, two observations favor a more active role for the OMCC in pilus polymerization. First, mutations in VirB10 that do not affect elaboration of a functional translocation channel were shown to selectively block pilus biogenesis (Jakubowski et al. 2009). The most noteworthy of these pilus-blocking mutations are deletions of part or all of the AP domains comprising the distal cap (Gordon et al. 2017). The essentiality of the cap domain suggests it could play an active role in pilus assembly. Second, although the chimeric T4SSs described above support DNA transfer (See Fig. 4), the heterologous OMCCs associated with pKM101's IMC did not support pilus biogenesis (Gordon et al. 2017). These findings suggest that the OMCC has evolved a specific role in assembly of cognate pili. Accordingly, following recruitment of VirB2–VirB5 complexes to VirB8, the pilin subunits might be shunted within the OMCC's chamber to the distal cap, which in turn mediates pilus nucleation. Another scenario depicted in Fig. 5 is that VirB2 begins to polymerize from an IM platform, but as a thin fiber that extends to the cell surface. Upon contacting the OMCC and, specifically, the distal cap, the fiber transitions to form the helically arrayed mature T-pilus.

7.3 Other Attachment Mechanisms

Whether pilus nucleation initiates from IM or OM platforms remains one of the central questions surrounding T-pilus biogenesis. A related question is whether the T-pilus contributes directly to substrate transfer. Although assembly of the channel

and T-pilus require nearly the same subunits, mutational studies have identified classes of mutations that selectively block substrate transfer without affecting pilus production (Tra^-, Pil^+) or conversely, permit substrate transfer but block pilus production (Tra^+, Pil^-) (Jakubowski et al. 2005, 2009; Sagulenko et al. 2001). It is possible that the latter mutations simply block pilus outgrowth and that *A. tumefaciens* cells harboring such mutations form short pili that cannot be detected by current assays. However, studies of the *E. coli* pKM101 system also have identified 'uncoupling' mutations that completely block pilus biogenesis as monitored by resistance to filamentous phages, which are known to bind the pilus tip (Gordon et al. 2017). The identification of 'uncoupling' mutations clearly argues against a direct role for the T-pilus in substrate transfer, suggesting instead that the T-pilus functions predominantly or exclusively as an attachment organelle. It is interesting to note that VirB2 pilin is required in Pil^-, Tra^+ 'uncoupling' mutants for successful T-DNA transfer. Furthermore, VirB2 interacts with various plant cytosolic proteins of established importance for efficient T-DNA transfer (Hwang and Gelvin 2004; Huang et al. 2018). These findings raise the possibility that in addition to their roles in T-pilus polymerization pilin monomers might alternatively be delivered via the T4SS into plant cells where contacts with certain plant proteins stimulate VirB/VirD4-mediated transformation. *A. tumefaciens* also elaborates several adhesive organelles or molecules on its cell surface, including unipolar polysaccharide (UPP), other exopolysaccharides (cyclic-β-1,2-D-glucan, succinoglycan, lipopolysaccharide), cellulose, calcium-binding surface proteins (rhicadhesin, RapA1), and type IV pili (Bash and Matthysse 2002; Heindl et al. 2014; Li et al. 2012; Matthysse 2014; Matthysse et al. 2005; Wang et al. 2014; Xu et al. 2013). It can be predicted that a combination of T-pili, possibly the VirB2 pilin and VirB5/Tzs, and other surface adhesins coordinate productive binding and colonization of *A. tumefaciens* on plant tissues. Importantly, even in the absence of T-pilus production, these alternative surface adhesins suffice to establish contacts necessary for VirB/VirD4-mediated interkingdom translocation.

8 Conclusions

Studies of the *A. tumefaciens* VirB/VirD4 T4SS have supplied key mechanistic and structural insights into the T4SS secretion superfamily. Compositionally, this is one of the simplest T4SSs, especially among systems elaborating both a trans-envelope translocation channel and an extracellular pilus. Early work defined the contributions and subcellular distributions of individual VirB and VirD4 subunits to machine and pilus assembly. Next, the development of creative in vivo assays, such as CRAfT and TrIP, refined our understanding of substrate recruitment and transfer reactions. Most recently, structural characterization of purified T4SS subassemblies, including OMCCs from the *A. tumefaciens* VirB/VirD4 system and several related conjugation machines, as well as the larger $VirB_{3–10}$/VirD4 substructure has

provided an architectural context for envisioning how these machines mediate substrate transfer across the cell envelope.

Strikingly, certain machine subassemblies, such as the VirD4 substrate receptor and the OMCC, are exchangeable between the VirB/VirD4 and closely related conjugation systems. Studies of such chimeric systems have been highly informative in two ways. First, investigations have supplied new mechanistic detail regarding substrate recruitment and pilus biogenesis pathways. Second, these studies have supplied a proof-of-principle for the engineering T4SSs with novel functions. Engineering of the *E. coli* pKM101 Tra T4SS, once considered a 'dedicated' conjugation machine, allowed for diversification of the substrate repertoire of this highly efficient nanomachine to include a potentially wide array of effector proteins. By extension, this conjugation system or other T4SSs normally functioning in infection processes, e.g., *A. tumefaciens* VirB/VirD4, *H. pylori* Cag, *Legionella* Dot/Icm, might be engineered to recruit and translocate novel substrates of potential therapeutic benefit to different eukaryotic target cells of interest. Indeed, as a first step in the application of T4SSs as substrate delivery platforms, T4SSs functioning in *Bartonella henselae*, *Legionella pneumophila*, and *Coxiella burnetii* were recently retailored to deliver DNA substrates to distinct mammalian cells (Guzman-Herrador et al. 2017).

Despite the remarkable progress in our studies of the VirB/VirD4 T4SS, there remain many questions: (1) How do the VirD4, VirB4, and VirB11 subunits coordinate their ATPase activities to recruit substrates and drive their translocation; (2) What is the architecture of the envelope-spanning channel, how is it gated at its entrance and exit points, and what signals regulate channel activity? (3) What is the structural and functional relationship of the translocation channel and the T-pilus? and (4) How does the T4SS establish productive pilus-dependent and -independent contacts with target cells? Answers to these questions will require continued commitment to basic studies of the paradigmatic VirB/VirD4 and related conjugation systems. Clearly, however, the pace of research on the T4SSs is accelerating rapidly at the ultrastructural level due to the implementation of high-resolution microscopy approaches. When combined with creative new in vivo approaches enabling structure–function assignments and definition of machine dynamics in real time, the incipient discoveries promise to refine and possibly completely reshape our views of these fascinatingly complex nanomachines.

Acknowledgements Work in the Christie laboratory is supported by NIH grant R01GM48476.

References

Akeda Y, Galan JE (2005) Chaperone release and unfolding of substrates in type III secretion. Nature 437:911–915. https://doi.org/10.1038/nature03992

Albright LM, Yanofsky MF, Leroux B, Ma DQ, Nester EW (1987) Processing of the T-DNA of *Agrobacterium tumefaciens* generates border nicks and linear, single-stranded T-DNA. J Bacteriol 169:1046–1055

Alvarez-Martinez CE, Christie PJ (2009) Biological diversity of prokaryotic type IV secretion systems. Microbiol Mol Biol Rev 73:775–808

Aly KA, Baron C (2007) The VirB5 protein localizes to the T-pilus tips in *Agrobacterium tumefaciens*. Microbiology 153:3766–3775

Aly KA, Krall L, Lottspeich F, Baron C (2008) The type IV secretion component VirB5 binds to the trans-zeatin biosynthetic enzyme Tzs and enables its translocation to the cell surface of *Agrobacterium tumefaciens*. J Bacteriol 190:1595–1604

Anderson LB, Hertzel AV, Das A (1996) *Agrobacterium tumefaciens* VirB7 and VirB9 form a disulfide-linked protein complex. Proc Natl Acad Sci USA 93:8889–8894

Atmakuri K, Cascales E, Burton OT, Banta LM, Christie PJ (2007) *Agrobacterium* ParA/MinD-like VirC1 spatially coordinates early conjugative DNA transfer reactions. EMBO J 26:2540–2551

Atmakuri K, Cascales E, Christie PJ (2004) Energetic components VirD4, VirB11 and VirB4 mediate early DNA transfer reactions required for bacterial type IV secretion. Mol Microbiol 54:1199–1211

Atmakuri K, Ding Z, Christie PJ (2003) VirE2, a type IV secretion substrate, interacts with the VirD4 transfer protein at cell poles of *Agrobacterium tumefaciens*. Mol Microbiol 49:1699–1713

Backert S, Fronzes R, Waksman G (2008) VirB2 and VirB5 proteins: specialized adhesins in bacterial type-IV secretion systems? Trends Microbiol 16:409–413

Banta LM et al (2011) An *Agrobacterium* VirB10 mutation conferring a type IV secretion system gating defect. J Bacteriol 193:2566–2574

Barden S, Lange S, Tegtmeyer N, Conradi J, Sewald N, Backert S, Niemann HH (2013) A helical RGD motif promoting cell adhesion: crystal structures of the *Helicobacter pylori* type IV secretion system pilus protein CagL. Structure 21:1931–1941

Baron C, Thorstenson YR, Zambryski PC (1997) The lipoprotein VirB7 interacts with VirB9 in the membranes of *Agrobacterium tumefaciens*. J Bacteriol 179:1211–1218

Bash R, Matthysse A (2002) Attachment to roots and virulence of a *chvB* mutant of *Agrobacterium tumefaciens* are temperature sensitive. Mol Plant Microbe Interact 15:160–163

Beaupre CE, Bohne J, Dale EM, Binns AN (1997) Interactions between VirB9 and VirB10 proteins involved in movement of DNA from *Agrobacterium tumefaciens* into plant cells. J Bacteriol 179:78–89

Beijersbergen A, Dulk-Ras AD, Schilperoort RA, Hooykas PJJ (1992) Conjugative transfer by the virulence system of *Agrobacterium tumefaciens*. Science 256:1324–1327

Beijersbergen A, Smith SJ, Hooykaas PJ (1994) Localization and topology of VirB proteins of *Agrobacterium tumefaciens*. Plasmid 32:212–218

Berger BR, Christie PJ (1994) Genetic complementation analysis of the *Agrobacterium tumefaciens virB* operon: *virB2* through *virB11* are essential virulence genes. J Bacteriol 176:3646–3660

Bhatty M, Laverde Gomez JA, Christie PJ (2013) The expanding bacterial type IV secretion lexicon. Res Microbiol 164:620–639. https://doi.org/10.1016/j.resmic.2013.03.012

Buchanan-Wollaston V, Passiatore JE, Cannon F (1987) The *mob* and *oriT* mobilization functions of a bacterial plasmid promote its transfer to plants. Nature 328:172–175

Bundock P, den DRA, Beijersbergen A, Hooykaas PJ (1995) Trans-kingdom T-DNA transfer from *Agrobacterium tumefaciens*. EMBO J 14:3206–3214

Cabezon E, Sastre JI, de la Cruz F (1997) Genetic evidence of a coupling role for the TraG protein family in bacterial conjugation. Mol Gen Genet 254:400–406

Cameron TA, Roper M, Zambryski PC (2012) Quantitative image analysis and modeling indicate the *Agrobacterium tumefaciens* type IV secretion system is organized in a periodic pattern of foci. PLoS ONE 7:e42219. https://doi.org/10.1371/journal.pone.0042219

Cascales E, Atmakuri K, Sarkar MK, Christie PJ (2013) DNA substrate-induced activation of the *Agrobacterium* VirB/VirD4 type IV secretion system. J Bacteriol 195:2691–2704. https://doi.org/10.1128/JB.00114-13

Cascales E, Christie PJ (2004a) *Agrobacterium* VirB10, an ATP energy sensor required for type IV secretion. Proc Natl Acad Sci U S A 101:17228–17233

Cascales E, Christie PJ (2004b) Definition of a bacterial type IV secretion pathway for a DNA substrate. Science 304:1170–1173

Chandran V, Fronzes R, Duquerroy S, Cronin N, Navaza J, Waksman G (2009) Structure of the outer membrane complex of a type IV secretion system. Nature 462:1011–1015

Christie PJ (1997) *Agrobacterium tumefaciens* T-complex transport apparatus: a paradigm for a new family of multifunctional transporters in eubacteria. J Bacteriol 179:3085–3094

Christie PJ, Atmakuri K, Krishnamoorthy V, Jakubowski S, Cascales E (2005) Biogenesis, architecture, and function of bacterial type IV secretion systems. Annu Rev Microbiol 59: 451–485

Das A, Xie YH (1998) Construction of transposon Tn3phoA: its application in defining the membrane topology of the *Agrobacterium tumefaciens* DNA transfer proteins. Mol Microbiol 27:405–414

de la Cruz F, Frost LS, Meyer RJ, Zechner EL (2010) Conjugative DNA metabolism in Gram-negative bacteria. FEMS Micro Rev 34:18–40. https://doi.org/10.1111/j.1574-6976.2009.00195.x

de Paz HD, Larrea D, Zunzunegui S, Dehio C, de la Cruz F, Llosa M (2010) Functional dissection of the conjugative coupling protein TrwB. J Bacteriol 192:2655–2669

de Paz HD, Sangari FJ, Bolland S, Garcia-Lobo JM, Dehio C, de la Cruz F, Llosa M (2005) Functional interactions between type IV secretion systems involved in DNA transfer and virulence. Microbiology 151:3505–3516

Durrenberger MB, Villiger W, Bachi TH (1991) Conjugational junctions: morphology of specific contacts in conjugating *Escherichia coli* bacteria. J Struct Biol 107:146–156

Eisenbrandt R, Kalkum M, Lai EM, Lurz R, Kado CI, Lanka E (1999) Conjugative pili of IncP plasmids, and the Ti plasmid T pilus are composed of cyclic subunits. J Biol Chem 274:22548–22555

Fernandez D, Spudich GM, Zhou XR, Christie PJ (1996) The *Agrobacterium tumefaciens* VirB7 lipoprotein is required for stabilization of VirB proteins during assembly of the T-complex transport apparatus. J Bacteriol 178:3168–3176

Filichkin SA, Gelvin SB (1993) Formation of a putative relaxation intermediate during T-DNA processing directed by the Agrobacterium tumefaciens Vir D1, D2 endonuclease. Mol Micribiol 8:915–926

Fronzes R, Schafer E, Wang L, Saibil HR, Orlova EV, Waksman G (2009) Structure of a type IV secretion system core complex. Science 323:266–268

Fullner KJ (1998) Role of *Agrobacterium virB* genes in transfer of T complexes and RSF1010. J Bacteriol 180:430–434

Fullner KJ, Lara JC, Nester EW (1996) Pilus assembly by *Agrobacterium* T-DNA transfer genes. Science 273:1107–1109

Gomis-Ruth FX, Moncalian G, Perez-Luque R, Gonzalez A, Cabezon E, de la Cruz F, Coll M (2001) The bacterial conjugation protein TrwB resembles ring helicases and F1-ATPase. Nature 409:637–641

Gordon JE et al (2017) Use of chimeric type IV secretion systems to define contributions of outer membrane subassemblies for contact-dependent translocation. Mol Microbiol 105:273–293. https://doi.org/10.1111/mmi.13700

Grohmann E, Christie PJ, Waksman G, Backert S (2018) Type IV secretion in gram-negative and gram-positive bacteria. Mol Microbiol 107:455–471

Guo M, Hou Q, Hew CL, Pan SQ (2007) *Agrobacterium* VirD2-binding protein is involved in tumorigenesis and redundantly encoded in conjugative transfer gene clusters. Mol Plant Microbe Interact 20:1201–1212

Guzman-Herrador DL, Steiner S, Alperi A, Gonzalez-Preito C, Roy CR, Llosa M (2017) DNA delivery and genomic integration into mammalian target cells through type IV A and B secretion systems of human pathogens. Front Microbiol 8:1503. https://doi.org/10.3389/fmicb.2017.01503

Hamilton CM et al (2000) TraG from RP4 and TraG and VirD4 from Ti plasmids confer relaxosome specificity to the conjugal transfer system of pTiC58. J Bacteriol 182:1541–1548

Heindl JE, Wang Y, Heckel BC, Mohari B, Feirer N, Fuqua C (2014) Mechanisms and regulation of surface interactions and biofilm formation in *Agrobacterium*. Front Plant Sci 5:176. https://doi.org/10.3389/fpls.2014.00176

Hodges LD, Vergunst AC, Neal-McKinney J, den Dulk-Ras A, Moyer DM, Hooykaas PJ, Ream W (2006) *Agrobacterium rhizogenes* GALLS protein contains domains for ATP binding, nuclear localization, and Type IV secretion. J Bacteriol 188:8222–8230

Hoppner C, Liu Z, Domke N, Binns AN, Baron C (2004) VirB1 orthologs from *Brucella suis* and pKM101 complement defects of the lytic transglycosylase required for efficient type IV secretion from *Agrobacterium tumefaciens*. J Bacteriol 186:1415–1422

Hormaeche I, Iloro I, Arrondo JL, Goni FM, de la Cruz F, Alkorta I (2004) Role of the transmembrane domain in the stability of TrwB, an integral protein involved in bacterial conjugation. J Biol Chem 279:10955–10961

Hormaeche I, Segura RL, Vecino AJ, Goni FM, de la Cruz F, Alkorta I (2006) The transmembrane domain provides nucleotide binding specificity to the bacterial conjugation protein TrwB. FEBS Lett 580:3075–3308

Hwang HH, Gelvin S (2004) Plant proteins that interact with VirB2, the *Agrobacterium tumefaciens* pilin protein, mediate plant transformation. Plant Cell 16:3148–3167

Huang FC, Fu BJ, Liu YT, Chang YR, Chi SF, Chien PR, Huang SC, Hwang HH (2018) *Arabidopsis* RETICULON-LIKE3 (RTNLB3) and RTNLB8 participate in *Agrobacterium*-mediated plant transformation. Int J Mol Sci 19(2). pii: E638. https://doi.org/10.3390/ijms19020638

Jakubowski SJ, Cascales E, Krishnamoorthy V, Christie PJ (2005) *Agrobacterium tumefaciens* VirB9, an outer-membrane-associated component of a type IV secretion system, regulates substrate selection and T-pilus biogenesis. J Bacteriol 187:3486–3495

Jakubowski SJ, Kerr JE, Garza I, Krishnamoorthy V, Bayliss R, Waksman G, Christie PJ (2009) *Agrobacterium* VirB10 domain requirements for type IV secretion and T pilus biogenesis. Mol Microbiol 71:779–794

Jakubowski SJ, Krishnamoorthy V, Cascales E, Christie PJ (2004) *Agrobacterium tumefaciens* VirB6 domains direct the ordered export of a DNA substrate through a type IV secretion system. J Mol Biol 341:961–977

Jakubowski SJ, Krishnamoorthy V, Christie PJ (2003) *Agrobacterium tumefaciens* VirB6 protein participates in formation of VirB7 and VirB9 complexes required for type IV secretion. J Bacteriol 185:2867–2878

Judd PK, Kumar RB, Das A (2005a) Spatial location and requirements for the assembly of the *Agrobacterium tumefaciens* type IV secretion apparatus. Proc Natl Acad Sci U S A 102:11498–11503

Judd PK, Kumar RB, Das A (2005b) The type IV secretion apparatus protein VirB6 of *Agrobacterium tumefaciens* localizes to a cell pole. Mol Microbiol 55:115–124

Kado CI (2000) The role of the T-pilus in horizontal gene transfer and tumorigenesis. Curr Opin Microbiol 3:643–648

Kato J, Lefebre M, Galan JE (2015) Structural features reminiscent of ATP-driven protein translocases are essential for the function of a Type III secretion-associated ATPase. J Bacteriol 197:3007–3014. https://doi.org/10.1128/JB.00434-15

Kerr JE, Christie PJ (2010) Evidence for VirB4-mediated dislocation of membrane-integrated VirB2 pilin during biogenesis of the *Agrobacterium* VirB/VirD4 type IV secretion system. J Bacteriol 192:4923–4934

Kuldau GA, De VG, Owen J, McCaffrey G, Zambryski P (1990) The *virB* operon of *Agrobacterium tumefaciens* pTiC58 encodes 11 open reading frames. Mol Gen Genetics 221:256–266

Kumar RB, Das A (2002) Polar location and functional domains of the *Agrobacterium tumefaciens* DNA transfer protein. VirD4. Mol Microbiol 43:1523–1532

Kunik T, Tzfira T, Kapulnik Y, Gafni Y, Dingwall C, Citovsky V (2001) Genetic transformation of HeLa cells by *Agrobacterium*. Proc Natl Acad Sci USA 98:1871–1876

Kwak MJ et al (2017) Architecture of the type IV coupling protein complex of *Legionella pneumophila*. Nat Microbiol 2:17114. https://doi.org/10.1038/nmicrobiol.2017.114

Lacroix B, Citovsky V (2011) Extracellular VirB5 enhances T-DNA transfer from *Agrobacterium* to the host plant. PLoS ONE 6:e25578. https://doi.org/10.1371/journal.pone.0025578

Lai EM, Kado CI (1998) Processed VirB2 is the major subunit of the promiscuous pilus of *Agrobacterium tumefaciens*. J Bacteriol 180:2711–2717

Lang S et al (2011) An activation domain of plasmid R1 TraI protein delineates stages of gene transfer initiation. Mol Microbiol 82:1071–1085. https://doi.org/10.1111/j.1365-2958.2011.07872.x

Larrea D, de Paz HD, Arechaga I, de la Cruz F, Llosa M (2013) Structural independence of conjugative coupling protein TrwB from its Type IV secretion machinery. Plasmid 70:146–153. https://doi.org/10.1016/j.plasmid.2013.03.006

Li G, Brown PJ, Tang JX, Xu J, Quardokus EM, Fuqua C, Brun YV (2012) Surface contact stimulates the just-in-time deployment of bacterial adhesins. Mol Microbiol 83:41–51. https://doi.org/10.1111/j.1365-2958.2011.07909.x

Llosa M, Zunzunegui S, de la Cruz F (2003) Conjugative coupling proteins interact with cognate and heterologous VirB10-like proteins while exhibiting specificity for cognate relaxosomes. Proc Natl Acad Sci U S A 100:10465–10470

Low HH et al (2014) Structure of a type IV secretion system. Nature 508:550–553. https://doi.org/10.1038/nature13081

Lu J, den Dulk-Ras A, Hooykaas PJJ, Glover JNM. Glover (2009) *Agrobacterium tumefaciens* VirC2 enhances T-DNA transfer and virulence through its C-terminal ribbon-helix-helix DNA-binding fold. Proc Natl Acad Sci USA 106:9643–0648

Matthysse AG (2014) Attachment of *Agrobacterium* to plant surfaces. Front Plant Sci 5:252. https://doi.org/10.3389/fpls.2014.00252

Matthysse AG, Marry M, Krall L, Kaye M, Ramey BE, Fuqua C, White AR (2005) The effect of cellulose overproduction on binding and biofilm formation on roots by *Agrobacterium tumefaciens*. Mol Plant Microbe Interact 18:1002–1010. https://doi.org/10.1094/MPMI-18-1002

Padavannil A et al (2014) Dimerization of VirD2 binding protein is essential for *Agrobacterium* induced tumor formation in plants. PLoS Pathog 10:e1003948. https://doi.org/10.1371/journal.ppat.1003948

Pena A, Arechaga I (2013) Molecular motors in bacterial secretion. J Mol Microbiol Biotechnol 23:357–369. https://doi.org/10.1159/000351360

Pena A et al (2012) The hexameric structure of a conjugative VirB4 protein ATPase provides new insights for a functional and phylogenetic relationship with DNA translocases. J Biol Chem 287:39925–39932

Piers KL, Heath JD, Liang X, Stephens KM, Nester EW (1996) *Agrobacterium tumefaciens*-mediated transformation of yeast. Proc Natl Acad Sci U S A 93:1613–1618

Rashkova S, Spudich GM, Christie PJ (1997) Characterization of membrane and protein interaction determinants of the *Agrobacterium tumefaciens* VirB11 ATPase. J Bacteriol 179:583–591

Redzej A et al (2017) Structure of a VirD4 coupling protein bound to a VirB type IV secretion machinery. EMBO J 36:3080–3095

Ripoll-Rozada J, Zunzunegui S, de la Cruz F, Arechaga I, Cabezon E (2013) Functional interactions of VirB11 traffic ATPases with VirB4 and VirD4 molecular motors in type IV secretion systems. J Bacteriol 195:4195–4201. https://doi.org/10.1128/JB.00437-13

Sagulenko E, Sagulenko V, Chen J, Christie PJ (2001) Role of *Agrobacterium* VirB11 ATPase in T-pilus assembly and substrate selection. J Bacteriol 183:5813–5825. https://doi.org/10.1128/JB.183.20.5813-5825.2001

Salmond GPC (1994) Secretion of extracellular virulence factors by plant pathogenic bacteria. Ann Rev Phytopathol 32:181–200

Samuels AL, Lanka E, Davies JE (2000) Conjugative junctions in RP4-mediated mating of *Escherichia coli*. J Bacteriol 182:2709–2715

Sardesai N, Lee LY, Chen Y, Yi H, Olbricht GR, Stirnberg A, Jeffries J, Xiong K, Doerge RW, Gelvin SB (2013) Cytokinins secreted by *Agrobacterium* promote transformation by repressing a plant myb transcription factor. Sci Signal 6:ra100 https://doi.org/10.1126/scisignal.2004518

Savvides SN (2007) Secretion superfamily ATPases swing big. Structure 15:255–257

Savvides SN et al (2003) VirB11 ATPases are dynamic hexameric assemblies: new insights into bacterial type IV secretion. EMBO J 22:1969–1980

Schmidt-Eisenlohr H, Domke N, Baron C (1999) TraC of IncN plasmid pKM101 associates with membranes and extracellular high-molecular-weight structures in *Escherichia coli*. J Bacteriol 181:5563–5571

Schrammeijer B, Dulk-Ras AdA, Vergunst AC, Jurado Jacome E, Hooykaas PJ (2003) Analysis of Vir protein translocation from *Agrobacterium tumefaciens* using *Saccharomyces cerevisiae* as a model: evidence for transport of a novel effector protein VirE3. Nucl Acids Res 31:860–868

Shirasu K, Kado CI (1993) The *virB* operon of the *Agrobacterium tumefaciens* virulence regulon has sequence similarities to B, C and D open reading frames downstream of the pertussis toxin-operon and to the DNA transfer-operons of broad-host-range conjugative plasmids. Nucleic Acids Res 21:353–354

Shirasu K, Morel P, Kado CI (1990) Characterization of the *virB* operon of an *Agrobacterium tumefaciens* Ti plasmid: nucleotide sequence and protein analysis. Mol Microbiol 4:1153–1163

Simone M, McCullen CA, Stahl LE, Binns AN (2001) The carboxy-terminus of VirE2 from *Agrobacterium tumefaciens* is required for its transport to host cells by the virB-encoded type IV transport system. Mol Microbiol 41:1283–1293

Spudich GM, Fernandez D, Zhou XR, Christie PJ (1996) Intermolecular disulfide bonds stabilize VirB7 homodimers and VirB7/VirB9 heterodimers during biogenesis of the *Agrobacterium tumefaciens* T-complex transport apparatus. Proc Natl Acad Sci U S A 93:7512–7517

Stachel SE, Nester EW (1986) The genetic and transcriptional organization of the *vir* region of the A6 Ti plasmid of *Agrobacterium tumefaciens*. EMBO J 5:1445–1454

Stachel SE, Timmerman B, Zambryski P (1987) Activation of *Agrobacterium tumefaciens vir* gene expression generates multiple single-stranded T-strand molecules from the pTiA6 T-region: requirement for 5' virD gene products. EMBO J 6:857–863

Stachel SE, Zambryski PC (1986) *Agrobacterium tumefaciens* and the susceptible plant cell: a novel adaptation of extracellular recognition and DNA conjugation. Cell 47:155–157

Toro N, Datta A, Carmi OA, Young C, Prusti RK, Nester EW (1989) The *Agrobacterium tumefaciens* virC1 gene product binds to overdrive, a T-DNA transfer enhancer. J Bacteriol 171:6845–6849

van Kregten M, Lindhout BI, Hooykaas PJ, van der Zaal BJ (2009) *Agrobacterium*-mediated T-DNA transfer and integration by minimal VirD2 consisting of the relaxase domain and a type IV secretion system translocation signal. Mol Plant-Microbe Inter: MPMI 22:1356–1365

Veluthambi K, Ream W, Gelvin SB (1988) Virulence genes, borders, and overdrive generate single-stranded T-DNA molecules from the A6 Ti plasmid of *Agrobacterium tumefaciens*. J Bacteriol 170:1523–1532

Vergunst AC, Schrammeijer B, den Dulk-Ras A, de Vlaam CM, Regensburg-Tuink TJ, Hooykaas PJ (2000) VirB/D4-dependent protein translocation from *Agrobacterium* into plant cells. Science 290:979–982

Vergunst AC, van Lier MC, den Dulk-Ras A, Grosse Stuve TA, Ouwehand A, Hooykaas PJ (2005) Positive charge is an important feature of the C-terminal transport signal of the VirB/D4-translocated proteins of *Agrobacterium*. Proc Natl Acad Sci USA 102:832–837

Vergunst AC, Van Lier MC, Den Dulk-Ras A, Hooykaas PJ (2003) Recognition of the *Agrobacterium tumefaciens* VirE2 translocation signal by the VirB/D4 transport system does not require VirE1. Plant Physiol 133:978–988

Wang Y, Haitjema CH, Fuqua C (2014) The Ctp type IVb pilus locus of *Agrobacterium tumefaciens* directs formation of the common pili and contributes to reversible surface attachment. J Bacteriol 196:2979–2988. https://doi.org/10.1128/JB.01670-14

Ward JE, Akiyoshi DE, Reiger D, Datta A, Gordon MP, Nester EW (1988) Characterization of the *virB* operon from an *Agrobacterium tumefaciens* Ti plasmid. J Biol Chem 263:5804–5814

Ward JE Jr, Dale EM, Binns AN (1991) Activity of the *Agrobacterium* T-DNA transfer machinery is affected by *virB* gene products. Proc Natl Acad Sci U S A 88:9350–9354

Whitaker N et al (2016) Chimeric coupling proteins mediate transfer of heterologous type IV effectors through the *Escherichia coli* pKM101-encoded conjugation machine. J Bacteriol 198:2701–2718. https://doi.org/10.1128/JB.00378-16

Whitaker N, Chen Y, Jakubowski SJ, Sarkar MK, Li F, Christie PJ (2015) The all-alpha domains of coupling proteins from the *Agrobacterium tumefaciens* VirB/VirD4 and E*nterococcus faecalis* pCF10-encoded type IV secretion systems confer specificity to binding of cognate DNA substrates. J Bacteriol 197:2335–2349. https://doi.org/10.1128/JB.00189-15

Winans SC, Burns DL, Christie PJ (1996) Adaptation of a conjugal transfer system for the export of pathogenic macromolecules. Trends Microbiol 4:64–68

Xu J, Kim J, Koestler BJ, Choi JH, Waters CM, Fuqua C (2013) Genetic analysis of *Agrobacterium tumefaciens* unipolar polysaccharide production reveals complex integrated control of the motile-to-sessile switch. Mol Microbiol 89:929–948. https://doi.org/10.1111/mmi.12321

Yang Q, Li X, Tu H, Pan SQ (2017) *Agrobacterium*-delivered virulence protein VirE2 is trafficked inside host cells via a myosin XI-K-powered ER/actin network. Proc Natl Acad Sci USA 114:2982–2987

Yanofsky MF, Porter SG, Young C, Albright LM, Gordon MP, Nester EW (1986) The *virD* operon of *Agrobacterium tumefaciens* encodes a site-specific endonuclease. Cell 47:471–477

Yeo HJ, Savvides SN, Herr AB, Lanka E, Waksman G (2000) Crystal structure of the hexameric traffic ATPase of the *Helicobacter pylori* type IV secretion system. Mol Cell 6:1461–1472

Young C, Nester EW (1988) Association of the VirD2 protein with the 5′ end of T strands in *Agrobacterium tumefaciens*. J Bacteriol 170:3367–3374

Yuan Q et al (2005) Identification of the VirB4-VirB8-VirB5-VirB2 pilus assembly sequence of type IV secretion systems. J Biol Chem 280:26349–26359

Zhao Z, Sagulenko E, Ding Z, Christie PJ (2001) Activities of *virE1* and the VirE1 secretion chaperone in export of the multifunctional VirE2 effector via an *Agrobacterium* type IV secretion pathway. J Bacteriol 183:3855–3865

Zhou XR, Christie PJ (1997) Suppression of mutant phenotypes of the *Agrobacterium tumefaciens* VirB11 ATPase by overproduction of VirB proteins. J Bacteriol 179:5835–5842

Zupan J, Hackworth CA, Aguilar J, Ward D, Zambryski P (2007) VirB1* promotes T-pilus formation in the vir-Type IV secretion system of *Agrobacterium tumefaciens*. J Bacteriol 189:6551–6563

Real-Time Trafficking of *Agrobacterium* Virulence Protein VirE2 Inside Host Cells

Haitao Tu, Xiaoyang Li, Qinghua Yang, Ling Peng and Shen Q. Pan

Contents

1 Introduction 262
2 Visualization of Delivered VirE2 Protein 263
2.1 A Split-GFP Approach to Visualize Delivered VirE2 263
2.2 Filamentous Structures of Delivered VirE2-GFP 265
2.3 VirE2-GFP11 Functions Similar to VirE2 266
2.4 VirE2-GFP11 Movement Inside Plant Cells 266
2.5 VirE2-GFP11 Does not Move Inside Yeast Cells 267
3 Internalization of Delivered VirE2 Protein into Host Cells 268
3.1 Delivery Site 268
3.2 Mode of VirE2 Internalization 269
3.3 Clathrin Adaptor AP2-Mediated Endocytosis 269
3.4 VirE2 Internalization Signals 271
3.5 VirE2 Internalization Signals Interact with Plant AP2M 272
4 Trafficking of Delivered VirE2 Protein Inside Host Cells 273
4.1 VirE2 Moves Along the ER 273
4.2 VirE2 Moves on an ER/F-Actin Network 274
4.3 VirE2 Movement Is Powered by Myosin 275
4.4 Trafficking Mode of Delivered VirE2 277
5 A Natural Model to Study Exogenous Protein Delivery 277
6 A Test System to Examine the Effect of a Molecule on Cells 279
7 Conclusions 281
References 282

Haitao Tu, Xiaoyang Li, Qinghua Yang, These authors contributed equally.

H. Tu
School of Stomatology and Medicine, Foshan Institute of Molecular Bio-Engineering, Foshan University, 528000 Foshan, China

X. Li · Q. Yang · L. Peng · S. Q. Pan (✉)
Department of Biological Sciences, National University of Singapore, 117543 Singapore, Singapore
e-mail: dbspansq@nus.edu.sg

Current Topics in Microbiology and Immunology (2018) 418:261–286
DOI 10.1007/82_2018_131

Published Online: 5 September 2018

Abstract *A. tumefaciens* delivers T-DNA and virulence proteins, including VirE2, into host plant cells, where T-DNA is proposed to be protected by VirE2 molecules as a nucleoprotein complex (T-complex) and trafficked into the nucleus. VirE2 is a protein that can self-aggregate and contains targeting sequences so that it can efficiently move from outside of a cell to the nucleus. We adopted a split-GFP approach and generated a VirE2-GFP fusion which retains the self-aggregating property and the targeting sequences. The fusion protein is fully functional and can move inside cells in real time in a readily detectable format: fluorescent and unique filamentous aggregates. Upon delivery mediated by the bacterial type IV secretion system (T4SS), VirE2-GFP is internalized into the plant cells via clathrin adaptor complex AP2-mediated endocytosis. Subsequently, VirE2-GFP binds to membrane structures such as the endoplasmic reticulum (ER) and is trafficked within the cell. This enables us to observe the highly dynamic activities of the cell. If a compound, a gene, or a condition affects the cell, the cellular dynamics shown by the VirE2-GFP will be affected and thus readily observed by confocal microscopy. This represents an excellent model to study the delivery and trafficking of an exogenously produced and delivered protein inside a cell in a natural setting in real time. The model may be used to explore the theoretical and applied aspects of natural protein delivery and targeting.

1 Introduction

Agrobacterium tumefaciens can deliver T-DNA into different eukaryotes, including plants (Chilton et al. 1977; Marton et al. 1979; Broothaerts et al. 2005), yeast (Bundock et al. 1995; Piers et al. 1996), algae (Kathiresan et al. 2009), and fungal cells (de Groot et al. 1998). During the transfer process, a single-stranded DNA (T-DNA) molecule is generated inside the bacterium by the VirD1–VirD2 endonuclease (Wang et al. 1984; Yanofsky et al. 1986; Scheiffele et al. 1995). Subsequently, VirD2 remains covalently attached to the 5′ end of the T-DNA (T-strand) (Yanofsky et al. 1986). As a nucleoprotein complex, the T-strand is then transferred into recipient cells via a VirB/VirD4 type IV secretion system (T4SS) (Cascales and Christie 2004) in a manner mechanistically similar to that of a conjugation process (Beijersbergen et al. 1992).

The same T4SS is also known to deliver protein substrates, including VirE2, VirD2, VirE3, VirD5, and VirF (Vergunst et al. 2000; Schrammeijer et al. 2003; Vergunst et al. 2005). These proteins are virulence effectors that interact with host factors in the recipient cells to facilitate transformation. As an abundant Vir protein, VirE2 may coat and protect the T-strand (Citovsky et al. 1988, 1992; Yusibov et al. 1994; Rossi et al. 1996). VirE2 may also mediate uptake of the T-DNA complex by forming a pore in the plant plasma membrane (Dumas et al. 2001). VirE2 could interact with plant VIP1, which is localized in the nucleus upon phosphorylation, and several import in α isoforms in the plant cells, suggesting that VirE2 might

contribute to nuclear targeting of T-DNA (Citovsky et al. 1992; Djamei et al. 2007; Bhattacharjee et al. 2008). In addition, VirE2 might facilitate chromatin targeting of the T-complex through association with host VIP2 (Anand et al. 2007). It appears that VirE2 participates in various steps of T-complex trafficking inside recipient cells, from the entry point to the final destination.

To elucidate how the nucleoprotein complex is trafficked inside recipient cells, it is important to directly visualize the DNA and protein molecules inside the recipient cells after translocation and in real time. A split-GFP system (Cabantous et al. 2005; Pedelacq et al. 2006) was adopted to successfully visualize VirE2, its aggregation forms, and its movement in the recipient cells during a natural transformation process (Li et al. 2014). The split-GFP approach enabled real-time visualization of VirE2 trafficking inside host cells. This system provides a new window to explore the molecular events of *Agrobacterium*-mediated transformation and natural protein delivery processes in real time.

2 Visualization of Delivered VirE2 Protein

2.1 A Split-GFP Approach to Visualize Delivered VirE2

To visualize *Agrobacterium*-delivered VirE2 protein in real time, a split-GFP approach (Cabantous et al. 2005; Pedelacq et al. 2006; Van Engelenburg and Palmer 2010) was adopted so that the delivered VirE2 protein could be detected in a functional form (Li et al. 2014). As shown in Fig. 1A, the split-GFP system is composed of two non-fluorescent GFP fragments: β-strands 1-10 of GFP (GFP1-10) containing 215 amino acid residues (positions 1-215) and β-strand 11 of GFP (GFP11) containing 16 amino acid residues (positions 216-231). GFP1-10 and GFP11 could bind each other spontaneously and restore the fluorescence of GFP_{comp} (Cabantous et al. 2005).

As indicated in Fig. 1, GFP11 was fused onto VirE2 at a permissive site (Zhou and Christie 1999) to create the VirE2-GFP11 fusion. The fusion was expressed inside *A. tumefaciens,* and GFP1-10 was expressed in the recipient cells. When VirE2-GFP11 was delivered into the recipient cells, GFP1-10 would be complemented by VirE2-GFP11 and the resulting VirE2-GFP_{comp} signals were detected (Li et al. 2014).

When GFP11 was fused onto the C-terminus of VirE2, no VirE2-GFP_{comp} signals were detected inside plant cells. This demonstrated that the position of GFP11 tagging was critical for the split-GFP experiments (Li et al. 2014).

Previously, different VirE2 tagging constructs were made to conduct various studies (Bhattacharjee et al. 2008; Aguilar et al. 2010). However, none of the tagged VirE2 proteins could be successfully translocated into recipient cells, presumably because the T4SS channel could not accommodate the enlarged size or a hindering structure of these fusion proteins (Bhattacharjee et al. 2008; Aguilar et al. 2010).

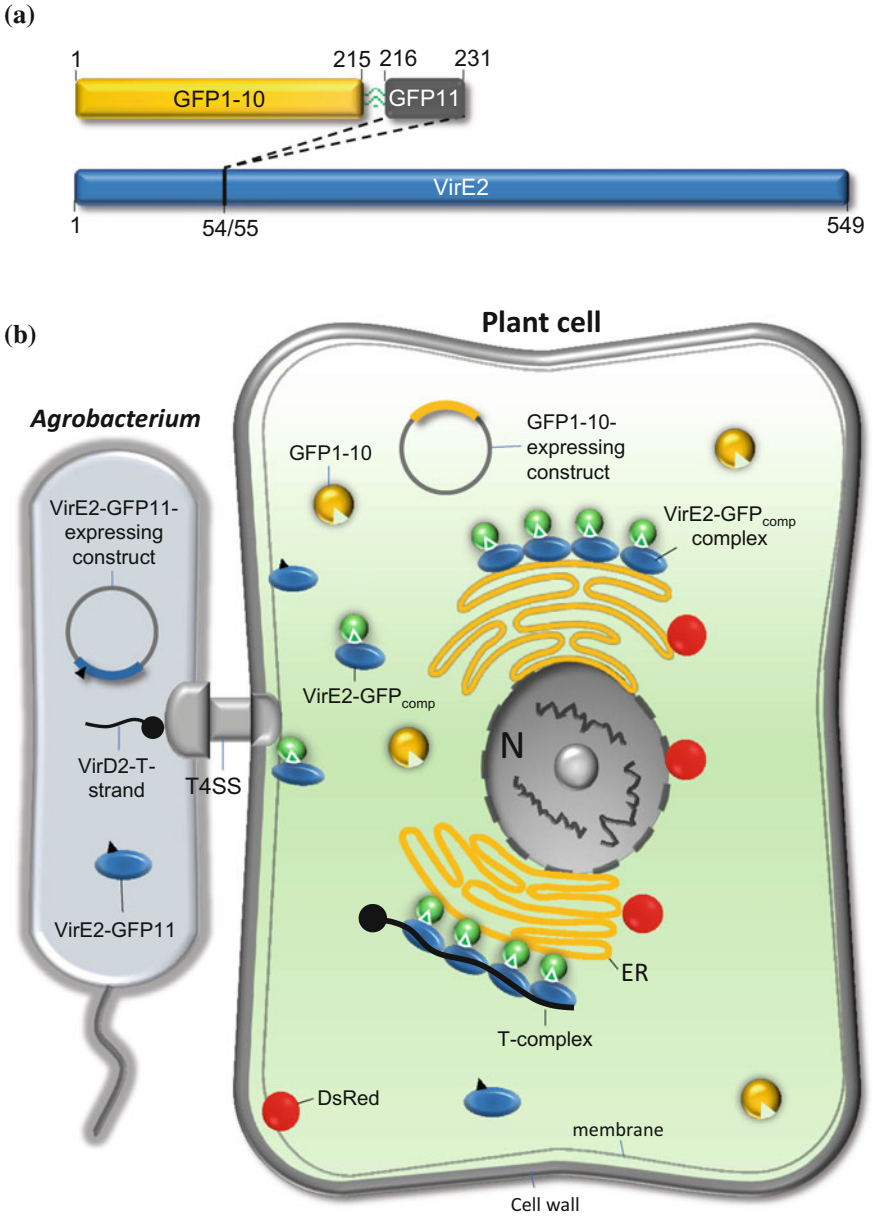


Fig. 1 A split-GFP method to visualize *Agrobacterium*-delivered VirE2 protein inside plant cells. **a** The split-GFP system is composed of two non-fluorescent GFP fragments: β-strands 1-10 of GFP (GFP1-10) containing 215 amino acid residues (positions 1-215) and β-strand 11 of GFP (GFP11) containing 16 amino acid residues (positions 216-231). GFP11 is inserted into VirE2 at amino acid position 54. **b** VirE2-GFP11 fusion is expressed inside *A. tumefaciens*; and GFP1-10 is expressed in the plant cells. When VirE2-GFP11 is delivered into the recipient cells, GFP1-10 would be complemented by VirE2-GFP11 and the resulting VirE2-GFP_{comp} signals are detected. DsRed is expressed inside plant cells to indicate membrane-related structures. Both the T-complex and VirE2-GFP_{comp} complex may be present in the same cell

Nevertheless, the subcellular localization of VirE2 has been investigated with the tagged VirE2 protein. Both cytoplasmic and nuclear localization of VirE2 have been reported (Citovsky et al. 1992; Rhee et al. 2000; Tzfira and Citovsky 2001; Bhattacharjee et al. 2008; Lee et al. 2008; Shi et al. 2014; Lapham et al. 2018). In those studies, VirE2 was artificially introduced into cells either by direct uptake or transgenic expression, which may differ from a natural *Agrobacterium*-mediated transfer process. In addition, VirE2 tagged with a full-length GFP either at the C-terminus or N-terminus may affect its translocation activity (Simone et al. 2001; Atmakuri et al. 2003; Schrammeijer et al. 2003; Bhattacharjee et al. 2008).

The split-GFP approach was also previously adopted to visualize *Agrobacterium*-delivered VirE2, although VirE2 was tagged with GFP11 at the N-terminus (Sakalis et al. 2014). However, N-terminal tagging at VirE2 might affect the role of VirE2 in transformation. Consequently, such a tagged VirE2 might not be suitable to represent the natural VirE2 trafficking inside host cells.

The studies have clearly demonstrated that GFP11 is an appropriate tag that can be fused onto a permissive site of VirE2 so that the VirE2-GFP11 is functional like VirE2 (Li et al. 2014). VirE2-GFP11 movement could represent VirE2 trafficking. Because GFP1-10 was expressed inside recipient cells, the GFP_{comp} signal was only detectable inside recipient cells so that VirE2-GFP11 trafficking signal could be readily monitored.

The monitoring process is straightforward: The bacterial cells expressing VirE2-GFP11 are infiltrated into transgenic *N. benthamiana* (Nb308A) leaves expressing both GFP1-10 and DsRed. The epidermal cells are examined at 2d post-agroinfiltration by confocal microscopy for VirE2-GFP_{comp} and DsRed signals. Images are taken in multiple focal planes (Z-stacks) and then assembled into a movie format. The green fluorescence signals represent VirE2-$GFP_{comp,}$ and the red fluorescence shows the subcellular structures associate with DsRed.

2.2 *Filamentous Structures of Delivered VirE2-GFP*

VirE2 is a non-specific ssDNA-binding protein (Citovsky et al. 1988) that can coat the entire length of T-strands in vitro with one VirE2 molecule covering 19 bases of T-DNA (Citovsky et al. 1997). In the presence or absence of T-DNA, numerous VirE2 molecules can form telephone cord-like multimers in vitro (Citovsky et al. 1997; Frenkiel-Krispin et al. 2007; Dym et al. 2008). This self-association capacity could thus amplify the VirE2-GFP_{comp} signals to facilitate direct visualization of VirE2 inside the recipient cells (Li et al. 2014), as VirE2 is an abundant Vir protein (Engstrom et al. 1987).

When the bacterial cells did not contain T-DNA, the filamentous structures of VirE2-GFP_{comp} signals were also observed, demonstrating that the aggregated form of VirE2-GFP complex could be free of any T-DNA (Li et al. 2014; Yang et al. 2017) (Fig. 1). As VirE2 transfer progressed, more filamentous structures were found and the filaments became even longer. This result suggests that VirE2

aggregation continued when more VirE2 was delivered. Under the same conditions, the negative controls did not generate any GFP fluorescence. These controls included *A. tumefaciens* strains which did not encode VirE2-GFP11 or VirD4. Therefore, naturally transferred VirE2-GFP11 protein and its aggregated form were successfully visualized inside live recipient cells.

2.3 VirE2-GFP11 Functions Similar to VirE2

To ensure that the VirE2-GFP11 movement represents VirE2 trafficking, the VirE2-GFP11 fusion should not disrupt VirE2 function. To achieve this, GFP11 was inserted at Pro54 of VirE2 (Accession No.: AAZ50538) (Li et al. 2014), a site that was shown to be tolerant for a 31-residue oligopeptide insertion (Zhou and Christie 1999). The *virE2* gene from EHA105 was used to generate the VirE2-GFP fusion, which was then utilized to replace the *virE2* gene of EHA105 that does not contain any T-DNA (Hood et al. 1993). Subsequently, VirE2 aggregation and trafficking were studied in the absence of T-DNA.

To test the virulence function of VirE2-GFP, the fusion construct was then used to replace the *virE2* gene of a tumorigenic strain A348 (Li et al. 2014). The resulting A348-*105virE2::GFP11* was inoculated onto roots of transgenic *Arabidopsis thaliana* (H16) expressing GFP1-10. A348-*105virE2::GFP11* caused tumors in a manner similar to the corresponding A348-*105virE2*, which is an A348 derivative with its *virE2* replaced by EHA105 *virE2*. As expected, the *virE2* deletion mutant A348Δ*virE2* was avirulent. The virulence function of VirE2-GFP11 was similar to that of wild-type VirE2, as the frequency and size of tumors caused by A348-*105virE2::GFP11* were similar to those of A348-*virE2*. The results suggest that VirE2-GFP11 was fully functional just like VirE2, even in the presence of GFP1-10 in the transgenic plants. Thus, the VirE2-GFP11 fusion was suitable for visualization of VirE2 and its trafficking upon delivery into recipient cells (Li et al. 2014).

2.4 VirE2-GFP11 Movement Inside Plant Cells

To visualize VirE2 inside a natural-host plant, *A. tumefaciens* EHA105*virE2::GFP11* cells were infiltrated into transgenic *Nicotiana benthamiana* (Nb308A) leaves expressing both GFP1-10 and DsRed. When VirE2-GFP11 was translocated into the plant cells, GFP1-10 bound to VirE2-GFP11 and the resulting VirE2-GFP_{comp} signals appeared as green fluorescence by confocal microscopy (Li et al. 2014). Two days after infiltration, VirE2-GFP_{comp} signals were found in the plant cells in both the cytoplasm and the nucleus. Most of the signals appeared as spots, but some appeared as filamentous structures.

Time course studies indicated that no VirE2-GFP$_{comp}$ signals were detected 16 h after agroinfiltration (Li et al. 2014). At 32 h after agroinfiltration, VirE2-GFP$_{comp}$ signals were detected as spots. At 48 h after agroinfiltration, both VirE2-GFP$_{comp}$ spots and filamentous structures were detected. Pre-induction of the bacteria by acetosyringone (AS) before agroinfiltration did not significantly speed up the appearance or increase the intensity of VirE2-GFP$_{comp}$ signals. This result suggests that *vir* gene induction was not a limiting factor, but it took time for VirE2 to be delivered to a detectable level in plant cells.

VirE2 was suggested to play a role in the T-complex trafficking by hijacking the plant MAPK-targeted VIP1 defense signaling pathway (Tzfira et al. 2002; Djamei et al. 2007). We thus monitored VirE2 movement in a time-lapse series and successfully captured the VirE2-GFP$_{comp}$ trafficking process (Li et al. 2014). The speed of VirE2-GFP movement varied; the majority ranged from 1.3 to 3.1 μm/sec. The movement was nearly linear and directional, suggesting that VirE2 movement was assisted by an active host process.

Both the shorter (dot) and longer (filament) forms of VirE2 aggregation moved inside the plant cells; some movements were directed toward the nucleus (Li et al. 2014). The VirE2-GFP filaments were found to be attached to the nucleus. The filamentous VirE2-GFP complex in the cytoplasm was also found to be linked to the VirE2-GFP complex inside the nucleus. These results suggest that the filamentous VirE2 complex was targeted for nuclear import. When the VirE2 nuclear localization signal 1 (NLS1) (Citovsky et al. 1992) was mutated, the VirE2-GFP complex was exclusively localized in the cytoplasm and nuclear import was not observed either for the VirE2-GFP$_{comp}$ spots or for the filamentous structures (Li et al. 2014). These experiments demonstrated that nuclear import of VirE2-GFP complex was dependent upon the nuclear localization signal.

2.5 *VirE2-GFP11 Does not Move Inside Yeast Cells*

This imaging approach was also applied to the non-natural-host species *Saccharomyces cerevisiae* (Li et al. 2014). A yeast strain encoding GFP1-10 was co-cultivated with *A. tumefaciens* EHA105*virE2::GFP11* that was induced by AS. The VirE2-GFP$_{comp}$ signals could appear as early as 2 h after co-cultivation. However, the signals did not move inside the yeast cells, and they were not localized in the nucleus. More often they were at the periphery of the yeast cells. This indicates that VirE2 is not actively trafficked into the yeast nucleus, presumably because yeast is a non-natural-host recipient and thus lacks the facilitator(s) for VirE2 trafficking. This is consistent with the previous observation that VirE2 was localized in the cytoplasm rather than the nucleus of yeast cells (Rhee et al. 2000).

We found that the transient transformation efficiency of the natural-host plant is 250–500-fold higher than is transformation of the non-natural-host yeast cells, whereas VirE2 delivery was 127–255-fold more efficient than the transient transformation for a non-natural-host recipient (Li et al. 2014). One limiting factor for

the non-natural-host yeast transformation was presumably T-complex trafficking, since the non-natural-host yeast cells could not facilitate active trafficking of the bacterial virulence factors, including VirE2 and perhaps the T-strand. In addition, the filamentous VirE2-GFP structures were not observed inside the yeast cells. It is not clear whether this is due to the limitation of yeast cellular space or the amount of VirE2 delivered into yeast cells.

3 Internalization of Delivered VirE2 Protein into Host Cells

3.1 Delivery Site

To study how VirE2 is internalized into host cells, VirE2 delivery into tobacco cells was observed under a confocal microscope (Li and Pan 2017). The T-DNA-free strain EHA105 was used to avoid any potential complication due to T-DNA trafficking. *A. tumefaciens* EHA105*virE2::GFP11* producing VirE2-GFP11 was infiltrated into transgenic *N. benthamiana* (Nb308A) leaves expressing both GFP1-10 and DsRed. VirE2 delivery into tobacco cells was examined at different time points. A small amount of VirE2 appeared at tobacco cell borders at 32 h after agroinfiltration (Li and Pan 2017). With increasing time, more VirE2 was observed at the cell borders and the VirE2 signals became filamentous. We found that VirE2 first appeared at tobacco cell borders and then moved into the nucleus (Li and Pan 2017).

Subsequently, we determined the spatial positioning of *A. tumefaciens* cells inside plant tissues (Li and Pan 2017). Bacterial cells were constructed to express GFP under the control of the *virB* promoter and thus they became fluorescently labeled naturally during agroinfiltration. After GFP-labeled *A. tumefaciens* cells EHA105 (pAT-GFP) were infiltrated into *N. benthamiana* leaves, most of the bacterial cells were observed to align in the intercellular space of agroinfiltrated tobacco cells (Li and Pan 2017). We observed that the bacterial cells tightly lined up in the intercellular spaces separately as single cells. These results suggest that the limited intercellular spaces of *N. benthamiana* epidermal cells can only accommodate single bacterial cells and the space limitation may allow only the lateral side of the bacterium to closely contact with the host cell.

At 48 h after agroinfiltration, VirE2 accumulated at the cytoplasmic sides of tobacco cells that are in a close contact with *A. tumefaciens* cells (Li and Pan 2017). Interestingly, VirE2 was delivered into plant cells from both sides of the bacterial cells. This observation suggests that a single bacterium could deliver VirE2 into two adjacent host cells simultaneously.

To determine the subcellular location of *Agrobacterium*-delivered VirE2 inside host cells, a specific plant plasma membrane tracker (Nelson et al. 2007) was expressed transiently inside plant cells by T-DNA delivered by the same bacterial cells delivering VirE2-GFP11. *Agrobacterium*-delivered VirE2 appeared to co-localize with the transiently expressed plasma membrane tracker (Li and Pan 2017).

These results suggest that VirE2 is associated with the plant cytoplasmic membrane at the delivery site.

3.2 Mode of VirE2 Internalization

To investigate how membrane-bound VirE2 moved into the cytoplasm, we used a fluorescent styryl dye FM4-64 (Geldner et al. 2003) to label the plant membranes and monitor membrane dynamics (Li and Pan 2017). This lipophilic dye can label membranes where it is applied, but it cannot penetrate the membranes by itself. This property allowed us to monitor the trafficking process of VirE2-bound membranes. *A. tumefaciens* EHA105*virE2::GFP11* cells were infiltrated into *N. benthamiana* leaves to start VirE2 delivery; 48 h later the FM4-64 dye was infiltrated into the same areas. VirE2 co-localized with FM4-64-labeled plasma membranes in a manner similar to that using the plasma membrane tracker. Interestingly, VirE2 co-localized with FM4-64-labeled endomembrane compartments that ranged from 0.8 μm to 4.5 μm in diameter, with an average of 2.2 μm (Li and Pan 2017).

The co-localization of VirE2 with FM4-64-labeled endomembrane compartments continued as FM4-64-labeled vesicles moved inside the cytoplasm (Li and Pan 2017). The speed of movement ranged from 0.4 to 2.1 μm/sec, which is consistent with endosome dynamics as reported in previous studies (Maizel et al. 2011). The data consistently suggest that VirE2 delivered onto host plasma membranes may utilize host endocytosis for cellular internalization and cytoplasmic movement.

3.3 Clathrin Adaptor AP2-Mediated Endocytosis

We examined whether the host endocytosis process was required for internalization of VirE2 protein (Li and Pan 2017). The plant endocytosis process is mediated by clathrin triskelions (McMahon and Boucrot 2011). Overexpression of a C-terminal part of clathrin heavy chain (Hub) that could bind to and deplete clathrin light chains would lead to strong dominant-negative effects on clathrin-mediated endocytosis (CME) (Liu et al. 1995; Kitakura et al. 2011; Dhonukshe et al. 2007).

The effect of Hub overexpression in *N. benthamiana* leaves was then tested and the FM4-64 dye was used to monitor the general endocytosis process. Transient expression of Hub under a CaMV 35S promoter dramatically decreased the internalization of FM4-64 dye (Li and Pan 2017). This result suggests that a dominant-negative strategy using Hub could indeed affect the endocytosis process in *N. benthamiana* epidermal cells. In addition, Hub overexpression increased VirE2 accumulation at cell borders. VirE2 stayed much longer at the cell borders in the tobacco cells overexpressing Hub as compared to the control, indicating that functional clathrin and active CME process were required for VirE2 departure from the plant cellular membrane.

To confirm that host endocytosis is important for VirE2 trafficking, the chemical inhibitor endosidin1 (ES1) was used to interfere with the endocytosis process (Li and Pan 2017), as ES1 affects the endocytosis pathway and causes aggregation of early endosomes in *Arabidopsis thaliana* (Robert et al. 2008). SYP61-mCherry was transiently expressed to label the highly dynamic round-shaped early endosomes (Robert et al. 2008; Foresti and Denecke 2008) in *N. benthamiana* epidermal cells. ES1 treatment caused abnormal VirE2 trafficking within the host cytoplasm; VirE2 accumulated inside the ES1-induced endosome aggregates (Li and Pan 2017). Estimation of co-localization through Pearson's correlation coefficient suggested VirE2 co-localization with the SYP61-mCherry marker, indicating VirE2 co-localization with the early endosomes. These results indicate that ES1 can interfere with host endocytosis and thus restrict VirE2 movement.

Early endosomes mainly function as the sorting hub for endocytic trafficking processes in plants; cargoes internalized from the plasma membrane are usually transported to late endosomes and vacuoles for degradation (Contento and Bassham 2012). To test whether VirE2 is trafficked to late endosomes, ARA6-DsRed (Ueda et al. 2004; Ebine et al. 2011) was transiently expressed to label the late endosomes in *N. benthamiana* epidermal cells (Li and Pan 2017. We did not observe any obvious association of VirE2 with the ARA6-DsRed-labeled late endosome structures (Li and Pan 2017), suggesting that VirE2 may escape from early endosomes and move to other locations to avoid degradation in the vacuoles.

Because *Agrobacterium*-delivered VirE2 was targeted to plant nuclei in a nuclear localization signal (NLS)-dependent manner (Li et al. 2014), we tested the effect of ES1 on nuclear targeting of VirE2 and found that ES1 treatment dramatically decreased the nuclear accumulation of VirE2 inside tobacco cells, while VirE2 accumulated at the cell borders or inside cytoplasm (Li and Pan 2017). This result indicates that ES1 affects VirE2 trafficking rather than delivery or oligomerization of VirE2. These findings suggest that host endocytosis plays an important role in cytoplasmic trafficking and subsequent nuclear targeting of VirE2 inside plant cells.

The importance of endocytosis for transformation was confirmed by studying the effects of chemical inhibitors (Li and Pan 2017). Tumorigenesis assays were conducted using *A. thaliana* roots treated with either ES1 or Tyrphostin A23, which is also a CME inhibitor for *A. thaliana* (Banbury et al. 2003). We found that treatment with ES1 or Tyrphostin A23 significantly attenuated tumorigenesis (Li and Pan 2017). These results suggest that interference with host endocytosis can attenuate the stable transformation of plant cells, presumably because blocked endocytosis affects VirE2 movement.

Endocytosis is a well-conserved process in eukaryotic cells, which is responsible for uptake of a variety of molecules from the outside environment. It participates in a great number of cellular functions such as nutrient uptake, signaling transduction, antigen detection, and cell differentiation (Wu et al. 2014). Although endocytosis is involved in viral entry into host cells (Mercer et al. 2010), it is not clear whether endocytosis is needed for cellular entry of a virulence protein transferred by a bacterial secretion apparatus. These studies demonstrated that *Agrobacterium*

hijacks the host CME pathway for VirE2 internalization into host cells. CME is the major pathway for the endocytosis process (McMahon and Boucrot 2011), which also plays important roles in antigen perception and initiation of plant defense responses upon pathogen infection (Bar and Avni 2014). It would be of interest to determine if *Agrobacterium* might hijack CME-related processes in plant defense to facilitate its infection.

As a highly selective process, recognition and binding of cargo proteins in CME is initiated by clathrin adaptors on the cytoplasmic side of a cell. It is not clear how VirE2 molecules reach the cytoplasmic side of the host cells so that VirE2 can interact with clathrin adaptors. One possibility is that *A. tumefaciens* T4SS can deliver VirE2 directly to the cytoplasmic side of host cells. Alternatively, VirE2 is translocated via additional bacterial or host factor(s) to the cytoplasmic side of host cells. This remains to be addressed.

3.4 VirE2 Internalization Signals

It is of interest to investigate how VirE2 is selected as a cargo for internalization. In general, selection of plasma membrane-associated cargo proteins for internalization depends upon the recognition of endocytic signals at the cytosolic side of cargo proteins by a variety of host adaptors (Bonifacino and Traub 2003; Traub 2009). Upon delivery into host plant cells through a T4SS, VirE2 might interact with a host adaptor protein at the plasma membrane. Sequence analysis indicated that the VirE2 (accession no. AAZ50538) contains five putative endocytic sorting motifs (Li and Pan 2017).

The potential critical leucine or tyrosine residue for each of the dileucine-based or tyrosine-based motifs, respectively, was mutated to alanine; a double mutant was constructed for the two tyrosine-based motifs in the C-terminus (Li and Pan 2017). Neither single mutation nor double mutation of the dual C-terminal tyrosine-based motifs affected VirE2 delivery to the host cellular membrane. However, the double mutation caused a significantly higher level of VirE2 accumulation at the membrane sites. Mutation of other putative endocytic motifs of VirE2 did not affect VirE2 delivery or internalization (Li and Pan 2017). These results suggest that the putative dual C-terminal tyrosine-based motifs are important for VirE2 trafficking.

By conducting assays for transient transformation, we found that both single and double mutations at the dual C-terminal endocytic signals significantly decreased transient transformation efficiency, although the effect of a double mutation (Y488A/Y494A) was more dramatic that was the single mutation Y494A, which affected the function more than Y488A (Li and Pan 2017). These results suggest that the dual C-terminal endocytic signals are required for VirE2 function and that the last endocytic signal at the VirE2 terminus is more important for this function.

Sequence alignment analysis indicated that the dual C-terminal tyrosine-based endocytic motifs are conserved among VirE2 proteins from different Ti plasmids, suggesting their conserved roles in different *Agrobacterium* strains (Li and

Pan 2017). Moreover, mutation of these conserved motifs of VirE2 from the virulent strain A348 also attenuated tumor formation on *Arabidopsis* root fragments (Li and Pan 2017). These results demonstrate that the dual tyrosine-based endocytic signals located at the VirE2 C-terminus are important for VirE2 function for both transient and stable transformation.

3.5 VirE2 Internalization Signals Interact with Plant AP2M

We hypothesize that the dual C-terminal tyrosine-based endocytic motifs of VirE2 might be recognized by a clathrin-associated sorting protein (Li and Pan 2017). The clathrin-mediated endocytosis process is facilitated by a group of host adaptors known as "clathrin-associated sorting proteins," which are responsible for endocytic signal recognition and cargo binding (Bonifacino and Traub 2003; Traub 2009). Among them, the adaptor protein 2 (AP-2) complex recognizes the tyrosine-based endocytic signal and binds to it through the C-terminal domain of the μ-subunit (AP2M) (Jackson et al. 2010).

When the C-terminal tail was fused to GST (GST-VirE2C), VirE2 interacted with the cargo-binding domain of AP2M that was fused to the maltose binding protein (MBP; MBP-AP2MC) (Li and Pan 2017). However, a double mutation at the dual tyrosine-based endocytic signals eliminated this interaction. These results suggest that AP2M recognizes and binds to the VirE2 C-terminal tail through the dual tyrosine-based sorting motifs.

The importance of the host AP-2 complex in transformation was confirmed by testing two insertional mutants of *A. thaliana* AP2M for tumorigenesis (Li and Pan 2017). These two mutants have been previously shown to display defects in endocytosis (Kim et al. 2013). We found that these mutants of AP2M displayed significantly attenuated tumor formation, as compared to the wild-type control (Li and Pan 2017). These results demonstrate that the host AP-2 complex is important for *Agrobacterium*-mediated transformation of plant cells.

Membrane-associated protein cargos are internalized to form clathrin-coated vesicles (CCVs) and are then transported to the other parts of the cell (Bonifacino and Traub 2003; Traub 2009). The AP-2 adaptor complex is composed of two large subunits (α and β2), one medium-sized subunit (μ2), and one small subunit (σ2); it specifically recognizes and binds to the tyrosine-based (YXXØ) and dileucine-based ([DE]XXXL[LI]) sorting motifs on cargo proteins (Traub and Bonifacino 2013). We demonstrated that two closely located tyrosine-based motifs on the VirE2 C-terminus were responsible for VirE2 interaction with the cargo-binding domain of μ subunit of the AP-2 adaptor complex (Li and Pan 2017). Mutation of these signal sequences decreased VirE2 internalization and impaired transformation, suggesting that the dual tyrosine-based motifs are important for VirE2 trafficking and function during AMT. The second motif appeared to play a more important role in the transformation process. These two motifs might have different binding affinity and accessibility to the AP-2 complex.

Previous studies showed that the AP-2 complex is responsible for cargo transportation between the plasma membrane and early endosomes in plant cells (McMahon and Boucrot 2011; Chen et al. 2011). Our data suggest that interaction of VirE2 with the AP-2 complex can facilitate VirE2 internalization from the host plasma membrane (Li and Pan 2017). However, the root transformation assays demonstrated that mutations at the critical dual motifs decreased the efficiency of transformation by only 30-fold (Li and Pan 2017), whereas VirE2-deletion mutants virtually did not generate any transformants (Li et al. 2014). These results suggest that VirE2 may also be trafficked via an alternative pathway(s) for its role in the transformation process.

We observed that the speed of VirE2 movement together with endomembrane compartments ranged from 0.4 to 2.1 μm/sec (Li and Pan 2017). After VirE2 departs from the host membranes via endomembrane compartments, VirE2 would be trafficked to subsequent locations. Early endosomes serve as the main sorting hub for both secretory and endocytic trafficking cargoes. Different pathogens and viruses have been reported to use host endocytic pathway to facilitate infection (Gruenberg and van der Goot 2006; Cossart and Helenius 2014). Pathogen effectors and viruses have evolved and developed a variety of approaches to escape from host endosomes after internalization (Spooner et al. 2006; Personnic et al. 2016). To reach the host nucleus, VirE2 would need to escape from endosomes to avoid degradation in the host cytoplasm. This is indeed supported by our observation that VirE2 was not associated with plant late endosomes (Li and Pan 2017). Thus, further studies are needed to investigate how VirE2 escapes from endosomes and moves to other parts of host cells.

4 Trafficking of Delivered VirE2 Protein Inside Host Cells

4.1 VirE2 Moves Along the ER

To determine the cellular structure facilitating VirE2 trafficking, we conducted double labeling experiments (Yang et al. 2017). *A. tumefaciens* cells EHA105*virE2::GFP11*, encoding a VirE2-GFP11 fusion, were infiltrated into the leaf tissues of transgenic tobacco (Nb308A) plants, which constitutively expressed GFP1-10 and free DsRed, which labeled the cellular structures and nucleus (Li et al. 2014). Upon delivery into plant cells by the bacterium, VirE2-GFP11 complemented GFP1-10. The resulting VirE2-GFP_{comp} signals started to appear inside plant cells 2d after agroinfiltration.

The VirE2-GFP_{comp} signals moved along a strand-like cellular structure labeled with free DsRed (Yang et al. 2017). Interestingly, direct entry of VirE2 into the nucleus was visualized. VirE2 moved faster along linear tracks, although the velocities varied on different linear tracks and moved slower along curved tracks.

Chemical inhibitors were used to study the potential cellular structure that might facilitate the movement of *Agrobacterium*-delivered VirE2 inside plant cells

(Yang et al. 2017). Cytochalasin D (CytoD) and brefeldin A (BFA) had a significant effect on VirE2 trafficking, whereas colchicine (Colc) had only a minor effect.

CytoD is a potent inhibitor of actin polymerization (Krucker et al. 2000). BFA inhibits protein transport from the endoplasmic reticulum (ER) to the Golgi apparatus by dilating the ER (Misumi et al. 1986). Colc inhibits microtubule polymerization (Skoufias and Wilson 1992). The effects of these inhibitors on the corresponding cellular structures were observed in tobacco leaves. Colc, BFA, and CytoD disrupted the microtubules, ER, and actin structures, respectively. Therefore, we hypothesized that VirE2 trafficking was facilitated by ER/actin structures (Yang et al. 2017).

We determined if VirE2 movement is associated with the ER (Yang et al. 2017) by using an ER-mCherry construct containing an ER-targeting sequence at the N-terminus and the tetrapeptide retrieval signal HDEL at the C-terminus (Nelson et al. 2007). We found that VirE2 aggregates appeared as dots and filaments inside tobacco cells. Both forms were co-localized with inter-connected ER tubules that were shown by ER-mCherry. Time-lapse imaging showed that VirE2 aggregates moved along the ER strands (Yang et al. 2017).

As ER membranes compartmentalize the intracellular space into ER lumen and cytosol, it was necessary to determine whether VirE2 was present on the cytosolic or luminal side. By using an ER-targeting sequence and conducting the split-GFP complementation experiments, we demonstrated that *Agrobacterium*-delivered VirE2 was on the cytosolic side of the ER after delivery into the plant cytoplasm (Yang et al. 2017).

4.2 VirE2 Moves on an ER/F-Actin Network

We determined whether VirE2 movement is associated with F-actin filaments (Yang et al. 2017). To visualize both F-actin and VirE2, an F-actin marker tdTomato-ABD2 (Nakano et al. 2009) was expressed by agroinfiltrating EHA105*virE2::GFP11* into transgenic *N. benthamiana* Nb307A leaves expressing GFP1-10. We found that VirE2-GFP_{comp} signals co-localized with F-actin filaments. Time-lapse imaging demonstrated that VirE2 moved along an F-actin filament (Yang et al. 2017).

VirE2 affects the fate of T-DNA in many ways (Ward and Zambryski 2001). Therefore, it is important to determine how VirE2 is trafficked through the cytoplasm and reaches the nucleus. VirE2 contains two bipartite NLS signals (Citovsky et al. 1992), which are present on the exterior side of the solenoidal structure (Dym et al. 2008). This structural arrangement may make the NLS signals available to interact with other host factors. When the NLS of VirE2 was mutated to be recognizable in animal cells, the "animalized" VirE2 was found to migrate along microtubules in cell-free *Xenopus* oocyte extracts, propelled by dynein motors (Salman et al. 2005). However, no plant dyneins have been found (Lawrence et al. 2001). It remains unknown whether VirE2 moves along microtubules in plant cells.

Therefore, the trafficking of "animalized" VirE2 in animal cells may not accurately represent the mechanism of VirE2 trafficking inside plant cells. Moreover, disruption of microtubules by colchicine did not affect VirE2 trafficking significantly (Yang et al. 2017). These results suggest that VirE2 uses a transport system other than microtubules when trafficked inside plant cells. Our study showed that *Agrobacterium*-delivered VirE2 was trafficked via the ER and F-actin network (Yang et al. 2017). The VirE2-associated T-complex may also use the same trafficking mode, as VirE2 can coat the surface of the T-complex.

We hypothesize that *Agrobacterium* has evolved to enable VirE2 to exploit ER streaming, which is part of the cytoplasmic streaming process. *Agrobacterium*-delivered VirE2 is associated with the ER (Yang et al. 2017), probably because of the high affinity of VirE2 for membranes (Dumas et al. 2001). However, it is possible that an unknown factor(s) is responsible for VirE2-ER association. The VirE2-associated ER might be thus driven primarily by the ER-associated myosin XI-K. The myosin-associated ER can move along actin filaments. Therefore, *Agrobacterium*-delivered VirE2 might be trafficked through plant cells via the myosin-powered ER/actin network, because of the dynamic three-way interactions between the ER, F-actin, and myosin (Ueda et al. 2010).

The ER stretches through the entire cytoplasm and continues to the outer membrane of the nucleus, which may provide VirE2 with a convenient path to reach the nucleus. Cytosolic facing of VirE2 on the ER seems to make the opening of the nuclear pore complex accessible for nuclear import of VirE2. Association of VirE2 with the ER also suggests that VirE2 may interact with other factors during the trafficking processes. Indeed, a SNARE-like protein was found to have a strong interaction with VirE2 (Lee et al. 2012). It has also been reported that reticulon domain proteins and a Rab GTPase, both involved in trafficking of proteins through endomembranes, are important for transformation (Hwang and Gelvin 2004). These findings suggest that vesicular budding or fusion processes may be involved in VirE2 trafficking inside the cytoplasm.

4.3 VirE2 Movement Is Powered by Myosin

The speed of VirE2-GFP movement ranged from 1.3 to 3.1 μm/sec, and the movement was linear and directional (Yang et al. 2017). These results suggest that VirE2 trafficking inside plant cells is an active process that may be powered by myosin.

We determined whether myosin plays a role in VirE2 movement inside plant cells (Yang et al. 2017), as ER/F-actin/myosins may exhibit a three-way interaction (Ueda et al. 2010). A selective myosin light-chain kinase inhibitor ML-7 (Saitoh et al. 1987) was used to inhibit plant myosin activity. Treatment with ML-7 inhibited VirE2 movement as the average velocity was reduced by 95% relative to that of the control.

A dominant-negative approach was adopted to identify the specific myosin responsible for VirE2 movement (Yang et al. 2017). Several dominant-negative mutants of plant myosin genes were overexpressed during *Agrobacterium*-mediated delivery of VirE2. *A. tumefaciens* cells containing T-DNA encoding the tail constructs were co-infiltrated with EHA105*virE2::GFP11* into tobacco (Nb308A) plants. The myosin tail expression took place later than VirE2 delivery, so that the myosin mutant constructs would not affect VirE2 delivery. Among the myosin mutants tested, only the XI-K tail inhibited VirE2 trafficking (Yang et al. 2017). These data suggest that myosins provide the driving force for VirE2 movement, and that myosin XI-K is the most important contributor.

To determine if the VirE2 movement observed during our study was directly related to *Agrobacterium*-mediated transformation, the effect of the selective myosin light-chain kinase inhibitor ML-7 was tested on *Arabidopsis* root transformation. We found that ML-7 significantly reduced the transformation efficiency (Yang et al. 2017), while ML-7 did not affect the growth of root segments or *Agrobacterium* growth. These results suggest that the inhibition of myosin activity might have reduced the transformation efficiency.

To confirm the specific effect of myosin inhibition on transformation, RNAi constructs containing a partial sequence of XI-2 and XI-K (Avisar et al. 2008) were used to silence the corresponding genes (Yang et al. 2017). The RNAi constructs used for these experiments generated specific but not off-target effects (Avisar et al. 2008). We found that silencing of XI-K attenuated tumor formation (Yang et al. 2017). These data clearly indicate that XI-K affects VirE2 movement, and thereby *Agrobacterium*-mediated transformation.

Our study also showed that VirE2 trafficking may require the plant-specific myosin XI family and XI-K in particular (Yang et al. 2017). Myosin XI family members are involved in cytoplasmic streaming (Yokota et al. 1999), ER motility (Ueda et al. 2010), and trafficking of organelles and vesicles (Avisar et al. 2008). Despite the conformational similarities with myosin V, myosin XI has a plant-specific binding mechanism (Li and Nebenführ 2007) and thus recognizes different cargos than myosin V. This may provide an explanation for the very significant difference in transformation efficiency between yeast and plant recipients (0.2% in *S. cerevisiae* vs. 100.0% in *N. benthamiana*). The efficiency of protein delivery is comparable between yeast and plants (50.9% in *S. cerevisiae* vs. 100.0% in *N. benthamiana*) (Li et al. 2014). The budding yeast *S. cerevisiae* lacks myosin XI-K, which would render VirE2 immobile in the yeast cells. Thus, the transformation efficiency is significantly reduced. Our study demonstrated that myosin XI-K plays a much more critical role in VirE2 trafficking than does XI-2 (Yang et al. 2017). Both XI-K and XI-2 are highly expressed inside plant cells (Peremyslov et al. 2011). However, myosin XI-K is the primary contributor to ER streaming (Ueda et al. 2010).

4.4 Trafficking Mode of Delivered VirE2

The studies demonstrated that VirE2 is trafficked inside plant cytoplasm via a myosin XI-K-powered ER/actin network (Yang et al. 2017). This may indicate how the T-complex is trafficked inside host cells, because its surface consists of VirE2 molecules. VirE2 was visualized in real time to be trafficked toward the plant nucleus along a linear cellular structure that was illustrated by free DsRed molecules. This linear structure was later determined to be part of the ER/actin network, based on experiments using chemical treatments and fluorescent marker labeling. Moreover, VirE2 was present on the cytosolic side of the ER. Myosin XI-K provided the driving force for VirE2 movement (Yang et al. 2017).

The presence of VirE2 on the cytoplasmic side of the ER enables us to speculate about the trafficking mode of delivered VirE2 protein. As the ER is linked to the outer membrane of the nuclear envelope, delivered VirE2 should reside on the same topological surface as the nuclear opening. Because the ER is a dynamic structure that is involved in continuous flow and movement of lipids and proteins, the associated VirE2 may move along the ER inside the cytoplasm. This might explain why VirE2 can readily reach the nuclear opening for efficient nuclear import.

As shown in Fig. 2, we hypothesize how myosin XI-K powers VirE2 trafficking. One possibility is that myosin XI-K recognizes VirE2 with its globular domains and drives its movement, although there is no experimental evidence for this interaction yet (Yang et al. 2018). Another possibility is that ER-associated VirE2 may be present in a vesicle form that may be recognized by myosin XI-K (Yang et al. 2018). In this model, myosin XI-K indirectly drives VirE2 movement by carrying VirE2-containing cargoes; indeed, some myosin XI-K-specific vesicle adaptors have been identified (Kurth et al. 2017) although it is not clear whether VirE2 is associated with such a vesicle. Alternatively, VirE2 might passively follow the flow of the ER because of its association with the ER, although VirE2 and myosin XI-K might not have any interactions (Yang et al. 2018).

5 A Natural Model to Study Exogenous Protein Delivery

Delivery of materials into cells is a critical component of genetic engineering, genome-editing, therapies, and a diversity of fundamental research applications. However, efficient intracellular delivery of exogenous compounds and macromolecular cargo remains a long-standing challenge. The limitations of established delivery technologies have hampered progress in multiple areas, as the potential of exciting new materials, insights into disease mechanism, and approaches to cell therapy have not been fully realized because of the delivery hurdles. There is an urgent need to develop next-generation approaches for intracellular delivery (Delvigne et al. 2015), which are safe and efficient.

The natural molecular tracking system VirE2-GFP (Li et al. 2014; Li and Pan 2017; Yang et al. 2017) may be particularly useful for studying protein delivery and

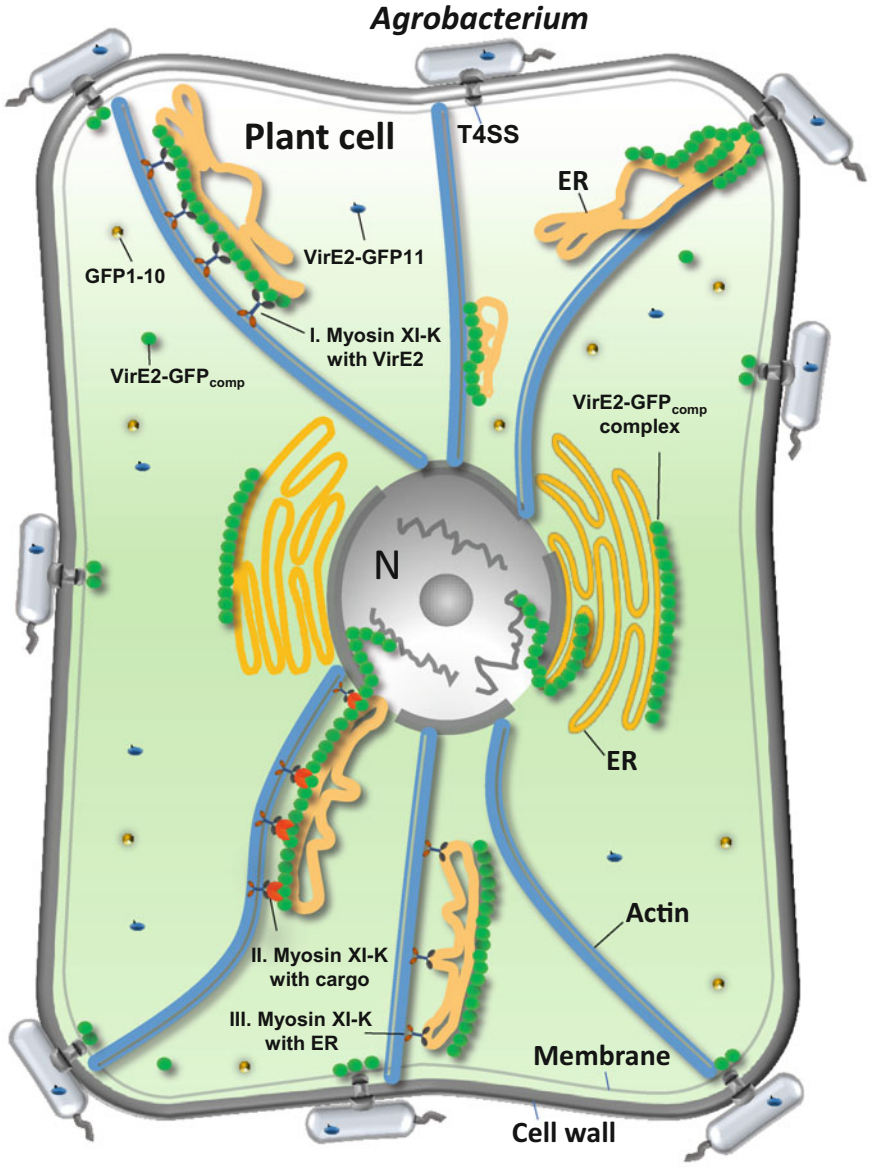


Fig. 2 A natural system to visualize the protein delivery process. VirE2-GFP11 fusion is expressed inside *A. tumefaciens*; and GFP1-10 is expressed in the plant cells. When VirE2-GFP11 is delivered into the plant cells, GFP1-10 would be complemented by VirE2-GFP11 and the resulting VirE2-GFP_{comp} signals can be detected. As VirE2-GFP11 is functional like VirE2, VirE2-GFP11 movement could represent the VirE2 trafficking. VirE2-GFP_{comp} complex starts to appear at the VirE2 delivery site and moves toward the nucleus. The movement is powered by myosin XI-K in three possible ways: (i) myosin XI-K interacts directly with VirE2; (ii) myosin XI-K interacts with a cargo associate with VirE2; and (iii) myosin XI-K interacts with ER associated with VirE2

trafficking (Fig. 2). VirE2 is a natural protein that can self-aggregate and contains the targeting sequences so that it can efficiently move all the way from the bacteria to the host plant nucleus. We adopted a split-GFP approach and generated the VirE2-GFP fusion (Li et al. 2014), which retains the self-aggregating property and the targeting sequences. The fusion protein is fully functional; it can be produced inside the bacterium and then delivered into plant cells and trafficked toward the plant nucleus in a readily detectable format: fluorescent and unique filamentous aggregates (Li et al. 2014). VirE2-GFP can track molecular events inside a cell in a natural setting in real time (Li et al. 2014). It is convenient to conduct experiments with this approach, as a large amount of VirE2-GFP can be delivered into the cells and the efficiency of delivery can be up to 100% of recipient cells that are in close contact with the bacterial cells (Li et al. 2014). This is useful to obtain more accurate insight into the delivery process.

Using the GFP-VirE2 system, we discovered that the exogenously produced protein VirE2 is naturally delivered into plant cells via clathrin-mediated endocytosis. We identified a VirE2 internalization signal that can target VirE2 (Li and Pan 2017) and potentially other macromolecules into cells. We also found that the delivered VirE2 is subsequently trafficked inside plant cells via a myosin XI-K-powered ER/actin network (Yang et al. 2017). We found that the VirE2-GFP system can be used as a new method to test drug toxicity, as described below. These findings demonstrate the usefulness of the VirE2-GFP-based imaging approach to study protein delivery and sorting.

6 A Test System to Examine the Effect of a Molecule on Cells

A biological activity is often studied in an unnatural setting, such as an in vitro system, and in a snapshot manner, for instance, by measuring a reporter activity at different time points. Furthermore, current reporters normally focus on one aspect of the biological process at a time; they are not suitable to study the overall wellbeing of the cell. These systems may therefore generate inaccurate conclusions and undesirable products.

The natural molecular tracking system VirE2-GFP11 may be defined as a bio-tracker, as the fusion protein is fully functional and can move inside cells in a readily detectable format: fluorescent and unique filamentous aggregates (Fig. 2). It is much easier to recognize the VirE2-GFP signals than GFP reporter activity. Importantly, VirE2-GFP can bind to membrane structures such as the endoplasmic reticulum (ER) that are trafficking inside a cell. This enables us to readily observe the highly dynamic activities of the cell. If a compound, a gene, or a condition affects the cell, the cellular dynamics shown by VirE2-GFP will be affected and thus readily observed under confocal microscopy. We found that the assay is highly sensitive and can readily detect minor effects on the cell.

Thus, VirE2-GFP can be used as a new research tool to label a cell, tissue, organ, organ system, or organism. This can facilitate real-time studies on the cell, tissue, organ, organ system, or organism in a natural setting. It enables the researcher to address new questions that cannot be addressed by current approaches.

The technology can study the wellbeing of a cell in a sensitive, rapid, and visually understandable manner. This will help improve the accuracy of biological research and develop quality products and services to meet pressing challenges. The bio-tracker-based toxicity testing method provides a new approach, which not only can reduce the use of animals, but also can present the toxicity in a visually understandable format. This is useful to improve the relations between the public and drug development.

As an example, we conduct non-specific toxicity tests as follows.

A. tumefaciens cells EHA105*virE2::GFP11*, encoding VirE2-GFP11 fusion, were infiltrated into the leaf tissues of transgenic tobacco (Nb308A) plants, which constitutively express GFP1-10 and free DsRed to indicate the cellular structures and the nucleus. Transgenic tobacco (Nb308A) plants were treated at 42 h post-agroinfiltration with the chemicals econazole and ketoconazole. The effects were observed 6 h later. As shown in Fig. 3, the effects of econazole and

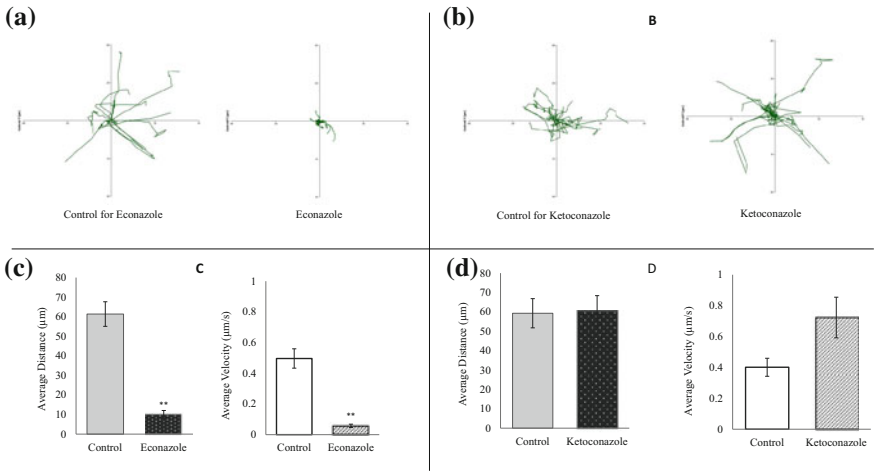


Fig. 3 Comparison of the non-specific toxicity between econazole and ketoconazole. *Agrobacterium* strain EHA105-VirE2::GFP11 was infiltrated into the leaves of *N. benthamiana* transgenic plant Nb308A. Chemical infiltration was performed 42 h after agroinfiltration. The control groups were set by infiltration with only the respective solvent aqueous dilutes. Six hours after chemical infiltration, images were taken in multiple focal planes (Z-stacks) with a step interval of 1.0 μm using a spinning disk confocal microscope system with Olympus UPLSAPO 60×/1.20 water. Time-lapse imaging was set at 3 min. The movement of VirE2 signals was tracked with "manual tracking" function provided by Volocity 3D Image Analysis Software (Ver. 6.2.1). Tracking plots, track lengths, and velocities were generated automatically after tracking was finished. Tracking plots (**a** and **b**) are presented in such a way that very movement track is assumed to start at the same origin. Mean track velocity and mean displacement of VirE2 movement (**c** and **d**) were measured; the data analysis was performed using ANOVA (*$P < 0.05$; **$P < 0.01$). The effects of econazole (**a** and **c**) at 45 μM and ketoconazole (**b** and **d**) at 94 μM were tested

ketoconazole on VirE2-GFP movement can be readily differentiated. Although both of them are imidazole derivatives and known to fight fungal infections, econazole caused much more toxicity effects on the plant cells than did ketoconazole, based on the pattern, distance and velocity of VirE2-GFP movement pattern. This is consistent with the fact that econazole indeed causes more harm in humans than does ketoconazole.

Because plant cells are very much different from the human cells, we consider the effect on VirE2-GFP movement as non-specific toxicity. Using plant cells to test non-specific toxicity should represent a new approach to assess the non-specific toxicity of a drug or compound.

7 Conclusions

Agrobacterium is widely used as a genetic vector to deliver DNA into various cells, whereas its capacity to deliver proteins is not fully explored. In a sense, *Agrobacterium* is regarded as a genetic engineer and not widely used as a vector for protein delivery. Direct visualization of *Agrobacterium*-delivered VirE2 protein, based on a split-GFP approach, indicated that *Agrobacterium* is more efficient in protein delivery than genetic transformation for a non-natural-host recipient. It should be of significance to further explore the capacity of *A. tumefaciens* to deliver proteins.

Visualization of VirE2 inside recipient cells may be also useful to study the trafficking pathway of T-strands, as VirE2 is a component of the proposed nucleoprotein complex. The abundance of VirE2 (Engstrom et al. 1987) is particularly suitable for its role to protect T-DNA by coating it with numerous molecules (Citovsky et al. 1988). It was estimated that about every 19 bases of T-DNA is coupled with one VirE2 molecule (Citovsky et al. 1997). In addition, VirE2 protein can also assemble without T-DNA to form homodimers and solenoids (Frenkiel-Krispin et al. 2007). These unique traits prompted us to use VirE2 as a model to study *Agrobacterium*-delivered molecules inside recipient cells.

The split-GFP approach enabled us to directly visualize *Agrobacterium*-delivered VirE2 in live recipient cells. Because VirE2-GFP11 is delivered in a natural setting, VirE2-GFP11 movement should represent the natural trafficking process inside recipient cells. This should be of use to further study how host factors facilitate VirE2 movement inside recipient cells.

As a natural genetic engineer, *A. tumefaciens* can cause crown gall disease on an exceptionally wide range of host plants in nature (De Cleene and De Ley 1976). The bacterium can achieve a high efficiency of transformation that can approach 100% (Li et al. 2014). Using VirE2-GFP as the model, we show that *Agrobacterium* uses the host endocytic process to facilitate delivery and trafficking for one of its virulence factors, VirE2, into host cells. Endocytosis is a well-conserved fundamental process in all eukaryotic cells. It may occur independently of cell types or differentiation. In addition, the bacterium hijacks a

conserved host network to move virulence factor VirE2 toward the nucleus. These may be important for *Agrobacterium* to achieve both a wide host range and a high efficiency.

VirE2-GFP may be used as a bio-tracker, because the VirE2-GFP fusion retains the self-aggregating property and the targeting sequences. In addition, the fusion protein is fully functional and can move inside cells in a readily detectable format: fluorescent and unique filamentous aggregates. VirE2-GFP can bind to some membrane structures such as the endoplasmic reticulum (ER) that are trafficking inside a cell. This enables us to readily observe the highly dynamic activities of the cell. If a compound, a gene, or a condition affects the cell, the cellular dynamics shown by the VirE2-GFP may be affected and thus readily observed under confocal microscopy. We found that the assay is highly sensitive and can readily detect minor effects on the cell.

Thus, VirE2-GFP can be used as a new research tool to label cells, tissues, organs, organ systems, or organisms. This system can facilitate real-time studies on the cell, tissue, organ, organ system, or organism in a natural setting. It enables the researcher to address new questions that cannot be addressed with the current approaches.

Acknowledgements The authors acknowledge grant supports from Singapore Ministry of Education (R-154-000-588-112 and R-154-000-685-112) and National Natural Science Foundation of China (Grant No. 31700118).

References

Aguilar J, Zupan J, Cameron TA, Zambryski PC (2010) *Agrobacterium* type IV secretion system and its substrates form helical arrays around the circumference of virulence-induced cells. Proc Natl Acad Sci USA 107:3758–3763

Anand A, Krichevsky A, Schornack S, Lahaye T, Tzfira T, Tang Y, Citovsky V, Mysore KS (2007) *Arabidopsis* VIRE2 INTERACTING PROTEIN2 is required for *Agrobacterium* T-DNA integration in plants. Plant Cell 19:1695–1708

Atmakuri K, Ding Z, Christie PJ (2003) VirE2, a type IV secretion substrate, interacts with the VirD4 transfer protein at cell poles of *Agrobacterium tumefaciens*. Mol Microbiol 49:1699–1713

Avisar D, Prokhnevsky AI, Makarova KS, Koonin EV, Dolja VV (2008) Myosin XI-K Is required for rapid trafficking of Golgi stacks, peroxisomes, and mitochondria in leaf cells of *Nicotiana benthamiana*. Plant Physiol 146:1098–1108

Banbury DN, Oakley JD, Sessions RB, Banting G (2003) Tyrphostin A23 inhibits internalization of the transferrin receptor by perturbing the interaction between tyrosine motifs and the medium chain subunit of the AP-2 adaptor complex. J Biol Chem 278:12022–12028

Bar M, Avni A (2014) Endosomal trafficking and signaling in plant defense responses. Current Opin Plant Biol 22:86–92

Beijersbergen A, Dulk-Ras AD, Schilperoort RA, Hooykaas PJ (1992) Conjugative transfer by the virulence system of *Agrobacterium tumefaciens*. Science 256:1324–1327

Bhattacharjee S, Lee LY, Oltmanns H, Cao H, Veena, Cuperus J, Gelvin SB (2008) IMPa-4, an *Arabidopsis* importin alpha isoform, is preferentially involved in *Agrobacterium*-mediated plant transformation. Plant Cell 20:2661–2680

Bonifacino JS, Traub LM (2003) Signals for sorting of transmembrane proteins to endosomes and lysosomes. Annu Rev Biochem 72:395–447

Broothaerts W, Mitchell H, Weir B, Kaines S, Smith L, Yang W, Mayer J, Roa-Rodríguez C, Jefferson RA (2005) Gene transfer to plants by diverse species of bacteria. Nature 433:629–633

Bundock P, den Dulk-Ras A, Beijersbergen A, Hooykaas PJ (1995) Trans-kingdom T-DNA transfer from *Agrobacterium tumefaciens* to *Saccharomyces cerevisiae*. EMBO J 14:3206–3214

Cabantous S, Pedelacq JD, Mark BL, Naranjo C, Terwilliger TC, Waldo GS (2005) Recent advances in GFP folding reporter and split-GFP solubility reporter technologies. Application to improving the folding and solubility of recalcitrant proteins from *Mycobacterium tuberculosis*. J Struc Funct Genom 6:113–119

Cascales E, Christie PJ (2004) Definition of a bacterial type IV secretion pathway for a DNA substrate. Science 304:1170–1173

Chen X, Irani NG, Friml J (2011) Clathrin-mediated endocytosis: the gateway into plant cells. Curr Opin Plant Biol 14:674–682

Chilton M-D, Drummond MH, Merlo DJ, Sciaky D, Montoya AL, Gordon MP, Nester EW (1977) Stable incorporation of plasmid DNA into higher plant cells: the molecular basis of crown gall tumorigenesis. Cell 11:263–271

Citovsky V, DEV G, Zambryski P (1988) Single-stranded DNA binding protein encoded by the virE locus of Agrobacterium tumefaciens. Science 240:501–504

Citovsky V, Guralnick B, Simon MN, Wall JS (1997) The molecular structure of *Agrobacterium* VirE2-single stranded DNA complexes involved in nuclear import. J Mol Biol 271:718–727

Citovsky V, Zupan J, Warnick D, Zambryski P (1992) Nuclear localization of *Agrobacterium* VirE2 protein in plant cells. Science 256:1802–1805

Contento AL, Bassham DC (2012) Structure and function of endosomes in plant cells. J Cell Sci 125:3511–3518

Cossart P, Helenius A (2014) Endocytosis of viruses and bacteria. Cold Spring Harb Perspec Biol 6:a016972

De Cleene M, De Ley J (1976) The host range of crown gall. Bot Rev 42:389–466

de Groot MJA, Bundock P, Hooykaas PJJ, Beijersbergen AGM (1998) *Agrobacterium tumefaciens*-mediated transformation of filamentous fungi. Nat Biotechnol 16:839

Delvigne F, Pêcheux H, Tarayre C (2015) Fluorescent reporter libraries as useful tools for optimizing microbial cell factories: a review of the current methods and applications. Frontier Bioengin Biotechnol 3:147

Dhonukshe P, Aniento F, Hwang I, Robinson DG, Mravec J, Stierhof YD, Friml J (2007) Clathrin-mediated constitutive endocytosis of PIN auxin efflux carriers in *Arabidopsis*. Curr Biol CB 17:520–527

Djamei A, Pitzschke A, Nakagami H, Rajh I, Hirt H (2007) Trojan horse strategy in *Agrobacterium* transformation: abusing MAPK defense signaling. Science 318:453–456

Dumas F, Duckely M, Pelczar P, Van Gelder P, Hohn B (2001) An *Agrobacterium* VirE2 channel for transferred-DNA transport into plant cells. Proc Natl Acad Sci USA 98:485–490

Dym O, Albeck S, Unger T, Jacobovitch J, Branzburg A, Michael Y, Frenkiel-Krispin D, Wolf SG, Elbaum M (2008) Crystal structure of the *Agrobacterium* virulence complex VirE1-VirE2 reveals a flexible protein that can accommodate different partners. Proc Natl Acad Sci USA 105:11170–11175

Ebine K, Fujimoto M, Okatani Y, Nishiyama T, Goh T, Ito E, Dainobu T, Nishitani A, Uemura T, Sato MH, Thordal-Christensen H, Tsutsumi N, Nakano A, Ueda T (2011) A membrane trafficking pathway regulated by the plant-specific RAB GTPase ARA6. Nat Cell Biol 13:853–859

Engstrom P, Zambryski P, Van Montagu M, Stachel S (1987) Characterization of *Agrobacterium tumefaciens* virulence proteins induced by the plant factor acetosyringone. J Mol Biol 197:635–645

Foresti O, Denecke J (2008) Intermediate organelles of the plant secretory pathway: identity and function. Traffic 9:1599–1612

Frenkiel-Krispin D, Wolf SG, Albeck S, Unger T, Peleg Y, Jacobovitch J, Michael Y, Daube S, Sharon M, Robinson CV, Svergun DI, Fass D, Tzfira T, Elbaum M (2007) Plant transformation by *Agrobacterium tumefaciens*: modulation of single-stranded DNA-VirE2 complex assembly by VirE1. J Biol Chem 282:3458–3464

Geldner N, Anders N, Wolters H, Keicher J, Kornberger W, Muller P, Delbarre A, Ueda T, Nakano A, Jurgens G (2003) The *Arabidopsis* GNOM ARF-GEF mediates endosomal recycling, auxin transport, and auxin-dependent plant growth. Cell 112:219–230

Gruenberg J, van der Goot FG (2006) Mechanisms of pathogen entry through the endosomal compartments. Nat Rev Mol Cell Biol 7:495

Hood EE, Gelvin SB, Melchers LS, Hoekema A (1993) New *Agrobacterium* helper plasmids for gene transfer to plants. Transgenic Res 2:208–218

Hwang H-H, Gelvin SB (2004) Plant proteins that interact with VirB2, the *Agrobacterium tumefaciens* pilin protein, mediate plant transformation. Plant Cell 16:3148–3167

Jackson LP, Kelly BT, McCoy AJ, Gaffry T, James LC, Collins BM, Honing S, Evans PR, Owen DJ (2010) A large-scale conformational change couples membrane recruitment to cargo binding in the AP2 clathrin adaptor complex. Cell 141:1220–1229

Kathiresan S, Chandrashekar A, Ravishankar GA, Sarada R (2009) *Agrobacterium*-mediated transformation in the green alga *Haematococcus Pluvialis* (Chlorophyceae, Volvocales) 1. J Phycol 45:642–649

Kim SY, Xu Z-Y, Song K, Kim DH, Kang H, Reichardt I, Sohn EJ, Friml J, Juergens G, Hwang I (2013) Adaptor protein complex 2–mediated endocytosis is crucial for male reproductive organ development in *Arabidopsis*. Plant Cell 25:2970–2985

Kitakura S, Vanneste S, Robert S, Lofke C, Teichmann T, Tanaka H, Friml J (2011) Clathrin mediates endocytosis and polar distribution of PIN auxin transporters in *Arabidopsis*. Plant Cell 23:1920–1931

Krucker T, Siggins GR, Halpain S (2000) Dynamic actin filaments are required for stable long-term potentiation (LTP) in area CA1 of the hippocampus. Proc Natl Acad Sci USA 97:6856–6861

Kurth EG, Peremyslov VV, Turner HL, Makarova KS, Iranzo J, Mekhedov SL, Koonin EV, Dolja VV (2017) Myosin-driven transport network in plants. Proc Natl Acad Sci USA 114: E1385–E1394

Lapham R, Lee LY, Tsugama D, Lee S, Mengiste T, Gelvin SB (2018) *VIP1* and its homologs are not required for *Agrobacterium*-mediated transformation, but play a role in *Botrytis* and salt stress responses. Front Plant Sci 9:749

Lawrence CJ, Morris NR, Meagher RB, Dawe RK (2001) Dyneins have run their course in plant lineage. Traffic 2:362–363

Lee LY, Fang MJ, Kuang LY, Gelvin SB (2008) Vectors for multi-color bimolecular fluorescence complementation to investigate protein-protein interactions in living plant cells. Plant Methods 4:24

Lee LY, Wu FH, Hsu CT, Shen SC, Yeh HY, Liao DC, Fang MJ, Liu NT, Yen YC, Dokladal L, Sykorova E, Gelvin SB, Lin CS (2012) Screening a cDNA library for protein-protein interactions directly in planta. Plant Cell 24:1746–1759

Li J-F, Nebenführ A (2007) Organelle targeting of myosin XI is mediated by two globular tail subdomains with separate cargo binding sites. J Biol Chem 282:20593–20602

Li X, Pan SQ (2017) *Agrobacterium* delivers VirE2 protein into host cells via clathrin-mediated endocytosis. Sci Adv 3:e1601528

Li X, Yang Q, Tu H, Lim Z, Pan SQ (2014) Direct visualization of *Agrobacterium*-delivered VirE2 in recipient cells. Plant J 77:487–495

Liu S-H, Wong ML, Craik CS, Brodsky FM (1995) Regulation of clathrin assembly and trimerization defined using recombinant triskelion hubs. Cell 83:257–267

Maizel A, Dv Wangenheim, Federici F, Haseloff J, Stelzer EHK (2011) High-resolution live imaging of plant growth in near physiological bright conditions using light sheet fluorescence microscopy. Plant J 68:377–385

Marton L, Wullems GJ, Molendijk L, Schilperoort RA (1979) In vitro transformation of cultured cells from *Nicotiana tabacum* by *Agrobacterium tumefaciens*. Nature 277:129

McMahon HT, Boucrot E (2011) Molecular mechanism and physiological functions of clathrin-mediated endocytosis. Nat Rev Mol Cell Biol 12:517–533

Mercer J, Schelhaas M, Helenius A (2010) Virus entry by endocytosis. Ann Rev Biochem 79:803–833

Misumi Y, Misumi Y, Miki K, Takatsuki A, Tamura G, Ikehara Y (1986) Novel blockade by brefeldin A of intracellular transport of secretory proteins in cultured rat hepatocytes. J Biol Chem 261:11398–11403

Nakano RT, Matsushima R, Ueda H, Tamura K, Shimada T, Li L, Hayashi Y, Kondo M, Nishimura M, Hara-Nishimura I (2009) GNOM-LIKE1/ERMO1 and SEC24a/ERMO2 are required for maintenance of endoplasmic reticulum morphology in *Arabidopsis thaliana*. Plant Cell 21:3672–3685

Nelson BK, Cai X, Nebenfuhr A (2007) A multicolored set of in vivo organelle markers for co-localization studies in *Arabidopsis* and other plants. Plant J 51:1126–1136

Pedelacq JD, Cabantous S, Tran T, Terwilliger TC, Waldo GS (2006) Engineering and characterization of a superfolder green fluorescent protein. Nat Biotechnol 24:79–88

Peremyslov VV, Mockler TC, Filichkin SA, Fox SE, Jaiswal P, Makarova KS, Koonin EV, Dolja VV (2011) Expression, splicing, and evolution of the myosin gene family in plants. Plant Physiol 155:1191–1204

Personnic N, Barlocher K, Finsel I, Hilbi H (2016) Subversion of retrograde trafficking by translocated pathogen effectors. Trends Microbiol 24:450–462

Piers KL, Heath JD, Liang XY, Stephens KM, Nester EW (1996) *Agrobacterium tumefaciens*-mediated transformation of yeast. Proc Natl Acad Sci USA 93:1613–1618

Rhee Y, Gurel F, Gafni Y, Dingwall C, Citovsky V (2000) A genetic system for detection of protein nuclear import and export. Nat Biotechnol 18:433–437

Robert S, Chary SN, Drakakaki G, Li S, Yang Z, Raikhel NV, Hicks GR (2008) Endosidin1 defines a compartment involved in endocytosis of the brassinosteroid receptor BRI1 and the auxin transporters PIN2 and AUX1. Proc Natl Acad Sci USA 105:8464–8469

Rossi L, Hohn B, Tinland B (1996) Integration of complete transferred DNA units is dependent on the activity of virulence E2 protein of *Agrobacterium tumefaciens*. Proc Natl Acad Sci USA 93:126–130

Saitoh M, Ishikawa T, Matsushima S, Naka M, Hidaka H (1987) Selective inhibition of catalytic activity of smooth muscle myosin light chain kinase. J Biol Chem 262:7796–7801

Sakalis PA, van Heusden GP, Hooykaas PJ (2014) Visualization of VirE2 protein translocation by the *Agrobacterium* type IV secretion system into host cells. Microbiol Open 3:104–117

Salman H, Abu-Arish A, Oliel S, Loyter A, Klafter J, Granek R, Elbaum M (2005) Nuclear localization signal peptides induce molecular delivery along microtubules. Biophys J 89:2134–2145

Scheiffele P, Pansegrau W, Lanka E (1995) Initiation of *Agrobacterium tumefaciens* T-DNA processing. Purified proteins VirD1 and VirD2 catalyze site- and strand-specific cleavage of superhelical T-border DNA in vitro. J Biol Chem 270:1269–1276

Schrammeijer B, den Dulk-Ras A, Vergunst AC, Jurado Jacome E, Hooykaas PJ (2003) Analysis of Vir protein translocation from *Agrobacterium tumefaciens* using *Saccharomyces cerevisiae* as a model: evidence for transport of a novel effector protein VirE3. Nucleic Acids Res 31:860–868

Shi Y, Lee LY, Gelvin SB (2014) Is VIP1 important for *Agrobacterium*-mediated transformation? Plant J. 79:848–860

Simone M, McCullen CA, Stahl LE, Binns AN (2001) The carboxy-terminus of VirE2 from *Agrobacterium tumefaciens* is required for its transport to host cells by the virB-encoded type IV transport system. Mol Microbiol 41:1283–1293

Skoufias DA, Wilson L (1992) Mechanism of inhibition of microtubule polymerization by colchicine: inhibitory potencies of unliganded colchicine and tubulin-colchicine complexes. Biochemistry 31:738–746

Spooner RA, Smith DC, Easton AJ, Roberts LM, Lord JM (2006) Retrograde transport pathways utilised by viruses and protein toxins. Virol J 3:26

Traub LM (2009) Tickets to ride: selecting cargo for clathrin-regulated internalization. Nat Rev Mol Cell Biol 10:583–596

Traub LM, Bonifacino JS (2013) Cargo recognition in clathrin-mediated endocytosis. Cold Spring Harb Perspect Biol 5:a016790

Tzfira T, Citovsky V (2001) Comparison between nuclear localization of nopaline and octopine-specific *Agrobacterium* VirE2 proteins in plant, yeast and mammalian cells. Mol Plant Pathol 2:171–176

Tzfira T, Vaidya M, Citovsky V (2002) Increasing plant susceptibility to *Agrobacterium* infection by overexpression of the *Arabidopsis* nuclear protein VIP1. Proc Natl Acad Sci USA 99:10435–10440

Ueda H, Yokota E, Kutsuna N, Shimada T, Tamura K, Shimmen T, Hasezawa S, Dolja VV, Hara-Nishimura I (2010) Myosin-dependent endoplasmic reticulum motility and F-actin organization in plant cells. Proc Natl Acad Sci USA 107:6894–6899

Ueda T, Uemura T, Sato MH, Nakano A (2004) Functional differentiation of endosomes in *Arabidopsis* cells. Plant J 40:783–789

Van Engelenburg SB, Palmer AE (2010) Imaging type-III secretion reveals dynamics and spatial segregation of *Salmonella* effectors. Nat Methods 7:325–330

Vergunst AC, Schrammeijer B, den Dulk-Ras A, de Vlaam CM, Regensburg-Tuink TJ, Hooykaas PJ (2000) VirB/D4-dependent protein translocation from *Agrobacterium* into plant cells. Science 290:979–982

Vergunst AC, van Lier MC, den Dulk-Ras A, Stuve TA, Ouwehand A, Hooykaas PJ (2005) Positive charge is an important feature of the C-terminal transport signal of the VirB/D4-translocated proteins of *Agrobacterium*. Proc Natl Acad Sci USA 102:832–837

Wang K, Herrera-Estrella L, Van Montagu M, Zambryski P (1984) Right 25 bp terminus sequence of the nopaline T-DNA is essential for and determines direction of DNA transfer from agrobacterium to the plant genome. Cell 38:455–462

Ward DV, Zambryski PC (2001) The six functions of *Agrobacterium* VirE2. Proc Natl Acad Sci USA 98:385–386

Wu LG, Hamid E, Shin W, Chiang HC (2014) Exocytosis and endocytosis: modes, functions, and coupling mechanisms. Ann Rev Physiol 76:301–331

Yang Q, Li X, Tu H, Pan SQ (2017) *Agrobacterium*-delivered virulence protein VirE2 is trafficked inside host cells via a myosin XI-K–powered ER/actin network. Proc Natl Acad Sci USA 114:2982–2987

Yang Q, Tu H, Pan SQ (2018) How does *Agrobacterium virulence* protein VirE2 hijack host actomyosin motility system? CommunIntegr Biol 11:e1415599

Yanofsky MF, Porter SG, Young C, Albright LM, Gordon MP, Nester EW (1986) The *virD* operon of *Agrobacterium tumefaciens* encodes a site-specific endonuclease. Cell 47:471–477

Yokota E, Yukawa C, Muto S, Sonobe S, Shimmen T (1999) Biochemical and immunocytochemical characterization of two types of myosins in cultured tobacco bright yellow-2 cells. Plant Physiol 121:525–534

Yusibov VM, Steck TR, Gupta V, Gelvin SB (1994) Association of single-stranded transferred DNA from *Agrobacterium tumefaciens* with tobacco cells. Proc Natl Acad Sci USA 91:2994–2998

Zhou XR, Christie PJ (1999) Mutagenesis of the Agrobacterium VirE2 single-stranded DNA-binding protein identifies regions required for self-association and interaction with VirE1 and a permissive site for hybrid protein construction. J Bacteriol 181:4342–4352

The Mechanism of T-DNA Integration: Some Major Unresolved Questions

Kamy Singer

Contents

1 Introduction 288
1.1 The Transfer of T-DNA 289
1.2 Stable and Transient T-DNA in the Plant Cell 291
2 The Mechanism of T-DNA Integration 292
3 The Major Unresolved Questions Related to the Mechanism of T-DNA Integration 294
3.1 Does T-DNA Integrate into the Plant Genome as a Single- or a Double-Stranded Intermediate? 294
3.2 What Are the Bacterial and Plant Factors Involved in T-DNA Integration? 297
3.3 What Is the Genomic Site Prerequisite for T-DNA Integration? 301
3.4 What Is the Spatial/Temporal Arrangement of T-DNA During Integration? 303
3.5 Why and How Do Complex T-DNA Insertions Form? 304
4 A Proposed Model 307
5 Conclusions 308
References 309

Abstract The mechanism of T-DNA integration into plant genomes during *Agrobacterium*-mediated genetic transformation is still not understood. As genetic transformation of plants via *Agrobacterium* has become a routine practice among plant biologists, understanding T-DNA integration remains important for several reasons. First, T-DNA is the final step in one of the unique cases of inter-kingdom horizontal gene transfer in nature. Second, understanding T-DNA integration is important for biotechnological applications. For example, better knowledge of this process may help develop methods to transform species that are currently not susceptible to *Agrobacterium*-mediated transformation. In addition, regulatory agencies usually require "clean" and "precise" transgenic insertion events, whereas

K. Singer
Department of Molecular, Cellular, and Developmental Biology, University of Michigan, Ann Arbor, MI 48109, USA

K. Singer (✉)
Department of Biological Sciences, Purdue University, West Lafayette, IN 47907, USA
e-mail: kamys@umich.edu

Current Topics in Microbiology and Immunology (2018) 418:287–317
DOI 10.1007/82_2018_98

Published Online: 11 July 2018

transgenic insertions are commonly complex unpredictable structures. Furthermore, whereas T-DNA integration under natural conditions occurs randomly, technology to direct T-DNA to specific sites in the genome is highly desired. A better understanding of T-DNA integration may help develop methods to achieve more desirable results. Finally, gene targeting methods that require a foreign DNA template for precise DNA modifications in plants often utilize *Agrobacterium* to deliver the DNA template. Better understanding of the fate of T-DNA in the plant nucleus may help utilize T-DNA for more efficient gene targeting. For introducing gene targeting reagents, efficient delivery of T-DNA without ectopic integration would be useful. The following review summarizes current knowledge related to T-DNA integration. Five major open questions related to T-DNA integration are being presented. Finally, different models for T-DNA integration are being discussed, and a revised model is proposed.

1 Introduction

Agrobacterium tumefaciens is a soil-borne bacterium well known for its unique ability of inter-kingdom horizontal gene transfer. In nature, this plant pathogen causes the disease crown gall (Smith and Townsend 1907). The disease is characterized by galls appearing at the plant's root, stem, or crown area. These galls are tumor growths that form as a result of a transfer of a region of the *Agrobacterium* tumor-inducing (Ti) plasmid into a plant cell and stable integration of this DNA into the plant genome (Zaenen et al. 1974; Chilton et al. 1977; Fig. 1). The transferred DNA (T-DNA) of the Ti plasmid carries genes that cause uncontrolled cell divisions by modifying the plant's hormonal balance, whereas other genes encode proteins involved in the production of opines, compounds that are utilized by the *Agrobacterium* colonies surrounding the galls.

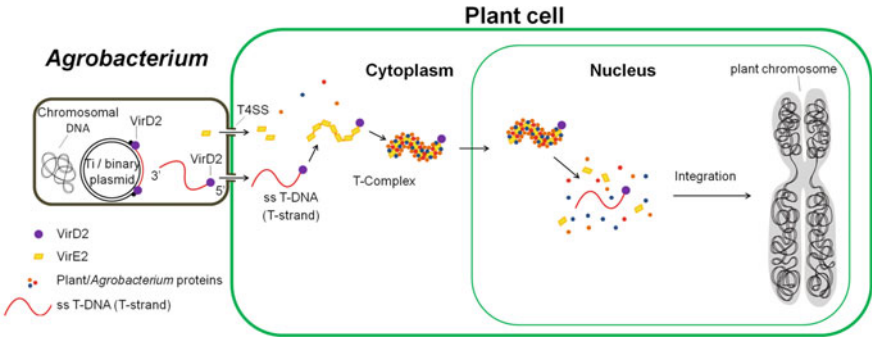


Fig. 1 Schematic illustration of a model describing DNA transfer from *Agrobacterium* to the plant cell (see relevant book chapters for further details). Adapted from Singer (2013) doctoral dissertation

Agrobacterium has been extensively studied because of its role as a "natural genetic engineer." Moreover, *Agrobacterium* has been harnessed by humans as a gene vector to genetically engineer plants. *Agrobacterium*-mediated transformation was the first method to generate transgenic plants (Barton et al. 1983; Zambryski et al. 1983). Three decades later, this bacterium is still a key player in many of the plant molecular genetics techniques used in agricultural biotechnologies (reviewed in Shiboleth and Tzfira 2012; Altpeter et al. 2016).

Whereas early events during *Agrobacterium* infection are relatively well studied, the final events in the plant nucleus are relatively less understood. Notably, the mechanism behind T-DNA integration into the plant genome is still unknown. Consequently, investigators often adopt different models to explain T-DNA integration. For instance, the process of T-DNA integration resulting in chromosome truncation has been explained in two different ways recently. In Teo et al. (2011), T-DNA is a single-stranded (ss) molecule during integration, whereas in Nelson et al. (2011) T-DNA is a double-stranded (ds) molecule. Moreover, according to the model in Teo et al. (2011), bacterial VirD2 is involved in integration, whereas according to Nelson et al. (2011), integration is mediated by plant host proteins. Thus, the literature may present conflicting models for T-DNA integration which are based on three decades of research evidence leading to somewhat conflicting conclusions. This review describes the various open questions that contribute to current models of T-DNA integration.

1.1 The Transfer of T-DNA

T-DNA integration requires the transfer of T-DNA from *Agrobacterium* into the plant cells. T-DNA transfer is a process related to bacterial conjugation (for review, see Lessl and Lanka 1994; Christie et al. 2005). Transfer begins inside the bacterium when T-DNA is separated from its parent plasmid, a Ti plasmid in natural strains or a binary plasmid in laboratory strains. The separation of T-DNA is initiated when a protein complex of VirD1, a helicase, and VirD2, an endonuclease, attaches to the left border (LB) and right border (RB) of a T-DNA region (Durrenberger et al. 1989; Scheiffele et al. 1995; Relic et al. 1998) (Fig. 2). The LB and RB are 25 base pairs (bp) of imperfect direct repeats (Yadav et al. 1982). VirD2 nicks the lower DNA strand between the third and fourth nucleotides of each of these repeats (Fig. 2) (Yanofsky et al. 1986; Wang et al. 1987). Consequently, a single-stranded (ss) T-DNA, termed the T-strand, is generated from the parent plasmid (Albright et al. 1987). A single VirD2 protein remains covalently attached to the 5′ end of the ss T-DNA (the RB side) (Herrera-Estrella et al. 1988; Ward and Barnes 1988; Young and Nester 1988; Vogel and Das 1992) (Fig. 2).

VirD2 protein pilots the ss T-DNA from its 5′ end through the *Agrobacterium* type IV protein secretion (T4S) system into the plant cytoplasm (Vergunst et al. 2005; van Kregten et al. 2009) (Fig. 1). In the plant cytoplasm, before entering the nucleus, the transported ss T-DNA is thought to be coated by multiple VirE2 ss

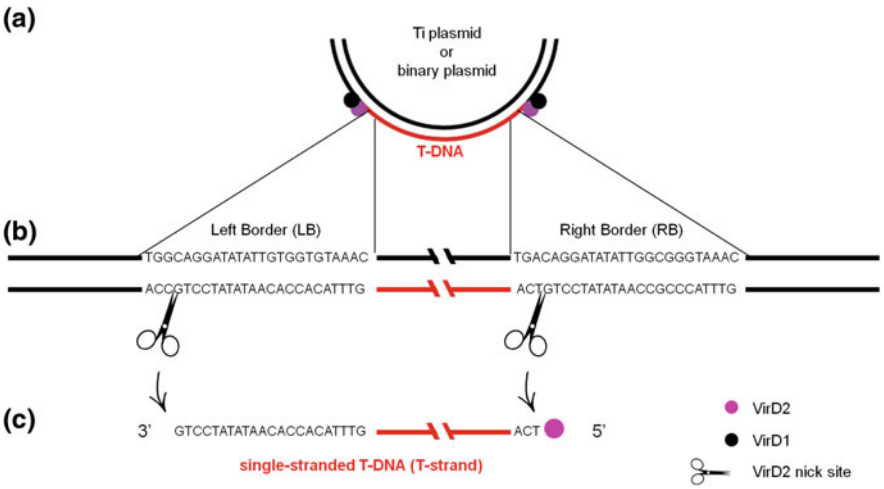


Fig. 2 Schematic illustration of T-DNA processing in *Agrobacterium*. **a** Only part of the plasmid (Ti or binary) is illustrated in the figure. The T-DNA region in the plasmid is marked in red (the lower DNA strand is processed and transferred). VirD1 (marked in black circle) and VirD2 (marked in purple circle) bind the left border (LB) and right border (RB) of the T-DNA region. **b** VirD2 nicks between the third and fourth nucleotide of each border (the 25 base pair DNA sequence of the LB and RB is illustrated; the nicking site is indicated by the scissors). **c** After the single-stranded T-DNA is separated from the parent plasmid, VirD2 protein remains attached to the 5′ end of the T-DNA (the RB of the T-DNA). Adapted from Singer (2013) doctoral dissertation

DNA binding proteins (Citovsky et al. 1988; Das 1988; Abu-Arish et al. 2004), which are secreted into the plant cell from *Agrobacterium* independent of T-DNA (Otten et al. 1984; Citovsky et al. 1992; Binns et al. 1995; Li and Pan 2017). In addition, it is believed that other bacterial and plant proteins interact with ViE2 and VirD2 in the formation of a "T-complex" (for reviews, see Lacroix et al. 2006; Gelvin 2010). The proposed role of the T-complex is to protect the ss T-DNA from degradation (Durrenberger et al. 1989; Tinland et al. 1995; Rossi et al. 1996) and to facilitate its transport into the nucleus. There are a number of facilitators of T-complex nuclear transport. VirD2 facilitates nuclear transport by its nuclear localization signal (NLS) domain (Herrera-Estrella et al. 1990; Shurvinton et al. 1992; Howard et al. 1992; Tinland et al. 1992; Ziemienowicz et al. 2001; van Kregten et al. 2009). Similarly, VirE2 may facilitate T-complex transport via its two NLS domains (Citovsky et al. 1992, 1994; Zupan et al. 1996; Ziemienowicz et al. 2001). However, a number of groups have reported that VirE2 remains mostly in the cytoplasm (Bhattacharjee et al. 2008; Grange et al. 2008; Lee et al. 2008; Sakalis et al. 2014; Shi et al. 2014). Another proposed facilitator of nuclear transport of the T-complex is the VirE2-interacting protein 1 (VIP1), a plant transcription factor that enters the nucleus upon activation of the defense response (Tzfira et al. 2001; Li et al. 2005a; Djamei et al. 2007; Pitzschke et al. 2009; Wang et al. 2017). VirE3, which binds VirE2 in the plant cytoplasm, may substitute for VIP1 to facilitate nuclear transport (Lacroix et al. 2005). Recently, however,

Shi et al. (2014) concluded that VIP1 is not important for *Agrobacterium*-mediated transformation. Finally, nuclear transport of the T-complex may be facilitated by additional host nuclear transporters that interact with T-complex components (Ballas and Citovsky 1997; Lacroix et al. 2005; Bako et al. 2003; Bhattacharjee et al. 2008).

1.2 Stable and Transient T-DNA in the Plant Cell

In the plant nucleus, T-complex proteins must be stripped off the ss T-DNA before integration. It has been hypothesized that stripping off these proteins from the T-complex is mediated by VirF and the host proteasomal degradation machinery (Schrammeijer et al. 2001; Tzfira et al. 2004a, b; Zaltsman et al. 2013). It is known that several T-DNA molecules can enter the plant cell simultaneously (Virts and Gelvin 1985). Whereas the number is unknown and likely varies under different conditions, it has been shown that the percentage of T-DNA molecules in the nucleus that eventually integrate into the plant genome is relatively low (Narasimhulu et al. 1996; Maximova et al. 1998; De Buck et al. 2000; Ghedira et al. 2013). Integration of T-DNA into the plant genome results in stable genetic transformation, whereas T-DNA that does not integrate into the plant genome results in transient genetic transformation.

Stable genetic transformation by *Agrobacterium*-mediated genetic transformation is the preferred method used by plant biologists to generate transgenic plants. Commonly, the desired DNA sequence is cloned between T-DNA borders. Consequently, after T-DNA integrates into the plant genome, a transgenic plant is produced. The *Agrobacterium* strains used for biotechnological applications are themselves genetically modified. This modification includes removal of the natural tumor-inducing genes from T-DNA so that the strains become "disarmed." However, the ability of disarmed strains to transfer a modified T-DNA is unaffected because the only elements on the T-DNA that are necessary for T-DNA transfer are the T-DNA left border (LB) and right border (RB) (Fig. 2) (Hoekema et al. 1983; Ream et al. 1983; Wang et al. 1984). Moreover, in order to make genetic engineering simpler, T-DNA is often placed on a smaller binary plasmid instead of the natural Ti plasmid because the former is easier to work with and can replicate in *E. coli* as well as in *Agrobacterium* (Hoekema et al. 1983; for review, see Tzfira and Citovsky 2006). In addition to introducing desired genes into the plant genome, T-DNA has also been instrumental for the creation of large mutant and enhancer trap libraries because integration occurs randomly in the plant genome. These T-DNA insertion collections have been especially important for studies of *Arabidopsis* and rice (Sessions et al. 2002; Sallaud et al. 2004; O'Malley and Ecker 2010).

Transient expression of foreign genes is another method often used for manipulating plant genomes. Gene expression of those T-DNA molecules lasts for a few days before the genes are silenced (Johansen and Carrington 2001). Manipulating plant genomes can be achieved by transient expression of engineered nucleases

such as meganucleases, ZFNs, TALENs, and Cas9. These nucleases have been used to target specific genomic sites and create double-strand breaks (DSBs) required for gene editing (for review, see Kumar and Jain 2015; Yin et al. 2017). In addition to its use for genomic modifications, *Agrobacterium*-mediated transient expression is an important investigative tool for plant biologists. For example, transient expression is commonly used to investigate cellular localization of proteins or to produce and isolate proteins *in planta* (Sparkes et al. 2006). Recently, transient expression by *Agrobacterium*-mediated transformation has been applied commercially using plants as factories for products such as vaccines and antibodies (for review, see Ko et al. 2009; Komarova et al. 2010).

In nature, T-DNA integration occurs during *Agrobacterium* infection of certain dicotyledonous plants and gymnosperms. However, under laboratory conditions, scientists have harnessed *Agrobacterium* to transform an increasing variety of plants, including monocotyledonous plant species. *Agrobacterium*-mediated transformation has also been successfully applied for transformation of non-plant eukaryotes (for review, see Soltani et al. 2010), such as yeast (Bundock et al. 1995; Bundock and Hooykaas 1996; Rolloos et al. 2014, 2015; Ohmine et al. 2016) and other fungi (de Groot et al. 1998; Korn et al. 2015), as well as for human cells (Kunik et al. 2001). Whereas studying T-DNA integration in non-plant organisms may contribute to understanding T-DNA integration in plants, T-DNA integration in non-plant organisms may involve mechanisms and enzymatic pathways that differ from that of T-DNA integration into plants. Therefore, the following review focuses on T-DNA integration in plants.

2 The Mechanism of T-DNA Integration

The evidence for much of our understanding of T-DNA integration has been facilitated by post-integration sequence analysis of T-DNA/plant genome junctions. This approach has been important for the development of the early models of T-DNA integration because it revealed the general patterns of T-DNA insertions (Mayerhofer et al. 1991; Gheysen et al. 1991). One of the earliest observations was that T-DNA integrates at random locations in the genome (Chyi et al. 1986; Gheysen et al. 1987). This topic is further discussed under the section "*Which is the genomic site prerequisite for T-DNA integration*".

DNA sequencing of the junctions between the integrated T-DNA and the surrounding plant genome also revealed that no homology, or only a few homologous nucleotides (nt) at the junction point, existed (Mayerhofer et al. 1991; Gheysen et al. 1991). Therefore, it was evident that homologous recombination is normally not involved in T-DNA integration, and the terms "illegitimate" recombination (IR) and nonhomologous recombination (NHR) have been used to describe T-DNA integration in plants (Mayerhofer et al. 1991; Gheysen et al. 1991; Bleuyard et al. 2006). More recently, the nonhomologous end-joining (NHEJ) DNA repair pathway of plants is frequently mentioned as the likely pathway responsible for T-DNA

integration. The NHEJ pathway is typically associated with a DNA repair pathway which is responsible for end joining between double-stranded DNA ends such as those present at genomic double-strand breaks (DSBs). Therefore, the NHEJ pathway may not describe well a model involving a single-stranded T-DNA intermediate. "*Does a T-DNA integrate into the plant genome as a single- or a double-stranded intermediate"* is discussed later on. It should be noted that the NHEJ repair pathway is usually associated with key enzymatic components such as the Ku70/Ku80 heterodimer (Critchlow and Jackson 1998). Nevertheless, recent studies have revealed the existence of additional DSBs repair pathways, often described as alternative NHEJ (A-NHEJ) or microhomology-mediated end joining (MMEJ), employing different enzymatic pathways and mechanisms (for review, see McVey and Lee 2008; Bleuyard et al. 2006). Recently, van Kregten et al. (2016) showed that DNA polymerase theta (pol θ) has an important role in T-DNA integration. This topic is further discussed under the section "*What are the bacterial and plant factors involved in T-DNA integration?"*

T-DNA integration is neither a "precise" nor a "clean" process (e.g., Kwok et al. 1985; Spielmann and Simpson 1986). It is not precise because T-DNA seldom preserves its two borders after integration in plants. It is not clean because insertions often include other DNA sequences from *Agrobacterium*. Commonly, the extra DNA sequences are derived from the parent plasmid (Ti or binary; Martineau et al. 1994; Kononov et al. 1997) but may also include DNA from unknown and known sources such as *Agrobacterium* chromosomal DNA (Ulker et al. 2008; Kleinboelting et al. 2015) and plant DNA (Kleinboelting et al. 2015). In addition, it is common for insertion sites to include two or more T-DNA molecules adjacent to each other (Cluster et al. 1996; Krizkova and Hrouda 1998; De Buck et al. 2000). The integration patterns differ under different experimental conditions and plant species (Grevelding et al. 1993; De Buck et al. 2009). In addition, transformed plants may contain more than a single T-DNA insertion site; each insertion may contain a single copy of T-DNA, or a cluster of T-DNA copies (e.g., Alonso et al. 2003; Rosso et al. 2003). Finally, major chromosomal aberrations may result from T-DNA integration (Nacry et al. 1998; Tax and Vernon 2001; Clark and Krysan 2010). The topic of "*Why and how do complex T-DNA insertions form*?" is integral to the question of T-DNA integration.

Finally, T-DNA integration models that include depictions of T-DNA vary in different publications. Often T-DNA is depicted as a straight line. In addition, the timing at which T-DNA ends interact with the target genome varies between the different models (for recent reviews of T-DNA integration, see Ziemienowicz et al. 2010; Windels et al. 2010; Magori and Citovsky 2011; Gelvin 2017). This topic is further discussed under the section "*What is the spatial and temporal arrangement of T-DNA during integration?*"

The following review attempts to explain the bigger question of T-DNA integration by presenting smaller questions. This has been done for the purpose of discussion. However, it should be emphasized that T-DNA integration may be mediated by different pathways under different conditions, and perhaps simultaneously. It is most likely a complex and multistep process. As such, the different

questions related to T-DNA integration are interrelated and may have more than a single answer. Correspondingly, whereas several major models have been proposed (reviewed in Tzfira et al. 2004a, b), when adopting different assumptions for different open questions, more models are possible.

3 The Major Unresolved Questions Related to the Mechanism of T-DNA Integration

3.1 Does T-DNA Integrate into the Plant Genome as a Single- or a Double-Stranded Intermediate?

T-DNA enters the plant nucleus as a single-stranded (ss) DNA but it is ultimately a double-stranded (ds) DNA when it becomes part of the host genome. However, without being able to visualize the integration process as it occurs, it is difficult to determine the timing of conversion from ss T-DNA to ds T-DNA. Mayerhofer et al. (1991) and Gheysen et al. (1991) discussed this question when proposing models for T-DNA integration via a mechanism of illegitimate recombination. According to the proposed ds T-DNA integration model, conversion from ss to ds T-DNA occurs extrachromosomally. Therefore, when T-DNA begins integration into the plant's genome, it is already a ds T-DNA intermediate (Mayerhofer et al. 1991). On the other hand, according to the proposed ss T-DNA integration model, the integration process begins with an ss T-DNA intermediate and the conversion to ds T-DNA happens during integration (Mayerhofer et al. 1991; Gheysen et al. 1991).

The ss T-DNA integration model was refined by Tinland et al. (1995) and became widely accepted soon thereafter (Tinland et al. 1995; Tinland 1996). According to this model, integration begins when the LB side of the ss T-DNA (the 3′ end) anneals to homologous sequences in the plant DNA, possibly by invading A-T-rich regions of melted chromosomal DNA (Brunaud et al. 2002). This annealing through homology may not include parts of the sequences at the 3′ distal end of the LB side, resulting in the loss of some of the 3′ side of the T-DNA due to exonuclease degradation. Next, the RB side of the ss T-DNA (the 5′ end) ligates to the 3′ end of the plant DNA. Unlike the 3′ end of the LB side, the RB is protected from exonuclease degradation by VirD2. VirD2 may also be involved in ligation of the 5′ ss T-DNA end to a 3′ end of the plant DNA. Several observations from different early studies support the ss T-DNA integration model: (a) T-DNA enters the nucleus as ss DNA molecule. Moreover, extrachromosomal recombination assays suggested that the T-DNA derivatives inside the plant nucleus are mainly ss T-DNA molecules (Tinland et al. 1994; Yusibov et al. 1994); (b) When ss DNA was introduced into plant protoplasts, the integration rate was comparable to (Furner et al. 1989) or higher than (Rodenburg et al. 1989) that of ds DNA; (c) The deletions of T-DNA post-integration are usually more severe at the LB side in comparison to the RB side (e.g., Tinland 1996; Kumar and Fladung 2002; Kim

et al. 2003; Zhang et al. 2008); d) The junctions between the T-DNA LB side and plant DNA after integration have been shown to contain higher microhomology levels compared to junctions involving the RB side of the T-DNA (e.g., Matsumoto et al. 1990; Tinland et al. 1995; Brunaud et al. 2002; Kim et al. 2003; Zhu et al. 2006; Thomas and Jones 2007); (e) There is evidence, although inconclusive, that VirD2 is involved in T-DNA integration (Pansegrau et al. 1993; Scheiffele et al. 1995; Mysore et al. 1998).

It should be noted that the ss T-DNA integration model could, in principle, apply to a T-DNA with ss DNA overhangs and a ds DNA internal body (Gheysen et al. 1991), although usually an ss T-DNA intermediate has been assumed (Gheysen et al. 1991; Tinland 1996; Brunaud et al. 2002; Meza et al. 2002). Moreover, the role of VirD2 in integration is inconclusive, and therefore the possible T-DNA intermediates depicted in Fig. 3 are shown with and without VirD2 attached to the 5′ end of the RB (Fig. 3).

The T-DNA integration model involving a ds intermediate is supported by evidence that T-DNA integration is linked to the repair of genomic DNA double-strand breaks (DSBs). The first evidence came from the observation that, when genomic DSBs are induced in protoplasts by X-ray irradiation, integration of foreign plasmid DNA is in enhanced (Kohler et al. 1989). Salomon and Puchta (1998) showed that when genomic site-specific DSBs are induced by *Agrobacterium*-mediated transient expression of the homing endonuclease I-*Sce*I, DSBs are often repaired with a T-DNA captured within the repaired break.

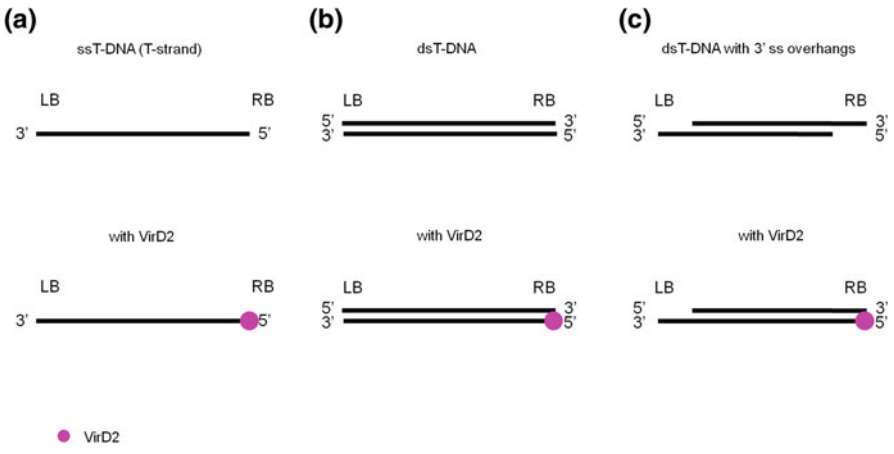


Fig. 3 Possible configurations of the T-DNA integration intermediate. **a** Single-stranded (ss) T-DNA (also termed T-strand). The 5′ end is always the RB side, while the 3′ end is always the LB side. Illustration underneath demonstrates VirD2 (in purple) attached to the 5′ end. **b** Double-stranded (ds) T-DNA with blunt ends. Illustration underneath demonstrates VirD2 (in purple) attached to the 5′ end of the RB side. **c** Double-stranded (ds) T-DNA internal body with 3′ single-stranded overhangs. Illustration underneath demonstrates VirD2 (in purple) attached to the 5′ end of the RB side and protecting the end from possible resection. Adapted from Singer (2013) doctoral dissertation

Moreover, early studies involving the sequencing of the junctions between T-DNA and plant DNA (e.g., Gheysen et al. 1991; Mayerhofer et al. 1991; Ohba et al. 1995; Takano et al. 1997) revealed patterns similar to those found in later studies of the mechanisms of DNA DSBs repair in plants (Gorbunova and Levy 1997; Salomon and Puchta 1998). The patterns of DSB repair in plants exhibited the characteristics of illegitimate/nonhomologous recombination. These included DNA deletions close to the breaks, repeated sequences or DNA from an unknown source ("filler" DNA), and little or no homology between DNA sequences forming the junctions. The notion that T-DNA integrates at genomic DSBs favors a model of a ds T-DNA as an intermediate because repair of DSBs involves end joining between two ds DNA ends. Moreover, Tzfira et al. (2003) and Chilton and Que (2003) provided evidence that T-DNAs captured at genomic DSBs were already ds intermediates prior to integration.

Evidence supporting the ds T-DNA model is also derived from the common formation of complex T-DNA insertions, in particular, complex insertions that include two T-DNAs ligated at their LB–LB sides or RB–RB sides without any microhomology within the ligated junction. The reason that this arrangement is difficult to explain via an ss model is that direct LB–LB end joining ("tail-to-tail" ligation) or RB–RB end joining ("head-to-head" ligation) cannot occur directly between the transferred ss T-DNA because there are always 3′ ends at the LB side and 5′ ends at the RB side (Gheysen et al. 1991; Mayerhofer et al. 1991; De Neve et al. 1997). In addition, extrachromosomal double-stranded circular T-DNA structures (T-circles) from *Agrobacterium*-infected plants have been isolated (Singer et al. 2012). By analyzing the DNA sequences of the extrachromosomal structures, it was found that the DNA junctions within the structures show the characteristic patterns of repaired DSBs. Importantly, it was possible to study the complete structure of the molecules (a feat more difficult to achieve in a genomic background when complex DNA repeats are involved). The structures included configurations such as multiple T-DNA copies arranged adjacent to each other or binary vector fragments attached to T-DNA sequences. Such structures are common post-T-DNA integration in transgenic plants. For example, according to different reports (e.g., Castle et al. 1993; Rios et al. 2002; De Buck et al. 2009), the integration of T-DNAs in clusters of two or more copies can account for about 50% of the integration events, and about 30–70% of events include sequences from the T-DNA parent binary plasmid (e.g., Martineau et al. 1994; Kononov et al. 1997; Nicolia et al. 2017). Therefore, these same structures were captured as ds DNA molecules before integration supports a notion that most T-DNA molecules integrate as ds T-DNA intermediates.

If T-DNA integrates as a ds T-DNA intermediate, then an important question is what mechanism accounts for the synthesis of the complementary strand. Liang and Tzfira (2013) showed that oligonucleotides can efficiently interact with the ss T-DNA and convert ss T-DNA to ds T-DNA molecules. Whereas the mechanism of this conversion is still unknown, it is has been shown that introduction of ss DNA into protoplasts using either electroporation or polyethylene glycol resulted in rapid synthesis of the complementary strand (Rodenburg et al. 1989; Furner et al. 1989). Therefore, this process can be mediated entirely by the plant DNA repair

machinery. There is also evidence for the existence of extrachromosomal ds T-DNA molecules after *Agrobacterium* infection. The first piece of evidence is the rapid and broad transient expression of T-DNA genes in infected leaves (Janssen and Gardner 1990), no matter if the transferred ss T-DNA is the coding or non-coding strand (Narasimhulu et al. 1996). In addition, experiments involving homologous recombination between extrachromosomal T-DNA constructs delivered as noncomplementary strands suggested that at least one of the ss T-DNA constructs must have been converted to ds T-DNA prior to recombination (Offringa et al. 1990). Recently, Dafny-Yelin et al. (2015) showed that blocking ss T-DNA to ds T-DNA conversion reduced T-DNA gene expression. Therefore, although there is no question that extrachromosomal ds T-DNAs exist in plants immediately after *Agrobacterium* infection, the question remains whether they are the only, or the predominant, intermediates in the integration process.

Identification of plant components that are important for T-DNA integration can provide more clues regarding the form, or the predominant form, of the T-DNA intermediate during integration. For example, evidence that Ku70/80 heterodimer is important for integration may support the ds T-DNA theory because Ku70/80 is involved in nonhomologous end joining (NHEJ) between ds DNA ends (Critchlow and Jackson 1998). However, studies to identify the plant components important for T-DNA integration are still ongoing (discussed below).

3.2 What Are the Bacterial and Plant Factors Involved in T-DNA Integration?

The major approaches to identifying the proteins involved in the process of *Agrobacterium*-mediated transformation are forward and reverse genetics. These approaches have led to the identification of bacterial and plant factors involved in the transformation process and its last step of T-DNA integration. Experimental assays can distinguish between a block in T-DNA transfer, a step prior to actual integration, from a block in T-DNA integration. The principle allowing this distinction is that mutants blocked in stable T-DNA integration but not T-DNA transfer will be able to transiently express genes in plant cells, but not generate stable transgenic plants or plant calli. Therefore, many of the proteins involved in T-DNA integration, including the specific protein domains important for this process, have been identified by this principle coupled with protein localization and protein–protein interaction studies. In addition, large-scale screens in *Arabidopsis* have been conducted to identify host proteins involved in *Agrobacterium* transformation (Zhu et al. 2003; Anand et al. 2007a; Gelvin 2010). When considering the commonly used experimental methods to identify genes involved in stable T-DNA integration, a possible scenario should be noted. If T-DNA randomly integrates into and mutates a gene leading to increased gene silencing, fewer stable transgenic event would be recovered through selection if the selection gene has

been silenced. Thus, a mutant may show reduced stable transformation, whereas stable T-DNA integration occurs at same or even higher rate as wild-type plants. This was demonstrated by Park et al. (2015) who analyzed T-DNA integration biochemically. Therefore, T-DNA integration does not necessarily equate to stable transformation.

Several lines of evidence suggest that plant factors mostly, if not entirely, mediate the process of T-DNA integration. First, there are not many *Agrobacterium* candidate proteins that can be involved in the process because T-DNA itself does not encode proteins that are required for T-DNA integration and only a few Vir proteins are known to be transported into the plant nucleus. Second, DNA sequencing of T-DNA/plant DNA junctions suggests that integration occurs through the same pathways responsible for DNA end-joining repair by the host plant cell (i.e., illegitimate/nonhomologous recombination). Another support for the notion that the host cell is responsible for T-DNA integration comes from results of *Agrobacterium*-mediated transformation of yeast (Bundock et al. 1995). In yeast, T-DNA can integrate via homologous recombination, a major pathway of DSBs repair that is used by this organism to repair DSBs, if sufficient homology between T-DNA and yeast sequences exists. Third, foreign DNA can be introduced into plant cells by other methods that do not include *Agrobacterium*, such as electroporation, polyethylene glycol, and particle bombardment transformation. By these methods, the introduced DNA integrates through illegitimate/nonhomologous recombination into the genome, demonstrating that the plant's own DNA repair machinery can potentially accomplish the task of T-DNA integration without the assistance of foreign genes (for review, see Somers and Makarevitch 2004). As ongoing studies are improving our understanding of the mechanisms and pathways behind DNA DSB repair in plants, we anticipate a better understanding of how plant factors facilitate T-DNA integration.

DNA end joining during DSBs repair is less understood in plants in comparison to yeast or mammalian cells. However, it is known that the major pathway of DNA DSBs repair in plants is the nonhomologous end joining (NHEJ) pathway, which is the major pathway for DSB repair in higher eukaryotes. The NHEJ pathway includes the key heterodimer Ku70/Ku80 that binds double-stranded DNA ends formed by the DSBs (Critchlow and Jackson 1998). Several studies investigated the role of Ku80, both in repairing DNA DSBs in plants and in T-DNA integration. Friesner and Britt (2003) reported that Ku80-deficient plants are more sensitive to DSB inducing gamma radiation and are reduced in T-DNA integration rates. The results of Friesner and Britt (2003) supported the involvement of the NHEJ repair pathway in T-DNA integration. Li et al. (2005b) further demonstrated that Ku80 is important for T-DNA integration. First, overexpression of Ku80 in plants enhanced T-DNA integration, whereas Ku80-deficient plants were deficient in T-DNA integration. In addition, Ku80 interacted with ds T-DNA *in planta*, as demonstrated by immunoprecipitation experiments. Both of these studies were done in *Arabidopsis* plants, and results by Jia et al. (2012) and Mestiri et al. (2014) also support the notion that Ku proteins are involved stable T-DNA integration in *Arabidopsis*. Moreover, in rice, knockdown of the Ku70/80 heterodimer also confirmed reduced

stable transformation rates (Nishizawa-Yokoi et al. 2012). On the other hand, contradictory results have been presented by other research groups. Gallego et al. (2003) found that whereas Ku80 has a role in NHEJ in *Arabidopsis* plants, a Ku80-deficient plant was not deficient in T-DNA integration. Park et al. (2015) examined a set of the NHEJ mutant genes in *Arabidopsis*, including *Ku80* and *Ku70*, and determined that deficiency in NHEJ proteins increased the rate of T-DNA integration. According to the authors of that study, the contradictory results can be explained by increased random DNA DSBs in the plant genome that results from deficiency in NHEJ proteins. This results in T-DNA having more available target sites for integration. Therefore, T-DNA integration rate could be affected either way from a deficiency in NHEJ factors: whereas it may be enhanced from increased availability of genomic DSBs as a result of deficiency in NHEJ factors, the integration rate may also be reduced because of reduced ligation ability of the T-DNA into DSBs.

Similar conflicting results have been obtained for Ligase IV, another key component of the NHEJ pathway. Whereas the importance of Ligase IV for NHEJ DNA repair has been demonstrated in plants (Friesner and Britt 2003; van Attikum et al. 2003), Ligase IV has been shown to be both dispensable (van Attikum et al. 2003; Park et al. 2015) and involved but nonessential (Friesner and Britt 2003; Nishizawa-Yokoi et al. 2012) for T-DNA integration. Other components of the NHEJ pathway have been shown to be either required (Jia et al. 2012) or dispensable (Park et al. 2015; Vaghchhipawala et al. 2012; Mestiri et at. 2014) for T-DNA integration (reviewed in Saika et al. 2014).

Alternative NHEJ pathways, such as the microhomology-mediated end joining (MMEJ), have received increasing attention in recent years (reviewed in Wang and Xu 2017). Mestiri et al. (2014) showed that mutations in several alternative NHEJ pathway genes reduced T-DNA integration. In addition, a quadruple *Arabidopsis* mutant disabling several end-joining pathways, including NHEJ, was severely compromised in *Agrobacterium*-mediated transformation. However, that T-DNA integration still occurred suggests additional pathways. On the other hand, Park et al. (2015) found that T-DNA integration was not reduced, but increased, in a *parp1* mutant. Finally, disabling another MMEJ component, polymerase theta (pol θ), completely eliminated T-DNA integration according to van Kregten et al. (2016). However, Gelvin's group found that these same mutants can still be transformed at $\sim$20% of wild-type levels (personal communication). Furthermore, a recent analysis of filler DNA at T-DNA junctions provided more support for an MMEJ mechanism acting at the LB end of the T-DNA (van Kregten et al. 2016).

The discrepancies of some of the results of these studies may be the result of different experimental conditions when measuring transient and stable T-DNA transformation, as well as conflating stable transformation with T-DNA integration. More interestingly, these discrepancies may point to other alternative pathways that are active under different conditions, such as tissue type and developmental stage.

Other plant proteins that have been identified as being involved or as affecting T-DNA integration include proteins that are involved in the chromatin structure or proteins that direct the T-DNA to the chromatin (for review, see Magori and

Citovsky 2011). In particular, evidence suggests that histones play an important role in T-DNA integration. Several studies demonstrated that plants deficient in different histones were reduced in T-DNA integration rate, whereas overexpression of histones resulted in increased stable transformation and T-DNA integration (Mysore et al. 2000; Yi et al. 2002, 2006; Anand et al. 2007b; Iwakawa et al. 2017). In addition, a domain in VIP1 has been shown to be important for the interaction with histone proteins and for stable transformation (Li et al. 2005a). Finally, VIP2 a transcriptional regulator influencing histone mRNA levels has been shown to be important for T-DNA integration (Anand et al. 2007b).

The bacterial factors that are potential candidates to be involved in T-DNA integration are limited to those that are secreted into the plant cells. They include VirE2, VirE3, VirF, VirD5, and VirD2 (Vergunst et al. 2000; Schrammeijer et al. 2003; Vergunst et al. 2005). Indirect evidence for involvement of Vir proteins in T-DNA integration can be derived from the patterns of T-DNA integration. If comparing T-DNA integration to integration of foreign DNA delivered by other non-*Agrobacterium* methods, T-DNA integration is usually much more efficient. Also, although T-DNA integration can result in complex insertions, these are usually considered more "simple" and precise compared to insertions produced via other methods (Hu et al. 2003; Makarevitch et al. 2003; Travella et al. 2005). These observations may suggest that T-DNA integration, in contrast to DNA delivered by other methods, uses another bacterial factor or factors in addition to the host DNA repair machinery to facilitate integration. However, it is also possible that this may be merely a result of more efficient nuclear localization due to VirD2 piloting the T-DNA, or the protection of T-DNA from degradation through VirE2 coating.

The bacterial candidate important for T-DNA integration that has been studied most extensively is VirD2, because it is transferred into the nucleus while attached to the 5′ end of the single-stranded T-DNA (T-strand). The earliest support for VirD2 involvement in integration was provided in an *in vitro* assay showing that VirD2 has an ability to rejoin ends from the cutting reaction (Pansegrau et al. 1993). Therefore, it has been suggested that the 5′ end of a T-DNA is ligated to the plant DNA via VirD2. Potentially supporting the notion that VirD2 has a ligase-like activity *in planta*, Tinland et al. (1995) reported a VirD2 mutant (R129G) that resulted in reduced precision of the RB side after T-DNA integration. However, this mutation did not reduce the efficiency of T-DNA integration, suggesting that the loss of precision may be only due to VirD2's role in protecting the 5′ end. Moreover, a different in vitro study rejected a general ligation activity of VirD2 (Ziemienowicz et al. 2000).

A better understanding of the potential role of VirD2 in T-DNA integration required investigating the different VirD2 domains. Whereas the N-terminal region of VirD2 contains a relaxase domain that is important for border nicking in *Agrobacterium* (Ward and Barnes 1988), the C-terminal domain contains three regions: a DUF domain, a bipartite NLS, and an omega (Ω) domain. The role of the DUF domain has been shown to be delivery of the ss T-DNA through the T4S system (van Kregten et al. 2009), whereas the role of the bipartite NLS domain is in nuclear transport (Howard et al. 1992; Shurvinton et al. 1992; Tinland et al. 1992,

1995; Rossi et al. 1993; Bravo-Angel et al. 1998; Mysore et al. 1998; van Kregten et al. 2009). The Ω domain has been shown to be important for tumorigenesis (Shurvinton et al. 1992; Bravo-Angel et al. 1998); however, its involvement in T-DNA integration is undetermined. Several reports have shown that a deletion or substitution mutation at the Ω domain reduced T-DNA integration to about 1–4% of the wild-type T-DNA rate (Shurvinton et al. 1992; Narasimhulu et al. 1996; Mysore et al. 1998), whereas the T-DNA transfer rate is reduced to only 20–30% of the wild-type rate (Narasimhulu et al. 1996; Bravo-Angel et al. 1998; Mysore et al. 1998). However, Bravo-Angel et al. (1998) and van Kregten et al. (2009) concluded that the Ω domain has no role in integration. Moreover, inducible expression of VirD2 in plants reduced the transformation efficiency (Hwang et al. 2006). Therefore, it is still controversial if VirD2 or any of the other bacterial Vir proteins have a direct role in T-DNA integration.

Recently, Zhang et al. (2017) showed that in yeast, VirD5 localizes to the centromeres/kinetochores in the nucleus and causes chromosome instability. The authors also showed that VirD5 inhibited cell growth in yeast and also in plants. Therefore, whether VirD5 is involved somehow in T-DNA integration is an interesting question.

3.3 What Is the Genomic Site Prerequisite for T-DNA Integration?

Large-scale analysis of T-DNA insertions has shown that insertions are distributed randomly among the plant chromosomes (e.g., Alonso et al. 2003; Sallaud et al. 2004). At the chromosome and gene level, there may be a distribution bias, although this is controversial. It has been suggested that T-DNA integrates preferably at genomic regions that are actively transcribed because T-DNA insertions are generally found more frequently at 5′ and 3′ regions of genes, but less frequent at regions closer to the centromeres and telomeres (Brunaud et al. 2002; Szabados et al. 2002; Alonso et al. 2003; Chen et al. 2003; An et al. 2003; Sallaud et al. 2004; Li et al. 2006; Zhang et al. 2007). A plausible explanation is that during transcription genomic DNA is more "open" and therefore more accessible to incoming T-DNA molecules. Indeed, it has been shown that T-DNA integration sites are preferably found in A-T-rich regions that have a relatively lower DNA duplex stability (e.g., Brunaud et al. 2002; Chen et al. 2003). In addition, a component of the T-complex may interact with host factors, such as a TATA-binding protein, that are involved in gene transcription (Bako et al. 2003). This way, they can guide the T-DNA to actively transcribed regions. Active regions may also be more prone to DNA damage, such as DSBs, and this may create "hot spots" for DNA repair factors and T-DNA integration.

On the other hand, results from previous large-scale studies may have been biased by the experimental method that relied on marker-based selection and

regeneration of plants. If not selected, T-DNA integration events may be excluded from a studied collection. In most studies, the analysis is based on selection via the T-DNA's own marker gene, such as antibiotic or herbicide resistance. However, if T-DNA integrates but the marker gene is not expressed, plants will not survive selection and therefore will not be included in the studied collection. In this regard, Francis and Spiker (2005) showed that in about 30% of transformed plants T-DNA genes are not transcribed. Furthermore, it has been shown that when not applying selection to detect T-DNA insertions, they are distributed randomly and are equally represented in centromeric and telomeric regions (Francis and Spiker 2005; Kim et al. 2007; Shilo et al. 2017). In addition, T-DNA integration can be mutagenic and therefore can disrupt genes that are essential for selection and recovery of plants. However, this likely occurs in a relatively small number of cases.

Different events have been proposed to stimulate the integration of T-DNA into specific sites in the plant DNA. The events include single-strand DNA nicks in the plant DNA, a relaxed duplex DNA forces that allow "invasion" of a T-DNA to the plant DNA, and genomic double-strand breaks (DSBs). Early models by Gheysen et al. (1991) and Mayerhofer et al. (1991) suggested that a nick in the plant DNA is first generated (Fig. 4a). This nick is later converted, via 5′ to 3′ exonuclease activity, into a gap (the "single-strand gap-repair" model). The LB and RB sides of a single-stranded T-DNA can anneal to the plant DNA at this gap through microhomologies and initiate integration. Revision of this model postulated that instead of annealing to DNA within a gap, the LB side invades and anneals to regions of microhomology at the plant DNA. This may happen more often at A-T-rich regions due to lower duplex stability (Tinland et al. 1995; Brunaud et al. 2002) (Fig. 4b). Recently, a link between T-DNA integration and genomic DSBs has become increasingly accepted (for review, see Magori and Citovsky 2011). It has been suggested that genomic DSBs are the prerequisite for T-DNA integration (Fig. 4c). The breaks may be spontaneous and may occur randomly in the genome under natural conditions. Extrachromosomal T-DNA molecules may be directed to DSBs, likely guided by host DNA repair proteins, and possibly also *Agrobacterium* proteins of the T-complex. Direct support for this model is that T-DNA can be directed to integrate into artificiality induced genomic DSBs (Salomon and Puchta 1998; Tzfira et al. 2003; Chilton and Que 2003; Zhang et al. 2018). Muller et al. (2007) reported that T-DNA insertions are found more frequently near or at palindromic sequences in the plant genome. This observation supports integration of T-DNA into genomic DSBs because palindromic regions are often found at sites of DSBs repair in plants (Muller et al. 1999) and, therefore, may be more susceptible to breaks due to their secondary structure. In addition, induction of genomic DSBs by irradiation increases integration of foreign DNA into plant genome (Kohler et al. 1989). Another possibility is that under natural conditions *Agrobacterium* can induce DSBs in order to facilitate T-DNA integration, as microbial pathogens have been shown to trigger host DNA DSBs in plants (Song and Bent, 2014). However, currently there is no evidence for such an activity induced by any of the *Agrobacterium* virulence factors.

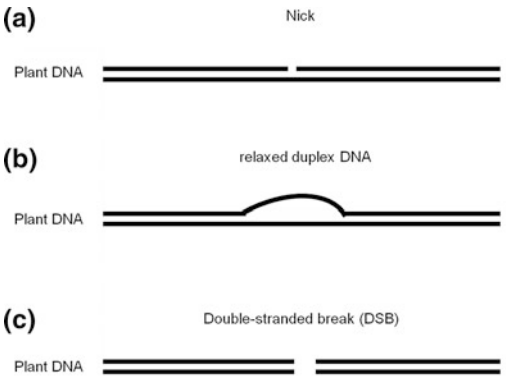


Fig. 4 Possible genomic pre-conditions for T-DNA integration. **a** Nick (later expanded into a gab). **b** Relaxed duplex DNA region. **c** Double-strand break (DSB). Adapted from Singer (2013) doctoral dissertation

3.4 What Is the Spatial/Temporal Arrangement of T-DNA During Integration?

The T-DNA region of natural *Agrobacterium* strains is 5-25 kbp in length (Barker et al. 1983; Suzuki et al. 2000). The size range of engineered T-DNA constructs used in laboratory strains may be similar, but also T-DNA constructs of up to 150 kbp can be successfully transferred and integrated into the plant genome (Hamilton et al. 1996). On the other hand, T-DNA can integrate into the plant genome with remarkably minimal damage to the plant DNA. For example, Meza et al. (2002), Windels et al. (2003), and Kleinboelting et al. (2015) reported that in the T-DNA collections they chose to analyze, most sequenced integration events (sites in which both LB and RB T-DNA junctions with the genomic DNA had been sequenced) showed a deletion of 100 bp or less of plant genomic DNA bordering the integration site. This result raises the question how T-DNA is spatially arranged during the process of integration, considering that T-DNA is a large molecule in comparison to the small integration site (Fig. 5a). The current popular models do not provide an explanation for this question. The temporal mode of T-DNA integration, on the other hand, has been discussed in early models for T-DNA integration. The model proposed by Tinland (1996) suggests that the LB side of T-DNA interacts with the plant DNA first, following which the RB side attaches to plant DNA (Fig. 5b). That the LB is the initiator of integration is based on the observations that T-DNA insertions share higher degree of microhomology at the T-DNA LB side with the plant DNA junctions, in comparison to the RB side (Tinland et al. 1995; Brunaud et al. 2002; Kim et al. 2003; Zhu et al. 2006; Thomas and Jones 2007). However, several studies reported similar frequencies of microhomologies at both ends (Meza et al. 2002; Windels et al. 2003; Forsbach et al. 2003; Kleinboelting et al. 2015). Therefore, Meza et al. (2002) proposed that in some cases the RB side is the first side to initiate integration into the plant DNA (Fig. 5c).

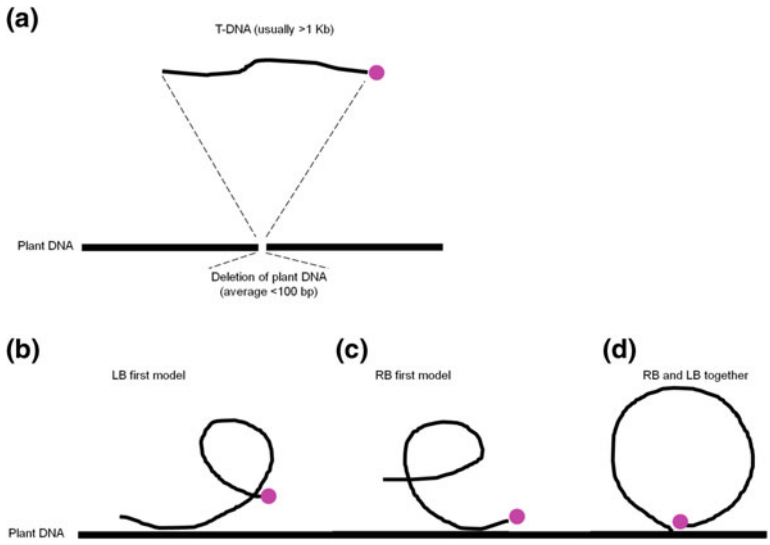


Fig. 5 Possible spatial arrangements of T-DNA during integration. **a** T-DNA, at a size of a few kbp DNA, integrates in most cases without causing major deletions at the target genomic site. **b** LB first model suggests that the LB anneals first via microhomology. **c** RB first model suggests that the RB anneals first via microhomology. **d** LB and RB are in close proximity during integration. Adapted from Singer (2013) doctoral dissertation

Interestingly, the Muller et al. (2007) analysis of T-DNA/plant junctions revealed that T-DNA integration also involves microhomologies in inverted orientation. Based on this finding, these authors proposed that the LB and RB sides of a T-DNA strand anneal to plant DNA simultaneously via microhomologies and that T-DNA ends are in close proximity during integration. However, the model of Muller et al. (2007) does not explain how and when the two ends of a T-DNA are brought into close proximity. The discovery of T-DNA circles (T-circles) provides a possible explanation (Singer et al. 2012), as circular double-stranded T-DNA in plants contains LB and RB sequences ligated extrachromosomally. These results suggest that the LB and RB ends are recognized by plant DNA repair factors and that these factors pull the two ends of a T-DNA toward each other before integration (Fig. 5c). Therefore, it is possible that double-stranded T-DNA approaches the plant genome with the LB and RB sides already in close proximity (Figs. 5d and 6).

3.5 Why and How Do Complex T-DNA Insertions Form?

As mentioned above, T-DNA integration is not a "precise" or "clean" process. Early studies indicated that T-DNA integration can often result in complex T-DNA insertions (Ooms et al. 1982; Kwok et al. 1985; Spielmann and Simpson 1986;

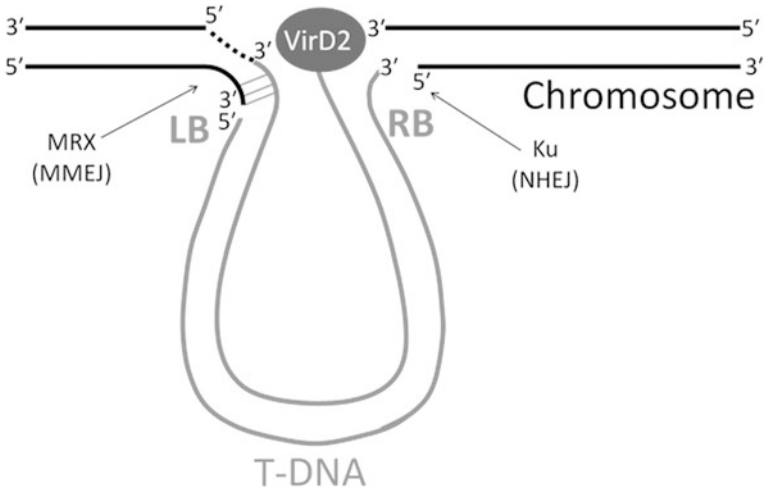


Fig. 6 A proposed model for T-DNA integration. Schematic illustration of double-stranded T-DNA (gray lines) and double-stranded plant DNA (black lines) during T-DNA integration into the plant genome. T-DNA is arranged in a looped mode in which the left border (LB) end and the right border (RB) end are in close proximity. The LB end has a 3′ single-stranded overhang that aligns via short microhomologies to the plant 3′ overhang through the Mre11, Rad50, and Xrs2 (MRX) complex of the MMEJ repair pathway. The dashed line represents a region of template-dependent DNA synthesis of the complementary strand. The RB end, with VirD2 covalently attached to the 5′ end, aligns to the plant DNA through the Ku70/80 mediated NHEJ DNA repair pathway. Illustration adapted from Singer (2013) Doctoral dissertation

Gheysen et al. 1987; Grevelding et al. 1993; Ohba et al. 1995). A complex T-DNA insertion can include, in addition to T-DNA, DNA sequences from various sources. The DNA sequences can be derived from the *Agrobacterium* binary or Ti plasmid (Martineau et al. 1994; Kononov et al. 1997) or even from bacterial chromosomal DNA (Ulker et al. 2008; Kleinboelting et al. 2015). In addition, plant DNA at the site of integration may be re-arranged and include duplications of plant DNA sequences that were not part of the original pre-integration genomic site (Gheysen et al. 1987; Takano et al. 1997; Windels et al. 2003; Kleinboelting et al. 2015). Also, several copies of T-DNA, or parts of the T-DNA sequence, are often clustered together at the integration site (Jorgensen et al. 1987; De Neve et al. 1997). In some cases, the additional DNA sequence that is found at the insertion site does not have any homology to a known DNA sequence. This kind of DNA is termed "filler" DNA, a term that is also used to describe additional DNA at DSB repair sites. The term "filler" is also used to describe additional DNA sequences that share homology with known DNA, such as DNA that is homologous to plant or *Agrobacterium* DNA (Gheysen et al. 1987; Gorbunova and Levy 1997; Windels et al. 2003; Kleinboelting et al. 2015).

The formation or appearance of the different DNA sequences that accompany complex T-DNA insertions can be explained in several ways. Therefore, each

complex insertion can be explained by a different mechanism or by a combination of mechanisms. Nevertheless, it is possible to distinguish between two general types of potential sources for the DNA that is found in complex T-DNA insertions. The first type includes DNA fragments that are present in the nucleus at the time of integration, which may also be described as "free-floating" DNA fragments. The free DNA fragments can ligate with T-DNA prior to or during integration and form complex insertions. The second type includes DNA that is synthesized during the process of DNA repair in the plant nucleus. During DNA repair and ligation, synthesis of DNA can occur using random DNA sequences as templates. This process is also known as synthesis-dependent strand annealing (SDSA). It involves a single-stranded DNA strand invading random DNA sequence *in cis* (DNA from the same molecule) or *in trans* (DNA from a different molecule), using it as a template, and often switching between different templates. Filler DNA also characterizes ds DNA end joining in higher eukaryotes (Gorbunova and Levy 1997; Salomon and Puchta 1998). It is difficult to determine whether a specific DNA sequence in a complex insertion is the result of ligation between free existing DNA fragments or the result of DNA synthesis. However, as discussed below, in many cases it is possible to surmise the origin of the DNA sequence from the sequence identity, length, and overall arrangement in the complex structure.

Ligation between free extrachromosomal DNA fragments is likely when the DNA sequence can be traced to *Agrobacterium* chromosomal DNA, pTi, or binary plasmid DNA sequences. In many instances, T-DNA is transferred together with the backbone of the parent plasmid, termed a "read- through" transfer, due to incorrect processing of the T-DNA borders in the *Agrobacterium* (Kononov et al. 1997; Wenck et al. 1997). However, non-read through *Agrobacterium* DNA often found in complex T-DNA insertions may be transferred from *Agrobacterium* independently and ligated to T-DNA molecules in plants before integration, or alternatively, transferred from *Agrobacterium* already linked to T-DNA. Clusters of two or more T-DNAs probably result from T-DNA molecules that were transferred independently, ligated into the plant nucleus, and then integrated. De Neve et al. (1997) provided compelling evidence supporting this notion by transforming plants simultaneously with different *Agrobacterium* strains that contained different T-DNA constructs. The authors showed that the two types of T-DNAs can integrate adjacent to each other. Similarly, Singer et al. (2012) isolated extrachromosomal T-DNA structures composed of T-DNA originating from two different *Agrobacterium* strains.

Synthesis-dependent strand annealing (SDSA) is likely the mechanism that accounts for other regions of DNA at the junctions between end-joined DNA fragments. This mechanism can sometimes generate a patchwork of short sequences resulting from consecutive template switches (Gorbunova and Levy 1997; Salomon and Puchta 1998; van Kregten et al. 2016). These sequences can be identical to those of T-DNA or plant DNA; therefore, it is a matter of debate whether a specific DNA fragment is a broken fragment of molecule patched together with another DNA or a new synthesis product. The recent discovery that DNA polymerase θ (pol θ) is involved in T-DNA integration (van Kregten et al. 2016) supports the latter, as pol θ is associated with microhomology annealing and low-fidelity DNA synthesis

(Wang and Xu 2017). The shorter the sequence and the more it is "scrambled", the more likely it is a synthesis product.

T-DNA insertions, where T-DNA copies are arranged adjacent to each other in clusters, may also result from T-DNA replication after transfer. In that case, the replicated T-DNA copy integrates adjacent to its template, as proposed by Van Lijsebettens et al. (1986) and Jorgensen et al. (1987) based on analyzing structures of integration events that include adjacent T-DNAs. In some cases, a pair of adjacent T-DNAs shared the same truncation point at their ends. Therefore, a truncated T-DNA replicated to produce another identical copy with the same truncation. T-DNA replication has been supported by statistical analysis of co-transformation and integration of different T-DNAs at the same locus (De Buck et al. 2009).

Truncations of T-DNA ends, especially at the LB side, are another common pattern of T-DNA integration. There may be several reasons for T-DNA insertion having more severe truncations at the LB side. First, T-DNA is transferred from its RB side piloted by VirD2; therefore, the LB side may be more prone to incorrect processing or breaks during the transfer process. VirD2 attached to the 5′ end of the T-DNA may protect the RB side from exonuclease activity, whereas the LB is exposed to such activity. Second, in the plant nucleus, some of the LB side of T-DNA may be lost during synthesis of a complementary strand (Liang and Tzfira 2013). Synthesis of the complementary strand cannot start from the end of the LB, because the LB side is the 3′ end, whereas synthesis is from the 5′ to 3′ end and requires priming. Third, the LB side may be lost in the process of integration when the single-stranded LB anneals through microhomologies to the plant genome (or another T-DNA). Microhomology usually resides in a region internal to the LB end; in that case, the remaining LB side that is not annealed may be degraded and lost (Tinland 1996).

4 A Proposed Model

A T-DNA that is transferred as a linear DNA molecule may circularize via end joining between its LB and RB ends, thus generating a T-circle with head-to-tail end joining (Singer et al. 2012). Throughout different experiments, the majority of the detected T-circles were cases of simple end joining between the LB and RB sides of a single T-DNA, with some small deletions or additions of DNA (Singer and Gelvin, unpublished data). On the other hand, in some cases, T-circles were multimers comprised of several T-DNA molecules or other complex structures. Interestingly, when two T-DNAs are involved in an end-joining event, T-DNA ends preferably ligate tail-to-tail and head-to-head (unpublished data). These results suggest that when the LB and RB sides of a T-DNA are brought into close proximity, a process that is likely mediated by the plant DNA repair pathways, the LB and RB ends are not favorably ligated to each other. This condition can favor T-DNA integration if the T-DNA is situated next to a plant DNA double-stranded

break, because the T-DNA ends may prefer to ligate to the plant DNA instead of ligating to themselves. A T-DNA LB–RB ligation event may not occur if the RB side is a blunt end that preferably utilizes an NHEJ DNA repair pathway, whereas the LB side is a 3′ overhang that preferably utilizes MMEJ DNA repair pathway.

It is unlikely that circularized (and ligated) T-DNAs are intermediates of T-DNA integration because T-DNAs after integration generally maintain the original linear left and right borders, whereas integration of a circular molecule will inevitably result in T-DNA with circularly permuted, random borders. On the other hand, an integration model involving a linear T-DNA in which the two ends are positioned at the opposite poles of molecule is also not likely, because precise and efficient T-DNA integration may require having the two ends of the T-DNA in close proximity.

Therefore, in the model presented T-DNA is proposed to integrate as a double-stranded DNA molecule that is spatially arranged in a looped form (Fig. 6). A looped configuration in which T-DNA ends are in close proximity can explain how T-DNA is often inserted into the genome without the target genomic sites suffering major deletions (Meza et al. 2002; Windels et al. 2003). The exposed T-DNA ends are likely brought together during the initial stage of the repair process by the DNA repair factors coating the ends. VirD2 may also be involved in bringing the T-DNA ends together, as purified VirD2 in vitro has been shown to catalyze end-joining reactions with single-stranded T-border DNA (Pansegrau et al. 1993). These factors may also facilitate the targeting of T-DNA ends to chromosomal sites where DNA repair occurs, such as sites of random genomic DSBs. Integration into these sites occurs when the ends of a T-DNA do not end join to each other to generate a T-circle, but instead end join with the chromosomal DNA (Fig. 6). It should be noted that the proposed model is simplified and does not explain other different outcomes of integration. For example, the frequent formation of filler DNA can be explained by synthesis activity of pol θ (van Kregten et al. 2016). Integration of other more complex structures can occur similarly following their formation extrachromosomally.

Whereas the proposed model suggests a spatial arrangement of T-DNA during integration, it also speculates that the two T-DNA ends utilize different DNA repair pathways for integration into the plant genome. The involvement of different repair pathways in T-DNA integration can explain the conflicting evidence regarding the importance of some key components of DNA repair pathways. It can also explain the tendency of T-DNA ends to generate LB–LB and RB–RB junctions. However, testing this model will require further biochemical and genetic experiments.

5 Conclusions

Agrobacterium tumefaciens remains the main vector used by plant biologists to genetically transform plants. However, there are still many questions to be answered in order to understand the mechanism of T-DNA integration. Because

most of the questions presented in this review are interrelated, understanding T-DNA integration will require different experimental approaches to answer the different questions. In particular, because T-DNA integration most likely relies mostly on plant host factors, a further understanding of pathways of DNA repair in plants is important for improving the understanding of T-DNA integration.

References

Abu-Arish A, Frenkiel-Krispin D, Fricke T et al (2004) Three-dimensional reconstruction of *Agrobacterium* VirE2 protein with single-stranded DNA. J Biol Chem 279:25359–25363

Albright LM, Yanofsky MF, Leroux B et al (1987) Processing of the T-DNA of *Agrobacterium tumefaciens* generates border nicks and linear, single-stranded T-DNA. J Bacteriol 16:1046–1055

Alonso JM, Stepanova AN, Leisse TJ et al (2003) Genome-wide insertional mutagenesis of *Arabidopsis thaliana*. Science 301:653–657

Altpeter F, Springer NM, Bartley LE et al (2016) Advancing crop transformation in the era of genome editing. Plant Cell 2:1510–1520

An S, Park S, Jeong DH et al (2003) Generation and analysis of end sequence database for T-DNA tagging lines in rice. Plant Physiol 133:2040–2047

Anand A, Krichevsky A, Schornack S et al (2007a) *Arabidopsis* VIRE2 INTERACTING PROTEIN2 is required for *Agrobacterium* T-DNA integration in plants. Plant Cell 19:1695–1708

Anand A, Vaghchhipawala Z, Ryu CM et al (2007b) Identification and characterization of plant genes involved in *Agrobacterium*-mediated plant transformation by virus-induced gene silencing. Mol Plant Microbe Interact 20:41–52

Bako L, Umeda M, Tiburcio AF et al (2003) The VirD2 pilot protein of *Agrobacterium*-transferred DNA interacts with the TATA box-binding protein and a nuclear protein kinase in plants. Proc Natl Acad Sci U S A 100:10108–10113

Ballas N, Citovsky V (1997) Nuclear localization signal binding protein from *Arabidopsis* mediates nuclear import of *Agrobacterium* VirD2 protein. Proc Natl Acad Sci U S A 94:10723–10728

Barker RF, Idler KB, Thompson DV et al (1983) Nucleotide sequence of the T-DNA region from the *Agrobacterium tumefaciens* octopine Ti plasmid pTi15955. Plant Mol Biol 2:335–350

Barton KA, Binns AN, Matzke AJ et al (1983) Regeneration of intact tobacco plants containing full length copies of genetically engineered T-DNA, and transmission of T-DNA to R1 progeny. Cell 32:1033–1043

Bhattacharjee S, Lee LY, Oltmanns H et al (2008) IMPa-4, an *Arabidopsis* importin alpha isoform, is preferentially involved in *Agrobacterium*-mediated plant transformation. Plant Cell 20:2661–2680

Binns AN, Beaupre CE, Dale EM (1995) Inhibition of VirB-mediated transfer of diverse substrates from *Agrobacterium tumefaciens* by the IncQ plasmid RSF1010. J Bacteriol 177:4890–4899

Bleuyard JY, Gallego ME, White CI (2006) Recent advances in understanding of the DNA double-strand break repair machinery of plants. DNA Repair 5:1–12

Bravo-Angel AM, Hohn B, Tinland B (1998) The omega sequence of VirD2 is important but not essential for efficient transfer of T-DNA by *Agrobacterium tumefaciens*. Mol Plant-Microbe Interact 11:57–63

Brunaud V, Balzergue S, Dubreucq B et al (2002) T-DNA integration into the *Arabidopsis* genome depends on sequences of pre-insertion sites. EMBO Rep 3:1152–1157

Bundock P, den Dulk-Ras A, Beijersbergen A et al (1995) Trans-kingdom T-DNA transfer from *Agrobacterium tumefaciens* to Saccharomyces cerevisiae. EMBO J 14:3206–3214

Bundock P, Hooykaas PJ (1996) Integration of *Agrobacterium tumefaciens* T-DNA in the *Saccharomyces cerevisiae* genome by illegitimate recombination. Proc Natl Acad Sci U S A 93:15272–15275

Castle LA, Errampalli D, Atherton TL et al (1993) Genetic and molecular characterization of embryonic mutants identified following seed transformation in *Arabidopsis*. Mol Gen Genet 241:504–514

Chen S, Jin W, Wang M et al (2003) Distribution and characterization of over 1000 T-DNA tags in rice genome. Plant J 36:105–113

Chilton MD, Drummond M, Merlo D et al (1977) Stable incorporation of plasmid DNA into higher plant cells: the molecular basis of crown gall tumorigenesis. Cell 11:263–271

Chilton MD, Que Q (2003) Targeted integration of T-DNA into the tobacco genome at double-stranded breaks: new insights on the mechanism of T-DNA integration. Plant Physiol 133:956–965

Christie PJ, Atmakuri K, Krishnamoorthy V et al (2005) Biogenesis, architecture, and function of bacterial type IV secretion systems. Annu Rev Microbiol 59:451–485

Chyi YS, Jorgensen RA, Goldstein D et al (1986) Locations and stability of *Agrobacterium*-mediated T-DNA insertions in the *Lycopersicon* genome. Mol Gen Genet 204:64–69

Citovsky V, DE Vos G, Zambryski P (1988) Single-stranded DNA binding protein encoded by the *virE* locus of *Agrobacterium tumefaciens*. Science 240:501–504

Citovsky V, Warnick D, Zambryski P (1994) Nuclear import of *Agrobacterium* VirD2 and VirE2 proteins in maize and tobacco. Proc Natl Acad Sci U S A 91:3210–3214

Citovsky V, Zupan J, Warnick D et al (1992) Nuclear localization of *Agrobacterium* VirE2 protein in plant cells. Science 256:1802–1805

Clark KA, Krysan PJ (2010) Chromosomal translocations are a common phenomenon in *Arabidopsis thaliana* T-DNA insertion lines. Plant J 64:990–1001

Cluster PD, O'Dell M, Metzlaff M et al (1996) Details of T-DNA structural organization from a transgenic *Petunia* population exhibiting co-suppression. Plant Mol Biol 32:1197–1203

Critchlow SE, Jackson SP (1998) DNA end-joining: from yeast to man. Trends Biochem Sci 23:394–398

Dafny-Yelin M, Levy A, Dafny R et al (2015) Blocking single-stranded transferred DNA conversion to double-stranded intermediates by overexpression of yeast DNA REPLICATION FACTOR A. Plant Physiol 167:153–163

Das A (1988) *Agrobacterium tumefaciens virE* operon encodes a single-stranded DNA-binding protein. Proc Natl Acad Sci U S A 85:2909–2913

De Buck S, De Wilde C, Van Montagu M et al (2000) Determination of the T-DNA transfer and the T-DNA integration frequencies upon cocultivation of *Arabidopsis thaliana* root explants. Mol Plant-Microbe Interact 13:658–665

De Buck S, Podevin N, Nolf J et al (2009) The T-DNA integration pattern in *Arabidopsis* transformants is highly determined by the transformed target cell. Plant J 60:134–145

de Groot MJ, Bundock P, Hooykaas PJ et al (1998) *Agrobacterium tumefaciens*-mediated transformation of filamentous fungi. Nat Biotechnol 16:839–842

De Neve M, De Buck S, Jacobs A et al (1997) T-DNA integration patterns in co-transformed plant cells suggest that T-DNA repeats originate from co-integration of separate T-DNAs. Plant J 11:15–29

Djamei A, Pitzschke A, Nakagami H et al (2007) Trojan horse strategy in *Agrobacterium* transformation: abusing MAPK defense signaling. Science 318:453–456

Durrenberger F, Crameri A, Hohn B et al (1989) Covalently bound VirD2 protein of *Agrobacterium tumefaciens* protects the T-DNA from exonucleolytic degradation. Proc Natl Acad Sci U S A 86:9154–9158

Forsbach A, Schubert D, Lechtenberg B et al (2003) A comprehensive characterization of single-copy T-DNA insertions in the *Arabidopsis thaliana* genome. Plant Mol Biol 52:161–176

Francis KE, Spiker S (2005) Identification of *Arabidopsis thaliana* transformants without selection reveals a high occurrence of silenced T-DNA integrations. Plant J 41:464–477

Friesner J, Britt AB (2003) Ku80- and DNA ligase IV-deficient plants are sensitive to ionizing radiation and defective in T-DNA integration. Plant J 34:427–440

Furner IJ, Higgins ES, Berrington AW (1989) Single-stranded DNA transforms plant protoplasts. Mol Gen Genet 220:65–68

Gallego ME, Bleuyard JY, Daoudal-Cotterell S et al (2003) Ku80 plays a role in non-homologous recombination but is not required for T-DNA integration in *Arabidopsis*. Plant J 35:557–565

Gelvin SB (2010) Plant proteins involved in *Agrobacterium*-mediated genetic transformation. Annu Rev Phytopathol 48:45–68

Gelvin SB (2017) Integration of *Agrobacterium* T-DNA into the plant genome. Annu Rev Genet 51:195–217

Ghedira R, De Buck S, Van Ex F et al (2013) T-DNA transfer and T-DNA integration efficiencies upon *Arabidopsis thaliana* root explant cocultivation and floral dip transformation. Planta 238:1025–1037

Gheysen G, Montagu MV, Zambryski P (1987) Integration of *Agrobacterium tumefaciens* transfer DNA (T-DNA) involves rearrangements of target plant DNA sequences. Proc Natl Acad Sci U S A 84:6169–6173

Gheysen G, Villarroel R, Van Montagu M (1991) Illegitimate recombination in plants: a model for T-DNA integration. Genes Dev 5:287–297

Gorbunova V, Levy AA (1997) Non-homologous DNA end joining in plant cells is associated with deletions and filler DNA insertions. Nucleic Acids Res 25:4650–4657

Grange W, Duckely M, Husale S et al (2008) VirE2: a unique ssDNA-compacting molecular machine. PLoS Biol 6(2):e44

Grevelding C, Fantes V, Kemper E et al (1993) Single-copy T-DNA insertions in *Arabidopsis* are the predominant form of integration in root-derived transgenics, whereas multiple insertions are found in leaf discs. Plant Mol Biol 23:847–860

Hamilton CM, Frary A, Lewis C et al (1996) Stable transfer of intact high molecular weight DNA into plant chromosomes. Proc Natl Acad Sci U S A 93:9975–9979

Herrera-Estrella A, Chen ZM, Van Montagu M et al (1988) VirD proteins of *Agrobacterium tumefaciens* are required for the formation of a covalent DNA-protein complex at the 5' terminus of T-strand molecules. EMBO J 7:4055–4062

Herrera-Estrella A, Van Montagu M, Wang K (1990) A bacterial peptide acting as a plant nuclear targeting signal: the amino-terminal portion of *Agrobacterium* VirD2 protein directs a beta-galactosidase fusion protein into tobacco nuclei. Proc Natl Acad Sci U S A 87:9534–9537

Hoekema A, Hirsch PR, Hooykaas PJJ et al (1983) A binary plant vector strategy based on separation of vir- and T-region of the *Agrobacterium tumefaciens* Ti-plasmid. Nature 303: 179–180

Howard EA, Zupan JR, Citovsky V et al (1992) The VirD2 protein of *A. tumefaciens* contains a C-terminal bipartite nuclear localization signal: implications for nuclear uptake of DNA in plant cells. Cell 68:109–118

Hu T, Metz S, Chay C et al (2003) *Agrobacterium*-mediated large-scale transformation of wheat (*Triticum aestivum* L.) using glyphosate selection. Plant Cell Rep 21:1010–1019

Hwang HH, Mysore KS, Gelvin SB (2006) Transgenic *Arabidopsis* plants expressing *Agrobacterium tumefaciens* VirD2 protein are less susceptible to *Agrobacterium* transformation. Mol Plant Pathol 7:473–484

Iwakawa H, Carter BC, Bishop BC, Ogas J, Gelvin SB (2017) Perturbation of H3K27me3-associated epigenetic processes increases *Agrobacterium*-mediated transformation. Mol Plant Microbe Interact 30:35–44

Janssen BJ, Gardner RC (1990) Localized transient expression of GUS in leaf discs following cocultivation with *Agrobacterium*. Plant Mol Biol 14:61–72

Jia Q, Bundock P, Hooykaas PJJ et al (2012) *Agrobacterium tumefaciens* T-DNA integration and gene targeting in *Arabidopsis thaliana* non-homologous end-joining mutants. J Bot 2012:1–13

Johansen LK, Carrington JC (2001) Silencing on the spot. Induction and suppression of RNA silencing in the *Agrobacterium*-mediated transient expression system. Plant Physiol 126:930–938

Jorgensen R, Snyder C, Jones JDG (1987) T-DNA is organized predominantly in inverted repeat structures in plants transformed with *Agrobacterium tumefaciens* C58 derivatives. Mol Gen Genet 207:471–477

Kim SI, Veena, Gelvin SB (2007) Genome-wide analysis of *Agrobacterium* T-DNA integration sites in the *Arabidopsis* genome generated under non-selective conditions. Plant J 51:779–791

Kim SR, Lee J, Jun SH et al (2003) Transgene structures in T-DNA-inserted rice plants. Plant Mol Biol 52:761–773

Kleinboelting N, Huep G, Appelhagen I et al (2015) The structural features of thousands of T-DNA insertion sites are consistent with a double-strand break repair-based insertion mechanism. Mol Plant 8:1651–1664

Ko K, Brodzik R, Steplewski Z (2009) Production of antibodies in plants: approaches and perspectives. Curr Top Microbiol Immunol 332:55–78

Köhler F, Cardon G, Pöhlman M et al (1989) Enhancement of transformation rates in higher plants by low-dose irradiation: Are DNA repair systems involved in the incorporation of exogenous DNA into the plant genome? Plant Mol Biol 12:189–199

Komarova TV, Baschieri S, Donini M et al (2010) Transient expression systems for plant-derived biopharmaceuticals. Expert Rev Vaccines 9:859–876

Kononov ME, Bassuner B, Gelvin SB (1997) Integration of T-DNA binary vector 'backbone' sequences into the tobacco genome: evidence for multiple complex patterns of integration. Plant J 11:945–957

Korn M, Schmidpeter J, Dahl M et al (2015) A genetic screen for pathogenicity genes in the hemibiotrophic fungus *Colletotrichum higginsianum* identifies the plasma membrane proton ump Pma2 required for host penetration. PLoS ONE 10:e0125960

Krizkova L, Hrouda M (1998) Direct repeats of T-DNA integrated in tobacco chromosome: characterization of junction regions. Plant J 16:673–680

Kumar S, Fladung M (2002) Transgene integration in aspen: structures of integration sites and mechanism of T-DNA integration. Plant J 31:543–551

Kumar V, Jain M (2015) The CRISPR-Cas system for plant genome editing: advances and opportunities. J Exp Bot 66:47–57

Kunik T, Tzfira T, Kapulnik Y et al (2001) Genetic transformation of HeLa cells by *Agrobacterium*. Proc Natl Acad Sci U S A 98:1871–1876

Kwok WW, Nester EW, Gordon MP (1985) Unusual plasmid DNA organization in an octopine crown gall tumor. Nucleic Acids Res 13:459–471

Lacroix B, Li J, Tzfira T et al (2006) Will you let me use your nucleus? How *Agrobacterium* gets its T-DNA expressed in the host plant cell. Can J Physiol Pharmacol 84:333–345

Lacroix B, Vaidya M, Tzfira T et al (2005) The VirE3 protein of *Agrobacterium* mimics a host cell function required for plant genetic transformation. EMBO J 24:428–437

Lee L-Y, Fang MJ, Kuang LY et al (2008) Vectors for multi-color bimolecular fluorescence complementation to investigate protein-protein interactions in living plant cells. Plant Methods 4:24–4811-4-24

Lessl M, Lanka E (1994) Common mechanisms in bacterial conjugation and Ti-mediated T-DNA transfer to plant cells. Cell 77:321–324

Li J, Krichevsky A, Vaidya M et al (2005a) Uncoupling of the functions of the *Arabidopsis* VIP1 protein in transient and stable plant genetic transformation by *Agrobacterium*. Proc Natl Acad Sci U S A 102:5733–5738

Li J, Vaidya M, White C et al (2005b) Involvement of KU80 in T-DNA integration in plant cells. Proc Natl Acad Sci U S A 102:19231–19236

Li X, Pan SQ (2017) *Agrobacterium* delivers VirE2 protein into host cells via clathrin-mediated endocytosis. Sci Adv 3:e1601528

Li Y, Rosso MG, Ülker B et al (2006) Analysis of T-DNA insertion site distribution patterns in *Arabidopsis thaliana* reveals special features of genes without insertions. Genomics 87:645–652

Liang Z, Tzfira T (2013) In vivo formation of double-stranded T-DNA molecules by T-strand priming. Nat Commun 4:2253

Magori S, Citovsky V (2011) Epigenetic control of *Agrobacterium* T-DNA integration. Biochim Biophys Acta 1809:388–394

Makarevitch I, Svitashev SK, Somers DA (2003) Complete sequence analysis of transgene loci from plants transformed via microprojectile bombardment. Plant Mol Biol 52:421–432

Martineau B, Voelker TA, Sanders RA (1994) On defining T-DNA. Plant Cell 6:1032–1033

Matsumoto S, Ito Y, Hosoi T et al (1990) Integration of *Agrobacterium* T-DNA into a tobacco chromosome: possible involvement of DNA homology between T-DNA and plant DNA. Mol Gen Genet 224:309–316

Maximova SN, Dandekar AM, Guiltinan MJ (1998) Investigation of *Agrobacterium*-mediated transformation of apple using green fluorescent protein: high transient expression and low stable transformation suggest that factors other than T-DNA transfer are rate-limiting. Plant Mol Biol 37:549–559

Mayerhofer R, Koncz-Kalman Z, Nawrath C et al (1991) T-DNA integration: a mode of illegitimate recombination in plants. EMBO J 10:697–704

McVey M, Lee SE (2008) MMEJ repair of double-strand breaks (director's cut): deleted sequences and alternative endings. Trends Genet 24:529–538

Mestiri I, Norre F, Gallego ME et al (2014) Multiple host-cell recombination pathways act in *Agrobacterium*-mediated transformation of plant cells. Plant J 77:511–520

Meza TJ, Stangeland B, Mercy IS et al (2002) Analyses of single-copy *Arabidopsis* T-DNA-transformed lines show that the presence of vector backbone sequences, short inverted repeats and DNA methylation is not sufficient or necessary for the induction of transgene silencing. Nucleic Acids Res 30:4556–4566

Muller AE, Atkinson RG, Sandoval RB et al (2007) Microhomologies between T-DNA ends and target sites often occur in inverted orientation and may be responsible for the high frequency of T-DNA-associated inversions. Plant Cell Rep 26:617–630

Muller AE, Kamisugi Y, Gruneberg R et al (1999) Palindromic sequences and A+T-rich DNA elements promote illegitimate recombination in *Nicotiana tabacum*. J Mol Biol 291:29–46

Mysore KS, Bassuner B, Deng X et al (1998) Role of the *Agrobacterium tumefaciens* VirD2 protein in T-DNA transfer and integration. Mol Plant-Microbe Interact 11:668–683

Mysore KS, Nam J, Gelvin SB (2000) An *Arabidopsis* histone H2A mutant is deficient in *Agrobacterium* T-DNA integration. Proc Natl Acad Sci U S A 97:948–953

Nacry P, Camilleri C, Courtial B et al (1998) Major chromosomal rearrangements induced by T-DNA transformation in *Arabidopsis*. Genetics 149:641–650

Narasimhulu SB, Deng XB, Sarria R et al (1996) Early transcription of *Agrobacterium* T-DNA genes in tobacco and maize. Plant Cell 8:873–886

Nelson AD, Lamb JC, Kobrossly PS et al (2011) Parameters affecting telomere-mediated chromosomal truncation in *Arabidopsis*. Plant Cell 23:2263–2272

Nicolia A, Ferradini N, Veronesi F et al (2017) An insight into T-DNA integration events in *Medicago sativa*. Int J Mol Sci 18. https://doi.org/10.3390/ijms18091951

Nishizawa-Yokoi A, Nonaka S, Saika H et al (2012) Suppression of Ku70/80 or Lig4 leads to decreased stable transformation and enhanced homologous recombination in rice. New Phytol 196:1048–1059

Offringa R, de Groot MJ, Haagsman HJ et al (1990) Extrachromosomal homologous recombination and gene targeting in plant cells after *Agrobacterium* mediated transformation. EMBO J 9:3077–3084

Ohba T, Yoshioka Y, Machida C et al (1995) DNA rearrangement associated with the integration of T-DNA in tobacco: an example for multiple duplications of DNA around the integration target. Plant J 7:157–164

Ohmine Y, Satoh Y, Kiyokawa K et al (2016) DNA repair genes RAD52 and SRS2, a cell wall synthesis regulator gene SMI1, and the membrane sterol synthesis scaffold gene ERG28 are important in efficient *Agrobacterium*-mediated yeast transformation with chromosomal T-DNA. BMC Microbiol 16:58-016-0672-0

O'Malley RC, Ecker JR (2010) Linking genotype to phenotype using the *Arabidopsis* unimutant collection. Plant J 61:928–940

Ooms G, Bakker A, Molendijk L et al (1982) T-DNA organization in homogeneous and heterogeneous octopine-type crown gall tissues of *Nicotiana tabacum*. Cell 30:589–597

Otten L, De Greve H, Leemans J, Hain R, Hooykaas P, Schell J (1984) Restoration of virulence of *Vir* region mutants of *Agrobacterium tumefaciens* strain B6S3 by coinfection with normal and mutant *Agrobacterium* strains. Mol Gen Genet 195:159–163

Pansegrau W, Schoumacher F, Hohn B et al (1993) Site-specific cleavage and joining of single-stranded DNA by VirD2 protein of *Agrobacterium tumefaciens* Ti plasmids: analogy to bacterial conjugation. Proc Natl Acad Sci U S A 90:11538–11542

Park SY, Vaghchhipawala Z, Vasudevan B et al (2015) *Agrobacterium* T-DNA integration into the plant genome can occur without the activity of key non-homologous end-joining proteins. Plant J 81:934–946

Pitzschke A, Djamei A, Teige M et al (2009) VIP1 response elements mediate mitogen-activated protein kinase 3-induced stress gene expression. Proc Natl Acad Sci U S A 106(43):18414–18419

Ream LW, Gordon MP, Nester EW (1983) Multiple mutations in the T region of the *Agrobacterium tumefaciens* tumor-inducing plasmid. Proc Natl Acad Sci U S A 80:1660–1664

Relic B, Andjelkovic M, Rossi L et al (1998) Interaction of the DNA modifying proteins VirD1 and VirD2 of *Agrobacterium tumefaciens*: analysis by subcellular localization in mammalian cells. Proc Natl Acad Sci U S A 95:9105–9110

Rios G, Lossow A, Hertel B et al (2002) Rapid identification of *Arabidopsis* insertion mutants by non-radioactive detection of T-DNA tagged genes. Plant J 32:243–253

Rodenburg KW, de Groot MJ, Schilperoort RA et al (1989) Single-stranded DNA used as an efficient new vehicle for transformation of plant protoplasts. Plant Mol Biol 13:711–719

Rolloos M, Dohmen MH, Hooykaas PJ et al (2014) Involvement of Rad52 in T-DNA circle formation during *Agrobacterium tumefaciens*-mediated transformation of *Saccharomyces cerevisiae*. Mol Microbiol 91:1240–1251

Rolloos M, Hooykaas PJ, van der Zaal BJ (2015) Enhanced targeted integration mediated by translocated I-*Sce*I during the *Agrobacterium* mediated transformation of yeast. Sci Rep 5:8345

Rossi L, Hohn B, Tinland B (1993) The VirD2 protein of *Agrobacterium tumefaciens* carries nuclear localization signals important for transfer of T-DNA to plant. Mol Gen Genet 239: 345–353

Rossi L, Hohn B, Tinland B (1996) Integration of complete transferred DNA units is dependent on the activity of virulence E2 protein of *Agrobacterium tumefaciens*. Proc Natl Acad Sci U S A 93:126–130

Rosso MG, Li Y, Strizhov N et al (2003) An *Arabidopsis thaliana* T-DNA mutagenized population (GABI-Kat) for flanking sequence tag-based reverse genetics. Plant Mol Biol 53:247–259

Saika H, Nishizawa-Yokoi A, Toki S (2014) The non-homologous end-joining pathway is involved in stable transformation in rice. Front Plant Sci 5:560

Sakalis PA, van Heusden GP, Hooykaas PJ (2014) Visualization of VirE2 protein translocation by the *Agrobacterium* type IV secretion system into host cells. Microbiologyopen 3(1):104–117

Sallaud C, Gay C, Larmande P et al (2004) High throughput T-DNA insertion mutagenesis in rice: a first step towards in silico reverse genetics. Plant J 39:450–464

Salomon S, Puchta H (1998) Capture of genomic and T-DNA sequences during double-strand break repair in somatic plant cells. EMBO J 17:6086–6095

Scheiffele P, Pansegrau W, Lanka E (1995) Initiation of *Agrobacterium tumefaciens* T-DNA processing. Purified proteins VirD1 and VirD2 catalyze site- and strand-specific cleavage of superhelical T-border DNA *in vitro*. J Biol Chem 270:1269–1276

Schrammeijer B, den Dulk-Ras A, Vergunst AC et al (2003) Analysis of Vir protein translocation from *Agrobacterium tumefaciens* using *Saccharomyces cerevisiae* as a model: evidence for transport of a novel effector protein VirE3. Nucleic Acids Res 31:860–868

Schrammeijer B, Risseeuw E, Pansegrau W et al (2001) Interaction of the virulence protein VirF of *Agrobacterium tumefaciens* with plant homologs of the yeast Skp1 protein. Curr Biol 11:258–262

Sessions A, Burke E, Presting G et al (2002) A high-throughput *Arabidopsis* reverse genetics system. Plant Cell 14:2985–2994

Shi Y, Lee LY, Gelvin SB (2014) Is VIP1 important for *Agrobacterium*-mediated transformation? Plant J 79:848–860

Shiboleth Y, Tzfira T (2012) *Agrobacterium*-mediated plant genetic transformation. In: Hasegawa PM (ed) Altman A. Plant Biotechnology and Agriculture Academic Press, San Diego, pp 99–116

Shilo S, Tripathi P, Melamed-Bessudo C et al (2017) T-DNA-genome junctions form early after infection and are influenced by the chromatin state of the host genome. PLoS Genet 13: e1006875

Shurvinton CE, Hodges L, Ream W (1992) A nuclear localization signal and the C-terminal omega sequence in the *Agrobacterium tumefaciens* VirD2 endonuclease are important for tumor formation. Proc Natl Acad Sci U S A 89:11837–11841

Singer K (2013) T-DNA integration during *Agrobacterium*-mediated transformation of plants (doctoral dissertation). Retrieved from: https://deepbluelibumichedu/handle/202742/99891?show=full University of Michigan, Horace H. Rackham School of Graduate Studies

Singer K, Shiboleth YM, Li J et al (2012) Formation of complex extrachromosomal T-DNA structures in *Agrobacterium tumefaciens*-infected plants. Plant Physiol 160:511–522

Smith EF, Townsend CO (1907) A plant-tumor of bacterial origin. Science 25:671–673

Soltani J, van Heusden GPH, Hooykaas PJJ (2010) *Agrobacterium*-mediated transformation of non-plant organisms. In: Tzfira T, Citovsky V (eds) *Agrobacterium*: from biology to biotechnology. Springer, New York, pp 650–675

Somers DA, Makarevitch I (2004) Transgene integration in plants: poking or patching holes in promiscuous genomes? Curr Opin Biotechnol 15:126–131

Song J, Bent AF (2014) Microbial pathogens trigger host DNA double-strand breaks whose abundance is reduced by plant defense responses. PLoS Pathog 10(4):e1004030

Sparkes IA, Runions J, Kearns A et al (2006) Rapid, transient expression of fluorescent fusion proteins in tobacco plants and generation of stably transformed plants. Nat Protoc 1:2019–2025

Spielmann A, Simpson R (1986) T-DNA structure in transgenic tobacco plants with multiple independent integration sites. Mol Gen Genet 205:34–41

Suzuki K, Hattori Y, Uraji M et al (2000) Complete nucleotide sequence of a plant tumor-inducing Ti plasmid. Gene 242:331–336

Szabados L, Kovacs I, Oberschall A et al (2002) Distribution of 1000 sequenced T-DNA tags in the *Arabidopsis* genome. Plant J 32:233–242

Takano M, Egawa H, Ikeda JE et al (1997) The structures of integration sites in transgenic rice. Plant J 11:353–361

Tax FE, Vernon DM (2001) T-DNA-associated duplication/translocations in *Arabidopsis*. Implications for mutant analysis and functional genomics. Plant Physiol 126:1527–1538

Teo CH, Ma L, Kapusi E et al (2011) Induction of telomere-mediated chromosomal truncation and stability of truncated chromosomes in *Arabidopsis thalia*na. Plant J 68:28–39

Thomas CM, Jones JD (2007) Molecular analysis of *Agrobacterium* T-DNA integration in tomato reveals a role for left border sequence homology in most integration events. Mol Genet Genom 278:411–420

Tinland B (1996) The integration of T-DNA into plant genomes. Trends Plant Sci 1:178–184

Tinland B, Hohn B, Puchta H (1994) *Agrobacterium tumefaciens* transfers single-stranded transferred DNA (T-DNA) into the plant cell nucleus. Proc Natl Acad Sci U S A 91:8000–8004

Tinland B, Koukolikova-Nicola Z, Hall MN et al (1992) The T-DNA-linked VirD2 protein contains two distinct functional nuclear localization signals. Proc Natl Acad Sci U S A 89:7442–7446

Tinland B, Schoumacher F, Gloeckler V et al (1995) The *Agrobacterium tumefaciens* virulence D2 protein is responsible for precise integration of T-DNA into the plant genome. EMBO J 14:3585–3595

Travella S, Ross SM, Harden J et al (2005) A comparison of transgenic barley lines produced by particle bombardment and *Agrobacterium*-mediated techniques. Plant Cell Rep 23:780–789

Tzfira T, Citovsky V (2006) *Agrobacterium*-mediated genetic transformation of plants: biology and biotechnology. Curr Opin Biotechnol 17:147–154

Tzfira T, Frankman LR, Vaidya M et al (2003) Site-specific integration of *Agrobacterium tumefaciens* T-DNA via double-stranded intermediates. Plant Physiol 133:1011–1023

Tzfira T, Li J, Lacroix B et al (2004a) *Agrobacterium* T-DNA integration: molecules and models. Trends Genet 20:375–383

Tzfira T, Vaidya M, Citovsky V (2001) VIP1, an *Arabidopsis* protein that interacts with *Agrobacterium* VirE2, is involved in VirE2 nuclear import and *Agrobacterium* infectivity. EMBO J 20:3596–3607

Tzfira T, Vaidya M, Citovsky V (2004b) Involvement of targeted proteolysis in plant genetic transformation by *Agrobacterium*. Nature 431:87–92

Ulker B, Li Y, Rosso MG et al (2008) T-DNA-mediated transfer of *Agrobacterium tumefaciens* chromosomal DNA into plants. Nat Biotechnol 26:1015–1017

Vaghchhipawala ZE, Vasudevan B, Lee S et al (2012) *Agrobacterium* may delay plant nonhomologous end-joining DNA repair via XRCC4 to favor T-DNA integration. Plant Cell 24:4110–4123

van Attikum H, Bundock P, Overmeer RM et al (2003) The *Arabidopsis* AtLIG4 gene is required for the repair of DNA damage, but not for the integration of *Agrobacterium* T-DNA. Nucleic Acids Res 31:4247–4255

van Kregten M, de Pater S, Romeijn R et al (2016) T-DNA integration in plants results from polymerase-theta-mediated DNA repair. Nat Plants 2:16164

van Kregten M, Lindhout BI, Hooykaas PJ et al (2009) *Agrobacterium*-mediated T-DNA transfer and integration by minimal VirD2 consisting of the relaxase domain and a type IV secretion system translocation signal. Mol Plant-Microbe Interact 22:1356–1365

Van Lijsebettens M, Inze D, Schell J et al (1986) Transformed cell clones as a tool to study T-DNA integration mediated by *Agrobacterium tumefaciens*. J Mol Biol 188:129–145

Vergunst AC, Schrammeijer B, den Dulk-Ras A et al (2000) VirB/D4-dependent protein translocation from *Agrobacterium* into plant cells. Science 290:979–982

Vergunst AC, van Lier MC, den Dulk-Ras A et al (2005) Positive charge is an important feature of the C-terminal transport signal of the VirB/D4-translocated proteins of *Agrobacterium*. Proc Natl Acad Sci U S A 102:832–837

Virts EL, Gelvin SB (1985) Analysis of transfer of tumor-inducing plasmids from *Agrobacterium tumefaciens* to *Petunia* protoplasts. J Bacteriol 162:1030–1038

Vogel AM, Das A (1992) Mutational analysis of *Agrobacterium tumefaciens* VirD2: tyrosine 29 is essential for endonuclease activity. J Bacteriol 174:303–308

Wang H, Xu X (2017) Microhomology-mediated end joining: new players join the team. Cell Biosci 7:6-017-0136-8. eCollection 2017

Wang K, Herrera-Estrella L, Van Montagu M et al (1984) Right 25 bp terminus sequence of the nopaline T-DNA is essential for and determines direction of DNA transfer from *Agrobacterium* to the plant genome. Cell 38:455–462

Wang K, Stachel SE, Timmerman B et al (1987) Site-specific nick in the T-DNA border sequence as a result of *Agrobacterium* vir gene expression. Science 235:587–591

Wang L, Lacroix B, Guo J et al (2017) The *Agrobacterium* VirE2 effector interacts with multiple members of the *Arabidopsis* VIP1 protein family. Mol Plant Pathol 19:1172–1183

Ward ER, Barnes WM (1988) VirD2 protein of *Agrobacterium tumefaciens* is very tightly linked to the 5′ end of T-strand DNA. Science 242:927–930

Wenck A, Czako M, Kanevski I et al (1997) Frequent collinear long transfer of DNA inclusive of the whole binary vector during *Agrobacterium*-mediated transformation. Plant Mol Biol 34:913–922

Windels P, De Buck S, Depicker A (2010) Patterns of T-DNA integration into the host genome. In: Tzfira T, Citovsky V (eds) *Agrobacterium*: from biology to biotechnology. Springer, New York, NY, pp 441–481

Windels P, De Buck S, Van Bockstaele E et al (2003) T-DNA integration in *Arabidopsis* chromosomes. Presence and origin of filler DNA sequences. Plant Physiol 133:2061–2068

Yadav NS, Vanderleyden J, Bennett DR et al (1982) Short direct repeats flank the T-DNA on a nopaline Ti plasmid. Proc Natl Acad Sci U S A 79:6322–6326

Yanofsky MF, Porter SG, Young C et al (1986) The *virD* operon of *Agrobacterium tumefaciens* encodes a site-specific endonuclease. Cell 47:471–477

Yi H, Mysore KS, Gelvin SB (2002) Expression of the *Arabidopsis* histone H2A-1 gene correlates with susceptibility to *Agrobacterium* transformation. Plant J 32:285–298

Yi H, Sardesai N, Fujinuma T et al (2006) Constitutive expression exposes functional redundancy between the *Arabidopsis* histone H2A gene HTA1 and other H2A gene family members. Plant Cell 18:1575–1589

Yin K, Gao C, Qiu JL (2017) Progress and prospects in plant genome editing. Nat Plants 3:17107

Young C, Nester EW (1988) Association of the VirD2 protein with the 5′ end of T strands in *Agrobacterium tumefaciens*. J Bacteriol 170:3367–3374

Yusibov VM, Steck TR, Gupta V et al (1994) Association of single-stranded transferred DNA from *Agrobacterium tumefaciens* with tobacco cells. Proc Natl Acad Sci U S A 91(8):2994–2998

Zaenen I, van Larebeke H, Teuchy H et al (1974) Supercoiled circular DNA in crown-gall inducing *Agrobacterium* strains. J Mol Biol 86:109–127

Zaltsman A, Lacroix B, Gafni Y et al (2013) Disassembly of synthetic *Agrobacterium* T-DNA-protein complexes via the host SCF(VBF) ubiquitin-ligase complex pathway. Proc Natl Acad Sci U S A 110:169–174

Zambryski P, Joos H, Genetello C et al (1983) Ti plasmid vector for the introduction of DNA into plant cells without alteration of their normal regeneration capacity. EMBO J 2:2143–2150

Zhang J, Cai L, Cheng J et al (2008) Transgene integration and organization in cotton (*Gossypium hirsutum* L.) genome. Transgenic Res 17:293–306. https://doi.org/10.1007/s11248-007-9101-3

Zhang J, Guo D, Chang Y et al (2007) Non-random distribution of T-DNA insertions at various levels of the genome hierarchy as revealed by analyzing 13 804 T-DNA flanking sequences from an enhancer-trap mutant library. Plant J 49:947–959

Zhang Q, Xing HL, Wang ZP et al (2018) Potential high-frequency off-target mutagenesis induced by CRISPR/Cas9 in Arabidopsis and its prevention. Plant Mol Biol 96(4–5):445–456

Zhang X, van Heusden GPH, Hooykaas PJJ (2017) Virulence protein VirD5 of *Agrobacterium tumefaciens* binds to kinetochores in host cells via an interaction with Spt4. Proc Natl Acad Sci U S A 114:10238–10243

Zhu QH, Ramm K, Eamens AL et al (2006) Transgene structures suggest that multiple mechanisms are involved in T-DNA integration in plants. Plant Sci 171:308–322

Zhu Y, Nam J, Humara JM et al (2003) Identification of *Arabidopsis* rat mutants. Plant Physiol 132:494–505

Ziemienowicz A, Merkle T, Schoumacher F et al (2001) Import of *Agrobacterium* T-DNA into plant nuclei: two distinct functions of VirD2 and VirE2 proteins. Plant Cell 13:369–383

Ziemienowicz A, Tinland B, Bryant J et al (2000) Plant enzymes but not *Agrobacterium* VirD2 mediate T-DNA ligation in vitro. Mol Cell Biol 20:6317–6322

Ziemienowicz A, Tzfira T, Hohn B (2010) Mechanisms of T-DNA Integration. In: Tzfira T, Citovsky V (eds) *Agrobacterium*: From Biology to Biotechnology Springer, New York, NY, p 395–440

Zupan JR, Citovsky V, Zambryski P (1996) *Agrobacterium* VirE2 protein mediates nuclear uptake of single-stranded DNA in plant cells. Proc Natl Acad Sci U S A 93:2392–2397

Transcriptome Profiling of Plant Genes in Response to *Agrobacterium tumefaciens*-Mediated Transformation

Christopher J. Willig, Kaixuan Duan and Zhanyuan J. Zhang

Contents

1 Introduction 320
2 Differential Gene Expression in Response to Inoculation with Various *Agrobacterium Tumefaciens* Strains 323
2.1 Transcriptomic Analysis of Model Plants During Agrobacterium-Mediated Transformation 323
2.2 Transcriptomic Analysis of Crop Species During Agrobacterium-Mediated Transformation 326
2.3 Evaluation of GO Terms and Individual Genes Common to Gene Sets from Multiple Studies 327
3 Transcriptional Profiling in Response to PAMP Exposure 328
4 Transcriptional Response to Heterologous Expression of *Vir* Genes in Host Plants 331
5 Advantages and Limitations of Using Transcriptomic Analysis as an Approach to Discover Plant Genes and Pathways Associated with *Agrobacterium*-Mediated Transformation 333
6 Summary and Outlook 334
References 346

Abstract *Agrobacterium tumefaciens* is a plant pathogen that causes crown gall disease. During infection of the host plant, *Agrobacterium* transfers T-DNA from its Ti plasmid into the host cell, which can then be integrated into the host genome. This unique genetic transformation capability has been employed as the dominant technology for producing genetically modified plants for both basic research and biotechnological applications. *Agrobacterium* has been well studied as a disease-causing agent. The *Agrobacterium*-mediated transformation process involves early attachment of the bacterium to the host's surface, followed by transfer of T-DNA and virulence proteins into the plant cell. Throughout this process, the host plants exhibit dynamic gene expression patterns at each infection

C. J. Willig · K. Duan · Z. J. Zhang (✉)
Division of Plant Sciences, University of Missouri, Columbia, MO, USA
e-mail: zhangzh@missouri.edu

Current Topics in Microbiology and Immunology (2018) 418:319–348
DOI 10.1007/82_2018_115

Published Online: 31 July 2018

stage or in response to *Agrobacterium* strains with varying pathogenic capabilities. Shifting host gene expression patterns throughout the transformation process have effects on transformation frequency, host morphology, and metabolism. Thus, gene expression profiling during the *Agrobacterium* infection process can be an important approach to help elucidate the interaction between *Agrobacterium* and plants. This review highlights recent findings on host plant differential gene expression patterns in response to *A. tumefaciens* or related elicitor molecules.

1 Introduction

Agrobacterium tumefaciens is a gram-negative soil-borne bacterium that is the causative agent of crown gall disease, which affects a wide range of host species (DeCleene and DeLay 1976). Through the course of its infection of a host plant, *Agrobacterium* mobilizes a single-strand DNA segment originating from a sequence located on its tumor-inducing (Ti) plasmid, referred to as transfer DNA (or T-DNA), into host plant cells (Chilton et al. 1977). Integration of T-DNA into the genome of the host and the subsequent expression of the gene products it codes for leads to the formation of tumorous growths that are characteristic of crown gall disease (Escobar and Dandekar 2003). This uncommon ability to transmit DNA sequences into hosts and modify their gene expression as part of its infection strategy has made *Agrobacterium* an important tool in the development of transgenic plants for crop breeding and basic research through the utilization of a method known as *Agrobacterium*-mediated transformation (Azpiroz-Leehan and Feldmann 1997). The close contact between pathogen and host throughout the different stages of infection gives rise to heavily intertwined and multi-layered molecular interactions among them. The successful genomic integration and expression of T-DNA in a host is dependent on a few different steps occurring during pathogenesis including signal perception, expression of pathogen virulence genes, attachment of *Agrobacterium* to the host cell, and T-DNA processing (the presence of a Ti plasmid is required for transfer of T-DNA), transport of T-DNA and accessory proteins into the host cell, nuclear import, and finally uncoating of the T-DNA strand and recombination with the host genome (Gelvin 2000). Molecular interactions at any one of these stages can affect the success of *Agrobacterium* at infecting the plant. To attempt to understand these interactions, we need to examine each stage in greater detail.

In nature, the infection process begins with the production of phenolic compounds by a potential host. This may be caused by environmental stress or damage due to wounding. Perception of plant-derived phenolic molecules can induce the expression of a suite of genes known as the *vir* regulon, which is housed on the Ti plasmid (Bhattacharya et al. 2010). The *Agrobacterium* cell attaches to the surface of the host plant, mediated by pathogen-secreted polysaccharide and protein adhesin compounds. Following attachment, *Agrobacterium* cells become sessile. Although the Ti plasmid is required for T-DNA transfer and tumorigenesis, it is not strictly required for this polar attachment process (Tomlinson and Fuqua 2009). It is at this stage that the plant has its first opportunity for defense against *Agrobacterium* infection. Some

bacteria-derived compounds known as pathogen-associated molecular patterns (PAMPs) may be perceived by the host, thereby invoking a quick, but relatively weak immune response termed PAMP-triggered immunity (PTI) (Zipfel and Robatzek 2010). The most well-studied *Agrobacterium*-derived PAMP that elicits a PTI response is EF-Tu, which was found to be recognized by the host receptor kinase protein EFR (EF-Tu receptor). In addition to EF-Tu, other yet to be characterized PAMPs may play a role in the molecular interactions at this stage (Zipfel et al. 2006).

Provided that a virulent strain of *Agrobacterium* has undergone attachment to the plant surface and the expression of *vir* genes has been induced, the T-DNA segment on the Ti plasmid is nicked by VirD2 at right and left border sequences and becomes covalently attached to this VirD2 protein at the 5' end of nicked T-strand (Mysore et al. 1998). The T-strand is shuttled through a type IV secretion system into the cytoplasm of a host plant cell where it is thought to be coated by molecules of the VirE2 protein (Ziemienowicz et al. 2001). Mutually bound VirD2, single-stranded T-DNA, and VirE2 are believed to form what is termed the "T-complex" (Vergunst et al. 2000). In addition to VirD2 and VirE2, other virulence proteins are known to be delivered into the host cell including VirE3, VirD5, and VirF. (Thompson et al. 1988; Ward et al. 1988, 2002; Kuldau et al. 1990; Shirasu et al. 1990; Beijersbergen et al. 1994; Christie and Vogel 2000; Schrammeijer et al. 2003). VirE2 contains a plant-active nuclear localization signal and has been shown to interact with the plant protein VIP1, which was initially thought to be involved in T-DNA integration into the genome (Tzfira and Citovsky 2001; Djamei et al. 2007). However, its participation in this process has been questioned in more recent studies (Shi et al. 2014; Lapham et al. 2018). The T-DNA from wild-type strains of *Agrobacterium* encodes the plant-active genes *iaaH*, *iaaM*, and *ipt*, which induce the biosynthesis of auxin and cytokinin. The expression of these genes gives rise to the formation of crown galls (Morris 1986; Binns and Costantino 1998). For the purpose of adapting T-DNA transfer for biotechnological applications, *Agrobacterium* strains which have had their Ti plasmids "disarmed" (i.e., non-tumorigenic) are used (Barton and Brill 1983). This means that the oncogenes present in the T-DNA segment of the plasmid have been deleted, whereas the *vir* genes remain. Therefore, non-tumorigenic strains of *Agrobacterium* often used in the laboratory still transfer proteins into host cells via a type IV secretion system, but do not lead to the formation of crown galls.

Bacterial pathogens other than *Agrobacterium* are known to transfer proteins into the plant cell during infection, which may serve to dampen the PTI response. However, host plants are often able to perceive these foreign proteins (termed effectors), thereby initiating a strong defense response called effector-triggered immunity (ETI) (Cui et al. 2009). This series of defenses include the hypersensitive response (HR) consisting of rapid programmed cell death, which prevents the spread of the disease to uninfected tissue. However, *Agrobacterium* is able to avoid eliciting HR in most plants despite its transfer of foreign Vir proteins. This is a key advantage *Agrobacterium* has over other plant bacterial pathogens (Pu and Goodman 1993; Staskawicz et al. 1995; Wood et al. 2001). Although the secreted Vir proteins seem not to elicit an immune response, they have been shown in several cases to have *in planta* functions that could modulate host gene expression patterns (Tzfira et al. 2004; Lacroix et al. 2005; García-Rodríguez et al. 2006; Magori and Citovsky 2011; Wang et al. 2014; Niu et al. 2015;

Zhang et al. 2017). In order to examine the differences between the effects caused by *Agrobacterium* PAMPs or other non-transferred factors and those caused by the *in planta* activities of the Vir proteins, some of the studies reviewed here utilized "cured" strains of *Agrobacterium*, meaning that they no longer contain a Ti plasmid and thus cannot express or transfer Vir proteins (Watson et al. 1975). Strains lacking Ti plasmids are often referred to in the literature as "avirulent" in contrast to "virulent" strains that contain a disarmed Ti plasmid (Veena et al. 2003). Both of these are considered non-tumorigenic as they cannot induce crown gall formation. To date, many functions of the *Agrobacterium* genes involved in the pathogen-host interaction are known and most of the Vir proteins have been well-characterized (Tzfira and Citovsky 2006). By contrast, many of the host plant factors that are directly involved in the transformation process, including attachment, pathogen recognition, T-DNA transfer, trafficking through the cell cytoplasm and integration of T-DNA into the host genome remain unidentified (Gelvin 2010, 2017; Lacroix and Citovsky 2013).

One of the many approaches that have been implemented to understand the molecular players and large-scale processes involved in transformation is differential gene expression analysis. The basic premise behind this analysis is that by contrasting the transcript levels of genes between two different conditions or treatments, one can elucidate which genes are altered by the treatment and thus might be involved in facilitating or suppressing a given process. The technology for evaluating transcript expression has advanced exponentially over the past 20 years, resulting in drastically decreased cost and labor and increased capacity for data generation. Each of the techniques that have been developed generally involves the production of a cDNA pool generated by reverse transcription of total mRNA extracts from plant tissues. The experimental methodologies utilized in the studies reviewed here include cDNA-AFLP (amplified fragment length polymorphism), RT-PCR (reverse transcription), suppression subtractive hybridization, DNA macroarrays, microarrays, quantitative real-time RT-PCR, and RNA-Seq. For a detailed review on these methods and more, see Casassola et al. (2013). The widespread use of microarrays and RNA-Seq has led to the accumulation of massive amounts transcriptional data as there can be coverage over nearly all of the genome in one experiment yielding hundreds or thousands of differentially expressed genes. As a result, it has become popular to analyze the data generated by such experiments by classifying them according to the probable molecular or biochemical function, biological process, or cellular compartment of their predicted gene product using tools such as Gene Ontology (GO), MapMan, and the Kyoto Encyclopedia of Genes and Genomes (KEGG). This analysis serves to give an overall picture of systems in the organism that are altered, dependent on the experimental treatment.

Here, we review the findings on the gene expression dynamics of plants during *Agrobacterium*-mediated transformation upon heterologous expression of *Agrobacterium*derived virulence proteins in the host, and upon treatment with PAMPs, with an emphasis on the commonalities and differences between *Agrobacterium* and other plant bacterial pathogens. We will then discuss the relative benefits and drawbacks of using modern transcriptomic methodologies to examine *Agrobacterium*–plant interactions. For ease of reference, Tables 1, 2, 3, and 4 summarizing the reviewed transcriptomic studies are listed at the end of this chapter.

2 Differential Gene Expression in Response to Inoculation with Various *Agrobacterium Tumefaciens* Strains

2.1 Transcriptomic Analysis of Model Plants During Agrobacterium-Mediated Transformation

Several studies have been carried out over the last 20 years that utilized different methods for detection of differential gene expression in response to inoculation with *Agrobacterium*. In a few cases, they have attempted to untangle the unique responses relating to particular steps of *Agrobacterium* infection and T-DNA transfer by various means. These include contrasting *Agrobacterium* response with that of other bacterial organisms, contrasting the responses to different *Agrobacterium* strains that are necessarily arrested at some point in the infection process owing to their genotype (including using tumorigenic versus non-tumorigenic strains), and temporally analyzing different infection stages during a time course. Other important differences exist among these studies with respect to plant species/genotype, tissue type sampled, culture conditions, inoculation method, *Agrobacterium* strain(s) used, and experimental design, which could account for some of the variability in results. Therefore, these details will be thoroughly outlined for each study.

In one of the earliest studies on the plant response to *Agrobacterium* (Ditt et al. 2001), researchers used a cell suspension culture of the tropical plant *Ageratum conyzoides* to construct a cDNA library of 16,000 sequences and then implemented cDNA-AFLP to identify cDNA sequences that were differentially regulated in response to the non-tumorigenic *Agrobacterium* strain EHA105. A total of 179 unique gene fragments were upregulated in response to *Agrobacterium*, whereas 72 were downregulated. Twenty of the most strongly induced of these sequences were used as queries to search for sequence similarity to other species. Top gene candidates in these similarity searches were involved in cellular functions such as signal perception, signal transduction, and defense. Using RT-PCR as an independent gene expression method, it was shown that four of the genes predicted to function in defense response were also induced by treatment with non-pathogenic *E. coli* cells, whereas two genes (one encoding a protein similar to a nodulin from *Oryza sativa*) and one encoding a protein similar to a lectin-like protein kinase from *Populus nigra* showed an *Agrobacterium*-specific response. This same research group later used the same experimental system to evaluate the expression of three of the putative defense genes identified in this study in response to the attachment of *Agrobacterium* to the host (Ditt et al. 2005). Plant cell cultures were infected with four different non-tumorigenic strains of *Agrobacterium*: EHA105 harboring the T-DNA binary vector pBISN1 (T-DNA transfer competent), LBA4404 (carrying a disarmed Ti plasmid, but no T-DNA), A136 (avirulent; no Ti plasmid) and *chvB* (contains an oncogenic Ti plasmid, but is attachment-deficient mutant). RT-PCR revealed that the ability of the *Agrobacterium* to attach to host cells had a negative effect on the expression levels of the three defense genes being measured compared to the attachment-deficient strain, which induced expression levels similar to that of

the non-pathogenic bacterial control. The authors suggested that exopolysaccharide (a product of the mutated gene in *chvB*) may play a role in suppressing the plant defense response.

In yet another published study by this group, the authors used a microarray to analyze wide-scale gene expression in suspension cell cultures of the *Arabidopsis thaliana* Ler-0 ecotype in response to infection by the tumorigenic *Agrobacterium* strain A348 along a time course (Ditt et al. 2006). Although this study used a 26,000 oligonucleotide array, surprisingly, no statistically significant differentially expressed genes were found at the first three time-point comparisons with the mock control (4, 12, and 24 h post-inoculation). In the comparison for the samples collected 48 h after infection, 303 differentially regulated genes were identified that showed consistency among the four experimental replications. Of these, 115 were upregulated and 188 were downregulated. GO biological process category analysis was performed on these two gene sets. The authors reported that genes falling into the "cell organization and biogenesis" and "protein metabolism" terms were overrepresented among the downregulated genes and "electron transport or energy pathways," "response to abiotic or biotic stimulus," and "response to stress" were overrepresented among the upregulated genes. They also compared their microarray data with other publicly available data sets evaluating the transcriptomic response of *Arabidopsis* to various other plant pathogens and to treatment with auxin compounds. Of these comparisons, the one that had the most overlapping differentially expressed genes with the *Agrobacterium* data set (at 53 genes) was one from mature *Arabidopsis* leaves that had been infiltrated with *Pseudomonas syringae* (Tao et al. 2003).

Working in a tobacco BY-2 suspension cell culture system, researchers in a 2003 study used suppression subtractive hybridization and macroarrays to determine changes in gene expression in host cells inoculated with *Agrobacterium* (Veena et al. 2003). Cells were exposed to both virulent and avirulent non-tumorigenic *Agrobacterium* strains with or without the capability to transfer Vir proteins or to transfer T-DNA containing a GUS-intron reporter gene. They found that a suite of genes related to defense responses including glutathione-S-transferase genes and alcohol dehydrogenase were induced in the earlier time-points following infection by *Agrobacterium*, regardless of its ability to transfer T-DNA. However, during the later time-points the relative expression of these genes was higher among cells infected by the avirulent *Agrobacterium* strain. Another interesting finding was that genes associated with cell division and plant growth such as the core histone gene family members encoding H2A, H2B, H3, and H4 along with ribosomal proteins were increased in their expression levels by exposure to the virulent strain at the later time-points after infection significantly above the levels induced by the avirulent strain at the same stage of infection. This timing coincided with the earliest point that expression of T-DNA could be detected in protoplasts, suggesting that these genes may play some important role in T-DNA integration specifically.

A study using mature *A. thaliana* plants examined the changes in both gene expression and phytohormone accumulation in response to infection by the tumorigenic *Agrobacterium* strain C58 or by the non-tumorigenic strain GV3101

(Lee et al. 2009). The lower part of inflorescence stalks of plants was inoculated just above the basal leaves. This location was chosen in order to allow the formation of crown gall tumors. Using microarray analysis to sample transcript levels at three different time-points (3 h and 6 days for both strains, and 35 days post-inoculation for the tumorigenic strain only), they found that over four times as many *Arabidopsis* genes were significantly altered in their expression levels upon exposure to the tumorigenic *Agrobacterium* strain as were altered by the disarmed strain using wounded, but uninoculated, inflorescence stalks as a control. It was found that both strains induced genes falling into the MapMan functional category of "stress," while hormone-related genes were affected by both, but with a stronger response to strain C58. Genes of functional classes associated with changes in host morphology were activated only by exposure to C58, but not by GV3101.

A recently published study that was conducted in our laboratory used next-generation sequencing to characterize the changes to the whole transcriptome through time in *A. thaliana* seedlings infected with either of two different non-tumorigenic strains of *Agrobacterium*: strain At804 (virulent) or A136 (avirulent). Seedlings were sampled across a time course from 0 to 48 h after infection. As was found in the two studies previously discussed in this section, substantially more genes were significantly altered in their expression levels by exposure to the T-DNA transfer competent bacterial strain than by the avirulent strain relative to a mock treatment. Because a greater number of transcripts can be detected and there is a greater sensitivity of measurement using RNA-Seq compared to the older differential gene expression techniques, we were able to find many more genes with altered expression levels than had been reported in previous studies. Using Gene Ontology (GO) enrichment analysis, Duan et al. (2018) found that categories such as "cell wall organization or biogenesis," "DNA replication," and "external encapsulating structure organization" were all overrepresented among downregulated genes and that "defense response," "response to stress," and "response to reactive oxygen species" were overrepresented among upregulated genes for the treatment for both strains. This agrees with the previous findings that indicate *Agrobacterium* seems to generally repress normal plant growth and cell division and to activate defense response pathways (at least in the early stages of infection) irrespective of its ability to transfer T-DNA and Vir proteins to the host. We found that in the latter time-points following infection that certain functional categories were unique to treatment with one strain or the other among upregulated genes. For instance, "cellular response to stress" and "secondary metabolite biosynthetic process" were found only in upregulated genes from the avirulent treatment, indicating that these responses may be attenuated by transfer of Vir proteins and/or T-DNA. On the other hand, categories having to do with cell growth, transcription and RNA metabolism, as well as "heterocycle metabolic process" and "response to abscisic acid," were found in upregulated genes only in the virulent treatment condition. It was suggested in Veena et al. (2003) that genes associated with cell division and growth processes were induced by successful *Agrobacterium*-mediated transformation. The results of our recent study support this

conclusion along with suggesting other biological and metabolic processes that may be activated specifically by Vir genes or T-DNA.

2.2 *Transcriptomic Analysis of Crop Species During Agrobacterium-Mediated Transformation*

The majority of gene expression studies that have been used to uncover the interaction between *Agrobacterium* and host plants have been carried out using *A. thaliana* and, to some extent, tobacco suspension cell cultures. However, there have been three studies published within the last few years that have made use of transcriptomic data to examine the effect of *Agrobacterium* on economically important crop species. In each of these studies, transcriptome profiling by microarray or RNA-Seq was one experimental approach of several that were used to gain insight into overcoming a crop plant's recalcitrance to transformation.

In Tie et al. (2012), the researchers used microarrays to measure gene expression throughout infection by *Agrobacterium* strain EHA105 in the two rice cultivars "Nipponbare" (Nip) and "Zhenshan 97" (ZS), representing *O. sativa* ssp. *japonica* and ssp. *indica*, respectively. Representatives from both subspecies were selected for comparison because *indica* varieties of rice are overall much more resistant to transformation than are *japonica* varieties. Embryogenic calli from either variety were sampled at 0, 1, 6, 12, and 24 h after inoculation. The highest number of differentially regulated genes occurred at the 1 or 6 h time-points, and of all unique locus identifiers that were differentially expressed in either type of callus over time (11,105 sequences) only 35% were shared between the two callus cultivars. GO enrichment analysis found some differences in biological process terms between the callus types when contrasted at the same time-point. Genes involved in "defense response" and "response to biotic stimulus" were overrepresented among the set of upregulated genes in the *indica* callus variety. Meanwhile, categories relating to "cell cycle," "cell division," and "DNA repair" were overrepresented among the downregulated genes in ZS callus at the earliest time-points. Some genes that are involved in ubiquitin-proteasome degradation were repressed in the more recalcitrant variety, leading the authors to suggest that degradation of the proteins coating the T-complex may be inhibited in ZS, giving rise to a deficiency in T-DNA integration and lowering the overall transformation frequency.

Another study combined transcriptomic analysis (using RNA-Seq) and proteomics (2-DE and MS) to uncover *Agrobacterium*–host interactions in transformed wheat immature embryo tissue (Zhou et al. 2013). Embryos of the Chinese commercial wheat variety "Yangmai12" were extracted from immature seed and pre-cultured for 4 days before being transformed with the non-tumorigenic *Agrobacterium* strain C58C1. Tissue samples were collected 36 h following inoculation. In total, the researchers found 4889 genes that showed significant differential

expression compared with the mock treatment. GO biological process term analysis showed categories such as "chromatin assembly or disassembly," "signal transduction," "biosynthesis of secondary metabolites," and "phenylpropanoid biosynthesis" were overrepresented among these genes. The molecular function terms showed that a large portion (20.5%) of the functionally classified genes fell into the "nucleic acid binding" category. The proteomic analysis uncovered 90 differentially expressed proteins (DEPs) between the two conditions. Notably, only 24 of these DEPs corresponded to gene sequences found in the transcriptomic data. Of these overlapping DEPs, half of them were predicted to play a role in response to stress or immunity. The differences observed between the results of the transcriptomic and proteomic analyses in this study demonstrate that measured mRNA levels do not perfectly correspond to final expression of a gene product.

In a study aimed at uncovering the mechanism of the observed improvement of *Agrobacterium*-mediated transformation in soybean by employment of sonication and α-aminooxyacetic acid (AOA), RNA-Seq was used to monitor transcriptomic changes in response to *Agrobacterium* infection combined with sonication of the explant tissue and media amended with AOA or to *Agrobacterium* infection without these additions to the protocol (standard *Agrobacterium* transformation) (Zhang et al. 2016). Cotyledonary nodes of the soybean genotype "Jidou17" were infected with *Agrobacterium* EHA105 with or without sonication plus AOA during inoculation and samples were collected five hours after infection. As in the previous studies, plant defense and immune responses were upregulated dependent on *Agrobacterium* infection. A total of 2158 differentially expressed genes (55.1% of them upregulated) were responsive to standard *Agrobacterium* transformation in contrast to mock treatment, whereas 5062 genes showed a difference in expression between standard transformation and transformation plus sonication and AOA with 69.6% of these representing downregulated genes. Upregulated and downregulated genes were categorized by their pathway function using KEGG. Interestingly, this study showed that in addition to defense genes related to PTI (which *Agrobacterium* has been long understood to stimulate) "R" genes known for participating in ETI such as RPM1, RPS2, RPS5, RIN4, and PBS1 were also induced. This is notable as *Agrobacterium* had not previously been shown to induce ETI. Consistent with the role of AOA as an inhibitor of phenylpropanoid biosynthesis, the authors showed that genes involved in this metabolic process are upregulated by *Agrobacterium* treatment (in agreement with the study discussed above), but their expression is ameliorated by sonication and AOA supplementation in tissue culture following transformation.

2.3 Evaluation of GO Terms and Individual Genes Common to Gene Sets from Multiple Studies

In spite of the variety of host plant species, explant tissues, culture systems, *Agrobacterium* strains, and timing of sample collection used in the above studies, there are common patterns seen in many of the data sets. Using GO biological

process terms as a representation of wide-scale gene expression changes, processes such as defense response, plant cell wall modification, cell growth and development and transcription have been shown to be consistently altered by *Agrobacterium* infection. For GO categories overrepresented among upregulated genes in response to disarmed, virulent *Agrobacterium* strains in four studies (Tie et al. 2012; Zhou et al. 2013; Duan et al. 2018; and Niu et al. 2015, discussed below) that had such data readily available, ones that showed the most consistency (called in three out of four datasets) were the broad classes "DNA-dependent," "response to stress," and "transport." Other categories of note included "regulation of gene expression," "regulation of macromolecule metabolic process," "M phase of mitotic cell cycle," and "response to oxidative stress." GO terms for the downregulated gene sets were less consistent between studies with only seven terms shared between at least two studies. These categories included "catabolic process," "cell wall polysaccharide metabolic process," and "signal transduction." For a complete list of GO terms overlapping between datasets from these studies, see Table 3. In the same way, comparing sets of individual transcripts that were upregulated in response to virulent disarmed *Agrobacterium* (or in response to expression of *vir* genes) across studies yielded some consensus genes worth evaluating. A total of 97 genes were identified in at least two of the three datasets compared (Lee et al. 2009; Niu et al. 2015; Duan et al. 2018). The three genes that were present in all three datasets encoded an oxidoreductase (FOX1), caffeoyl CoA 3-O-methyltransferase (CCOAMT; a component of the phenylpropanoid biosynthesis pathway), and a cytochrome P450 protein (CYP71A12). Other notable genes shared by two of the datasets coded for the VirE2-binding F-box protein (VBF), a non-symbiotic hemoglobin (HB1; GLB1), several peroxidase proteins (including PER37, PER71, and PER4), and a regulator of the chromosome condensation family protein (AT4G14368). These concurring upregulated genes could serve as potential candidates for future functional validation of their role in the *Agrobacterium*–plant interaction. For a complete list of these genes, see Table 4.

3 Transcriptional Profiling in Response to PAMP Exposure

A plant's first line of defense against pathogen attack depends on its ability to perceive and respond to pathogen-associated molecular patterns (PAMPs). These usually consist of evolutionarily conserved structural molecules (or epitopes within these molecules), and thus hosts are able to recognize broad ranges of potential pathogens for each PAMP. PAMP perception induces a series of responses known as PAMP-triggered immunity (PTI), which includes reactive oxygen species (ROS) burst, calcium flux, MAP kinase activation, ethylene production, callose deposition, and transcriptional reprogramming (Zipfel and Robatzek 2010). One of

the most well-known bacteria-derived PAMPs is flg22, derived from flagellin, which is perceived by the protein receptor FLS2 in *Arabidopsis* (Gómez-Gómez and Boller 2000; Chinchilla et al. 2006). Recognition of flg22 confers resistance to several bacterial pathogens; however, *Agrobacterium* is able to evade the perception of its flagellin due to a mutation near the N-terminus of the protein (Felix et al. 1999; Kunze et al. 2004). Another bacteria-derived PAMP, Elongation Factor Tu (EF-Tu), is recognized by *Arabidopsis* plants upon exposure to *Agrobacterium*. This response is mediated through the recognition of receptor protein EFR, and there is some convergence between the signaling responses initiated by either flagellin or EF-Tu perception. Furthermore, *Arabidopsis efr* mutants show increased susceptibility to *Agrobacterium* (Zipfel et al. 2006). Clearly, PTI has an important role to play during the *Agrobacterium* transformation process. Several studies have used transcriptomic profiling to investigate the response to PAMP perception. Although some of these PAMPs were derived from bacteria other than *Agrobacterium*, these studies can yield insight into which genes form part of a common defense pathway and which, if any, are more of an *Agrobacterium*-specific response.

Zipfel et al. (2004) explored the gene expression response of *Arabidopsis* Ler-0 seedlings 30 min after treatment with flg22 using an Affymetrix whole-genome array chip. In total, they found that 966 genes were upregulated and 202 genes were downregulated. Treatment of *Arabidopsis fls2* mutants with flg22, however, only altered the expression patterns of six genes when compared with the control, none of which were contained in the set of those that showed a response to flg22 in the wild-type plant. This result showed that FLS2 is solely required for the flg22-dependent defense response to take place. Of the upregulated genes with known functions, a large fraction of them were predicted to be involved in signal perception, signal transduction, or transcriptional regulation. This included 155 receptor-like kinase (RLK) proteins, one of which was FLS2. Treatment of *fls2* mutant seedlings with crude extracts of pathogenic bacteria including *Agrobacterium* gave similar physiological responses to that of wild-type plants, indicating that there were PAMPs other than flg22 being perceived by the plant immune system. Another paper published by this group, in addition to identifying the *Arabidopsis* receptor-like kinase EFR as the receptor for EF-Tu and demonstrating that EF-Tu recognition occurs during *Agrobacterium*-mediated transformation, compared the transcriptomic responses to EF-Tu or flagellin perception (Zipfel et al. 2006). *Arabidopsis* seedlings were this time treated with the EF-Tu-derived peptides elf26 and elf18 and sampled at 30 and 60 min after induction. A total of 866 genes were called as upregulated and 83 were downregulated after 60 min. There was a high degree of overlap demonstrated between the genes affected by either EF-Tu or flagellin. For instance, over 100 of the 610 predicted RLK proteins in the *Arabidopsis* genome were induced by both PAMPs. Treatment with either PAMP also increased the abundance of receptor sites for the other PAMP after only a few hours, indicating a positive feedback in signaling shared between the two. Another noteworthy result is that *efr* mutant plants treated

with *Agrobacterium* crude extracts still showed a defense response, suggesting that there are yet other PAMPs from *Agrobacterium* that plants can perceive.

Other studies have explored, along similar lines, the transcriptomic effects of treatment with types of bacterial PAMPs other than flagellin and EF-Tu on *Arabidopsis* plants. One study looked at the response of mature *Arabidopsis* Col-0 leaves to infiltration with *Staphylococcus aureus*-derived peptidoglycan (PGN) after 4 h (Gust et al. 2007). When contrasted with a water-treated control, 236 genes were shown to be upregulated greater than twofold. Once again, many of these genes overlapped with those induced by flg22 at the same time after treatment. Some of the specific classes of these genes included chitinases, peroxidases, and the phenylpropanoid pathway enzyme phenylalanine ammonia lyase 1 (PAL1), along with WRKY and AP2 family transcription factors. Researchers in another study looked at differential gene expression induced by treatment with lipopolysaccharide (LPS), which is a cell membrane component of gram-negative bacteria including *Agrobacterium*, and harpin, a component of the type III secretion system which *Agrobacterium* does not possess (Livaja et al. 2008). They used *Arabidopsis* cell suspension cultures sampled at six time-points up to 24 h after treatment and analyzed gene expression via microarrays. A total of 1573 genes had altered expression levels greater than twofold by either PAMP with 313 responding to both, 309 unique to LPS, and 915 unique to harpin. Harpin showed a much quicker and stronger response than did LPS, with the highest amount of DE genes overlapping between the two PAMPs being observed with harpin 30 min after treatment and LPS 24 h after treatment. There was a substantial difference in the induction of several gene categories between treatment with harpin and LPS, including WRKY family transcription factors, genes associated with ROS burst, and MAP kinase signaling components.

These studies that explore the host transcriptomic response to treatment by bacterial PAMPs can give clues into the mechanism of plant defense against *Agrobacterium*-mediated transformation. Even though *Agrobacterium* does not induce PTI through flagellin or harpin perception, we can see from these studies that there is a high degree of commonality in the response to each of these PAMPs, and thus, *Agrobacterium* must overcome some of the same barriers to infection as do other bacterial plant pathogens. Comparing the sets of genes upregulated or downregulated by PAMP treatment (Zipfel et al. 2004, 2006; Gust et al. 2007; Livaja et al. 2008), treatment with avirulent *Agrobacterium* strain A136 (Duan et al. 2018) yielded 400 genes that were shared between at least three of the sets. Of these, 212 of them were not upregulated upon treatment with any virulent *Agrobacterium* strain or upon VirE3 expression. Some notable genes included in this set were the transcription factor *WRKY33*, *MILDEW RESISTANCE LOCUS O 12* (*MLO12*), a PBS1-like gene (*AT3G55450*), *ETHYLENE RESPONSIVE ELEMENT BINDING FACTOR 4* (*ERF4*), and a paralog of the *Arabidopsis* defense signaling gene *NPR1* (*NPR3*). There were only 10 genes represented in three or more of the datasets for PAMP-downregulated genes. These included *PHOSPHOENOLPYRUVATE*

CARBOXYLASE KINASE (*PPCK*), *EXPANSIN-LIKE A1* (*EXLA1*), and a type A cytokinin signaling gene *Arabidopsis RESPONSE REGULATOR 6* (*ARR6*). Some of the genes that are upregulated by PAMP treatment, but not by virulent *Agrobacterium* or Vir proteins, could be those that *Agrobacterium* is able to suppress the expression of in order to ameliorate the plant defense response. See Table 2 for a complete list of these consensus PAMP-responding genes. It is plausible that the variations in the expression of the some of the key genes involved in this common defense response may account for differences seen in the amenability of different cell types, explant tissue, or plant genotypes/species to T-DNA transfer. Transcriptomic profiling, in combination with other molecular techniques, will help to shed light on the key molecules and pathways that could potentially be targeted for improving plant transformation capabilities.

4 Transcriptional Response to Heterologous Expression of *Vir* Genes in Host Plants

There have been two studies published so far that used transgenic *Arabidopsis thaliana* plants expressing *Agrobacterium*-derived *vir* genes in order to examine differences in gene expression compared with wild-type plants. In Niu et al. (2015), the investigators transformed *Arabidopsis* Col-0 plants with the *virE3* gene sequence under the control of a tamoxifen-inducible promoter. The VirE3 protein had been previously implicated as a possible plant-active transcription factor because it could induce transcription in yeast when fused with a DNA-binding domain (García-Rodríguez et al. 2006). Fourteen-day-old seedlings of these *virE3* transgenic plants were treated with tamoxifen to induce transgene expression and RNA-Seq was used to profile the differential gene expression compared to wild-type and mock-treated transgenic plants. They found 607 genes that were upregulated and 132 that were repressed specifically by *virE3* expression (using a fold-change cutoff of 3 or 0.33). Among the upregulated genes were those encoding the plant protein VBF (a functional homolog of the *Agrobacterium* VirF protein), which had been previously shown to be induced by *Agrobacterium* infection and to play a role in destabilizing VirE2 (Zaltsman et al. 2010). Another gene that they found to be highly induced by VirE3 was that coding for NIMIN1, which binds the salicylic acid signaling protein NPR1 and reduces expression of the pathogenesis-related (PR) genes. (Weigel et al. 2005). They showed through additional experiments that the promoters of both *VBF* and *NIMIN1* could be bound by VirE3. The authors performed GO term enrichment analysis on their datasets of differentially expressed genes. Biological process categories including "signal transduction," "response to stress," "DNA-dependent transcription," and "developmental processes" were overrepresented among upregulated genes in tamoxifen-induced VirE3 plants. Examples of the molecular

function categories that were highlighted are “kinase activity,”, “transporter activity,” and “Transcription factor activity.”

Duan et al. (2018) produced transgenic *virE3* as well as *virE2 Arabidopsis* Col-0 plants with their expression under the control of the CaMV 35S promoter. These plants were tested for expression of *virE3* and *virE2* mRNA, respectively, using qRT-PCR. As determined by qRT-PCR that for both transgenic plants, the expression levels of defense genes, including *PR3*, *CRK41*, and *CRK18*, were elevated compared to wild-type controls upon treatment with the avirulent (cured) *Agrobacterium* strain A136. Meanwhile, the defense genes *FRK1*, *PR2*, and *PR4* showed increased expression only in the *virE3*-overexpression plants. These limited gene expression data can be interpreted to indicate that overexpression of VirE2, and to a greater extent VirE3, enhances the plant defense response to *Agrobacterium*, though an alternative interpretation is that constitutive expression of the virulence transgenes has interrupted the normal functioning of the defense pathways in the plant.

Heterologous expression of *vir* genes in plants is an attractive method for investigating their specific functions as it allows the gene expression changes they induce in the host to be decoupled from expression changes caused by the presence of the *Agrobacterium* itself (i.e., defense responses due to induction of PTI). It would be far more difficult to evaluate gene expression from the same angle using modified *Agrobacterium* strains to induce gene expression changes. However, there are necessarily caveats when attempting to compare the expression of a virulence gene in the host cell to what actually occurs during the infection process. For example, delivery of the molecule from *Agrobacterium* to the host may be precisely controlled, and thus, expression levels of the transgene from the plant may not correlate with the level of protein that occurs in *Agrobacterium*-mediated transformation. It is also possible that differential expression of some genes may represent secondary transcriptional effects following induction of the virulence gene, which do not represent the normal function of the *vir* gene during T-DNA transfer. Taking this into consideration, it is obviously advantageous to design transgene cassettes using an inducible system such as that used by Niu et al. (2015) rather than constitutive expression. This way at least the timing of expression can be controlled, potentially allowing for temporal isolation of primary transcriptional effects from those of more indirect responses. The *Arabidopsis* genes *VBF* and *NIMIN1*, which Niu et al. (2015) demonstrated could be transcriptionally regulated by VirE3, were both included in a set of genes of upregulated in response to both tamoxifen-induced VirE3 expression and treatment with virulent *Agrobacterium* strain At804 (see Table 2). Surprisingly, however, both of these genes were almost equally upregulated in seedlings treated with either At804 or the avirulent strain A136 relative to the mock treatment (Duan et al. 2018). This result suggests either that some *Agrobacterium*-derived PAMP, or at least an extracellular signal, initiates a VirE3-independent route by which these genes can be induced, or that VirE3 is dispensable not only for overall T-DNA transfer efficiency, but also for its role in

transcriptional activation, and thus is functionally redundant with some host protein. This question could be resolved by evaluating the expression levels of these genes in plants treated with a *virE3* mutant *Agrobacterium* strain.

5 Advantages and Limitations of Using Transcriptomic Analysis as an Approach to Discover Plant Genes and Pathways Associated with *Agrobacterium*-Mediated Transformation

There are obvious advantages to employing a transcriptomic approach for the purpose of gene discovery and pathway elucidation. Transcriptome-wide analysis allows for a quick survey of nearly all expressed genes in a given organism under specific conditions at a relatively low cost. The assays used are very sensitive and can give an accurate reflection of gene expression levels at specific times. Relative to older techniques, these methods also have a high degree of reproducibility. The advent of next-generation sequencing technology has made transcriptomic analysis the preferred choice for gene discovery and, as demonstrated above, it has been instrumental for uncovering the molecular players in specific pathways such as those involved in the *Agrobacterium*-mediated transformation process. Some of the main reasons for the rapid adoption and wide use of RNA-Seq include the low background signal, its capability for detecting a wide dynamic range of expression, and its generation of novel sequence data at a single-base resolution. Unlike microarray technology, it is not reliant on preexisting sequence information and is not susceptible to false signals generated by potential cross-hybridization of similar transcript species to the same sequence probes. All of this allows a much greater amount of much higher quality data to be generated than was previously possible for gene expression studies.

As a tool for gene discovery, transcriptomic analysis has intrinsic limitations. The sequences detected by these methods, after all, reflect only mRNA transcripts and not the final products of the respective gene's expression (i.e., protein or metabolites). Thus, a survey of the transcriptome is blind to the effect of translational or posttranslational controls on gene expression. Another drawback to RNA-Seq is that the amount of data generated can be overwhelming and requires significant storage space and some specialty to manage and analyze the data. Too much data availability can be an obstacle to gene discovery as it is often difficult to decide which genes to study further out of a large pool of candidates. Furthermore, the degree of differential expression does not necessarily indicate which genes have the most critical functions in a given pathway because genes that are likely induced through multiple steps in a regulatory cascade often display the highest fold changes, whereas genes that function earlier in a pathway may show comparatively subtler effects. Additionally, some portion of the genes called as differentially

expressed may have had their expression patterns altered as secondary effects from the original treatment and may not, therefore, have any relationship to the process being studied. These drawbacks limit the conclusions that can be drawn about the genes involved in *Agrobacterium*-mediated transformation from transcriptomic data alone. However, by using other molecular biology techniques to evaluate characteristics of pathways such as protein–protein interactions and protein-DNA binding, as well as using transformation experiments to validate the importance of specific genes, we can compensate for the constraints inherent in transcriptomic analysis in order to develop a more complete picture of the molecular mechanism underlying this process.

6 Summary and Outlook

Taken together, the studies that have used differential gene expression to examine the host response to *Agrobacterium* or elicitor compounds have given us a fuller picture of which genes could be important in *Agrobacterium*–plant interactions. However, many questions still need to be addressed in order to apply our understanding of the process toward the development or improvement of *Agrobacterium*-mediated transformation techniques for the benefit of crop breeding and basic research (Altpeter et al. 2016). The interaction between *Agrobacterium* and plants is complex and multi-layered. More focused transcriptomic analyses of specific pathways known to be involved would be desirable to maximize the utility of this approach. Next-generation sequencing technology has allowed greater sensitivity when measuring gene expression and offers attractive novel possibilities for experimental designs, which could give a more refined view of molecular processes. For instance, in order to explore gene regulatory interactions that occur in response to *Agrobacterium*, RNA-Seq could be combined with ChIP-Seq to determine the direct induction of genes by key transcription factors.

With respect to biotechnological applications, one could make improvements to the transformation process in various ways informed by differential gene expression data in response to *Agrobacterium*. For example, an *Agrobacterium* strain was modified to elicit a weaker plant defense in a study on potato (Vences-Guzman et al. 2013). In other studies, the composition of tissue culture medium and/or growth conditions was modified in order to minimize plant defense responses (Zhang et al. 2013, 2016). Because of the similarity of the plant defense response between *Agrobacterium* and other bacterial pathogens, effector proteins originating from other bacteria have even been utilized to repress plant defense gene induction. In one instance, the AvrPto protein from *Pseudomonas syringae* was conditionally expressed in *Arabidopsis* plants leading to higher efficiency of *Agrobacterium*-mediated transformation due to its suppression of plant defenses (Tsuda et al. 2012). As more of the mechanisms involved in *Agrobacterium*–plant interactions are uncovered, plant transformation researchers will be able to apply some of these discoveries to make new improvements to transformation techniques.

Table 1 Summary of transcriptomic analysis of *Agrobacterium*-mediated plant transformation or PAMP exposure

Authors	Analysis method	*Agrobacterium* strain (tumorigenic or disarmed) or PAMP	Plant materials/treatment/time frame	Key discovery
Ditt et al. (2001)	cDNA-AFLP	EHA105 (pBISN1); disarmed	*Ageratum conyzoides* cell cultures or Tobacco BY-2 cell cultures 24 or 48 h post-inoculation	Plant cells rapidly responded to *Agrobacterium* infection. About 250 unique gene fragments differentially regulated
Veena et al. (2003)	Suppressive subtractive hybridization and DNA macroarrays	At739, At1221 (avirulent), At543, At804, At1222 (virulent); disarmed	Tobacco BY-2 cell suspension cultures at 0 to 36 h post-inoculation	Transfer of T-DNA/Vir proteins modulated the expression of host genes differently than transfer-deficient *Agrobacterium*
Zipfel et al. (2004)	Affymetrix ATH1 whole-genome microarray chip	None; flg22	*Arabidopsis* Ler-0 seedlings 30 min after flg22 treatment	1168 genes differentially regulated including 155 RLKs
Ditt et al. (2006)	*Arabidopsis* 26,000-gene oligonucleotide microarray	A348; tumorigenic	*Arabidopsis* cell cultures at 4 to 48 h post-inoculation	Differential expression only observed at 48 h time-point. Induced genes encode known defense proteins and repressed genes encode proteins involved in cell proliferation
Zipfel et al. (2006)	Affymetrix ATH1 whole-genome microarray chip	None; elf26	*Arabidopsis* Ler-0 seedlings 30 and 60 min after elf26 treatment	427 genes were upregulated and 7 were suppressed at 30 min following treatment with elf26. 866 were upregulated and 83 were downregulated at 60 min. A large portion overlap with those induced by flg22
Gust et al. (2007)	Affymetrix ATH1 whole-genome microarray chip	None; Peptidoglycan (PGN)	*Arabidopsis* Col-0 mature plants at 4 h following infiltration with PGN	236 genes were upregulated. A large involved in defense response including several *WRKY* transcription factors
Livaja et al. (2008)	Agilent *Arabidopsis* cDNA microarray	None; Lipopolysaccharide (LPS), harpin	*Arabidopsis* cell suspension culture at 30 min to 24 h following treatment with either LPS or harpin	1573 genes showed altered expression pattern in response in total with 313 responding to both PAMPs. LPS showed much weaker and later defense induction than harpin

(continued)

Table 1 (continued)

Authors	Analysis method	*Agrobacterium* strain (tumorigenic or disarmed) or PAMP	Plant materials/treatment/time frame	Key discovery
Lee et al. (2009)	Affymetrix ATH1 whole-genome microarray chip	C58 (tumorigenic), GV3101 (disarmed)	*Arabidopsis* inflorescence stalks at 3 h and 6 d post-inoculation, and 35-day-old tumors	Pathogen defense genes responded predominantly to oncogenic strain C58. Host responses were much stronger toward the oncogenic strain C58 than to the disarmed strain GV3101
Tie et al. (2012)	Microarray	EHA105; disarmed	*Oryza sativa* ssp. *japonica* genotype "Nipponbare" and ssp. *indica* genotype "Zhenshan 97" callus tissue at 0 to 24 h post-inoculation	Representative genotypes of japonica and indica rice showed divergent transcriptional responses to transformation with EHA105. Defense response was stronger in Zhenshan 97, cell division and growth was repressed
Zhou et al. (2013)	RNA-Seq	C58C1; disarmed	*Triticum aestivum* genotype "Yangmai12" 4-day pre-cultured immature embryo tissue at 36 h post-inoculation	4889 wheat genes showed significant differential expression. Major categories included stress or immunity response and secondary metabolite biosynthesis
Niu et al. (2015)	RNA-Seq	None; *virE3* tamoxifen-inducible transgenic line	*Arabidopsis* seedlings at 4 h after tamoxifen treatment	Expression of *virE3* upregulated 607 genes > three-fold and repressed 132 genes > three-fold
Zhang et al. (2016)	RNA-Seq	EHA105; disarmed	*Glycine max* genotype "Jidou17" cotyledonary nodes with or without sonication and AOA treatment at 5 h post-inoculation	AOA and sonication treatment, which leads to greater transformation efficiency, suppresses expression of some defense-related genes and isoflavone biosynthesis enzymes relative to *Agrobacterium* alone
Duan et al. (2018)	RNA-Seq	A136, At804; disarmed	*Arabidopsis* Col-0 seedlings at 0 to 48 h post-inoculation	*Agrobacterium* infection suppressed genes important for plant growth and development but induced defense response genes

Table 2 *Arabidopsis* genes upregulated in at least three studies by PAMPs or avirulent *Agrobacterium* strain, but not upregulated by any virulent *Agrobacterium* treatment, and genes downregulated by PAMPs

	Function	Gene names
Upregulated PAMP genes	Carbohydrate metabolism	UGT74B1 (AT1G24100), AT1G78820, AT2G15480, AT2G27500, PEN2 (AT2G44490), AT3G10720, CHIV (AT3G54420), AT3G62720, AT4G02330, AT4G30290, AT5G03700, AT5G18470
	Cell growth related	EXLA1 (AT3G45970), PSK4 (AT3G49780)
	Defense signaling	MLO2 (AT1G11310), PP2-A5 (AT1G65390), RLP12 (AT1G71400), PSKR1 (AT2G02220), AT2G38870, MLO12 (AT2G39200), NHL2 (AT3G11650), SYP122 (AT3G52400), EXO70H1 (AT3G55150), MPK5 (AT4G11330), RAF27 (AT4G18950), NHL3 (AT5G06320), NPR3 (AT5G45110), EXO70H7 (AT5G59730), BON1 (AT5G61900), PROPEP3 (AT5G64905), MAPKKK5 (AT5G66850)
	DNA modification	PEARLI4 (AT2G20960), AT4G35110
	Hormone metabolism	GA2OX6 (AT1G02400)
	Hormone signaling	JAZ1 (AT1G19180), AT5G35735
	Metabolic enzyme—phenylpropanoid biosynthesis	CCR2 (AT1G80820)
	Other metabolic enzyme	KCS1 (AT1G01120), AT1G17420, AT1G30370, AT1G55450, AT1G71697, AT2G39400, SDR5 (AT2G47140), AT3G15530, AGK2 (AT3G57550), APS3 (AT4G14680), AT4G24160, AT4G24380, GGT1 (AT4G39640), NADP-ME2 (AT5G11670)
	Protein folding	J20 (AT4G13830)
	Protein metabolism	PUB20 (AT1G66160), AT2G35930, PUB22 (AT3G52450), AT3G59080, SGT1A (AT4G23570), PUB51 (AT5G61560)
	Protein modification	PAPP2C (AT1G22280), AT3G59350, AT5G36250, AT5G46080
	Receptor-like kinase	LRR XI-23 (AT1G09970), LRK10L1.2 (AT1G18390), AT1G25390, AT1G51790, AT1G51800, AT1G51850, AT1G61370, SD1-29 (AT1G61380), AT1G70530, RIPK (AT2G05940), SOBIR1 (AT2G31880), AT2G33580, BIK1 (AT2G39660), PTI1-4 (AT2G47060), AT3G02880, CERK1 (AT3G21630), PBL1 (AT3G55450), AT4G08850, CRK10 (AT4G23180), CRK11 (AT4G23190), LecRKA4.1 (AT5G01540), AT5G20050, AT5G22690, AT5G35370, AT5G40170, AT5G41550, AT5G46330, AT5G58120, AT5G60270

(continued)

Table 2 (continued)

	Function	Gene names
	Stress response—cytochrome P450	CYP707A3 (AT5G45340)
	Stress response—glutathione-S-transferase	GSTF8 (AT2G47730)
	Other stress response	HEMA2 (AT1G09940), ACA.l (AT1G13210), PHO1;H3 (AT1G14040), CAD1 (AT1G29690), AT1G33590, AT1G33600, AT1G61360, AT1G72900, AT1G72920, NUDT21 (AT1G73540), UCP5 (AT2G22500), HSPRO2 (AT2G40000), ATL6 (AT3G05200), TH9 (AT3G08710), ATSIK (AT3G08760), ACA.k (AT3G22910), DTX18 (AT3G23550), ERD5 (AT3G30775), CAF1a (AT3G44260), ZAR1 (AT3G50950), CNGC13 (AT4G01010), AT4G02200, NUDT7 (AT4G12720), ATL17 (AT4G15975), CBL1 (AT4G17615), CML42 (AT4G20780), AT4G31000, AT4G39830, AT5G12940, CAF1b (AT5G22250), AT5G24430, FC1 (AT5G26030), AT5G39020, CAD1 (AT5G44070), AT5G44910, RDUF2 (AT5G59550), ACHT5 (AT5G61440)
	Transcription factor	MYB51 (AT1G18570), SCL1 (AT1G21450), AT1G21910, AT1G28370, AT1G35350, BZIP60 (AT1G42990), ERF8 (AT1G53170), WRKY6 (AT1G62300), RAV2 (AT1G68840), WRKY17 (AT2G24570), AT2G42360, MYB2 (AT2G47190), ATL2 (AT3G16720), AT3G49530, GATA-8 (AT3G54810), WRKY22 (AT4G01250), HAT1 (AT4G17460), ERF1 (AT4G17500), WRKY28 (AT4G18170), SHB1 (AT4G25350), WRKY11 (AT4G31550), RHA3B (AT4G35480), MYB73 (AT4G37260), ZAT6 (AT5G04340), WRKY72 (AT5G15130), ERF5 (AT5G47230), BZIP01 (AT5G49450), MYB96 (AT5G62470), AZF1 (AT5G67450)
	Uncategorized	AT1G03740, AT1G18380, AT1G19380, AT1G21010, AT1G22890, AT1G51920, AT1G52200, ZCF37 (AT1G59590), FH7 (AT1G59910), AT1G65400, AT1G70740, AT1G72060, AT1G80450, J8 (AT1G80920), AT2G16900, AT2G18680, VQ12 (AT2G22880), AT2G24600, AT2G25735, AT2G27660, AT2G27830, AT2G32140, PHT1;4 (AT2G38940), AT2G44500, AT3G13430, AT3G18560, AT3G19010, AT3G25600, AT3G46110, CaM9 (AT3G51920), AT3G55470, SIB1 (AT3G56710), AT3G56880, AT3G57450, AT3G60420, SAG21 (AT4G02380), AMT1;1 (AT4G13510), AT4G20000, TET9 (AT4G30430), LHT7 (AT4G35180), SEN1 (AT4G35770), AGP18 (AT4G37450), AT5G08350, EXL4 (AT5G09440), AT5G14700, AT5G18150, AT5G20400, AT5G28630, TCH2 (AT5G37770), AT5G42010, AT5G47710, AT5G53110, EXL2 (AT5G64260)

(continued)

Table 2 (continued)

	Function	Gene names
Downregulated PAMP genes	Carbohydrate metabolism	PPCK1 (AT1G08650), XTH12 (AT5G57530)
	Cell growth related	EXLA1 (AT3G45970)
	Hormone signaling	ARR6 (AT5G62920)
	Other stress response	GRXS8 (AT4G15660)
	Transcription factor	DEWAX (AT5G61590), LBD39 (AT4G37540)
	Transport	UMAMIT29 (AT4G01430), MFS1 (AT4G34950), G3Pp1 (AT3G47420)

Table 3 GO biological process categories differentially regulated by *Agrobacterium* infection or Vir proteins in more than one of the reviewed studies

	Studies represented	GO categories	Number of studies represented
Upregulated categories	Tie et al. (2012) (Nipponbare); Niu et al. (2015); Duan et al. (2018) (At804)	Response to stress	3
		Transport	
	Zhou et al. (2013); Niu et al. (2015); Duan et al. (2018) (At804)	DNA-dependent	3
	Tie et al. (2012) (Nipponbare); Duan et al. (2018) (At804)	Nitrogen compound metabolic process	2
		Biological regulation	
		Carboxylic acid metabolic process	
		Cellular metabolic process	
		Cellular nitrogen compound metabolic process	
		Cellular process	
		Cellular protein metabolic process	
		Cellular response to chemical stimulus	
		Cellular response to organic substance	
		Cellular response to stimulus	
		Establishment of localization	
		Hormone-mediated signaling pathway	
		Metabolic process	
		Monocarboxylic acid metabolic process	
		Nucleic acid metabolic process	
		Organic acid metabolic process	
		Oxoacid metabolic process	
		Regulation of RNA metabolic process	

(continued)

Table 3 (continued)

	Studies represented	GO categories	Number of studies represented
		Regulation of biological process	
		Regulation of biosynthetic process	
		Regulation of cellular biosynthetic process	
		Regulation of cellular metabolic process	
		Regulation of cellular process	
		Regulation of gene expression	
		Regulation of macromolecule biosynthetic process	
		Regulation of macromolecule metabolic process	
		Regulation of metabolic process	
		Regulation of nitrogen compound metabolic process	
		Regulation of primary metabolic process	
		Regulation of transcription	
		Response to endogenous stimulus	
		Response to organic substance	
		Response to oxidative stress	
		Response to stimulus	
	Zhou et al. (2013); Duan et al. (2018) (At804)	mRNA splicing, via spliceosome	2
		RNA biosynthetic process	
		RNA metabolic process	
		RNA processing	
		RNA splicing	

(continued)

Table 3 (continued)

	Studies represented	GO categories	Number of studies represented
	Niu et al. (2015); Duan et al. (2018) (At804)	Transcription	2
	Tie et al. (2012); Zhou et al. (2013)	DNA-dependent DNA replication	2
		L-serine metabolic process	
		M phase	
		M phase of meiotic cell cycle	
		M phase of mitotic cell cycle	
		DNA metabolic process	
		Nucleoside	
		Nucleotide and nucleic acid metabolic process	
		DNA repair	
		DNA replication	
Downregulated categories	Tie et al. (2012) (Nipponbare); Duan et al. (2018) (At804)	Catabolic process	2
		Cell wall macromolecule metabolic process	
		Cell wall polysaccharide metabolic process	
		Hemicellulose metabolic process	
		Polysaccharide metabolic process	
	Tie et al. (2012) (Nipponbare); Niu et al. (2015)	Response to stress	2
		Signal transduction	

Table 4 Genes upregulated in response to virulent *Agrobacterium* strain or Vir proteins in more than one of the reviewed studies

Studies represented in	Function	Gene names	Number of studies represented
Lee et al. (2009) (GV3101); Niu et al. (2015); Duan et al. (2018) (At804)	Metabolic enzyme—phenylpropanoid biosynthesis	CCoAOMT6 (AT1G67980)	3
	Other stress response	FOX1 (AT1G26380)	
	Stress response—cytochrome P450	CYP71A12 (AT2G30750)	
Lee et al. (2009) (GV3101), Duan et al. (2018) (At804)	Stress response—cytochrome P450	CYP71A13 (AT2G30770), CYP71B15 (AT3G26830), CYP82C2 (AT4G31970), CYP79B2 (AT4G39950)	2
	Stress response—glutathione S-transferase	AtGSTF7 (AT1G02920), AtGSTF6 (AT1G02930)	
	Stress response—peroxidase	PER37 (AT4G08770), PER71 (AT5G64120)	
	Other stress response	ATBBE10 (AT1G30720), HB1/GLB1 (AT2G16060), ARD3 (AT2G26400), AT2G43570, AT2G43620, AT4G33560	
	Other metabolic enzyme	AT1G43800, PXMT1 (AT1G66700), ALDH2C4 (AT3G24503)	
	Uncategorized	MBL1 (AT1G78850), AT3G29970, AT4G12500, AT4G13180, AT5G42830	
Niu et al. (2015); Duan et al. (2018) (At804)	Cell growth related	AT5G27420	2
	Defense signaling	MPK11 (AT1G01560), NIMIN1 (AT1G02450), FMO1 (AT1G19250), GRX480 (AT1G28480), NHL25 (AT5G36970)	

(continued)

Table 4 (continued)

Studies represented in	Function	Gene names	Number of studies represented
	DNA modification	EFD (AT3G54150)	
	Hormone metabolism	ACS7 (AT4G26200)	
	Metabolic enzyme—phenylpropanoid biosynthesis	COMT-like3 (AT1G21120)	
	Protein metabolism	AT3-MMP (AT1G24140), VBF (AT1G56250)	
	Protein modification	AtDSP13 (AT3G02800), PIMT2 (AT5G50240)	
	Receptor-like kinase	AT1G21240, AT1G21241, AT1G21242, AT1G21243, AT1G21244, RMG1 (AT1G21245), CRK31 (AT1G21246), AT1G21247, AT1G21248, AT1G21249, LECRK-VI.3 (AT1G21250), AT1G21251, AT1G21252	
	Stress response—cytochrome P450	CYP706A2 (AT4G22710), CYP81F2 (AT5G57220)	
	Stress response—glutathione S-transferase	AtGSTU3 (AT2G29470), AT5G44990	
	Stress response—peroxidase	PER4 (AT1G14540), PER5 (AT1G14550), PER62 (AT5G39580)	
	Other stress response	MC8 (AT1G16420), ATBBE4 (AT1G26390), ATBBE6 (AT1G26410), AT1G57630, AT1G66090, ATGLR2.8 (AT2G29110), PMAT2 (AT3G29670), AT3G60140, CBP60G (AT5G26920), ATGLR1.2 (AT5G48400)	

(continued)

Table 4 (continued)

Studies represented in	Function	Gene names	Number of studies represented
	Transcription factor	AT2G43000, AT3G23230, AT3G46080, ZAT72 (AT3G46090), BHLH041 (AT5G56960)	
	Other metabolic enzyme	ASB1 (AT1G25220), AT3G44830, AT4G16820, AT4G39670	
	Uncategorized	AT1G05675, ARI12 (AT1G05880), AT1G32350, AGP5 (AT1G35230), AT1G53620, AT1G53625, AT1G56060, AT1G76650, AT2G23270, AT2G38790, AT3G01830, AT3G02840, AT3G19615, AT3G25240, AT4G14368, AT4G37290, VQ29 (AT4G37710), AT5G22530, AT5G25260, AT5G42380	

Acknowledgements We want to thank members of the Zhanyuan Zhang's laboratory for their technical assistance during the course of the RNA-Seq study. This review work was supported by a University of Missouri Research Board grant and the Life Science Fellowship.

References

Altpeter F, Springer NM, Bartley LE et al (2016) Advancing crop transformation in the era of genome editing. Plant Cell 28:1510–1520

Azpiroz-Leehan R, Feldmann KA (1997) T-DNA insertion mutagenesis in *Arabidopsis*: going back and forth. Trends Genet 13:152–156

Barton KA, Brill WJ (1983) Prospects in plant genetic engineering. Science 219:671–682

Beijersbergen A, Smith SJ, Hooykaas PJ (1994) Localization and topology of VirB proteins of *Agrobacterium tumefaciens*. Plasmid 32:212–218

Bhattacharya A, Priyanka S, Citovsky V (2010) The roles of plant phenolics in defence and communication during *Agrobacterium* and *Rhizobium* infection. Mol Plant Pathol 11:705–719

Binns AN, Costantino P (1998) The *Agrobacterium* oncogenes. In: Spaink HP, Kondorosi A, Hooykaas PJJ (eds) The Rhizobiaceae. Springer, Netherlands, pp 251–266

Casassola A, Brammer SP, Chaves MS et al (2013) Gene expression: a review on methods for the study of defense-related gene differential expression in plants. Am J Plant Sci 4:64–73

Chilton MD, Drummond MH, Merio DJ et al (1977) Stable incorporation of plasmid DNA into higher plant cells: the molecular basis of crown gall tumorigenesis. Cell 11:263–271

Chinchilla D, Bauer Z, Regenass M et al (2006) The *Arabidopsis* receptor kinase FLS2 binds flg22 and determines the specificity of flagellin perception. Plant Cell 18:465–476

Christie PJ, Vogel JP (2000) Bacterial type IV secretion: conjugation systems adapted to deliver effector molecules to host cells. Trends Microbiol 8:354–360

Cui H, Xiang T, Zhou JM (2009) Plant immunity: a lesson from pathogenic bacterial effector proteins. Cell Microbiol 11:1453–1461

DeCleene M, DeLay J (1976) The host range of crown gall. Bot Rev 42:389–466

Ditt RF, Nester EW, Comai L (2001) Plant gene expression response to *Agrobacterium tumefaciens*. Proc Natl Acad Sci U S A 98:10954–10959

Ditt RF, Nester E, Comai L (2005) The plant cell defense and *Agrobacterium tumefaciens*. FEMS Microbiol Lett 247:207–213

Ditt F, Kerr KF, de Figueiredo P et al (2006) The *Arabidopsis thaliana* transcriptome in response to *Agrobacterium tumefaciens*. Mol Plant Microbe Interact 19:665–681

Djamei A, Pitzschke A, Nakagami H, Rajh I, Hirt H (2007) Trojan horse strategy in *Agrobacterium* transformation—abusing MAPK defense signaling. Science 318:453–456

Duan K, Willig C, de Tar JR et al (2018) Transcriptomic analysis of *Arabidopsis* seedlings in response to *Agrobacterium*-mediated transformation process. Mol Plant Microbe Interact 31:445–459

Escobar MA, Dandekar AM (2003) *Agrobacterium tumefaciens* as an agent of disease. Trends Plant Sci 8:380–386

Felix G, Duran JD, Volko S et al (1999) Plants have a sensitive perception system for the most conserved domain of bacterial flagellin. Plant J 18:265–276

García-Rodríguez FM, Schrammeijer B, Hooykaas PJ (2006) The *Agrobacterium* VirE3 effector protein: a potential plant transcriptional activator. Nucl Acids Res 34:6496–6504

Gelvin SB (2000) *Agrobacterium* and plant genes involved in T-DNA transfer and integration. Annu Rev Plant Physiol Plant Mol Biol 51:223–256

Gelvin SB (2010) Plant proteins involved in *Agrobacterium*-mediated genetic transformation. Annu Rev Phytopathol 48:45–68

Gelvin SB (2017) Integration of *Agrobacterium* T-DNA into the plant genome. Annu Rev Genet 51:195–217

Gómez-Gómez L, Boller T (2000) FLS2: An LRR receptor–like kinase involved in the perception of the bacterial elicitor flagellin in *Arabidopsis*. Mol Cell 5:1003–1011

Gust AA, Biswas R, Lenz HD et al (2007) Bacteria-derived peptidoglycans constitute pathogen-associated molecular patterns triggering innate immunity in *Arabidopsis*. J Biol Chem 282:32338–32348

Kuldau GA, DeVos G, Owen J et al (1990) The *virB* operon of *Agrobacterium tumefaciens* pTiC58 encodes 11 open reading frames. Mol Gen Genet 221:256–266

Kunze G, Zipfel C, Robatzek S et al (2004) The N terminus of bacterial elongation factor Tu elicits innate immunity in *Arabidopsis* plants. Plant Cell 16:3496–3507

Lacroix B, Citovsky V (2013) The roles of bacterial and host plant factors in *Agrobacterium*-mediated genetic transformation. Int J Dev Biol 57:467–481

Lacroix B, Vaidya M, Tzfira T, Citovsky V (2005) The VirE3 protein of *Agrobacterium* mimics a host cell function required for plant genetic transformation. EMBO J 24:428–437

Lapham R, Lee LY, Tsugama D et al (2018) VIP1 and its homologs are not required for *Agrobacterium*-mediated transformation, but play a role in *Botrytis* and salt stress responses. Front Plant Sci 9:749

Lee CW, Efetova M, Engelmann JC et al (2009) *Agrobacterium tumefaciens* promotes tumor induction by modulating pathogen defense in *Arabidopsis thaliana*. Plant Cell 21:2948–2962

Livaja M, Zeidler D, von Rad U et al (2008) Transcriptional responses of *Arabidopsis thaliana* to the bacteria-derived PAMPs harpin and lipopolysaccharide. Immunobiology 213:161–171

Magori S, Citovsky V (2011) *Agrobacterium* counteracts host-induced degradation of its effector F-box protein. Sci Signal 4:ra69

Morris RO (1986) Genes specifying auxin and cytokinin biosynthesis in phytopathogens. Annu Rev Plant Physiol 37:509–538

Mysore KS, Bassuner B, Deng XB et al (1998) Role of the *Agrobacterium tumefaciens* VirD2 protein in T-DNA transfer and integration. Mol Plant Microbe Interact 11:668–683

Niu X, Zhou M, Henkel CV et al (2015) The *Agrobacterium tumefaciens* virulence protein VirE3 is a transcriptional activator of the F-box gene VBF. Plant J 84:914–924

Pu XA, Goodman RN (1993) Attachment of agrobacteria to grape cells. Appl Environ Microbiol 59:2572–2577

Schrammeijer B, den Dulk-Ras A, Vergunst AC et al (2003) Analysis of Vir protein translocation from *Agrobacterium tumefaciens* using *Saccharomyces cerevisiae* as a model: Evidence for transport of a novel effector protein VirE3. Nucl Acids Res 31:860–868

Shi Y, Lee LY, Gelvin SB (2014) Is VIP1 important for *Agrobacterium*-mediated transformation? Plant J 79:848–860

Shirasu K, Morel P, Kado CI (1990) Characterization of the *virB* operon of an *Agrobacterium tumefaciens* Ti plasmid: nucleotide sequence and protein analysis. Mol Microbiol 4:1153–1163

Staskawicz BJ, Ausubel FM, Baker BJ et al (1995) Molecular genetics of plant disease resistance. Science 268:661–667

Tao Y, Xie Z, Chen W et al (2003) Quantitative nature of *Arabidopsis* responses during compatible and incompatible interactions with the bacterial pathogen *Pseudomonas syringae*. Plant Cell 15:317–330

Thompson DV, Melchers LS, Idler KB et al (1988) Analysis of the complete nucleotide sequence of the *Agrobacterium tumefaciens virB* operon. Nucleic Acids Res 16:4621–4636

Tie W, Zhou F, Wang L et al (2012) Reasons for lower transformation efficiency in indica rice using *Agrobacterium tumefaciens*-mediated transformation: lessons from transformation assays and genome-wide expression profiling. Plant Mol Biol 78:1–18

Tomlinson AD, Fuqua C (2009) Mechanisms and regulation of polar surface attachment in *Agrobacterium tumefaciens*. Curr Opin Microbiol 12:708–714

Tsuda K, Qi Y, le Nguyen V et al (2012) An efficient *Agrobacterium*-mediated transient transformation of *Arabidopsis*. Plant J 69:713–719

Tzfira T, Citovsky V (2001) VIP1, an *Arabidopsis* protein that interacts with *Agrobacterium* VirE2, is involved in VirE2 nuclear import and *Agrobacterium* infectivity. EMBO J 20:3596–3607

Tzfira T, Citovsky V (2006) *Agrobacterium*-mediated genetic transformation of plants: biology and biotechnology. Curr Opin Biotechnol 17:147–154

Tzfira T, Vaidya M, Citovsky V (2004) Involvement of targeted proteolysis in plant genetic transformation by *Agrobacterium*. Nature 431:87–92

Veena JH, Doerge RW, Gelvin SB (2003) Transfer of T-DNA and Vir proteins to plant cells by *Agrobacterium tumefaciens* induces expression of host genes involved in mediating transformation and suppresses host defense gene expression. Plant J 35:219–236

Vences-Guzman MA, Guan Z, Bermudez-Barrientos JR et al (2013) Agrobacteria lacking ornithine lipids induce more rapid tumour formation. Environ Microbiol 15:895–906

Vergunst AC, Schrammeijer B, den Dulk-Ras A et al (2000) VirB/D4-dependent protein translocation from *Agrobacterium* into plant cells. Science 290:979–982

Wang Y, Peng W, Zhou X et al (2014) The putative Agrobacterium transcriptional activator-like virulence protein VirD5 may target T-complex to prevent the degradation of coat proteins in the plant cell nucleus. New Phytol 203:1266–1281

Ward DV, Draper O, Zupan JR, Zambryski PC (2002) Peptide linkage mapping of the *Agrobacterium tumefaciens vir*-encoded type IV secretion system reveals protein subassemblies. Proc Natl Acad Sci USA 99:11493–11500

Ward JE, Akiyoshi DE, Regier D et al (1988) Characterization of the *virB* operon from an *Agrobacterium tumefaciens* Ti plasmid. J Biol Chem 263:5804–5814

Watson B, Currier TC, Gordon MP et al (1975) Plasmid required for virulence of *Agrobacterium tumefaciens*. J Bacteriol 123:255–264

Weigel RR, Pfitzner UM, Gatz C (2005) Interaction of NIMIN1 with NPR1 modulates PR gene expression in *Arabidopsis*. Plant Cell 17:1279–1291

Wood DW, Setubal JC, Kaul R et al (2001) The genome of the natural genetic engineer *Agrobacterium tumefaciens* C58. Science 294:2317–2323

Zaltsman A, Krichevsky A, Loyter A et al (2010) *Agrobacterium* induces expression of a host F-box protein required for tumorigenicity. Cell Host Microbe 7:197–209

Zhang WJ, Dewey RE, Boss W et al (2013) Enhanced *Agrobacterium*-mediated transformation efficiencies in monocot cells is associated with attenuated defense responses. Plant Mol Biol 81:273–286

Zhang YM, Liu ZH, Yang RJ et al (2016) Improvement of soybean transformation via *Agrobacterium tumefaciens* methods involving α-aminooxyacetic acid and sonication treatments enlightened by gene expression profile analysis. Plant Cell Rep 35:1259–1271

Zhang X, van Heusden GPH, Hooykaas PJ (2017) Virulence protein VirD5 of *Agrobacterium tumefaciens* binds to kinetochores in host cells via an interaction with Spt4. Proc Natl Acad Sci 114:10238–10243

Zhou X, Wang K, Lv D et al (2013) Global analysis of differentially expressed genes and proteins in the wheat callus infected by *Agrobacterium tumefaciens*. PLoS ONE 8:e79390

Ziemienowicz A, Merkle T, Schoumacher F et al (2001) Import of *Agrobacterium* T-DNA into plant nuclei: two distinct functions of VirD2 and VirE2 proteins. Plant Cell 13:369–383

Zipfel C, Robatzek S (2010) Pathogen-associated molecular pattern-triggered immunity: veni, vidi…? Plant Physiol 154:551–554

Zipfel C, Robatzek S, Navarro L et al (2004) Bacterial disease resistance in *Arabidopsis* through flagellin perception. Nature 428:764–767

Zipfel C, Kunze G, Chinchilla D et al (2006) Perception of the bacterial PAMP EF-Tu by the receptor EFR restricts *Agrobacterium*-mediated transformation. Cell 125:749–760

Agrobacterium-Mediated Transformation of Yeast and Fungi

Paul J. J. Hooykaas, G. Paul H. van Heusden, Xiaolei Niu, M. Reza Roushan, Jalal Soltani, Xiaorong Zhang and Bert J. van der Zaal

Contents

1 Introduction 350
2 Yeasts and Fungi Transformed by *Agrobacterium* 353
3 T-DNA 355
3.1 T-DNA Structure in Yeast and Fungi 355
3.2 Integration of T-DNA by Homologous Recombination 356
3.3 Extrachromosomal T-DNA 357
4 Role of Virulence Proteins in AMT of Yeast and Fungi 358
5 Use of Yeast to Study the *Agrobacterium* Virulence System 359
5.1 Visualization of Effector Protein Translocation 359
5.2 Functional Analysis of Translocated Effector Proteins in Yeast 360
6 Host Factors 363
6.1 The Role of Host Proteins During Agrobacterium-Mediated Transformation 363
6.2 Role of Host DNA Repair Factors in Non-homologous T-DNA Integration 364
7 Conclusions 366
References 368

Abstract Two decades ago, it was discovered that the well-known plant vector *Agrobacterium tumefaciens* can also transform yeasts and fungi when these microorganisms are co-cultivated on a solid substrate in the presence of a phenolic inducer such as acetosyringone. It is important that the medium has a low pH (5–6) and that the temperature is kept at room temperature (20–25 °C) during co-cultivation. Nowadays, *Agrobacterium*-mediated transformation (AMT) is the method of choice for the transformation of many fungal species; as the method is simple, the transformation efficiencies are much higher than with other methods,

P. J. J. Hooykaas (✉) · G. P. H. van Heusden · X. Niu · M. Reza Roushan
J. Soltani · X. Zhang · B. J. van der Zaal
Sylvius Lab, Department of Molecular and Developmental Genetics, Institute of Biology, Leiden University, Sylviusweg 72, 2333 BE Leiden, The Netherlands
e-mail: p.j.j.hooykaas@biology.leidenuniv.nl

Current Topics in Microbiology and Immunology (2018) 418:349–374
DOI 10.1007/82_2018_90

Published Online: 17 May 2018

and AMT leads to single-copy integration much more frequently than do other methods. Integration of T-DNA in fungi occurs by non-homologous end-joining (NHEJ), but also targeted integration of the T-DNA by homologous recombination (HR) is possible. In contrast to AMT of plants, which relies on the assistance of a number of translocated virulence (effector) proteins, none of these (VirE2, VirE3, VirD5, VirF) are necessary for AMT of yeast or fungi. This is in line with the idea that some of these proteins help to overcome plant defense. Importantly, it also showed that VirE2 is not necessary for the transport of the T-strand into the nucleus. The yeast *Saccharomyces cerevisiae* is a fast-growing organism with a relatively simple genome with reduced genetic redundancy. This yeast species has therefore been used to unravel basic molecular processes in eukaryotic cells as well as to elucidate the function of virulence factors of pathogenic microorganisms acting in plants or animals. Translocation of *Agrobacterium* virulence proteins into yeast was recently visualized in real time by confocal microscopy. In addition, the yeast 2-hybrid system, one of many tools that have been developed for use in this yeast, was used to identify plant and yeast proteins interacting with the translocated *Agrobacterium* virulence proteins. Dedicated mutant libraries, containing for each gene a mutant with a precise deletion, have been used to unravel the mode of action of some of the *Agrobacterium* virulence proteins. Yeast deletion mutant collections were also helpful in identifying host factors promoting or inhibiting AMT, including factors involved in T-DNA integration. Thus, the homologous recombination (HR) factor Rad52 was found to be essential for targeted integration of T-DNA by HR in yeast. Proteins mediating double-strand break (DSB) repair by end-joining (Ku70, Ku80, Lig4) turned out to be essential for non-homologous integration. Inactivation of any one of the genes encoding these end-joining factors in other yeasts and fungi was employed to reduce or totally eliminate non-homologous integration and promote efficient targeted integration at the homologous locus by HR. In plants, however, their inactivation did not prevent non-homologous integration, indicating that T-DNA is captured by different DNA repair pathways in plants and fungi.

1 Introduction

Agrobacterium tumefaciens causes crown gall disease on many dicotyledonous plant species and some gymnosperms (De Cleene and De Ley 1976). Below we present a short introduction to the molecular mechanisms underlying this disease, but for more details about *Agrobacterium* biology and literature references, see the following reviews: Nester et al. 1984; Winans 1991; Zhu et al. 2000; Tzfira et al. 2000; Gelvin 2003, 2010; and the other chapters in this volume.

Crown galls consist of cells that have been transformed into tumor cells by the transfer of an oncogenic piece of DNA, transferred DNA or T-DNA, from the bacterium. T-DNA is a segment of DNA that is naturally present in a large Ti plasmid in *Agrobacterium*. It contains a number of oncogenes (*onc*-genes) that

encode enzymes involved in the production of plant growth regulators. Transfer of T-DNA to plant cells leads to their uncontrolled growth and thus to tumor formation. None of the T-DNA genes is involved in T-DNA transfer. Rather, a set of genes (the virulence genes), which are located elsewhere in the Ti plasmid, are needed for the mobilization of T-DNA into plant cells. These *vir* genes act *in trans* to process and transfer T-DNA, which is surrounded by direct repeat (border repeat) sequences of 24 bp. This has led to the development of the binary vector system consisting of an *Agrobacterium* strain containing a Ti plasmid from which the T-DNA has been removed (helper strain) and a separate cloning vector containing a plant selection marker between 24 bp border repeats into which genes of interest can be cloned (binary vector). Nowadays, *Agrobacterium* is often the preferred vector for plant transformation in plant biotechnology and plant research. This is because of the ease of handling, the use of plant tissues as targets for transformation rather than protoplasts, and the relatively low cost associated with the use of *Agrobacterium* as a vector in comparison to other methods requiring expensive equipment such as an electroporator or a particle gun.

The virulence (*vir)* genes are activated in an acidic environment (pH 5–6) when the bacteria sense the presence of phenolic compounds such as acetosyringone, which are released from wounded plant cells. The VirA chemoreceptor becomes activated by autophosphorylation when the proper inducing conditions are met. Subsequently, VirA activates the transcriptional activator VirG by phosphorylation, which then mediates transcription of the other *vir* genes. These include the *vir*B operon (with 11 genes) and the *virD* operon (with 4 or 5 genes; intact *virD3* is absent in some Ti plasmids), which are essential for transformation. The *virB* operon encodes a type four secretion system (T4SS) which is the nanomachine for delivery of T-DNA and a number of virulence effector proteins into host cells. The *virD* operon encodes the VirD2 relaxase and its associated protein VirD1, which initiate T-DNA transfer by nicking the border repeats. This leads to release of single-stranded DNA copies of the T-DNA (T-strands) that are translocated into plant cells. The VirC1 and VirC2 proteins are accessory factors which enhance nicking of the border repeats by VirD2 and thus potentiate transformation. The VirD4 protein is a coupling protein which forms the interface between the relaxase and the T4SS. Some other virulence proteins do not act in the bacterium, but are translocated by the T4SS into the host cells, where they assist in transformation. VirE2 protein is especially important, as plant transformation occurs with a 1000–10,000-fold lower efficiency in its absence. The VirE2 protein encodes a single-stranded DNA binding protein that is thought to coat the T-strand in the plant cell and thus protects it against nucleases. VirE2 may also assist in the delivery of the T-strand into the nucleus. The VirD2 protein, which remains covalently attached to the T-strand during the nicking reaction, contains a nuclear localization sequence, which is essential for nuclear delivery. Besides VirE2, effector proteins transferred by *A. tumefaciens* into host cells by the T4SS include VirE3, VirF, and VirD5. Their localization in plant cells and more recently also in yeast cells (Fig. 1) has been determined, but the functions of these proteins are still not fully understood.

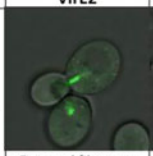
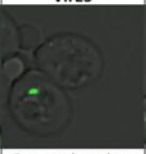
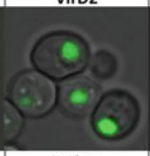
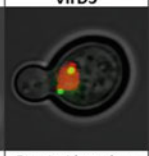
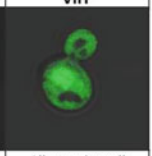
	VirE2	VirE3	VirD2	VirD5	VirF
Localization after ectopic expression					
	Dots and filaments outside nucleus	Dots inside nucleus and near spindle pole bodies	Nucleus	Dots inside nucleus	All over the cell
Possible function	ssDNA binding protein	Transcriptional enhancer	Relaxase; guides T-strand through T4SS	Disruption of kinetochore-microtubule interaction	F-box protein involved in protein degradation
Essential for *S. cerevisiae* transformation?	No	No	Yes	No	No

Fig. 1 Visualization of the *Agrobacterium* virulence proteins fused to GFP (green color) expressed from a yeast promoter in yeast cells

In plants, exogenous DNA integrates with high efficiency by non-homologous recombination. This is also the case for T-DNA, whether it contains homology with the plant genome or not. The ends of T-DNA are reasonably well protected during integration, with sometimes full preservation of the right border end and with usually only a small truncation of the left border end. Integration may be accompanied by the formation of small deletions in the host genome at the integration site. Data regarding the preferred integration site are likely to be biased by the demand for expression of T-DNA in the cell lines that are selected as T-DNA transformants. In the transformants obtained, however, integration seems to be random, and therefore, T-DNA integration can and has been successfully used as a mutagen leading to (T-DNA) tagged mutations.

Although tumors are not formed on monocots, infection with *Agrobacterium* can still lead to transformation of these plants, which include important food crops such as the cereals (Hooykaas-van Slogteren et al. 1984; Ishida et al. 1996). This prompted us to determine whether other organisms could be transformed by *Agrobacterium* as well. In view of the resemblance of AMT to bacterial conjugation, we first tested and found evidence that the *Agrobacterium* virulence system could mobilize plasmids to other bacteria (Beijersbergen et al. 1992). Some twenty years ago, we could show that *Agrobacterium* can also transform the yeast *S. cerevisiae* (Bundock et al. 1995) and fungi (De Groot et al. 1998) under laboratory conditions. Since then, *Agrobacterium*-mediated transformation (AMT) of many other yeasts and fungi has been demonstrated (reviewed in: Michielse et al. 2005a; Soltani et al. 2008). The ability of *Agrobacterium* to transform the yeast *S. cerevisiae* offered the possibility to use the many experimental tools available for this organism to study the transformation process in detail. In this chapter, we will review the molecular mechanisms underlying AMT of yeast and fungi and compare the requirements with those for AMT of plants. We shall not, however, discuss

detailed practical requirements of AMT for yeast and fungi as these have already been discussed in earlier reviews (Michielse et al. 2005a, 2008; Soltani et al. 2008). Instead, we rather focus on the more recent discoveries using yeast as a model to study *Agrobacterium* virulence effectors.

2 Yeasts and Fungi Transformed by *Agrobacterium*

In the early 1990s, it was discovered that the yeast *S. cerevisiae* can be transformed by *Agrobacterium* during co-cultivation on plates with *vir*-induction medium (Bundock et al. 1995; Piers et al. 1996). A few years later, the transformation of mycelium-forming fungi, including *Aspergillus awamori*, *Colletotrichum gloeosporioides*, *Fusarium venenatum*, *Neurospora crassa*, *Trichoderma reesei*, and the edible white button mushroom *Agaricus bisporus*, was reported (De Groot et al. 1998). Since then, *Agrobacterium*-mediated transformation (AMT) has been demonstrated for numerous other yeasts and fungi (reviewed in: Michielse et al. 2005a; Soltani et al. 2008), including species from the phyla Ascomycota, Basidiomycota, Glomeromycota, and Zygomycota. These encompass fungi important for industry (De Groot et al. 1998), plant and animal pathogens (Almeida et al. 2007; Betts et al. 2007; Blaise et al. 2007; Bourras et al. 2012; Jeon et al. 2007; Nemecek et al. 2006; Zhang et al. 2015), and fungal species living in a symbiosis with plants (mycorrhiza) or algae (lichens) (Murata et al. 2006; Pardo et al. 2002; Park et al. 2013). Edible species such as the mushrooms *A. bisporus* (de Groot et al. 1998; Mikosch et al. 2001), *Flammulina velutipes*, *Grifola frondosa*, and *Hypsizygus marmoreus* (Hatoh et al. 2013; Zhang et al. 2014), *Pleurotus eryngii* (Wang et al. 2016), *Tricholoma matsutake* (Murata et al. 2006), and the truffle *Tuber borchii* (Grimaldi et al. 2005) have also been transformed by AMT.

For fungi, AMT has several advantages over conventional transformation methods. First, the use of protoplasts, required in most other methods of transformation, can mostly be omitted in AMT (de Groot et al. 1998). The isolation of protoplasts is not only laborious, but also depends on the quality of the cell wall-degrading enzymes, which are not always commercially available (Gardiner and Howlett 2004; Rolland et al. 2003). It is therefore a great advantage that intact yeast cells, germinating conidia, or even vegetative and fruiting body mycelia can often be used as starting material for AMT. Secondly, some fungal species including *Agaricus bisporus*, *Armillaria mellea*, *Calonectria morganii*, *Ceratocystis resinifera*, and *Helminthosporium turcicum*, that could not be stably transformed by any of the traditional methods, could be transformed by *Agrobacterium* (de Groot et al. 1998; Loppnau et al. 2004; Malonek and Meinhardt 2001; Degefu and Hanif, 2003; Baumgartner et al. 2010). Sometimes, however, setting up a system for AMT can meet with difficulties, and AMT was not successful for the species *Sclerotinia sclerotiorum* (Rolland et al. 2003) and the black yeast *Knufia petricola* (Noack-Schönmann et al. 2014). Thirdly, AMT is preferred in many fungal species because it leads to less complex DNA integration patterns than does polyethylene

glycol (PEG) transformation or electroporation and also results in a much higher frequency of single-copy events (Betts et al. 2007; Blaise et al. 2007; Campoy et al. 2003; Kilaru et al. 2009; Meyer et al. 2003; Rogers et al. 2004; Sugui et al. 2005; Tanguay and Breuil 2003).

The binary vectors used for AMT of yeast and fungi have a selectable marker between the T-DNA borders, just like those used for plants. As with plants, antibiotic resistance genes and herbicide resistance genes have been used as selection markers in yeast and fungi (Frandsen 2011). It is important that these markers are controlled by a promoter active in the host organism. In some fungi, the Cauliflower Mosaic Virus 35S promoter, which is frequently used for plant transformation, can be employed to drive expression of the selectable marker (Mullins et al. 2001). However, an endogenous promoter or a promoter from a related fungus confers better growth in the selection medium (Mullins et al. 2001; White and Chen 2006). Also, sometimes a 5′ intron is required for sufficient expression (Kilaru et al. 2009). Genes such as *URA3*, *TRP1*, and *LEU2* have also been used as selection markers, especially for the transformation of auxotrophic mutants of the yeast *S. cerevisiae* (Bundock et al. 1995; Piers et al. 1996).

Various *Agrobacterium* helper strains have been used for the transformation of yeast and fungi. The most popular are the strains AGL1, EHA105, LBA1100, LBA1126, and LBA4404. Systematic comparisons of different strains in relation to transformation frequencies have not been published, but the use of *Agrobacterium* strains containing helpers derived from the supervirulent pTiBo542 plasmid (AGL1, EHA105) or with mutations leading to higher virulence gene expression (LBA1126) often compared favorably in their transformation performance (Campoy et al. 2003; Park and Kim 2004; Piers et al. 1996). The introduction of a construct carrying the *virG* mutant gene coding for the constitutively active VirGN54D transcriptional activator can also sometimes considerably improve transformation efficiency (Betts et al. 2007).

Transformation efficiency is influenced by many variables, with each fungus requiring slightly different conditions to obtain an optimal transformation frequency. Factors affecting the transformation efficiency include: the starting material (mycelium, conidiospores, fruiting bodies, protoplasts); the ratio between *Agrobacterium* and recipient cells; the length of the co-cultivation period, whereby a longer period generally yields more transformants, but these are usually more difficult to select from the co-cultivation mixtures; concentration of the inducer acetosyringone; a temperature of between 20 and 25 °C is usually optimal, but cold-adapted fungus *Pseudogymnoascus destructans* is only transformed by *Agrobacterium* at temperatures between 15 and 18 °C (Zhang et al. 2015); pH between 5.0 and 5.3, and the choice of the solid support (nitrocellulose, cellophane, Hybond, etc.) (Almeida et al. 2007; Betts et al. 2007; Flowers and Vaillancourt 2005; Leclerque et al. 2004; Tsuji et al. 2003; White and Chen 2006; Yousefi-Pour et al. 2013; Zhang et al. 2014). The optimal pH (usually between pH 5.0 and 5.3) also depends on the *Agrobacterium* strain used, as the pH requirements for optimal *vir* gene induction are slightly different for different *Agrobacterium* strains (Turk et al. 1991). It has been reported that AMT of *S. cerevisiae* mutants deficient in

purine biosynthesis was more efficient than that of the wild-type (Roberts et al. 2003; Soltani 2009) and that addition of purine synthesis inhibitors during AMT of tobacco resulted in increased transformation (Roberts et al. 2003). However, it was subsequently found that the effects of nucleobase auxotrophies and purine synthesis inhibitors are rather species- or even strain-specific. For instance, adenine auxotrophy did not increase transformation in *Cryptococcus neoformans* and *C. gattii*, but uracil auxotrophy gave larger numbers of transformants in some strains (McClelland et al. 2005). Addition of a purine synthesis inhibitor even led to a strongly decreased transformation frequency in the fungus *Paracoccidioides brasiliensis* (Almeida et al. 2007).

3 T-DNA

3.1 T-DNA Structure in Yeast and Fungi

The T-strand is converted into a dsDNA molecule upon arrival in the host nucleus. Whether this occurs before or during the integration process is still uncertain, although data indicate that both may be possible (Chilton and Que 2003; Tzfira et al. 2003; Van Kregten et al. 2016). In *S. cerevisiae*, AMT occurred only at a low frequency when T-DNA lacked homology with the yeast genome. Integrants revealed that, as in plants, the T-DNA ends were relatively well preserved during integration in the yeast genome, sometimes ending exactly at the site where the Ti plasmid DNA had been nicked by VirD2 (Bundock and Hooykaas 1996; Bundock et al. 2002). Small genomic deletions were often found at the T-DNA insertion sites, and filler sequences may be present, which was similarly seen in other yeasts and fungi (Choi et al. 2007; Li et al. 2007; Meng et al. 2007). Sometimes, as in plants, T-DNA integration in fungi is accompanied by gross genomic rearrangements such as large deletions, inversions, or translocations (Choi et al. 2007; Li et al. 2007; Michielse et al. 2009). Remarkably, whereas *S. cerevisiae* T-DNA transformants only have a single copy of the (non-homologous) T-DNA inserted in the genome (Bundock et al. 2002), the integration of multiple copies of T-DNA has been reported for other yeasts and fungi. These may be present in an inverted or direct repeat at one locus or located at a few different loci in the genome (Betts et al. 2007; Campoy et al. 2003; Combier et al. 2003; de Groot et al. 1998; Degefu and Hanif 2003; Flowers and Vaillancourt 2005; Li et al. 2007; Meng et al. 2007; Malonek and Meinhardt 2001; Michielse et al. 2004a, 2009; Mullins et al. 2001; Park et al. 2013; Rho et al. 2001; Sullivan et al. 2002; Tanguay and Breuil 2003; Tsuji et al. 2003). Whether single copy or multi-copy integration, predominates may also depend on the transformation conditions, i.e., the type of tissue transformed, the presence of inducer in the pre-culture medium, duration of the co-cultivation, and the ratio of the numbers of *Agrobacterium*:fungus cells used (Combier et al. 2003; Mikosch et al. 2001; Rho et al. 2001; Sullivan et al. 2002). As mentioned above, T-DNA integration appears to occur at fairly random positions in the plant genome. Without provided homology to

the genomic DNA, the same is observed in yeasts and fungi, although there may sometimes be some bias toward intergenic/regulatory regions, possibly because selection of transformants requires expression of the selection marker on the T-DNA (Blaise et al. 2007; Bourras et al. 2012; Choi et al. 2007; Li et al. 2007; Meng et al. 2007; Walton et al. 2005). By integration, T-DNA can inactivate the gene at the insertion site, resulting in an insertion mutation. Therefore, when efficient, AMT has been used to create collections of T-DNA insertion mutants useful to identify tagged mutations both in plants (Krysan et al. 1999) and in fungi, including the human pathogens *Cryptococcus neoformans* (Walton et al. 2005) and *Blastomyces dermatitidis* (Nemecek et al. 2006), and the phytopathogens *Fusarium oxysporum* (Michielse et al. 2009), *Leptosphaeria maculans* (Bourras et al. 2012), and the rice blast fungus *Magnaporthe oryzae* (Betts et al. 2007; Jeon et al. 2007). In plants, mutant phenotypes seen are not always due to T-DNA insertion, but may be due to other genetic or epigenetic changes brought about by the transformation procedure. This was also observed in T-DNA mutant collections in fungi: From a low percentage up to about a third to half of the mutant phenotypes were not due to a T-DNA disruption of a particular gene, depending on the species involved (Blaise et al. 2007; Idnurm et al. 2004: Walton et al. 2005). This lack of linkage between the T-DNA disruption and phenotype is, however, not unique to AMT, but has also been observed in transformants obtained in other ways (Mullins et al. 2001).

3.2 Integration of T-DNA by Homologous Recombination

In contrast to plants and many fungi, the yeast *S. cerevisiae* integrates exogenous DNA preferentially by homologous recombination (HR). This turned out to be the case also for T-DNA: T-DNAs embracing a segment homologous to the yeast genome gave 100–1000-fold higher transformation frequencies than did T-DNAs lacking such homology (Bundock et al. 1995; Bundock and Hooykaas 1996). This is not the case in plants, where T-DNAs with and without homology are integrated with similar efficiency and almost exclusively by non-homologous recombination (Offringa et al. 1990). This showed that the host cell largely determines the fate of T-DNA. When replacement vectors were used, which harbor a T-DNA where a selectable marker is surrounded by sequences homologous to an endogenous genomic sequence of the recipient host cell, not only HR-directed replacement events, but also HR-directed insertions of the complete T-DNA were found, both in about equal frequency (Bundock et al. 1995; van Attikum and Hooykaas 2003). In the case of insertion of the complete T-DNA, the transferred T-DNA must have formed a circular molecule before integration. DNA sequencing revealed that the circles had formed by a precise fusion of the parts of the right and left border repeats that entered the host as part of the T-strand, reconstituting a complete (mixed) border repeat (Bundock et al. 1995). This can be ascribed to the strand transferase activity of the VirD2 protein, which is the reversal of its nicking activity (Pansegrau et al. 1993). T-circle formation has also been observed in plants, but

here circles had deletions of the ends and were also sometimes accompanied by insertions of filler DNA, thus more resembling non-homologous T-DNA integration (Singer et al. 2012). With T-DNA insertion vectors (with the border repeats not at the outside flanks of the homologous sequences, but with the homology disrupted internally by the border repeats), evidence for integration by gap-repair was obtained, although also both insertion and replacement events were seen (Risseeuw et al. 1996). Integration of T-DNA by HR was fully dependent on the action of the *RAD52* gene (van Attikum and Hooykaas 2003).

In other yeasts such as *Kluyveromyces lactis* and fungi, the frequency of AMT is similar whether the vector contains homology or not. When homology is present in the vector, integration may preferentially occur by homologous recombination (Amey et al. 2003; Lee and Bostock 2006; Michielse et al. 2005a; Sugui et al. 2005; Yu et al. 2015; Zhang et al. 2003; Zwiers and de Waard 2001) or still by non-homologous recombination (Gardiner and Howlett 2004; Loppnau et al. 2004; McClelland et al. 2005), depending on the species. This may also depend on the length of homology that is offered by the vector, with only longer segments of homology promoting integration by HR in some species (Michielse et al. 2005a; Zhang et al. 2003). In contrast to *S. cerevisiae*, where the insertion events were all insertions of a single copy of the re-circularized T-DNA, in other yeasts and fungi such as *K. lactis* and the brown rot pathogen *Monilinia fructicola*, multiple copies of T-DNA had sometimes integrated in tandem at the homologous locus in the chromosome (Bundock et al. 1999; Lee and Bostock 2006).

3.3 Extrachromosomal T-DNA

High transformation frequencies were seen in yeast when a T-DNA was transferred that could be maintained as a plasmid or mini-chromosome without the need for integration into any of the chromosomes. This was accomplished in several ways. Firstly, by the addition of the replication unit of the yeast 2μ plasmid to T-DNA: Such T-DNAs were stably maintained as circular autonomous plasmids, which had the transferred parts of the left and right border repeat fused back to a complete (mixed) border repeat (Bundock et al. 1995); secondly, by the inclusion of a chromosomal autonomous replicating sequence (ARS) in T-DNA (Piers et al. 1996; Rolloos et al. 2014; Ohmine et al. 2016). When such T-DNA also contained telomeric repeats adjacent to both border repeats, it could be maintained as an unstable mini-chromosome (Piers et al. 1996). In the absence of telomeric repeats, the transfer frequency of T-DNA dropped 500-fold and only colonies with a chromosomally integrated T-DNA were obtained (Piers et al. 1996). Yeast is unique in having very small chromosomal centromeres (CENs), and addition of such a CEN in addition to an ARS to T-DNA resulted in T-DNAs that were transferred at high frequency and were stably maintained as a single-copy plasmid (Rolloos et al. 2014; Ohmine et al. 2016). It is remarkable that the homologous repair protein Rad52, but not the crucial end-joining factor yKu70, turned out to be

important for T-circle formation: In the *rad52* mutant, transformation was reduced by 75% (Rolloos et al. 2014; Ohmine et al. 2016). When in a control experiment an identical linearized T-DNA vector was introduced by lithium acetate transformation, a similar transformation frequency was seen in both the wild-type yeast and the *rad52* mutant. Taken together, these results suggest that concatemers of T-strands were formed upon AMT by a process of strand-transfer catalyzed by VirD2, for which such activity has been demonstrated in in vitro experiments (Pansegrau et al. 1993). These concatemers could subsequently be resolved into T-circles by HR. Only at most 25% of the T-circles would then be formed immediately from single T-strands by the VirD2 strand transferase activity linking its left and right border parts. This would lead to stable transformants in the absence of HR (Rolloos et al. 2014). In a similar series of experiments, a T-DNA vector was used that, in addition to ARS and CEN sequences, also had telomeric repeats at both ends of T-DNA. After transfer by AMT, this T-DNA could be maintained in yeast either as a linear mini-chromosome or as a T-circle. In the *rad52* mutant, the transfer was strongly reduced, in line with what was described above. Interestingly in the remaining transformants only T-circles, but no linear mini-chromosomes, were found, suggesting that *RAD52* is needed for conversion of the introduced T-strands into a mini-chromosome (Ohmine et al. 2016).

4 Role of Virulence Proteins in AMT of Yeast and Fungi

Based on the requirement of acetosyringone to obtain transformants, and on the structure of integrated T-DNA, it was inferred that transformation of yeasts and fungi by *A. tumefaciens* was mediated by the virulence system (Bundock et al. 1995; de Groot et al. 1998; Piers et al. 1996). Indeed, mutation of *Agrobacterium* genes in one of the key components required for the induction of the T-DNA transfer system (VirA, VirG), or the subsequent generation (VirD1, VirD2) and transport of the T-strand (VirD4, VirB1-11), completely abolished the ability to transform either the yeast *S. cerevisiae* or the fungus *A. awamori.* These results showed that the *virA*, *virB*, *virD*, and *virG* genes that are essential for plant transformation are likewise needed for transformation of yeast (Bundock et al. 1995; Piers et al. 1996) and the fungus *A. awamori* (Michielse et al. 2004b). The products of the *virC* genes, VirC1 and VirC2, are DNA-binding proteins which enhance the nicking reaction of VirD2 at the border repeats and thus the formation of the T-strands (Atmakuri et al. 2007). In plants, *virC* mutants are attenuated in virulence. Similarly, a tenfold reduced transformation is seen with *virC* mutants in yeast and fungi. Fungal transformants in this case were characterized by the presence of complex T-DNA structures containing multicopy and truncated T-DNAs and vector backbone sequences (Michielse et al. 2004b). This is in line with reduced T-DNA border processing in the *virC* mutants, and thus reduced T-DNA transfer and increased left border skipping. That AMT occurs by a similar process in plants, yeasts, and fungi is also in line with the observation that similar pH and temperature conditions are required for both

plant and fungal transformation. A low pH of 5–6 is required for the induction of the *vir* genes, whereas temperatures of around 22 °C are optimal for the T-DNA transfer machinery (Fullner and Nester 1996). The *Agrobacterium* virulence system has also been used to transfer into yeast a modified CloDF13 plasmid, in which a yeast selectable marker and the replication unit of the yeast 2μ plasmid had been inserted. As this CloDF13 plasmid encoded its own relaxase and its own coupling protein and had its own nick site, transfer did not require any of the Agrobacterium *virD* genes, but only relied on expression of the *virB* encoded transport system (Escudero et al. 2003).

Some important differences between the AMT requirements of plants and fungi were noticed regarding the effector proteins that are delivered by *Agrobacterium* into host cells concomitantly with T-DNA. As a single-strand DNA-binding protein, VirE2 is thought to play an important role in the protection of the T-strand from nucleases in the host cell. In the absence of VirE2, there is almost no transformation of most plants (Rossi et al. 1996). Transformation of yeast and fungi by *virE2* mutants still occurs, but at 2–10-fold reduced frequency (Bundock et al. 1995; Michielse et al. 2004a). The *A. awamori* transformants obtained after transformation with a *virE2* mutant had more pronounced left border truncations (Michielse et al. 2004b), indicating that VirE2 in fungi, as in plants, may help to protect the T-strand against nucleases. The translocated effector proteins VirE3, VirF, and VirD5, necessary for optimal infection of plants, were both alone and in combination dispensable for transformation of yeast (Bundock et al. 1995; our unpublished results) and *A. awamori* (Michielse et al. 2004a).

5 Use of Yeast to Study the *Agrobacterium* Virulence System

5.1 Visualization of Effector Protein Translocation

Although the absence of the translocated effector proteins does not prevent AMT of yeasts or fungi, these proteins are still translocated efficiently into yeast and, by inference, into fungal cells. To study protein transfer from *Agrobacterium* to yeast, the Cre recombinase reporter assay for translocation (CRAfT) was used (Vergunst et al. 2000). To this end, fusions between the Cre recombinase and Vir proteins were expressed in *Agrobacterium*. Transfer of the Cre–Vir fusion proteins from *Agrobacterium* to yeast was subsequently monitored by the selectable excision of a floxed *URA3* marker gene (i.e., *URA3* surrounded by *lox*-sequences in a direct repeat) from the yeast genome by the Cre recombinase. In this way, the translocation of the VirE2, VirE3, and VirF proteins into yeast cells was demonstrated (Schrammeijer et al. 2003). More recently, the translocation of VirE2 protein into yeast cells could be visualized by using bimolecular fluorescence complementation (BiFC) and split GFP strategies (Li et al. 2014; Sakalis et al. 2014). To this end, *Agrobacterium* strains expressing VirE2 tagged with one part of a fluorescent protein

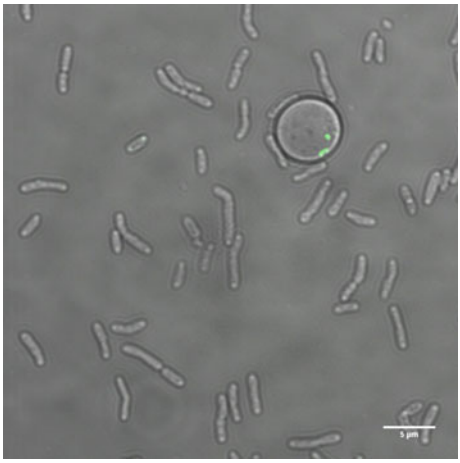


Fig. 2 Co-cultivation of *Agrobacterium* (rods) and yeast (the globular cell). Green color shows translocation of VirE2 from *Agrobacterium* into a yeast cell and is visualized by a split GFP approach

were co-cultivated with yeast cells expressing the complementary part, either fused to VirE2 (for BiFC) or not (Split GFP). Fluorescent dots and filaments were visible in recipient cells 20–25 h after the start of the co-cultivation, indicative of VirE2 protein translocation (Fig. 2). Under optimal conditions, a fluorescent signal was sometimes seen already 8 h after the start of co-cultivation. Evidence was obtained that the fluorescent filaments co-localized with microtubules, as they disappeared after treatment with benomyl (Sakalis et al. 2014). Interestingly, Salman et al. (2005) showed that "animalized VirE2" is able to move along microtubules in Xenopus cells. Formation of these fluorescent structures in the yeast cell was independent of T-DNA transfer. By a similar strategy, the translocation of the other Vir effector proteins (VirE3, VirF, VirD2, and VirD5) could be followed in real time (Sakalis 2013; Roushan, Hooykaas and van Heusden, unpublished).

5.2 *Functional Analysis of Translocated Effector Proteins in Yeast*

The yeast 2-hybrid system has invariably been the starting point for the identification of plant interaction partners of *Agrobacterium* virulence proteins. In this way, it was found that VirD2 protein could interact with a set of plant cyclophilins (peptidyl-prolyl *cis-trans* isomerases), which also function as chaperones (Deng et al. 1998). In addition, VirD2 interacted with importin α/karyopherin α via its C-terminal nuclear localization sequence (NLS), which is necessary for nuclear import (Ballas and Citovsky 1997). Bhattacharjee et al. (2008) showed that both VirD2 and VirE2 could interact in yeast with multiple *Arabidopsis* importin α isoforms. By means of a yeast 2-hybrid screen, two interactors were also identified for VirE2, which were called VIP1 and VIP2 (Tzfira et al. 2000). Both VIP1 (Lacroix and Citovsky 2013) and VIP2 (Anand et al. 2007) act as transcription

factors in plant cells. It was proposed that VIP1, by binding to VirE2 molecules coating the T-strand, plays an important role in the transport of the T-complex into the nucleus (Tzfira et al. 2001). However, recently it was found that *vip1* mutants of *Arabidopsis thaliana* are equally well transformed as the wild-type, indicating that VIP1 is not essential for transformation of plants (Shi et al. 2014). Nevertheless, both the VIP1 and the VIP2 protein may assist in transformation by mediating binding of the T-complex to the chromatin (Lacroix et al. 2008). Regarding targeting the T-complex to chromatin, VirD2 may also play a role here as it can bind to the core histone proteins in yeast after entry into the yeast cell during AMT (Wolterink-van Loo et al. 2015).

Arabidopsis Skp1-like ASK proteins were identified by a 2-hybrid screen as interactors of VirF (Schrammeijer et al. 2001). The Skp1-like proteins are essential components of SCF-complexes, which have an important role in eukaryotic cells in the ubiquitination and proteolytic degradation of specific target proteins, which often need to be phosphorylated first (Skowyra et al. 1997). The Skp1-like subunit connects the Cul1 scaffold of the SCF-complex to the F-box subunit which specifies the target proteins to be degraded. After the discovery of the plant Skp1-like proteins as interactors of VirF, the VirF protein sequence was inspected for the presence of an F-box, and indeed such an F-box turned out to be present and essential for the biological function of VirF (Schrammeijer et al. 2001). The hunt for target proteins, which are degraded in the host plant upon the introduction of VirF, has not yet been completed. Initially, it was discovered that VirF can interact with the defense transcription factor VIP1, one of the interactors of VirE2 (Tzfira et al. 2004). Yeast cells expressing GFP-VIP1 lost fluorescence when VirF was expressed. Moreover, yeast cells expressing GFP-VirE2 lost fluorescence when both VIP1 and VirF were co-expressed, suggesting that VirE2 may be degraded under the direction of VirF in the presence of VIP1 (Tzfira et al. 2004). It was hypothesized that VirF may have an important function in the transformation process by the degradation of the VirE2 coat that may be formed on the T-strand in the host cell and which eventually could be inhibitory to T-DNA integration (Tzfira et al. 2004). The VirF protein is important for transformation of plants of the Solanaceae family such as tobacco and tomato, but not of many other plants (Hooykaas et al. 1984). These latter plants may have a host F-box protein which can compensate for the absence of *virF* in the bacterium. A host gene encoding such an F-box protein was identified in *A. thaliana*, and called *VBF* (Zaltsman et al. 2010).

By a yeast 2-hybrid screen, three *Arabidopsis* interactors were identified with VirE3 as a bait. The first was importin α/karyopherin α, with which VirE3 interacts through its NLSs to gain entry into the nucleus (Garcia-Rodriguez et al. 2006). Secondly, the Csn5 subunit of the COP9 signalosome was identified as an interactor. Thirdly, VirE3 also interacts with the host TFIIB-like protein pBrp (Garcia-Rodriguez et al. 2006). While Brp is normally located at the outside of the plastids, it moves to the nucleus in the presence of VirE3, and together with VirE3 it activates the transcription of a set of host genes (Niu et al. 2015). These induced genes include *VBF* (Niu et al. 2015), which was already known to be induced during transformation (Zaltsman et al. 2010). This result explained why *virE3 virF*

double mutants are much more attenuated in virulence on some host plants that are single mutants (Garcia-Rodriguez et al. 2006; Niu et al. 2015).

Transformation of yeast and fungi can occur at high frequencies even in the absence of VirF and/or VirE3 (Bundock et al. 1995; Michielse et al. 2004a; our unpublished results). Just as in plants, as mentioned above, the yeast and fungal genomes may also encode an F-box protein which can compensate for the absence of VirF in AMT. However, mutation of all the individual yeast genes encoding F-box proteins (except for the few essential genes) did not lead in any case to a reduced AMT (Niu 2013). Also, no reduced accumulation of VIP1 or VirE2 was observed in yeast in the presence of VirF in our laboratory, and whereas we observed an interaction of VirF with the *Arabidopsis* Skp1-like proteins, we did not observe a clear interaction with the yeast Skp1 protein (Niu 2013). This may be due to strain differences, but further work is needed to clarify the biological role of VirF. Using a yeast 2-hybrid screen with the VirF protein lacking the F-box as a bait, several new putative interactors have been identified recently (Garcia-Cano et al. 2015, 2018). These include the trihelix-domain transcription factors VFP3 and VFP5 and the GLABROUS1 enhancer/binding protein-like transcription factor VFP4, involved in activation of the defense response. Targeted degradation of these new transcription factors under control of VirF may help to suppress plant defense and increase transformation (Garcia-Cano et al. 2015, 2018). Nevertheless, it is clear that none of the translocated effector proteins VirE2, VirE3, and VirF plays an essential role in the transformation of yeasts and fungi. Therefore, it is likely that these proteins have a plant-specific function, such as in the suppression of plant defense, rather than in a process invariably linked to the transformation process such as uncoating of the T-strand.

The function of VirD5 is still largely unknown. An interaction between VirD5 and VirF was shown by bimolecular fluorescence complementation; further results indicated that VirD5 may protect VirF from proteolytic degradation (Magori and Citovsky 2011). On the other hand, it has been reported that VirD5 is a nuclear competitor of VBF for binding to VIP1 to stabilize VIP1 and VirE2 (Wang et al. 2014). In order to study the function of VirD5, a gene construct that would express VirD5 constitutively in plants was introduced by AMT, but no stable transformants were obtained. Subsequently, a construct was introduced into plants in which VirD5 was expressed from an inducible promoter; this time, transformants were obtained. Induction of VirD5, however, led to an inhibition of growth and death of the seedlings. Stable expression in yeast was also not possible, and therefore, it was not possible to screen for interactors in a yeast 2-hybrid screen. When *virD5* was present behind the inducible *GAL1* promoter no growth occurred on galactose induction medium, but growth was normal in glucose repression medium (Zhang et al. 2017; our unpublished results). Transient expression of a GFP-VirD5 fusion in yeast revealed that VirD5 was present at a few specific sites, (seen as fluorescent dots) in the nucleus (Fig. 1). A yeast deletion library, consisting of about 5000 mutants, was transformed with the p*GAL1*-VirD5 construct to find mutant strains with deletions of genes that normally mediate the toxicity of VirD5; these strains would now survive VirD5 induction by galactose. Three mutants were found that could grow in the

presence of galactose. Two of these were defective in inducing the expression of VirD5 by having deletions in the *GAL3* and *GAL4* genes, respectively. The third one lacked the *SPT4* gene. Interestingly, Spt4 protein had a subcellular localization similar to that of VirD5, showing a limited number of specific dots in the nucleus that represented the centromeres/kinetochores (Crotti and Basrai 2004). Indeed, VirD5 was similarly found to co-localize with proteins such as Ndc10 that form part of the kinetochores, and this localization of VirD5 was dependent on the presence of the Spt4 protein (Zhang et al. 2017). We found that most cells expressing VirD5 displayed a large elongated bud and failed to segregate their chromosomes equally to daughter cells at anaphase. As a consequence, many cells became aneuploid. VirD5 expressing cells also showed a more than tenfold higher loss of mini-chromosomes than did control cells. Subsequently, VirD5 was found by BiFC to interact at the centromeres/kinetochores with the essential mitotic regulatory Ipl1/Aurora kinase (Zhang 2016). This serine/threonine protein kinase plays an essential role in the sensing and correction of erroneous kinetochore–microtubule attachments during mitosis. It phosphorylates key substrates involved in the kinetochore–spindle binding and contributes to the activation of the spindle checkpoint (Biggins et al. 1999). Both loss and overexpression of the Ipl1/Aurora kinases lead to chromosome mis-segregation and aneuploidy in yeast cells. In in vitro experiments, VirD5 stimulated the kinase activity of the Ipl1/Aurora kinase (Zhang 2016). The results obtained with VirD5 in yeast were corroborated in plants: VirD5 was found to interact with the three plant Aurora kinases and to cause chromosome mis-segregation in plant cells (Zhang 2016; Zhang et al. 2017). These results suggest that stimulation of the Aurora kinase by VirD5 leads to a temporary spindle checkpoint, allowing T-DNA more time for integration. Otherwise, aneuploidy is a hallmark of tumor cells and may contribute to crown gall tumor formation.

6 Host Factors

6.1 The Role of Host Proteins During Agrobacterium-Mediated Transformation

By screening large mutant collections for mutants with reduced and increased AMT, many genes affecting AMT have been identified in the model plant *A. thaliana* (Zhu et al. 2003). These include genes involved in chromatin structure and remodeling, cytoskeletal functions, and cell wall structure. As the requirements for AMT may be somewhat different for yeasts and fungi than for plants, Soltani (2009) screened a genome-wide yeast deletion collection for mutants with reduced or enhanced AMT using both a replacement T-DNA vector integrating by HR and a T-DNA vector that could autonomously replicate in yeast by a 2μ replication unit. As in plants, many different host factors were found to affect AMT. Most striking was that deletion strains lacking components of SAGA, SLIK, ADA, and NuA4

histone acetyltransferase complexes including *EAF7*, *GCN5*, *NGG,1* and *YAF9* had a strongly enhanced AMT efficiency, whereas strains lacking components of histone deacetylase complexes such as *HDA2*, *HDA3,* and *HST4* had a strongly diminished AMT efficiency (Soltani et al. 2009). The Yaf9 protein also forms part of the SWR1 chromatin-remodeling complex, and mutation of any of the other components, such as *ARP6* and *EAF6* increases AMT to some extent (Luo et al. 2015; Soltani 2009; Soltani et al. 2009). However, mutation of *ARP6* also leads to disrupted microtubule structures, and this suggests that binding of the T-strand or virulence proteins such as VirE2 to the microtubules may in fact be inhibitory for transformation (Luo et al. 2015). A role of chromatin-related genes in plants has also been reported (Crane and Gelvin 2007).

Genes involved in HR such as *RAD52* were also important for AMT by these two types of T-DNA vector (van Attikum and Hooykaas 2003; Rolloos et al. 2014; Soltani et al. 2009). A yeast deletion collection was screened for factors negatively affecting AMT by a T-DNA vector that contained not only an ARS and CEN sequence, but also telomeric repeats near both border repeats (Ohmine et al. 2016). In this screen reduced AMT was seen again in the *rad52* mutant, but also in mutants with deletions of *SRS2*, encoding a DNA helicase, the cell wall regulator *SMI1*, and the membrane sterol synthesis scaffold gene *ERG28*. The lower efficiency of the *erg28* mutant was probably due to less growth inhibition of the mutant yeast as compared to the wild-type by the presence of *Agrobacterium*, resulting in lower AMT frequencies per recipient (Ohmine et al. 2016). How *SMI1* affects AMT is not yet clear, but virulence protein translocation as measured by the CRAfT assay is also severely affected in this mutant (Ohmine et al. 2016).

6.2 Role of Host DNA Repair Factors in Non-homologous T-DNA Integration

In *S. cerevisiae,* the integration of exogenous DNA by homologous recombination is very efficient, in contrast to integration by non-homologous end-joining. However, in plants and certain fungi, the insertion of exogenous DNA mainly occurs by non-homologous recombination, even when the DNA fragment has extensive sequence homology to the host chromosome (Offringa et al. 1990; Loppnau et al. 2004; Gardiner and Howlett 2004; McClelland et al. 2005). This is not a peculiarity of T-DNA, but is also the case when DNA is introduced by other means. In fact, delivery of a single-stranded DNA molecule protected by virulence proteins, as is the case with AMT, may be beneficial for integration by HR. Introduction of a gene disruption construct by AMT into *K. lactis* gave a large increase in targeted integration as compared to delivery of an identical construct by electroporation (Bundock et al. 1999). A similar increase in gene targeting frequency was seen in *A. awamori* when delivery of a disruption cassette was done by AMT in comparison with PEG transformation; this also allowed the use of shorter

DNA flanking sequences to obtain an adequate gene targeting frequency (Michielse et al. 2005b). In the taxol-producing fungus *Pestalotiopsis microspore*, AMT gave highly efficient gene targeting, whereas the same construct introduced by PEG transformation did not integrate at all, but rather formed unstable extrachromosomal DNAs with telomeric repeats (Yu et al. 2015). For the application of AMT in biotechnology, it is of great importance to improve the efficiency of integration via homologous recombination over non-homologous recombination, as this would not only favor the targeted integration of transgenes at desired safe havens in the genome, but would ultimately also allow the directed modification of any endogenous gene at its locus in the genome. Our goal was therefore to find the genes encoding the proteins involved in non-homologous integration with the aim to inactivate these to obtain host cells that would exclusively integrate transgenes by HR. Using the yeast *S. cerevisiae* as a model we found that the proteins essential for the repair of DSBs in genomic DNA by non-homologous end-joining (NHEJ) (i.e. yKu70, yKu80, Lig4) are essential for non-homologous T-DNA integration (van Attikum et al. 2001; van Attikum and Hooykaas 2003). Binding of the Ku70-Ku80 heterodimer to the ends of a DSB protects these and attracts other proteins such as Nej1 and Lif1 to the DNA ends, eventually leading to ligation of the ends by Ligase 4 (Critchlow and Jackson 1998; Lewis and Resnick 2000). In yeast, the MRX complex, consisting of Mre11-Rad50-Xrs2, also plays a role in end-joining. In the *mre11*, *rad50*, and *xrs2* mutants non-homologous T-DNA integration was 20–50-fold reduced. Residual T-DNA integration in these mutants occurred predominantly at the (sub)telomeric repeats or at the rDNA repeats (van Attikum et al. 2001). Mutations in *RAD51* or *RAD52* did not negatively affect non-homologous T-DNA integration; in the *rad52,* mutant integration was in fact twofold enhanced (van Attikum et al. 2001). In *S. cerevisiae,* the Rad52 and yKu70 proteins play a critical role in determining whether T-DNA is integrated via HR or via NHEJ; when both are inactivated, no T-DNA integration occurs at all (van Attikum and Hooykaas 2003). Histone modifiers and ATP-dependent chromatin-remodeling complexes are recruited to sites of DNA damage (reviewed by Hauer and Gasser 2018; Smeenk and van Attikum 2013) and are necessary for optimal repair. The presence or absence of specific components of such complexes is therefore likely to play an important role in T-DNA integration and may explain the effects of such mutations on T-DNA integration.

In line with the finding that the factors involved in DSB repair by NHEJ are involved in T-DNA integration, T-DNA can be captured at a unique DSB in the plant genome made by the homing endonuclease I-*Sce*I (Salomon and Puchta 1998). Even without selection for this event, integration at such DSB occurred with a frequency of 1–2% of the transformants (Chilton and Que 2003; Tzfira et al. 2003). In our laboratory, we have studied the integration of *Agrobacterium* T-DNA at a single chromosomal DSB created by the HO endonuclease at the *MAT* locus in the yeast genome (van Attikum 2003). We found T-DNA insertions at the DSB with a 1% frequency, when we expressed the HO endonuclease in yeast cells during AMT. We did not obtain any transformants when we performed such experiments with the *yKu70* mutant, showing that yKu70 is essential for T-DNA integration by

NHEJ at a preformed DSB. Conversely, in similar co-cultivation experiments with the *rad52* mutant expressing the HO endonuclease, the percentage of T-DNA insertions at the DSB increased to 16%. Therefore, T-DNA integration at the DSB is likely suppressed by the presence of the Rad52 protein (van Attikum 2003).

Inactivation of Ku70 or Ku80 prevents non-homologous integration and can thus be used to obtain transformants that have integrated the transgenes by homologous recombination (Hooykaas et al. 2000; van Attikum et al. 2001). This idea has been used to promote gene targeting in many different yeasts and fungi, including the yeast *K. lactis* (Kooistra et al. 2004) and the filamentous fungus *Neurospora crassa* (Ninomiya et al. 2004). Gene targeting can be enhanced by the introduction of a DSB in the target locus. We found that the signal for transport by the T4SS of the *Agrobacterium* virulence system lies in the 30 C-terminal amino acids of translocated virulence proteins. Coupling of this transport signal to the C-terminus of heterologous proteins enables their mobilization from *Agrobacterium* to eukaryotic target cells (Vergunst et al. 2005). In this way, recombinases such as Cre and nucleases such as the homing endonuclease I-*Sce*I can be transferred by *Agrobacterium* into host cells together with the T-DNA vector (Vergunst et al. 2000; van Kregten et al. 2009). Translocation of I-*Sce*I together with a gene-targeting vector leads to an increased frequency of gene targeting in yeast if there is a nuclease recognition site at the target locus (Rolloos et al. 2015).

7 Conclusions

AMT has become a widely used tool for transformation of various fungi because AMT has important advantages over other transformation methods. The transformation protocols are relatively simple, frequencies of transformation are relatively high, single-copy integration events are more numerous than with other methods, and gene targeting is sometimes achieved much more easily with AMT. Non-homologous T-DNA integration at random positions in the genome of fungi has made *Agrobacterium* a useful tool for mutagenesis and gene tagging in fungi. Conversely, for biotechnology purposes targeted integration to a safe haven in the genome is preferred, thus guaranteeing stability of the transgene and preventing unwanted mutagenesis. This can easily be realized in the yeast *S. cerevisiae* but is much less efficient or virtually impossible in other yeasts and fungi. Identification of the genes involved in DSB repair by NHEJ as the key factors that control (T-)DNA integration by non-homologous integration in *S. cerevisiae* (van Attikum et al. 2001) enabled the development of strains with strongly improved gene targeting (e.g., Kooistra et al. 2004; Ninomiya et al. 2004). Such targeting is potentiated by inactivating or disrupting one of the key players of NHEJ (i.e. Ku70, Ku80 or Lig4).

The presence of natural T-DNAs has been discovered in the genome of some plant species, including the edible sweet potato (Kyndt et al. 2015). The ability of *Agrobacterium* to transfer T-DNA to yeasts and fungi may also occur in nature, at sites where these microorganisms are living in close proximity and where AMT may

become possible when phenolic *vir* inducers are abundantly secreted by wounded plant cells. Such "natural" transformation has been observed in in vitro experiments, whereby *Agrobacterium* and the fungus *Verticillium albo-atrum* were co-cultivated on plant tissues (potato and carrot slices, tobacco leaves and stems) in the absence of acetosyringone (Knight et al. 2010). Where horizontal transfer of T-DNA may have contributed to the evolution of plant species (Kyndt et al. 2015), it is not unreasonable to suspect that AMT may have likewise contributed to the evolution of some fungi.

The use of the yeast *S. cerevisiae* as a model for the studies of AMT has not only led to the development of an efficient new method for the transformation of various fungi, but also generated new insight in the process of AMT itself. Yeast cells are transparent and do not contain large amounts of endogenous fluorescent compounds, such as chlorophyll and are, therefore, very suitable for microscopical analysis. These attributes enabled the development of a system based on split GFP by which virulence protein translocation into host cells could be visualized in real time; the results of which subsequently could be applied in plants (Li et al. 2014; Sakalis et al. 2014). The finding that AMT of yeast and fungi is still efficient in the absence of the translocated virulence proteins VirD5, VirE3, and VirF indicated that these proteins are not essential for the transformation process itself, but rather have plant-specific functions. The translocated VirE2 protein is slightly different as some reduction in the transformation of yeasts and fungi was seen. Nevertheless, we conclude that even in the absence of VirE2, the T-strand can reach the fungal nucleus and integrate into the genome. As expected, the NLS of VirD2 is essential for AMT and is probably the main factor essential for nuclear targeting. As to (non-homologous) T-DNA integration, studies in yeasts and fungi revealed that this occurred (almost) exclusively by NHEJ. Whereas T-DNA integration is abolished in fungal NHEJ-mutants, *A. thaliana* mutants in *AtKU70*, *AtKU80*, and *AtLIG4* are still efficiently transformed by AMT (Park et al. 2015). Recently, it was found that T-DNA integration in plants occurs through a process of alternative end-joining, which is independent of Ku70, Ku80, and Lig4, but which is mediated by the polymerase PolQ, which is conserved in animals and plants, but not present in yeasts and fungi (van Kregten et al. 2016). This underscored that (T)-DNA integration is determined by the host cell and the DNA recombination enzymes available in the host cell, and thus may occur by entirely different means in different host cells. We conclude that the yeast *S.cerevisiae* is an excellent model organism to start the study of AMT and virulence gene function, but that one should never forget to put the findings to the test in fungi and plants, for which AMT is a preferred method of transformation.

Acknowledgements We acknowledge the contributions of Amke den Dulk-Ras, Alice Beijersbergen, Paul Bundock, Haico van Attikum, Jesus Escudero, Marcel de Groot, Carolien Michielse, Barbara Schrammeijer, Martijn Rolloos, and Suzanne Wolterink to our research regarding AMT of yeast and fungi. We had to restrict the numbers of references. Therefore, we apologize to those whose work related to the topic we did not refer to. The work in our laboratory was supported over the years by grants from the divisions Earth and Life Sciences (ALW) and Chemical Sciences (CW) of the Netherlands Organization of Scientific Research NWO, the organization for Applied Research (STW), and the Royal Netherlands Academy of Arts and Sciences (KNAW).

References

Almeida AJ, Carmona JAC, Cunha C et al (2007) Towards a molecular genetic system for the pathogenic fungus *Paracoccidioides brasilensis*. Fungal Genet Biol 44:1387–1398

Amey RC, Mills PR, Bailey A, Foster GD (2003) Investigating the role of a *Verticillium fungicola* beta-1, 6-glucanase during infection of *Agaricus bisporus* using targeted gene disruption. Fungal Genet Biol 39:264–275

Anand A, Krichevsky A, Schornack S et al (2007) *Arabidopsis* VIRE2 INTERACTING PROTEIN2 is required for *Agrobacterium* T-DNA integration in plants. Plant Cell 19:1695–1708

Atmakuri K, Cascales E, Burton OT et al (2007) *Agrobacterium* ParA/MinD-like VirC1 spatially coordinates early conjugative DNA transfer reactions. EMBO J 26:2540–2551

Ballas N, Citovsky V (1997) Nuclear localization signal binding protein from *Arabidopsis* mediates nuclear import of *Agrobacterium* VirD2 protein. Proc Natl Acad Sci U S A 94:10723–10728

Baumgartner K, Fujiyoshi P, Foster GD, Bailey AM (2010) *Agrobacterium tumefaciens*-mediated transformation for investigation of somatic recombination in the fungal pathogen *Armillaria mellea*. Appl Environ Microbiol 76:7990–7996

Betts MF, Tucker SL, Galadima N et al (2007) Development of a high throughput transformation system for insertional mutagenesis in *Magnaporthe oryzae*. Fungal Genet Biol 44:1035–1049

Beijersbergen A, Dulk-Ras AD, Schilperoort RA, Hooykaas PJ (1992) Conjugative transfer by the virulence system of *Agrobacterium tumefaciens*. Nature 256:1324–1327

Bhattacharjee S, Lee LY, Oltmanns H, Cao H, Veena Cuperus J, Gelvin SB (2008) IMPa-4, an Arabidopsis importin α isoform, is preferentially involved in Agrobacterium-mediated plant transformation. Plant Cell 20:2661–2680

Biggins S, Severin FF, Bhalla N et al (1999) The conserved protein kinase Ipl1 regulates microtubule binding to kinetochores in budding yeast. Genes Devel 13:532–544

Blaise F, Rémy E, Meyer M et al (2007) A critical assessment of *Agrobacterium tumefaciens*-mediated transformation as a tool for pathogenicity gene discovery in the phytopathogenic fungus *Leptosphaeria maculans*. Fungal Genet Biol 44:123–138

Bourras S, Meyer M, Grandaubert J et al (2012) Incidence of genome structure, DNA asymmetry, and cell physiology on T-DNA integration in chromosomes of the phytopathogenic fungus *Leptosphaeria maculans*. G3 (Bethesda) 2:891–904

Bundock P, Dulk-Ras A, Beijersbergen A, Hooykaas PJ (1995) Trans-kingdom T-DNA transfer from *Agrobacterium tumefaciens* to *Saccharomyces cerevisiae*. EMBO J 14:3206–3214

Bundock P, Hooykaas PJ (1996) Integration of *Agrobacterium tumefaciens* T-DNA in the *Saccharomyces cerevisiae* genome by illegitimate recombination. Proc Natl Acad Sci U S A 93:15272–15275

Bundock P, Mroczek K, Winkler AA et al (1999) T-DNA from *Agrobacterium tumefaciens* as an efficient tool for gene targeting in *Kluyveromyces lactis*. Mol Gen Genet 261:115–121

Bundock P, van Attikum H, Dulk-Ras A, Hooykaas PJ (2002) Insertional mutagenesis in yeasts using T-DNA from *Agrobacterium tumefaciens*. Yeast 19:529–536

Campoy S, Perez F, Martin JF et al (2003) Stable transformants of the azaphilone pigment-producing *Monascus purpureus* obtained by protoplast transformation and *Agrobacterium*-mediated DNA transfer. Curr Genet 43:447–452

Chilton MD, Que Q (2003) Targeted integration of T-DNA into the tobacco genome at double-stranded breaks: new insights on the mechanism of T-DNA integration. Plant Physiol 133:956–965

Choi J, Park J, Jeon J et al (2007) Genome-wide analysis of T-DNA integration into the chromosomes of *Magnaporthe oryzae*. Mol Microbiol 66:371–382

Combier JP, Melayah D, Raffier C et al (2003) *Agrobacterium tumefaciens*-mediated transformation as a tool for insertional mutagenesis in the symbiotic ectomycorrhizal fungus *Hebeloma cylindrosporum*. FEMS Microbiol Lett 220:141–148

Crane YM, Gelvin SB (2007) RNAi-mediated gene silencing reveals involvement of *Arabidopsis* chromatin-related genes in *Agrobacterium*-mediated root transformation. Proc Natl Acad Sci U S A 104:15156–15161

Critchlow SE, Jackson SP (1998) DNA end-joining: from yeast to man. Trends Biochem Sci 23:394–398

Crotti LB, Basrai MA (2004) Functional roles for evolutionarily conserved Spt4p at centromeres and heterochromatin in *Saccharomyces cerevisiae*. EMBO J 23:1804–1814

De Cleene M, De Ley J (1976) The host range of crown gall. Botan Rev 42:389–466

Degefu Y, Hanif M (2003) *Agrobacterium tumefaciens*-mediated transformation of *Helminthosporium turcicum*, the maize leaf-blight fungus. Arch Microbiol 180:279–284

de Groot MJ, Bundock P, Hooykaas PJ, Beijersbergen AG (1998) *Agrobacterium tumefaciens*-mediated transformation of filamentous fungi. Nat Biotechnol 16:839–842

Deng W, Chen L, Wood DW et al (1998) *Agrobacterium* VirD2 protein interacts with plant host cyclophilins. Proc Natl Acad Sci U S A 95:7040–7045

Escudero J, den Dulk-Ras H, Regensburg-Tuïnk TJG, Hooykaas PJJ (2003) VirD4-independent transformation by CloDF13 evidences an unknown factor required for the genetic colonization of plants via *Agrobacterium*. Mol Microbiol 47:891–901

Flowers JL, Vaillancourt LJ (2005) Parameters affecting the efficiency of *Agrobacterium tumefaciens*-mediated transformation of *Colletotrichum graminicola*. Curr Genet 15:1–9

Frandsen RJ (2011) A guide to binary vectors and strategies for targeted genome modification in fungi using *Agrobacterium tumefaciens*-mediated transformation. J Microbiol Methods 87:247–262

Fullner KJ, Nester EW (1996) Temperature affects the T-DNA transfer machinery of *Agrobacterium tumefaciens*. J Bacteriol 178:1498–1504

García-Cano E, Magori S, Sun Q et al (2015) Interaction of *Arabidopsis* trihelix-domain transcription factors VFP3 and VFP5 with *Agrobacterium* virulence protein VirF. PLoS ONE 10:e0142128

García-Cano E, Hak H, Magori S et al (2018) The *Agrobacterium* F-box protein effector VirF destabilizes the *Arabidopsis* GLABROUS1 enhancer/binding protein-like transcription factor VFP4, a transcriptional activator of defense response genes. Mol Plant Microbe Interact 31:576–586

Garcia-Rodriguez FM, Schrammeijer B, Hooykaas PJJ (2006) The *Agrobacterium* VirE3 effector protein: a potential plant transcriptional activator. Nucleic Acids Res 34:6496–6504

Gardiner DM, Howlett BJ (2004) Negative selection using thymidine kinase increases the efficiency of recovery of transformants with targeted genes in the filamentous fungus *Leptosphaeria maculans*. Curr Genet 45:249–255

Gelvin SB (2003) *Agrobacterium*-mediated plant transformation: the biology behind the "gene-jockeying" tool. Microbiol Mol Biol Rev 67:16–37

Gelvin SB (2010) Plant proteins involved in *Agrobacterium*-mediated genetic transformation. Annu Rev Phytopathol 48:45–68

Grimaldi B, de Raaf MA, Filetici P et al (2005) *Agrobacterium*-mediated gene transfer and enhanced green fluorescent protein visualization in the mycorrhizal ascomycete *Tuber borchii*: a first step towards truffle genetics. Curr Genet 48:69–74

Hatoh K, Izumitsu K, Morita A et al (2013) Transformation of the mushroom species *Hypsizigus marmoreus*, *Flammulina velutipes*, and *Grifola frondosa* by an *Agrobacterium*-mediated method using a universal transformation plasmid. Mycoscience 54:8–12

Hauer MH, Gasser SM (2018) Chromatin and nucleosome dynamics in DNA damage and repair. Genes Devel 31:2204–2221

Hooykaas PJJ, van Attikum H, Bundock P (2000) Nucleic acid integration in eukaryotes. EP00204693.6

Hooykaas PJ, Hofker M, den Dulk-Ras H, Schilperoort RA (1984) A comparison of virulence determinants in an octopine Ti plasmid, a nopaline Ti plasmid, and an Ri plasmid by complementation analysis of *Agrobacterium tumefaciens* mutants. Plasmid 11:195–205

Hooykaas-van Slogteren GMS, Hooykaas PJJ, Schilperoort RA (1984) Expression of Ti plasmid genes in monocotyledonous plants infected with *Agrobacterium tumefaciens*. Nature 311:763–764

Idnurm A, Reedy JL, Nussbau JC, Heitman J (2004) *Cryptococcus neoformans* virulence gene discovery through insertional mutagenesis. Eukaryot Cell 3:420–429

Ishida Y, Saito H, Ohta S et al (1996) High efficiency transformation of maize (*Zea mays* L.) mediated by *Agrobacterium tumefaciens*. Nat Biotech 14:745–750

Jeon J, Park SY, Chi MH et al (2007) Genome-wide functional analysis of pathogenicity genes in the rice blast fungus. Nature Genet 39:561–565

Kilaru S, Collins CM, Hartley AJ et al (2009) Establishing molecular tools for genetic manipulation of the pleuromutilin-producing fungus *Clitopilus passeckerianus*. Appl Environ Microbiol 75:7196–7204

Knight CJ, Bailey AM, Foster GD (2010) Investigating *Agrobacterium*-mediated transformation of *Verticillium albo-atrum* on plant surfaces. PLoS ONE 5:e13684

Kooistra R, Hooykaas PJ, Steensma HY (2004) Efficient gene targeting in *Kluyveromyces lactis*. Yeast 21:781–792

Krysan PJ, Young JC, Sussman MR (1999) T-DNA as an insertional mutagen in *Arabidopsis*. Plant Cell 11:2283–2290

Kyndt T, Quispe D, Zhai H et al (2015) The genome of cultivated sweet potato contains *Agrobacterium* T-DNAs with expressed genes: An example of a naturally transgenic food crop. Proc Natl Acad Sci U S A 112:5844–5849

Lacroix B, Citovsky V (2013) Characterization of VIP1 activity as a transcriptional regulator *in vitro* and *in planta*. Sci Rep 3:2440

Leclerque A, Wan H, Abschütz A et al (2004) *Agrobacterium*-mediated insertional mutagenesis (AIM) of the entomopathogenic fungus *Beauveria bassiana*. Curr Genet 45:111–119

Lee MH, Bostock RM (2006) *Agrobacterium* T-DNA-mediated integration and gene replacement in the brown rot pathogen *Monilinia fructicola*. Curr Genet 49:309–322

Lewis LK, Resnick MA (2000) Tying up loose ends: nonhomologous end-joining in *Saccharomyces cerevisiae*. Mutat Res 451:71–89

Li G, Zhou Z, Liu G et al (2007) Characterization of T-DNA insertion patterns in the genome of rice blast fungus *Magnaporthe oryzae*. Curr Genet 51:233–243

Li X, Yang Q, Tu H et al (2014) Direct visualization of *Agrobacterium*-delivered VirE2 in recipient cells. Plant J 77:487–495

Loppnau P, Tanguay P, Breuil C (2004) Isolation and disruption of the melanin pathway polyketide synthase gene of the softwood deep stain fungus *Ceratocystis resinifera*. Fungal Genet Biol 41:33–41

Luo Y, Chen Z, Zhu D et al (2015) Yeast actin-related protein ARP6 negatively regulates *Agrobacterium*-mediated transformation of yeast cell. BioMed Res Int article ID 275092. http://dx.doi.org/10.1155/2015/275092

Magori S, Citovsky V (2011) *Agrobacterium* counteracts host-induced degradation of its effector F-box protein. Sci Signal 4(195):ra69

Malonek S, Meinhardt F (2001) *Agrobacterium tumefaciens*-mediated genetic transformation of the phytopathogenic ascomycete *Calonectria morganii*. Curr Genet 40:152–155

McClelland CM, Chang YC, Kwon-Chung KJ (2005) High frequency transformation of *Cryptococcus neoformans* and *Cryptococcus gattii* by *Agrobacterium tumefaciens*. Fungal Genet Biol 42:904–913

Meng Y, Patel G, Heist M et al (2007) A systematic analysis of T-DNA insertion events in *Magnaporthe oryzae*. Fungal Genet Biol 44:1050–1064

Meyer V, Mueller D, Strowig T, Stahl U (2003) Comparison of different transformation methods for *Aspergillus giganteus*. Curr Genet 43:371–377

Michielse CB, Ram AFJ, Hooykaas PJ, van den Hondel CAMJJ (2004a) *Agrobacterium*-mediated transformation of *Aspergillus awamori* in the absence of full length VirD2, VirC2 or VirE2 leads to insertion of aberrant T-DNA structures. J Bacteriol 186:2038–2045

Michielse CB, Ram AFJ, Hooykaas PJ, van den Hondel CAMJJ (2004b) Role of bacterial virulence proteins in *Agrobacterium*-mediated transformation of *Aspergillus awamori*. Fungal Genet Biol 45:571–578

Michielse CB, Hooykaas PJ, van den Hondel CAMJJ, Ram AFJ (2005a) *Agrobacterium*-mediated transformation as a tool for functional genomics in fungi. Curr Genet 48:1–17

Michielse CB, Arentshorst M, Ram AF, van den Hondel CAMJJ (2005b) *Agrobacterium*-mediated transformation leads to improved gene replacement efficiency in *Aspergillus awamori*. Fungal Genet Biol 42:9–19

Michielse CB, Hooykaas PJ, van den Hondel CA, Ram AF (2008) *Agrobacterium*-mediated transformation of the filamentous fungus *Aspergillus awamori*. Nat Protocols 3:1671–1678

Michielse CB, van Wijk R, Reijnen L et al (2009) Insight into the molecular requirements for pathogenicity of *Fusarium oxysporum* f. sp. lycopersici through large-scale insertional mutagenesis. Genome Biol 10:R4

Mikosch TS, Lavrijssen B, Sonnenberg AS, van Griensven LJ (2001) Transformation of the cultivated mushroom *Agaricus bisporus* (Lange) using T-DNA from *Agrobacterium tumefaciens*. Curr Genet 39:35–39

Mullins ED, Chen X, Romaine P at al. (2001) *Agrobacterium*-mediated transformation of *Fusarium oxysporum*: an efficient tool for insertional mutagenesis and gene transfer. Phytopathol 91:173–180

Murata H, Sunagawa M, Igasaki T, Shishido K (2006) *Agrobacterium*-mediated transformation of the ectomycorrhizal basidiomycete *Tricholoma matsutake* that produces commercially valuable fruit bodies, matsutake. Mycoscience 47:228–231

Nemecek JC, Wüthrich M, Klein BS (2006) Global control of dimorphism and virulence in fungi. Science 312:583–588

Nester EW, Gordon MP, Amasino RM, Yanofsky MF (1984) Crown gall: a molecular and physiological analysis. Annu Rev Plant Physiol 35:387–413

Ninomiya Y, Suzuki K, Ishii C, Inoue H (2004) Highly efficient gene replacements in *Neurospora* strains deficient for nonhomologous end-joining. Proc Natl Acad Sci U S A 101:12248–12253

Niu X (2013) Functional analysis of Agrobacterium virulence genes. Ph.D. thesis. Leiden University, Leiden, The Netherlands

Niu X, Zhou M, Henkel CV et al (2015) The *Agrobacterium tumefaciens* virulence protein VirE3 is a transcriptional activator of the F-box gene VBF. Plant J 84:914–924

Noack-Schonmann S, Bus T, Banasiak R et al. (2014) Genetic transformation of *Knufia petricola* A95—a model organism for biofilm-material interactions AMB Express 4:80

Offringa R, de Groot MJA, Haagsman HJ et al (1990) Extrachromosomal homologous recombination and gene targeting in plant cells after *Agrobacterium* mediated transformation. EMBO J 9:3077–3084

Ohmine Y, Satoh Y, Kiyokawa K et al (2016) DNA repair genes RAD52 and SRS2, a cell wall synthesis regulator gene SMI1, and the membrane sterol synthesis scaffold gene ERG28 are important in efficient *Agrobacterium*-mediated yeast transformation with chromosomal T-DNA. BMC Microbiol 16:58

Pansegrau W, Schoumacher F, Hohn B, Lanka E (1993) Site-specific cleavage and joining of single-stranded DNA by VirD2 protein of *Agrobacterium tumefaciens* Ti plasmids: analogy to bacterial conjugation. Proc Natl Acad Sci U S A 90:11538–11542

Pardo AG, Hanif M, Raudaskoski M, Gorfer M (2002) Genetic transformation of ectomycorrhizal fungi mediated by *Agrobacterium tumefaciens*. Mycol Res 106:132–137

Park SM, Kim DK (2004) Transformation of a filamentous fungus *Cryphonectria parasitica* using *Agrobacterium tumefaciens*. Biotechnol Bioprocess Eng 9:217–222

Park SY, Jeong MH, Wang HY et al (2013) *Agrobacterium tumefaciens*-mediated transformation of the lichen fungus, *Umbilicaria muehlenbergii*. PLoS ONE 8:e83896

Park SY, Vaghchhipawala Z, Vasudevan B et al (2015) *Agrobacterium* T-DNA integration into the plant genome can occur without the activity of key non-homologous end-joining proteins. Plant J 81:934–946

Piers KL, Heath JD, Liang X et al (1996) *Agrobacterium tumefaciens*-mediated transformation of yeast. Proc Natl Acad Sci U S A 93:1613–1618

Rho HS, Kang S, Lee YH (2001) *Agrobacterium tumefaciens*-mediated transformation of the plant pathogenic fungus, *Magnaporthe grisea*. Mol Cells 12:407–411

Risseeuw E, Franke-van Dijk ME, Hooykaas PJ (1996) Integration of an insertion-type transferred DNA vector from *Agrobacterium tumefaciens* into the *Saccharomyces cerevisiae* genome by gap repair. Mol Cell Biol 16:5924–5932

Roberts RL, Metz M, Monks DE et al (2003) Purine synthesis and increased *Agrobacterium tumefaciens* transformation of yeast and plants. Proc Natl Acad Sci U S A 100:6634–6639

Rogers CW, Challen MP, Green JR, Whipps JM (2004) Use of REMI and *Agrobacterium*-mediated transformation to identify pathogenicity mutants of the biocontrol fungus, *Coniothyrium minitans*. FEMS Microbiol Lett 241:207–214

Rolland S, Jobic C, Fevre M, Bruel C (2003) *Agrobacterium*-mediated transformation of *Botrytis cinerea*, simple purification of monokaryotic transformants and rapid conidia-based identification of the transfer-DNA host genomic DNA flanking sequences. Curr Genet 44:164–171

Rolloos M, Dohmen MHC, Hooykaas PJJ, van der Zaal BJ (2014) Involvement of Rad52 in T-circle formation during *Agrobacterium tumefaciens*-mediated transformation of *Saccharomyces cerevisiae*. Mol Microbiol 91:1240–1251

Rolloos M, Hooykaas PJJ, van der Zaal BJ (2015) Enhanced targeted integration mediated by translocated I-*Sce*I during the *Agrobacterium* mediated transformation of yeast. Sci Rep 5:8345

Rossi L, Hohn B, Tinland B (1996) Integration of complete transferred DNA units is dependent on the activity of virulence E2 protein of *Agrobacterium tumefaciens*. Proc Natl Acad Sci U S A 93:126–130

Sakalis PA, van Heusden GPH, Hooykaas PJJ (2014) Visualization of VirE2 protein translocation by the *Agrobacterium* type IV secretion system into host cells. Microbiol Open 3:104–117

Salomon S, Puchta H (1998) Capture of genomic and T-DNA sequences during double-strand break repair in somatic plant cells. EMBO J 17:6086–6095

Schrammeijer B, Risseeuw E, Pansegrau W et al (2001) Interaction of the virulence protein VirF of *Agrobacterium tumefaciens* with plant homologs of the yeast Skp1 protein. Curr Biol 11:258–262

Schrammeijer B, Dulk-Ras A, Vergunst AC et al (2003) Analysis of Vir protein translocation from *Agrobacterium tumefaciens* using *Saccharomyces cerevisiae* as a model: evidence for transport of a novel effector protein VirE3. Nucleic Acids Res 31:860–868

Shi Y, Lee LY, Gelvin SB (2014) Is VIP1 important for *Agrobacterium*-mediated transformation? Plant J 79:848–860

Singer K, Shiboleth YM, Li J, Tzfira T (2012) Formation of complex extrachromosomal T-DNA structures in *Agrobacterium tumefaciens*-infected plants. Plant Physiol 160:511–522

Skowyra D, Craig KL, Tyers M et al (1997) F-box proteins are receptors that recruit phosphorylated substrates to the SCF ubiquitin-ligase complex. Cell 91:209–219

Smeenk G, van Attikum H (2013) The chromatin response to DNA breaks: leaving a mark on genome integrity. Annu Rev Biochem 82:55–80

Soltani J (2009) Host genes involved in *Agrobacterium*-mediated transformation. Ph.D. thesis. Leiden University, Leiden, The Netherlands

Soltani J, van Heusden GPH, Hooykaas PJJ (2008) *Agrobacterium*-mediated transformation of non-plant organisms. In: Tzfira T, Citovsky V (eds) *Agrobacterium*: from biology to biotechnology, pp 649–675. Springer Press, New York

Soltani J, van Heusden GPH, Hooykaas PJJ (2009) Deletion of host histone acetyltransferases and deacetylases strongly affects *Agrobacterium*-mediated transformation of *Saccharomyces cerevisiae*. FEMS Microbiol Lett 298:228–233

Sugui JA, Chang YC, Kwon-Chung KJ (2005) *Agrobacterium tumefaciens*-mediated transformation of *Aspergillus fumigatus*: an efficient tool for insertional mutagenesis and targeted gene disruption. Appl Environ Microbiol 71:1798–1802

Sullivan TD, Rooney PJ, Klein BS (2002) *Agrobacterium tumefaciens* integrates transfer DNA into single chromosomal sites of dimorphic fungi and yields homokaryotic progeny from multinucleate yeast. Eukaryot Cell 1:895–905

Tanguay P, Breuil C (2003) Transforming the sapstaining fungus *Ophiostoma piceae* with *Agrobacterium tumefaciens*. Can J Microbiol 49:301–304

Tsuji G, Fujii S, Fujihara N et al (2003) *Agrobacterium tumefaciens*-mediated transformation for random insertional mutagenesis in *Colletotrichum lagenarium*. J Gen Plant Pathol 69:230–239

Turk SC, Melchers LS, den Dulk-Ras H et al (1991) Environmental conditions differentially affect vir gene induction in different *Agrobacterium* strains. Role of the VirA sensor protein. Plant Mol Biol 16:1051–1059

Tzfira T, Rhee Y, Chen MH et al (2000) Nucleic acid transport in plant-microbe interactions: the molecules that walk through the walls. Annu Rev Microbiol 54:187–219

Tzfira T, Vaidya M, Citovsky V (2001) VIP1, an *Arabidopsis* protein that interacts with *Agrobacterium* VirE2, is involved in VirE2 nuclear import and *Agrobacterium* infectivity. EMBO J 20:3596–3607

Tzfira T, Frankman LR, Vaidya M, Citovsky V (2003) Site-specific integration of *Agrobacterium tumefaciens* T-DNA via double-stranded intermediates. Plant Physiol 133:1011–1023

Tzfira T, Vaidya M, Citovsky V (2004) Involvement of targeted proteolysis in plant genetic transformation by *Agrobacterium*. Nature 431:87–92

van Attikum H, Bundock P, Hooykaas PJ (2001) Non-homologous end-joining proteins are required for *Agrobacterium* T-DNA integration. EMBO J 20:6550–6558

van Attikum H (2003) Genetic requirements for the integration of *Agrobacterium* T-DNA in the eukaryotic genome. PhD thesis. Leiden University, Leiden, The Netherlands

van Attikum H, Hooykaas PJ (2003) Genetic requirements for the targeted integration of Agrobacterium T-DNA in *Saccharomyces cerevisiae*. Nucleic Acids Res 31:826–832

van Kregten M, Lindhout BI, Hooykaas PJJ, van der Zaal BJ (2009) *Agrobacterium*-mediated T-DNA transfer and integration by minimal VirD2 consisting of the relaxase domain and a type IV secretion system translocation signal. Mol Plant-Micr Int 22:1356–1365

van Kregten M, de Pater S, Romeijn R, van Schendel R, Hooykaas PJJ, Tijsterman M (2016) T-DNA integration in plants results from polymerase-θ-mediated DNA repair. Nature Plants 2:16164

Vergunst AC, Schrammeijer B, Den Dulk A et al (2000) VirB/D4-dependent protein translocation from *Agrobacterium* into plant cells. Science 290:979–982

Vergunst AC, van Lier MCM, den Dulk-Ras A et al (2005) Positive charge is an important feature of the C-terminal teransport signal of the VirB/D4-translocated proteins of *Agrobacterium*. Proc Natl Acad Sci U S A 102:832–837

Walton FJ, Idnurm A, Heitman J (2005) Novel gene functions required for melanization of the human pathogen *Cryptococcus neoformans*. Mol Microbiol 57:1381–1396

Wang Y, Peng W, Zhou X, Huang F, Shao L, Luo M (2014) The putative *Agrobacterium* transcriptional activator-like virulence protein VirD5 may target T-complex to prevent the degradation of coat proteins in the plant cell nucleus. New Phytol 203:1266–1281

Wang Y, Wang R, Mao WJ et al (2016) A simple and efficient transformation system for the edible mushroom *Pleurotus eryngii*. Mycoscience 57:356–360

White D, Chen W (2006) Genetic transformation of *Ascochyta rabiei* using *Agrobacterium*-mediated transformation. Curr Genet 49:272–280

Winans SC (1991) An *Agrobacterium* two-component regulatory system for the detection of chemicals released from plant wounds. Mol Microbiol 5:2345–2350

Wolterink-van Loo S, Ayala AAE, Hooykaas PJJ, van Heusden GPH (2015) Interaction of the *Agrobacterium tumefaciens* virulence protein VirD2 with histones. Microbiology 161:401–410

Yousefi-Pour HM, Soltani J, Nazeri S (2013) A survey on optimization of *Agrobacterium*-mediated genetic transformation of the fungus *Colletotrichum gloeosporioides*. J Cell Mol Res 5:35–41

Yu X, Huo L, Liu H et al (2015) Melanin is required for the formation of the multi-cellular conidia in the endophytic fungus *Pestalotiopsis microspora*. Microbiol Res 179:1–11

Zaltsman A, Krichevsky A, Loyter A, Citovsky V (2010) *Agrobacterium* induces expression of a host F-box protein required for tumorigenicity. Cell Host Microbe 7:197–209

Zhang A, Lu P, Dahl-Roshak AM et al (2003) Efficient disruption of a polyketide synthase gene (*pks1*) required for melanin synthesis through *Agrobacterium*-mediated transformation of *Glarea lozoyensis*. Mol Genet Genomics 268:645–655

Zhang J, Shi L, Chen H et al (2014) An efficient *Agrobacterium*-mediated transformation method for the edible mushroom *Hypsizygus marmoreus*. Microbiol Res 169:741–748

Zhang T, Ren P, Chaturvedi V et al (2015) Development of an *Agrobacterium*-mediated transformation system for the cold-adapted fungi *Pseudogymnoascus destructans* and *P. pannorum*. Fungal Genet Biol 81:73–81

Zhang X (2016) Functional analysis of *Agrobacterium tumefaciens* virulence protein VirD5. Ph.D. thesis. Leiden University, Leiden, The Netherlands

Zhang X, van Heusden GPH, Hooykaas PJJ (2017) Virulence protein VirD5 of *Agrobacterium tumefaciens* binds to kinetochores in host cells via an interaction with Spt4. Proc Natl Acad Sci U S A 114:10238–10243

Zhu J, Oger PM, Schrammeijer B et al (2000) The bases of crown gall tumorigenesis. J Bacteriol 182:3885–3895

Zhu Y, Nam J, Humara JM et al (2003) Identification of *Arabidopsis rat* mutants. Plant Physiol 132:494–505

Zwiers LH, de Waard MA (2001) Efficient *Agrobacterium tumefaciens*-mediated gene disruption in the phytopathogen *Mycosphaerella graminicola*. Curr Genet 39:388–393

The *Agrobacterium* Phenotypic Plasticity (*Plast*) Genes

Léon Otten

Contents

1 Introduction 376
2 An Updated Version of the Plast Protein Family 379
2.1 Who Qualifies as a Plast Family Member? 379
2.2 A Plast Protein Tree 380
3 Biological Effects of *Plast* Genes 385
3.1 General Considerations 385
3.2 *rolB* and Its Closest Relatives 387
3.3 *rolC* 393
3.4 Gene *6b* 396
3.5 *orf13* 401
3.6 *orf14* 403
3.7 Gene *5* 403
3.8 *lso* 404
3.9 Gene *6a* 404
3.10 Genes *3'*, *4'*, *7'*, *b*, *c'*, *d*, and *e* 405
4 What Could Be the Ancestral *Plast* Gene Mechanism? 405
5 Future Research 406
5.1 Search for the Fundamental Mechanism of *Plast* Genes 406
5.2 Role of *Plast* Genes in Tumors, Hairy Roots, and Natural Transformants 407
5.3 Possible Applications 408
6 Conclusions 408
References 409

Abstract The transfer of T-DNA sequences from *Agrobacterium* to plant cells is a well-understood process of natural genetic engineering. The expression of T-DNA genes in plants leads to tumors, hairy roots, or transgenic plants. The transformed cells multiply and synthesize small molecules, called opines, used by Agrobacteria for their growth. Several T-DNA genes stimulate or influence plant growth. Among these, *iaaH* and *iaaM* encode proteins involved in auxin synthesis, whereas *ipt*

L. Otten (✉)
IBMP, Strasbourg, France
e-mail: leon.otten@ibmp-cnrs.unistra.fr

Current Topics in Microbiology and Immunology (2018) 418:375–419
DOI 10.1007/82_2018_93

Published Online: 17 May 2018

encodes a protein involved in cytokinin synthesis. Growth can also be induced or modified by other T-DNA genes, collectively called *plast* genes (for phenotypic plasticity). The *plast* genes are defined by their common ancestry and are mostly found on T-DNAs. They can influence plant growth in different ways, but the molecular basis of their morphogenetic activity remains largely unclear. Only some *plast* genes, such as *6b*, *rolB*, *rolC*, and *orf13*, have been studied in detail. *Plast* genes have a significant potential for applied research and may be used to modify the growth of crop plants. In this review, I summarize the most important findings and models from 30 years of *plast* gene research and propose some outlooks for the future.

1 Introduction

The transformation process whereby *Agrobacterium* introduces part of its DNA into plant cells is a well-known process (Gelvin 2012; Kado 2014; Nester 2015). Under both natural conditions and in the laboratory, Agrobacteria can infect a large range of plants and transfer part of their DNA from a large tumor- or root-inducing plasmid (pTi or pRi) into plant cells. Expression of transferred DNA (T-DNA) causes abnormal growth and leads to the production of small molecules (opines) which can be used by the bacteria for their growth. Thus, using genetic transformation, the bacteria redirects the plant's metabolism for their own benefit (Schell et al. 1979). This unique natural phenomenon has been used extensively in plant biotechnology. In order to adapt the Ti plasmid for gene transfer, the original T-DNA genes had to be removed (the "disarming" of the Ti/Ri plasmid). Unfortunately, "taming" *Agrobacterium* for use as a gene transfer tool for biotechnology purposes also reduced the original interest in determining the function of T-DNA genes. In this review, I will argue that many interesting aspects of the natural T-DNA genes remain to be studied and could be of great interest for crop improvement.

Different types of Agrobacteria have been described, and hundreds of natural isolates have been collected. On the basis of their phytopathogenic properties, Agrobacteria have been divided into *Agrobacterium radiobacter* (avirulent), *Agrobacterium tumefaciens*, *Agrobacterium rhizogenes* (Riker 1930), *Agrobacterium vitis* (Kerr and Panagopoulos 1977; Ophel and Kerr 1990), and a few additional groups, such as *Agrobacterium rubi* (Hildebrand 1940) and *Agrobacterium larrymoorii* (Bouzar and Jones 2001). *A. tumefaciens* and *A. vitis* mostly induce tumors (crown galls) or shoot-like structures called teratomata. *A. rhizogenes* induces the neoformation of modified, transformed roots (hairy roots), and can also create new transgenic plants (natural transformants) with cellular T-DNA (cT-DNA) sequences. Such natural transformants most likely arose by spontaneous regeneration of hairy roots (White et al. 1983; Suzuki et al. 2002; Chen et al. 2014; Chen and Otten 2017). As some of these plants produce opines (Chen et al. 2016), it is possible that natural transformants are not just evolutionary

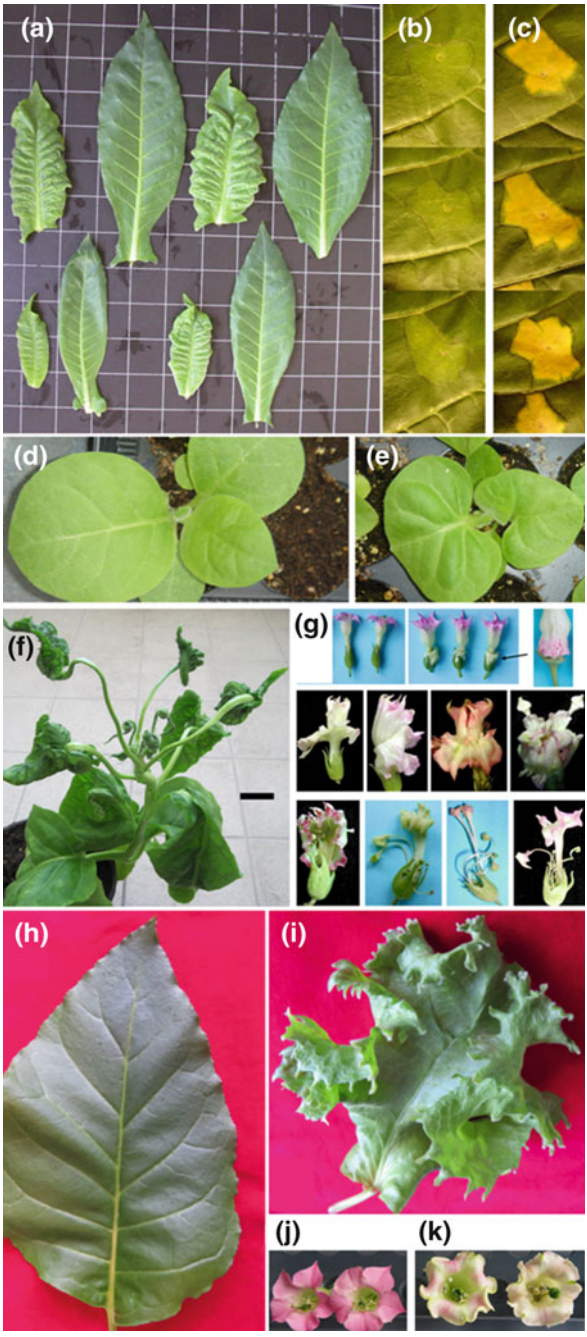
accidents but are of benefit to Agrobacteria and were the real "purpose" of the initial transformation event.

The T-DNAs of different *Agrobacterium* strains are highly variable in size and structure and can be mosaic, resulting from DNA exchange among strains. Different functional groups of T-DNA genes can be distinguished (Chen 2016). One group of genes encodes opine synthesis, such as the nopaline synthase gene *nos*. A second group encodes hormone synthesis enzymes. These are the tryptophan monooxygenase (*iaaM*) and indoleacetamide hydrolase (*iaaH*) genes, which generally occur together and encode the synthesis of indoleacetic acid, and the isopentenyl transferase (*ipt*) gene for the synthesis of cytokinins. A third group is the *plast* gene family (with genes such as *rolB*, *rolC,* and *6b*). The *plast* gene nomenclature is far from standardized. In order to avoid confusion, I will use a single symbol for each *plast* gene or protein, based on the most commonly used forms.

The remaining T-DNA genes (such as *orf13a*, *orf511,* or *rolA*) have no known function and may be placed in a fourth group of "orphan genes."

In this Chapter, I will explore and discuss the properties of the *plast* genes. They were identified by David Tepfer and his team at the French Institut National de Recherche Agronomique (INRA) at Versailles and defined on the basis of weak but significant protein similarity (Levesque et al. 1988). The name *plast* stands for "developmental plasticity" and was coined to highlight the capacity of the *plast* genes, when introduced into wild-type plants, to change their development in various remarkable ways. A few examples are shown in Fig. 1. The original paper cites eleven Plast proteins (from both *A. rhizogenes* and *A. tumefaciens*) with similarity values ranging between 13 and 34%. The *A. rhizogenes* Plast protein genes were from the TL-DNA of strain A4: RolB, RolC, Orf13, and Orf14, and the N-terminal part of Orf8 (Orf8 is similar to IaaM, but not accompanied by an IaaH protein, see also below). *A. tumefaciens* Plast proteins were initially described from strain 15955. The 15955 TL-DNA codes for p5, p7, 6a, 6b, and the N-terminal part of IaaM; the TR-DNA encodes p3'. The authors divided the Plast proteins into acidic and basic groups, according to the calculated pI values. Most importantly, Levesque et al. (1988) suggested that *plast* genes could have similar functions because of their common ancestry, and that their diversification could be an adaptation to different plant species. The proposal for a common basic function of the *plast* genes was a very logical one but, surprisingly, it was largely ignored in subsequent studies on individual *plast* genes. As an exception, one may cite a study on Orf13 (Kiyokawa et al. 1994) which noted a structural similarity with RolC, and suggested that Orf13 may be related to cytokinin function, "…as proposed for RolC." However, practically no attempts have been made to compare experimentally the effects of different Plast proteins and to detect common properties, for example by studying hybrid proteins, or by defining essential amino acid residues.

Several studies were carried out on plants carrying combinations of *plast* genes (e.g., Aoki and Syono 2000). Although they can be very interesting, I will not discuss such results because they do not provide much information on the functions of the individual genes. Nevertheless, some of these studies contain interesting approaches that might be useful in studies on single *plast* genes. One example is a


(a)
(b)
(c)
(d)
(e)
(f)
(g)
(h)
(i)
(j)
(k)

◂**Fig. 1** Various modifications of tobacco growth induced by *plast* genes. **a** Tobacco with a tetracycline-inducible *rolB* gene (Röder et al. 1994) and control leaves. The induced leaves are wrinkled and smaller. **b** Tobacco leaves infiltrated with an empty vector construct. **c** Tobacco leaves infiltrated with a 35S::A4-*rolB* construct. The *rolB* gene induces leaf necrosis (Mohajjel-Shoja 2010). **d** Control tobacco seedlings. **e** Tobacco seedlings transformed with a dex-inducible A4-*orf13* construct. After induction, the leaves become heart-shaped (Otten, unpublished). **f** A tobacco plant transformed with a dex-inducible T-*6b* gene. After induction at the 6-leaf stage, the plant develops numerous abnormalities (Chen and Otten 2015). **g** Flowers of different tobacco plants transformed with a 35S::AB4-*6b* gene (Helfer et al. 2003). Top left: normal tobacco flowers. The other forms show doubling of the corolla and separation of the petals to different extents. Arrow: flower with sepals removed, the second corolla is orientated upside down, forming a catacorolla. **h** Normal tobacco leaf, with a smooth leaf edge. **i** Leaf from a tobacco plant transformed with a 35S::TE-*6b* construct (Chen et al. 2018) derived from the TE cT-DNA of a natural transformant, *N. otophora*. The 35S::TE-*6b* leaf shows numerous leaflets growing from the leaf edge. **j** Flowers from normal tobacco. **k** Flowers from the 35S::TE-*6b* tobacco plant, the color of the corolla is different, petals are thicker and do not form catacorollas (Otten, unpublished)

detailed study on gravitropic responses of rapeseed hairy roots carrying multiple T-DNA genes (Legué et al. 1996). This type of study could easily be extended to roots carrying individual *plast* genes.

Before discussing the individual *plast* genes, I will give an overview on their identification and classification.

2 An Updated Version of the Plast Protein Family

2.1 Who Qualifies as a Plast Family Member?

Because the Plast proteins are highly diverged, it is not easy to delimit this family. Several homologs are not detectable by the commonly used blastp method (NCBI), but can be found by using the position-iterated protein BLAST (PSI-BLAST) blast software. We have chosen an arbitrary e-value of 0.02 as a limit to include a Plast protein candidate in the family.

Automatic annotation programs identify Plast-like proteins as members of the EMBL-EBI Pfam RolB_RolC glucosidase family (PF02027). This family is also mentioned on the site of the NCBI CDD Conserved Protein Domain Family (Marchler-Bauer et al. 2017). EMBL-EBI describes PF02027 as follows: "This family of proteins includes RolB and RolC. RolC releases cytokinins from glucoside conjugates (Estruch et al. 1991b), whereas RolB hydrolyses indole glucosides (Estruch et al. 1991c)."

In a study on protein families (Studholme et al. 2005), the PF02027 family is mentioned as being "restricted to plant-associated alphaproteobacteria and to plants of the genus *Nicotiana*" and they cite the two Estruch studies from 1991. However, the Estruch papers have been seriously challenged (see below). The PF02027 description should therefore be modified to mention the doubts regarding the Estruch conclusions and to cite the occurrence of *plast* genes in natural

transformants such as *Ipomoea* (Kyndt et al. 2015) and *Linaria* (Matveeva et al. 2012), as well as in bacteria outside the *Agrobacterium* group, and in fungi (see below).

The most commonly studied *Agrobacterium* strains and natural transformants mentioned in this Chapter are the following: *A. rhizogenes*: agropine strains A4 (the T-DNA of which is highly similar to those of strains 1855, 15834, HRI), mannopine strain 8196, cucumopine strain 2659, mikimopine strain 1724; *A. tumefaciens*: octopine strain B6 (similar to A6, 15955, Ach5, B6S3), nopaline strain C58 (similar to pTiSAKURA), Lippia strain AB2/73; *A. vitis*: octopine/cucumopine strains Tm4, AB3, and AKE-10, and the vitopine strain S4; Natural transformants: *Nicotiana tabacum*, *Nicotiana tomentosiformis*, *Nicotiana otophora*, *Nicotiana glauca*, *Linaria vulgaris*, and *Ipomoea batatas*. Other *Nicotiana* species have been reported to contain cT-DNA genes (Intrieri and Buiatti 2001): *N. cordifolia* (*rolB*, *rolC*, *orf13*, *orf14*), *Nicotiana miersi* (*rolB*), and *Nicotiana debneyi* (*rolC*), but this report remains to be confirmed.

2.2 A Plast Protein Tree

After the initial description of the Plast protein family (Levesque et al. 1988), additional *plast* genes were discovered. These included an isolated *rolB* variant, *rolBTR*, found on the TR-DNA of *A. rhizogenes* pRiA4 (Bouchez and Camilleri 1990; Lemcke and Schmülling 1998b), gene *e* from the *A. tumefaciens* nopaline strain C58 (protein e is weakly similar to p5 and RolB (Broer et al. 1995), and the *Lippia* strain oncogene *lso* from *A. tumefaciens* AB2/73 (Otten and Schmidt 1998). Use of the PSI-BLAST method permitted the discovery of further *plast* genes, including genes *b*, *c'*, and *d* from C58, and genes *7* and *4'* from the 15955 TR-DNA (Otten et al. 1999). An alignment of 50 *Agrobacterium* Plast proteins (Helfer et al. 2002) showed only very few conserved residues and two subgroups with Orf14, RolC, Orf13, 6a, and 6b on the one hand, and N-IaaM, N-Orf8, p4', p7, RolB, RolBTR, Lso, c', e, p5, b, d, and p3' on the other hand. Interestingly, the TD-Orf14 protein has only little similarity (33%) to other Orf14 proteins and seems to be an intermediate between Orf14 and RolC. Variants of *plast* genes have also been described (for *6b*, six alleles with different biological activities were reported; Helfer et al. 2002). Later, Plast-like proteins were detected outside the genus *Agrobacterium*. These included proteins from an ectomycorrhizal fungus, *Laccaria bicolor* (Mohajjel-Shoja et al. 2011), and from *Rhizobium mesoamericanum* (Chen et al. 2014). Later, two further fungal species: *Laccaria amethystina* and *Pisolithus microcarpus* (which forms associations with pines and Eucalyptus), and several more bacterial species: *Bradyrhizobium* sp., *Mesorhizobium plurifarium*, *Rhizobium leguminosarum*, and *Burkholderia* sp. (Chen 2016) were found to contain *plast*-like genes. Since then, other Fungal and Bacterial Plast proteins

Table 1 Representative Plast proteins and their accession numbers

	Plast protein	Accession number	Bacterial origin
1	C58-p5	AAD30487.1	At
2	CG474-p5	AAB41867.1	Av
3	Bo542-p5	AAZ50393.1	At
4	Tm4-p5	AAB41873.1	Av
5	Tm4-TB-b	AAD30490.1	Av
6	C58-b	AAD30482.1	At
7	C58-d	AAD30485.1	At
8	Bo542-d	AAZ50418.1	At
9	AB4-p3'	CAA54542.1	Av
10	oct-p3'	CAA25183.1	At
11	Chry-e	AAK08598.1	At
12	SAK-e	BAA87804.1	At
13	C58-c'	AAD30484.1	At
14	K599-N-Orf8	ABS11822.1	Ar
15	A4-N-Orf8	ABI54188.1	Ar
16	TA-N-Orf8	KJ599826	Ntf
17	2659-RolB	CAA82552.1	Ar
18	Ng-RolB	CAA27161.1	Ng
19	A4-RolBTR	CAA34077.1	Ar
20	TD-Orf14	AIM40184.1	Ntf
21	Lso	AAC25913.1	At
22	Ng-Orf13R	BAB85946.1	Ng
23	TE-2-Orf13	AWOL wgs	No
24	8196-Orf13	AAA22097.1	Ar
25	A4-Orf13	ABI54192.1	Ar
26	1724-Orf13	BAA22337.1	Ar
27	t-Orf13-1	CAA07584.1	Nt
28	C58-6b	AAK90972.1	At
29	CG474-6b	AAB41871.1	Av
30	Chry5-6b	AAB49454.1	At
31	Bo542-6b	AAA98501.1	At
32	Tm4-6b	CAA39648.1	Av
33	S4-6b	AAA25043.1	Av
34	AB4-6b	CAA54541.1	Av
35	TE1-6b-1	AWOL wgs	No
36	TE1-6b-2	AWOL wgs	No
37	TE2-6b	AWOL wgs	No
38	Ach5-6a	P04030.1	At
39	C58-6a	AAK90971.1	At
40	NCPPB3554-6a	KWT91792.1	Av

(continued)

Table 1 (continued)

	Plast protein	Accession number	Bacterial origin
41	2659-RolC	CAA82553.1	Ar
42	Ng-RolC	P07051.2	Ng
43	TE-2-RolC	AWOL wgs	No
44	A4-RolC	P20403.1	Ar
45	Lv-RolC	ACD81987.1	Lv
46	8196-RolC	AAA22096.1	Ar
47	t-RolC	CAA62988.1	Nt
48	A4-Orf14	ABI54193.1	Ar
49	8196-Orf14	AAA22099.1	Ar
50	1724-Orf14	BAA22339.1	Ar
51	2659-Orf14	CAB65899.1	Ar
52	Ng-Orf14	BAB85948.1	Ng
53	t-Orf14	CBJ56561.1	At
54	TE-1-Orf14-1	AWOL wgs	No
4155	TE-1-Orf14-2	AWOL wgs	No
56	TE-2-Orf14	AWOL wgs	No
57	TA-Orf14	KJ599826	Ntf
58	S4-N-IaaM	AAA98149.1	Av
59	C58-N-IaaM	CAB44640.1	At
60	15955-N-IaaM	CAA25167.1	At
61	Tm4-TA-N-IaaM	P25017.1	Av
62	Tm4-TB-N-IaaM	AAD30493.1	Av
63	15834-N-IaaM	ABI15642.1	Ar
64	Ag162-N-IaaM	AAC77909.1	Av
65	oct-p7	AAF77121.1	At
66	Bo542-p7	AAZ50396.1	At
67	Bo542-p4’	AAZ50416.1	At
68	15955-p4’	CAA25180.1	At

Origins At: *A. tumefaciens*, Ar: *A. rhizogenes*, Av: *A. vitis*, Ng: *N. glauca*, Nt: *N. tabacum*, Ntf: N. tomentosiformis, No: *N. otophora*, Lv: *Linaria vulgaris*

(FBPs) were discovered (Otten, unpublished). The functions of these FBPs are unknown. The existence of FBPs in such highly diverse organisms might be explained by a very ancient, common evolutionary origin. However, this seems unlikely because they are only found in a few bacteria and fungi. This patchy distribution suggests horizontal gene transfer. In the case of fungi, FBPs could have been introduced by the *Agrobacterium* transformation process, as fungi are easily transformed by *Agrobacterium* (de Groot et al. 1998). Transfer to other bacteria could have occurred by bacterial conjugation. The occurrence of *plast* genes in *Laccaria* and *Pisolithus* is particularly intriguing because these fungi strongly interact with plants. Their *plast* genes could facilitate this interaction by modifying

Table 2 Representative Fungal and Bacterial Plast proteins (FBP proteins) with accession number, origin, and size (number of amino acids)

	Accession number	Organism	Size
1	SEP38629.1	*R. tibeticum*	201
2	WP_063702303.1	*Bradyrhizobium sp.* BR 10245	272
3	SEP38081.1	*R. tibeticum*	296
4	WP_007538668.1	*R. mesoamericanum*	246
5	WP_007538785.1	*R. mesoamericanum*	246
6	WP_027584846	*Bradyrhizobium* sp. Ai1a-2	88
7	WP_063702296.1	*Bradyrhizobium sp.* BR 10245	223
8	CCM79785.1	*R. mesoamericanum*	191
9	WP_063676423.1	*Bradyrhizobium neotropicale*	160
10	WP_027556322	*Bradyrhizobium* sp. Cp5.3	215
11	WP_007538654	*R. mesoamericanum*	214
12	WP_028371611	*Burkholderia* sp. UYPR1.413	215
13	WP_037076395	*R. mesoamericanum*	216
14	CDX20869	*Mesorhizobium plurifarium*	248
15	WP_007538800	*R. mesoamericanum*	211
16	CCM79787	*R. mesoamericanum* STM3625	229
17	WP_041010497	*Mesorhizobium plurifarium*	186
18	WP_035685339.1	*Bradyrhizobium* sp. Cp5.3	195
19	WP_027556320	*Bradyrhizobium* sp. Cp5.3	196
20	XP_001884962	*Laccaria bicolor* S238N-H82	273
21	XP_001884964	*Laccaria bicolor* S238N-H82	409
22	XP_001884963	*Laccaria bicolor* S238N-H82	284
23	XP_001881215	*Laccaria bicolor* S238N-H82	491
24	KIK02568	*Laccaria amethystina* LaAM-08-1	441
25	XP_001884861	*Laccaria bicolor* S238N-H82	451
26	KIK12364	*Pisolithus microcarpus*	152
27	SDN78772.1	*Ensifer sp.* YR511	340

Proteins much larger or smaller than typical Plast proteins may represent annotation errors or truncated versions. R: *Rhizobium*

the fungus or the host plants. The same holds for the bacterial *plast* genes. It has been suggested that the mycosphere can favor horizontal gene transfer between bacteria and fungi (Zhang et al. 2014), and this might be a possible explanation for the presence of *plast* genes in plant-associated fungi. Representative Plast proteins from *Agrobacterium* and natural transformants identified by the end of 2017 are listed in Table 1, and FBP's in Table 2. Their sequences were used to construct a phylogenetic tree by the ClustalW method (Fig. 2). The FBPs cluster separately from the *Agrobacterium* Plast proteins, suggesting that the FBPs were not recently derived from *Agrobacterium*. The function of the FBPs remains to be tested by mutation and functional assays under different conditions.

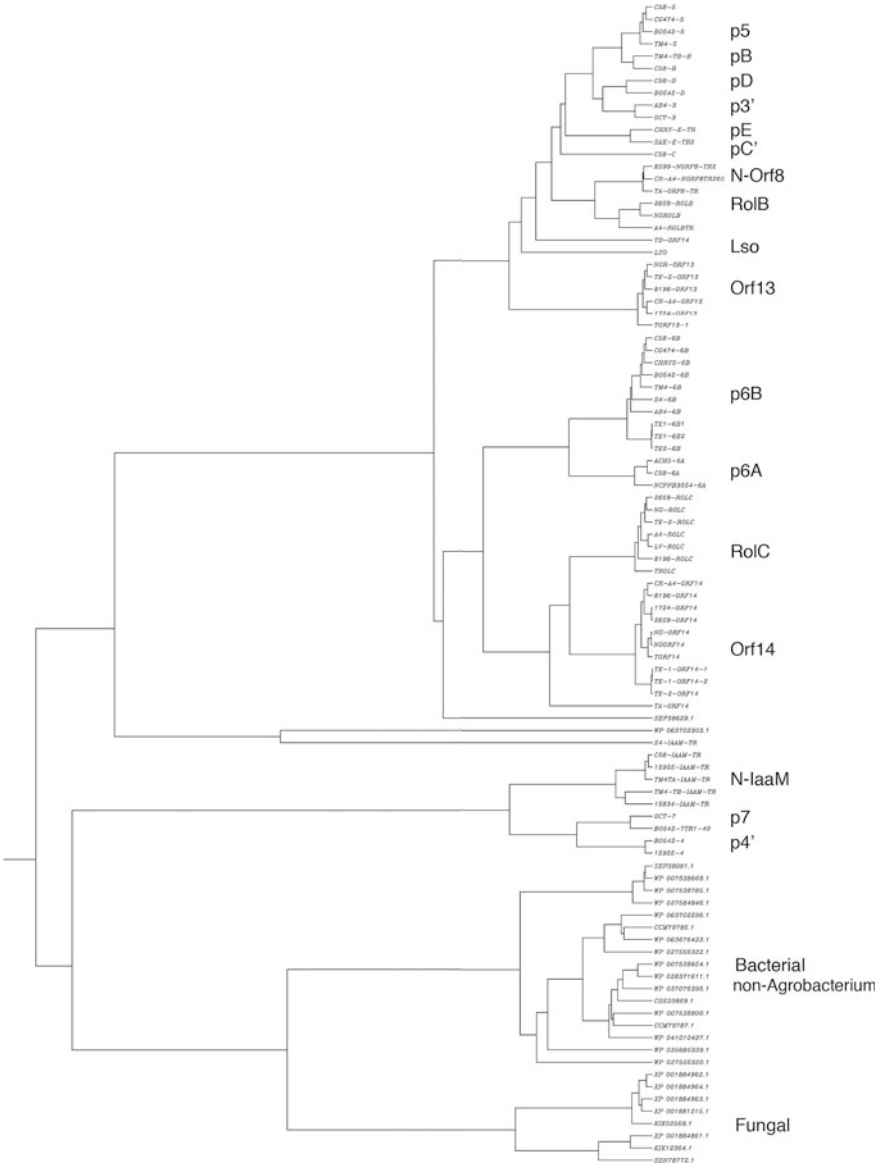


Fig. 2 Phylogenetic tree of representative Plast proteins. Proteins were analyzed using Clustal Omega, and a rooted phylogenetic tree with branch length (UPGMA) was constructed. Plants are indicated by the following prefixes: t: *Nicotiana tabacum*, Ng: *Nicotiana glauca*, Lv: *Linaria vulgaris*. Details of the *Agrobacterium* and plant Plast proteins and accession numbers are given in Table 1, details of the Bacterial and Fungal Plast proteins (FBPs) in Table 2

3 Biological Effects of *Plast* Genes

3.1 General Considerations

Several authors have noted that the expression of a particular Plast protein leads to remarkably similar types of phenotypic modification in a wide range of plant species. For example, in the case of RolC plants, one often finds dwarf growth and light green lanceolate leaves (Faiss et al. 1996). These findings indicate that Plast proteins probably target processes which are conserved across plant species. At the same time, introducing a *plast* gene into a given plant species may generate weak and strong forms of these typical phenotypes because T-DNA gene expression is strongly influenced by the neighboring plant DNA (the "position effect"). Transgene silencing (Dehio and Schell 1994) and transgenic chimeras (Schmülling and Schell 1993; Oono et al. 1993) further add to this variability. Therefore, the description of a particular *plast* gene phenotype should ideally include the full range of phenotypes, from barely visible to highly modified. Strong expression might even prevent regeneration so that no full plant phenotype can be obtained.

So far, *plast* genes have not been studied outside Angiosperms except in a curious study of sea urchin embryos transformed with 35S-*rolB* and 35S-*rolC* constructs (Bulgakov et al. 2006), where these genes stimulate the formation of teratomas.

In order to get an idea of *plast* gene targets, it would be worthwhile to express them not only in different Angiosperms, but also in other organisms such as mosses, algae, fungi, yeast, or even bacteria, for two main reasons. First, *plast* genes exist and possibly function in both fungi and bacteria outside the *Agrobacterium* group (see above). Thus, their expression in closely related fungi or bacteria devoid of FSB genes may lead to measurable effects. Second, *plast* genes show phenotypic effects across a wide range of Angiosperm species, including monocots. If a *plast* gene targets a basic plant mechanism, such as induction and growth of the vascular system, it could change the growth of higher and lower plants, but not of organisms such as fungi or bacteria. If a *plast* gene would affect a more general target such as amino acid metabolism, it may modify a broader range of organisms. Therefore, by studying the range of species affected by *plast* genes, one may get a better idea about possible *plast* gene targets.

Although studies using individual *plast* genes will be essential to understand their basic activity, it is important to remember that *plast* genes evolved in combination with other T-DNA genes. In the course of T-DNA evolution, selection must have occurred to obtain optimal growth of tumors, hairy roots, and transgenic plants, and more efficient ways to produce and secrete opines. Therefore, the action of *plast* genes can only be fully understood by studying them in the presence of other T-DNA genes and in those tissues where they are normally expressed. *Plast* gene effects have often been artificially enhanced by replacing their native promoters by inducible or strong, constitutive promoters. However, *plast* genes have their own regulation patterns (both temporally and spatially), in order to ensure

correct functioning during the tumor or hairy root induction phase, and during the subsequent opine production phase. These patterns are mostly unknown. In the case of natural transformants, precise *plast* gene regulation may have been required at the initial stages to ensure hairy root induction, but also at a later stage to allow regeneration of fertile plants from the hairy roots, or once the new plants were fully established. Possibly, there is a connection between *plast* genes and opine production (synthesis and secretion), but this has not been investigated, with the notable exception of gene *6a* (see below). This possibility concerns not only tumors and hairy roots, but also natural transformants. The TB-*mas2'* cT-DNA gene of tobacco encodes a protein involved in synthesis of desoxyfructosyl-glutamine (DFG) in roots (Chen et al. 2016). DFG production in tobacco roots may be influenced in one way or another by one or several of the following intact tobacco *plast* genes: TA-*rolCL*, TA-*rolCR*, TA-*orf13L*, TA-*orf13R*, TB-*orf14L*, TB-*orf14R*, and TD-*orf14* (Chen et al. 2014).

Plast genes are often closely associated on T-DNAs with other genes, which might indicate some type of functional connection. For two typical T-DNA structures and their *plast* genes, see Fig. 3. *rolB* and *rolC* are often linked with *rolA*, and the *rolC-orf13* region often contains *orf13a* (a non-*plast* gene of unknown function). *6b* is generally connected to either *6a* or *3'*, suggesting that the *6a* and *3' plast* genes play similar roles. The *orf13* and *orf14* genes also often occur together. The *orf8*/*iaaM* genes of *Agrobacterium* contain a *rolB* part at their 5' end. They could therefore theoretically code for bifunctional proteins with RolB and IaaM activity (see below), suggesting a functional link between the well-known enzymatic activity of IaaM and the unknown activity of RolB. The common *rolA-rolB-rolC-orf13-orf13a-orf14* T-DNA segment and the replacement of *orf14* by other, highly diverged *orf14* variants within this segment (Chen et al. 2014, 2018), also indicate that some gene combinations may have a special function.

A very general (and often neglected) problem in *plast* gene studies is the occurrence of secondary *plast* gene effects. Because *plast* genes can cause large changes in plant growth, it is not surprising to find secondary changes in metabolism, transcription patterns, stress responses, and defense reactions. Secondary

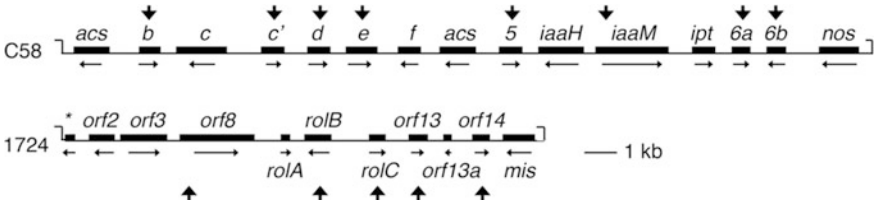


Fig. 3 Two representative T-DNA structures with *plast* genes, Top: T-DNA from *A. tumefaciens* nopaline strain C58 (NC_004972.1). Bottom: T-DNA from *A. rhizogenes* mikimopine strain 1724 (NC_002575.1). Horizontal arrows: direction of transcription. *acs*: agrocinopine synthase, *iaaH*: indoleacetamide hydrolase, *iaaM*: tryptophane monooxygenase, *ipt*: isopentenyl transferase, *nos*: nopaline synthase, *mis*: mikimopine synthase. *: remnant of *acs* gene. Vertical arrows indicate the *plast* genes; in the case of the *iaaM* and *orf8* genes, they constitute the 5' part of a larger gene

growth modifications can occur by adaptation to a basic growth change because of the inherent plasticity of plant development. Such morphogenetic "chain reactions" can be seen in tobacco plants that express an inducible T-*6b* gene (Chen and Otten 2015). Therefore, one cannot conclude that a *plast* gene *P* directly regulates or induces product A before the full chain of events has been elucidated. The use of inducible promoters may allow such types of analysis because they can show *plast* gene effects in the context of a normal plant, at early stages of induction, before secondary effects develop. Another way to distinguish late from early effects is to block late effects with specific inhibitors, leaving early effects unmodified (Clément et al. 2007).

Plants expressing *plast* genes may show initial changes in one part, with subsequent changes in other parts (such as a modification of leaf growth leading to a reduction in root growth). This can be detected by testing the effects of inducible *plast* genes in isolated organs. In this way, it was shown that the *6b* gene strongly influences root growth activity, independent of its effects on leaves (Clément et al. 2007).

Plast protein stability and degradation mechanisms could be very important for their activity. In the T-6b protein, amino acid residues were identified that reduced protein stability and led to loss of tumor induction (Helfer et al. 2002). It is possible that natural Plast protein variants differ in stability, and that steady-state levels depend on tissue type, intracellular location, or external conditions.

We will now turn to what is known about the individual *plast* genes. Unfortunately, in many cases the conclusions are not clear. It would be tempting to select from the various results those that seem most trustworthy, according to type of data, reproducibility, and coherence with other data. I have however chosen to present a larger number of data, because one of the lesser models might actually be correct. More work will hopefully provide simple and convincing explanations for the mode of action of *plast* genes. The most intensively studied *plast* genes, *rolB*, *rolC* and *6b*, are presented in the following way: description of their biological effects, description of promoter properties, models for activity, and use in practical applications.

3.2 rolB *and Its Closest Relatives*

The *rol* (root locus) genes *rolA*, *rolB*, *rolC*, and *rolD* (also called *orf10*, *orf11*, *orf12*, and *orf15*) were initially defined on the basis of the capacity of *A. rhizogenes* A4 T-DNA mutants to induce hairy roots on *Kalanchoe daigremontiana* leaves (White et al. 1985). An A4 *rolB* mutation abolished hairy root growth, showing its importance in the hairy root phenomenon. The *rolABC* combination is sufficient to induce the typical hairy root (HR) phenotype. A4-*rolB* induces roots by itself (although these differ from *rolABC* roots), suggesting an auxin-like effect (Schmülling et al. 1988). A4-*rolB* roots grow faster than do normal roots (Schmülling et al. 1988; Altabella et al. 1995). These differences need further anatomical and molecular analysis, preferably with inducible genes, coupled with

confocal analysis and cell mapping, as has been done for the T-*6b plast* gene (Pasternak et al. 2017). The effects of *rolB* on root growth depend on the plant species and culture conditions. Apple rootstock M.9/29 *rolB* tissues rooted efficiently on hormone-free medium (contrary to untransformed rootstocks), but quite unexpectedly, root growth, length, and morphology of the regenerated plants were the same as for untransformed apple (Zhu et al. 2001).

The first *rolB*-expressing transformants (Cardarelli et al. 1987; Spena et al. 1987) showed wrinkled, dark green leaves and adventitious roots. Wrinkled leaves are typical for HR regenerants and are assumed to result from differences in growth between the leaf lamina and the veins, although this has not been analyzed in detail. Others concluded that *rolA* is the main wrinkling factor in HR plants, not *rolB* (Sinkar et al. 1988). *rolB* expression in potato (van Altvorst et al. 1992) leads to wider, shorter leaves, without wrinkling, and a reduction in apical dominance. The dark green color of *rolB* tomato plants was shown to be due to an increase in chlorophyll content (Bettini et al. 2016).

Tobacco plants carrying a native A4-*rolB* gene (Schmülling et al. 1988) have larger stigmata and flowers and form adventitious roots on stems. 35S-A4-*rolB* roots tend to form callus in vitro, and 35S-A4-*rolB* plants are less efficient in adventitious root induction than are plants with the native A4-*rolB* gene, suggesting that *rolB* overexpression inhibits root growth and that precise regulation of *rolB* expression plays an important role in hairy root growth (Spena et al. 1987). Tobacco plants expressing 35S-A4-*rolB* showed ovoid, round-edged leaves (Nilsson et al. 1993b). It was also noted that 35S-A4-*rolB*-induced chlorosis and necrosis in tobacco. This phenotype starts in the intercostal areas, expands to the whole leaf, and is most obvious at flowering (Schmülling et al. 1988; Nilsson et al. 1993b; Röder et al. 1994). Necrosis could be counteracted by *rolC* (Röder et al. 1994). Leaf necrosis has not been observed in indoleacetic acid (IAA) overproducing plants (Klee et al. 1987; Sitbon et al. 1992; Kares et al. 1990; Nilsson et al. 1993b), indicating that *rolB* plants do not mimick plants with high IAA levels (see also below).

Tetracycline-inducible *rolB* tobacco plants (Röder et al. 1994) showed leaf wrinkling, and cut leaf veins form roots in the absence of auxin. Such leaf fragments produce roots over a wide range of NAA concentrations, contrary to wild-type tobacco. Therefore, they do not show a "high auxin" phenotype as obtained by overexpression of *iaa* auxin synthesis genes.

In *Arabidopsis thaliana*, 35S-*rolB* causes necrosis as in tobacco; this impairs shoot regeneration from calli. However, somatic silencing allows the growth of normal shoots. The occurrence of normal and senescing sectors led to the conclusion that *rolB* acts cell autonomously (Dehio and Schell 1994).

A more recent *rolB* study in *Arabidopsis* (Kodahl et al. 2016) showed dwarfing, early necrosis in rosette leaves, a change in leaf and flower morphology, and more inflorescences per rosette area. Leaves of such plants are slightly wrinkled and light green. Loss of chlorophyll may be the first stage in the necrotic process.

An important study found that *rolB* not only induced roots, but also flowers and shoots on thin cell layers (TCLs) from tobacco. This may be due to growth stimulation of small meristems, rather than initiation de novo, because normal tobacco

TCLs form meristemoids on hormone-free medium (Altamura et al. 1994). On the facultative apomictic plant *Hieracium piloselloides* Vill., both a native and 35S-controlled *rolB* gene induce ectopic flower meristems (Koltunow et al. 2001).

Contrary to the *orf13* (Hansen et al. 1993) and *6b* (Helfer et al. 2003) phenotypes (see below), the *rolB* phenotype is not transmissible by grafting (Hansen et al. 1993). However, graft transmission may require high levels of *plast* gene expression. In the case of *rolB,* this leads to necrosis, which may prevent graft transmissibility.

rolB genes show some structural and functional heterogeneity. 1855-*rolB* can induce roots on tobacco stems and leaves (Cardarelli et al. 1987; Spena et al. 1987), but not on *K. daigremontiana* leaves where it requires coinoculation with *iaa* genes, contrary to A4-*rolB* (Spena et al. 1987). 1724-RolB protein is longer than are A4-RolB or 8196-RolB by a 17 AA N-terminal extension (Tanaka et al. 1994; Satuti et al. 2005). 8196-*rolB* (Hansen et al. 1991) or 2659-*rolB* (Serino et al. 1994) are less dependent on auxin for root induction on carrots than is A4(1855)-*rolB*.

An inactive Ng-*rolB* gene from *N. glauca* was re-activated by removal of two stop codons (Ng-*rolB**) and induces roots on leaf discs with its native promoter (Aoki and Syono 1999c). 35S-Ng-*rolB** does not induce necrosis, unlike 35S-A4-*rolB* (Aoki and Syono 1999c; Aoki 2004). However, the sequence of the original Ng-*rolB* gene, as introduced into the ancestor plant, is unknown, and the Ng-*rolB* reconstruction might be incomplete.

The *rolB* promoter has been analyzed by different authors. Early experiments showed A4-*rolB* expression in stems, but not in leaves, and only little in roots (Spena et al. 1987). Subsequently, A4-*rolB* was found to be expressed in the root cap and in apical and lateral root meristems, but not in leaf meristems (Schmülling et al. 1989). *rolB* expression is strongly induced by auxin, but rather slowly (Maurel et al. 1990; Capone et al. 1991), and by sucrose (Nilsson and Olsson 1997). In *rolB* tobacco protoplasts, a P*rolB-GUS* construct is transiently induced by auxin.

A 2659-*rolB-GUS* construct was expressed in the root apex and vascular system of aerial organs, as was 1855-*rolB-GUS*. Contrary to 1855-*rolB-GUS,* 2659-*rolB-GUS* is not active in protoderm and root cap cells (Altamura et al. 1991; Capone et al. 1991, 1994).

1724-*rolB* is highly expressed in tobacco main and lateral root meristems, in veins of cotyledons, and regions for root-specific and auxin-induced expression were identified (Satuti et al. 2005).

In carrot, *rolB* is expressed in pericycle cells (Capone et al. 1991) and in phloem parenchyma cells (Altamura et al. 1991; Nilsson et al. 1997). Expression is also found in groups of root pericycle cells prior to and during lateral root initiation (Sugaya et al. 1989; Altamura et al. 1991; Nilsson et al. 1997), and in shoot and root meristems (Baumann et al. 1999). *rolB* expression patterns suggested that *rolB* could play a more general role in growth regulation, perhaps by promoting meristem formation (Altamura et al. 1994).

rolB is activated in tobacco embryogenesis at the advanced globular stage, when auxin polar transport starts (Chichiriccò et al. 1992). In carrot, embryos *rolB* is

activated at the preglobular stage in the central region, at the onset of auxin synthesis (Di Cola et al. 1997).

A *rolB* promoter domain (called B) binds the zinc finger Dof transcription factor (tobacco RolB binding factor, NtBBF1) at ACTTTA and allows auxin regulation (De Paolis et al. 1996; Baumann et al. 1999). Five different regulatory regions were identified in the *rolB* promoter, with different expression patterns in different root tissues (Capone et al. 1994). The Ng-*rolB* promoter allows expression in meristematic regions of roots and shoots (Nagata et al. 1995, 1996) and its expression is increased in genetic tumors (Ichikawa et al. 1990; Aoki et al. 1994), by auxin (Nagata et al. 1995), and by wounding of leaves, ahead of Ng-*rolC* induction. The biological relevance of Ng-*rolB* expression is unclear because Ng-*rolB* is interrupted by two stop codons. As with Ng-*rolB*, the t-*rolB* gene of *N. tabacum* is expressed (Meyer et al. 1995), but contains stop codons (Chen et al. 2014).

Root induction in *rolB* tobacco explants can be prevented with oligogalacturonides, which inhibit auxin-induced expression of *rolB* (Bellincampi et al. 1996). A role for extracellular H_2O_2 in this process could be excluded (Bellincampi et al. 2000).

Several models were proposed to explain the different *rolB* effects. 1. A model in which *rolB* affects auxin metabolism or auxin perception. 2. A model in which RolB acts like a phosphatase enzyme. 3. A model in which RolB acts as a transcription factor.

1. The *rolB*-auxin model. After the early discovery of the mode of action of the *A. tumefaciens* auxin (*iaa*) and cytokinin (*ipt*) synthesis genes, it was hypothesized that other T-DNA genes such as gene *6b*, gene *5*, *rolB*, or *rolC* influence or modulate the effects of these hormone genes, either at the level of hormone synthesis or hormone sensitivity. Leaf explants of HR and *rolB* plants initiate roots in the absence of auxin, contrary to wild-type plants. It was reported (Estruch et al. 1991c) that the RolB protein has β-glucosidase activity toward indoxyl-β-glucoside (an artificial substrate), suggesting that it can hydrolyze endogenous auxin conjugates and increase the concentration of free auxin. This activity requires a factor from added plant sap. Unfortunately, this attractive model could not be confirmed by further research (Nilsson et al. 1993b; Schmülling et al. 1993; Delbarre et al. 1994; Nilsson and Olsson 1997). No change in auxin concentration (Delbarre et al. 1994), metabolism, or transport (Nilsson et al. 1993b) was found in *rolB* plants.

   A significant effort was made to test auxin sensitivity in tobacco mesophyll protoplasts (Shen et al. 1988, 1990; Barbier-Brygoo et al. 1990; Maurel et al. 1991, 1994). The native A4-*rolB* gene strongly increases auxin-induced hyperpolarization of the plasma membrane (at 10,000 times lower auxin concentrations than normal). This could conceivably occur by interaction of RolB with membrane proteins (Shen et al. 1988).

   Remarkably and unexpectedly, the cell division properties of A4-*rolB* protoplasts remain unchanged, and 1855-*rolB* does not show a higher sensitivity to auxin as measured by root induction on leaf explants (Spano et al. 1988). Auxin

sensitivity measurements are complicated by the fact that A4-*rolB* is induced by auxin, leading to a positive feedback. Another study showed that RolB stimulates auxin binding to tobacco plant membranes, a binding which can be blocked by RolB antibodies. Such antibodies also block auxin binding on wild-type membranes, making the results somewhat difficult to interpret (Filippini et al. 1994). These initial auxin-binding studies have not been followed up and require confirmation and extension.

2. In 1996, it was reported that RolB has tyrosine phosphatase activity and is localized in the plasma membrane (Filippini et al. 1996). No protein similarity exists between RolB and tyrosine phosphatase proteins (TPP). The TTP CX5R motif is present in RolB but lacking in the biologically active RolBTR protein (Lemcke and Schmülling 1998b). Tyrosine phosphorylation by RolB has been assumed to explain the resistance of *Rubia cordifolia rolB* cell cultures to cantharidin (a protein phosphatase 2A inhibitor). These cultures are also resistant to calcium channel blockers such as verapamil. This was proposed to be due to the indirect activation of calcium channels by tyrosine phosphorylation (Bulgakov et al. 2002). Mutation of the A4-RolB CX5R motif does not abolish RolB-induced necrosis (Mohajjel-Shoja 2010). It is difficult to reconcile this finding with the RolB-phosphatase model.
3. In 2004, it was reported (Moriuchi et al. 2004) that 1724-RolB interacts with the tobacco 14-3-3 protein Nt14-3-3ωII and is localized in the nucleus. 14-3-3 proteins have different functions, one being importing proteins into the nucleus, the other blocking export from the nucleus by interacting with a nuclear export signal (NES). No nuclear localization signal (NLS) was found in RolB, and a conventional phosphoserine/threonine motif is lacking. Point mutants lose localization in the nucleus, binding to Nt14-3-3ωII, and capacity to induce roots, with one exception (L33F) which partially localizes in the nucleus without interacting with Nt14-3-3ωII. Thus, other factors might substitute for Nt14-3-3ωII in nuclear transport. 1724-*rolB* and Nt14-3-3ωII are both expressed in the external phloem parenchyma and in the internal phloem, and to a lesser extent in the cambium and leaf midrib. This study is the most detailed molecular study on RolB. The role of RolB in the nucleus remains unknown, and the mechanisms leading to root induction, necrosis, and other phenomena such as leaf wrinkling are not elucidated. A role for RolB in auxin perception at the plasma membrane is difficult to reconcile with a role as a transcription factor in the nucleus.

rolB has been used in a number of practical applications. In one study, *rolB* was specifically expressed in ovules and young fruits of tomato, using the tomato TPRP-F1 promoter. This leads to parthenocarpic, seedless fruits (Carmi et al. 2003), mimicking *iaa* gene effects (Shabtai et al. 2007). The *rolB* gene with its native promoter does not produce such effects in tomato (van Altvorst et al. 1992). In another study, specific expression of *rolB* in male and female organs of tobacco with the DMC1 promoter from *A. thaliana* was used to modify stamen and pistil development in order to prevent self-pollination. The growth-modifying effect was considered to result from an increase in auxin sensitivity (Cecchetti et al. 2004).

In potato, *rolB* placed under patatin promoter control (Wasserman et al. 2015) leads to changes in starch composition, but the mechanism for this is unknown.

rolB stimulates secondary plant metabolism, as do *rolC* and *6b* (Bulgakov et al. 1998; Kiselev et al. 2006; Shkryl et al. 2008), and it does so in a specific way (Bulgakov et al. 2016). *rolB* greatly increases resveratrol production in *Vitis amurensis*, but also leads to necrosis, as in tobacco and *Arabidopsis* (Kiselev et al. 2007). Its expression increases tolerance to biotic and abiotic stresses (Veremeichik et al. 2012; Bulgakov et al. 2013; Arshad et al. 2014). In tomato, *rolB* enhances nutritional quality and resistance against fungi (Arshad et al. 2014).

The following paragraph will review some Plast proteins that are closely related to RolB.

RolBTR and the RolB-like N-terminal parts from Orf8 (N-Orf8) and IaaM (N-IaaM) (Fig. 2) may be considered as RolB variants. The *rolBTR* gene is found on the TR-DNA of *A. rhizogenes* A4 and is not associated with the *rolA* and *rolC* genes (Bouchez and Camilleri 1990). It has only 40% DNA identity with *rolB* and does not induce roots on tobacco leaf disks as does *rolB*. Overexpression in tobacco leads to shorter plants with smaller, epinastic, oval, and slightly wrinkled leaves, with off-shoots at the base. No change was detected in seedling auxin sensitivity (Lemcke and Schmülling 1998b). A 14 amino acid C-terminal fragment of RolBTR is essential for morphogenetic activity, but is lacking in RolB. No CX5R tyrosine phosphatase motif was found (see above). The *rolBTR* expression pattern in hairy roots, its role in HR initiation or maintenance, and its relation to *rolB* activity are unknown.

A4-*orf8* from *A. rhizogenes* A4 is a homolog of the T-DNA *iaaM* gene (Levesque et al. 1988; Otten and Helfer 2001). *iaaM* and *iaaH* genes were initially found in *Pseudomonas savastanoi* (Comai and Kosuge 1982) and allow this bacterium to stimulate plant growth by secretion of auxin. The IaaM tryptophan monooxygenase enzymes encoded by the *Agrobacterium* T-DNAs catalyze the synthesis of indoleacetamide (IAM) from tryptophan (Van Onckelen et al. 1985, 1986). IAM is converted by the indoleacetamide hydrolase (IaaH) enzyme into IAA (Schröder et al. 1984; Thomashow et al. 1984, 1986). A4-Orf8 was reported to stimulate IAM synthesis similar to its homolog IaaM (Lemcke et al. 2000), but this could not be confirmed (Otten and Helfer 2001). It was also reported that 35S-A4-*orf8* tobacco grows on inhibitory NAA concentrations, whereas *iaaM* plants do not (Lemcke and Schmülling 1998a). The reason for this is not clear. Interestingly, the 550 AA C-terminal part of the T-DNA-encoded IaaM/Orf8 proteins is similar to the *Pseudomonas* IaaM sequence, but contrary to the *Pseudomonas* sequence, the T-DNA proteins contain a 200 AA N-terminal extension with homology to other Plast proteins. The apparent fusion of a RolB-like protein to an IaaM enzyme suggests that RolB proteins (and by extension, the Plast proteins) might be involved in auxin metabolism, or that RolB function was functionally linked to auxin synthesis by placing both genes under control of the same promoter. In order to investigate the function of these RolB-like sequences, the N and C parts of A4-Orf8 and A4-IaaM were separately expressed in tobacco under 35S promoter control. A4-N-Orf8 induces dwarfing with massive hexose and starch accumulation in

source leaves, leading to white leaves (Otten and Helfer 2001). This is caused by inhibition of sucrose export from source leaves (Umber et al. 2002). Thus, the RolB-like part of Orf8 functionally differs from RolB. 35S-A4-N-*iaaM* expression does not change the phenotype of transgenic tobacco (Otten and Helfer 2001). The N-terminal part of IaaM proteins is quite different from RolB and rather resembles p4' and p7 (Fig. 2). It could therefore result from another fusion event. IaaM proteins from *Agrobacterium* strains S4, C58, and Tm4 have not yet been investigated and could have a biologically active N-terminal Plast-like part.

The C-terminal part of A4-IaaM encodes a protein involved in IAM synthesis, similar to the intact A4-IaaM protein, but A4-C-Orf8 does not, nor does the complete A4-Orf8 protein (Otten and Helfer 2001; Umber et al. 2002). However, 35S-A4-*orf8* expression in tobacco causes dark green leaves and local leaf cell expansion, pointing to a role for A4-*orf8* in hairy roots which is unrelated to IAM synthesis. Crosses between A4-*N-orf8* and A4-*C-orf8* plants showed that A4-*C-orf8* reduces the A4-*N-orf8*-encoded accumulation of sucrose and starch, but the mechanism for this is unclear (Umber et al. 2005). The role of A4-*orf8* in hairy root induction also remains to be tested.

3.3 rolC

rolC is another well-known *plast* gene. Similar to *rolB*, *rolC* has been defined on the basis of *A. rhizogenes* A4 mutants (White et al. 1985). *rolC* is not essential for root induction on *K. daigremontiana* leaves, but in its absence, root growth is retarded. Unlike *rolB*, the wild-type *rolC* gene will not induce roots on tobacco leaves (Spena et al. 1987) nor on *K. daigremontiana* (Schmülling et al. 1988). However, 35S-*rolC* can induce roots on tobacco leaf explants (Spena et al. 1987). These roots grow on hormone-free medium, grow faster than normal roots, and are more branched (Schmülling et al. 1988; Altabella et al. 1995; Faiss et al. 1996; Palazon et al. 1998). Wild-type, *rolB*, and *rolC* root growth patterns need a much more detailed analysis. This should include anatomical studies and cell division and expansion patterns, preferably starting from the embryonic stage or from roots with inducible *plast* genes.

Tobacco *rolC* transformants show dwarf growth, loss of apical dominance, and short corollas (Oono et al. 1987; Schmülling et al. 1988). Later it was reported that *rolC* causes early flowering and size reduction of stem epidermal cells (Oono et al. 1990; Winefield et al. 1999). 35S-*rolC* tobacco plants are dwarf, bushy, with tiny flowers, and small, lanceolate, pale green leaves (Schmülling et al. 1988; Schmülling and Schell 1993). In line with the enzymatic studies (Estruch et al. 1991b, see below), it was proposed that *rolC* has a cytokinin-like effect. However, tobacco leaves expressing *ipt* are dark green, contrary to *rolC* leaves. Also, cytokinins inhibit root growth. *rolC*-induced chlorosis is cell-autonomous, as shown with a *rolC* gene construct interrupted by the transposable element Ac from maize (which is also active in tobacco). Spontaneous somatic Ac transposition activates

the *rolC* gene and creates pale green sectors in transgenic tobacco plants (Spena et al. 1989; Fladung and Ahuja 1997). These spots show changes in cell size, and the RolC protein is only found in the altered sectors (Estruch et al. 1991b). Experimental 35S-*rolC* tobacco chimeras were also generated with a FLP/FRT recombination system (Gidoni et al. 2001). A potato *rolC* phenotype is not graft transmissible (Fladung 1990), but others concluded that *rolC* acts at a distance (Guivarc'h et al. 1996). In potato, 35S-*rolC* induces a decrease in chlorophyll, small leaves of normal form, normal-size flowers, and dwarfism, up to moss-like growth in vitro. These plants produce more and smaller tubers, with more eyes (Fladung 1990). 35S-*rolC* tobacco leaves have higher turgor, indicating higher hexose content (Nilsson et al. 1993a). In line with this observation, *rolC* potato leaves contain more glucose and fructose, but sucrose levels are not modified (Fladung and Gieffers 1993). 35S-A4-*rolC* and 35S-Nt-*rolC* tobacco plants also show higher levels of sugars and starch (Mohajjel-Shoja et al. 2011). Another 35S-*rolC* tobacco study (Guivarc'h et al. 1996) found no altered phenotype until early flowering, when precocious entry into flowering and anatomical changes in the shoot meristem was detected. Young leaves show a more radial growth of palissade and spongy parenchyma, and earlier vacuolation. Stem pith cells are larger, with differences becoming evident about 1.5 mm below the apex. There is a strong delay in fiber lignification in internal and external phloem, and total resorption of pith after flower formation, leading to hollow stems.

rolC counteracts the necrotic effects of the *rolB* gene (see above, Röder et al. 1994). The mechanism for this antagonistic activity remains unexplained.

In *Dianthus caryophyllus* (carnation), *rolC* enhances root and shoot formation (Casanova et al. 2004). In *Populus tremula* x *Populus tremuloides* (hybrid aspen), the most affected *rolC* plants had fasciated stems, with an increased meristem and smaller cells (Nilsson et al. 1996a). In *Panax, ginseng rolC* induces callus which form embryos in medium without hormones. These embryos show an abnormal development with fasciated stems and multiple meristems. *Panax ginseng* callus spontaneously forms embryos, so that the *rolC* gene may enhance this process, rather than initiate it (Gorpenchenko et al. 2006).

An Ng-*rolC* gene from *N. glauca* expressed in tobacco under 35S promoter control (Aoki 2004) leads to the same phenotype as a *rolC* gene from *A. rhizogenes*: dwarfed, lanceolate, and pale green leaves, and small floral organs. Some 35S-Ng-*rolC* plants show pale green inner blades and wrinkled dark green margins. This could result from chimeric tissues, partly transformed and partly wild-type (Schmülling and Schell 1993; Aoki and Syono 1999c).

Dexamethasone-inducible versions of A4-*rolC* and t-*rolC* (the latter from *N. tabacum*) induce leaf expansion, chlorosis, starch accumulation, enations, and sucrose uptake from root fragments (Mohajjel-Shoja et al. 2011). The Ng-*rolC* and t-*rolC* results strongly suggest that t-*rolC* plays a role in the growth of the normal tobacco plant, but this requires further analysis, for example by CRISPR-Cas9 or RNA-silencing studies.

A4-*rolC* is expressed in the phloem, not in meristems (Schmülling et al. 1989). Its expression is higher in roots than in leaves (Leach 1991). The *rolC* promoter is

expressed in all root cell types, especially in the elongation zone, in the vascular system (Nagata et al. 1995), in the phloem (Sugaya et al. 1989), and in embryos (Sugaya and Uchimiya 1992). *rolC* is activated by sucrose (Yokoyama et al. 1994; Nilsson et al. 1996b) and moderately induced by auxin (Maurel et al. 1990). In *rolC* tobacco, expression is restricted to protophloem and companion cells (Guivarc'h et al. 1996). These authors noted that detection of RolC by antibodies does not produce the same results as do RolC-GUS constructs. A *rolC*-GUS reporter in tobacco is expressed in large trichome glandular cells, in the vascular system, and in root tips (Hu et al. 2003). In potato, *rolC* is expressed in the phloem tissue, bundle sheet cells, and vascular parenchyma, but not in xylem or non-vascular tissues. No differences were seen between source and sink leaves (Graham et al. 1997). In rice, the *rolC* promoter allows expression in vascular tissue (Matsuki et al. 1989).

In aspen, the *rolC* promoter causes specific expression in companion cells, but later in development a shift to other cell types occurred. *rolC* is induced by sucrose, not by sorbitol (Nilsson et al. 1996a). This led to the idea that *rolC* is involved in some aspect of sucrose metabolism or transport (Nilsson and Olsson 1997).

Strong *rolC* induction occurs when carrot cells are induced to form somatic embryos by transfer to medium lacking 2,4-D (Fujii and Uchimiya 1991; Fujii et al. 1994; Fujii 1997; Suzuki et al. 1992). In tobacco and chickpea, the *rolC* promoter allows expression in phloem, epidermis, and trichomes. The *rolC* promoter contains several sequence motifs and gel-retardation assays have been used to characterize these (Saha et al. 2007). A 43 kd nuclear tobacco protein interacts with the *rolC* gene promoter (Matsuki and Uchimiya 1994).

Crosses between *N. glauca* and *Nicotiana langsdorfii* lead to spontaneous tumors, and it has been speculated that the *N. glauca* cT-DNA genes play a role in this. Ng-*rolC* is expressed in such genetic tumors, but not in leaves (Aoki et al. 1994). Expression occurs in the vascular system, similar to *rolC*, and is induced by wounding, but slower than Ng-*rolB* (Nagata et al. 1995, 1996). Tobacco t-*rolC* is expressed in young leaves and shoot tips, not in older leaves or roots. The gene is down-regulated by auxin and induced by cytokinin (Meyer et al. 1995).

Several models have been proposed for *rolC* action, but none has been definitively adopted: 1. an enzymatic role; 2. a role in auxin perception; 3. a role related to sucrose.

1. It has been reported that RolC has cytokinin-β-glucosidase activity, is localized in the cytosol (Estruch et al. 1991a), and is not associated with membranes (Oono et al. 1991). The glucosidase activity would liberate free cytokinins from conjugates and lead to cytokinin effects. The use of an inducible *rolC* gene (Faiss et al. 1996) made it possible to study changes in cytokinin content, starting from a normal plant. No changes were found in cytokinin glucosides, nor in free cytokinins. Data from *rolC*-expressing hybrid aspen (Nilsson and Olsson 1996a) show the same results. A cross between a *rolC* plant and an *ipt* plant did not lead to cleavage of the main product of the Ipt enzyme, zeatin-O-β-glucoside (Schmülling et al. 1993). Using root inhibition by cytokinins as a bioassay, *rolC* tobacco roots show the same sensitivity as do normal

roots (Faiss et al. 1996). Therefore, the cytokinin-β-glucosidase hypothesis has become relatively unlikely.
2. Tobacco protoplasts transformed with A4-*rolC* or 35S-*rolC* show a weak shift in hyperpolarization in response to NAA, and their cell division is not changed compared to that of normal protoplasts (Maurel et al. 1991). These studies have not been followed up.
3. Based on the induction of *rolC* genes by sucrose, and the increased leaf turgor of *rolC* plants, RolC activity might create a sink for sucrose (Nilsson and Olsson 1997). Recent studies on dex-A4-*rolC* and dex-t-*rolC* plants (Mohajjel-Shoja et al. 2011) are consistent with this hypothesis, but do not provide a mechanism.

rolC genes may serve in different applications. In potato, attempts were made to change tuber growth. Smaller tubers and changes in carbohydrate composition were noted (Fladung et al. 1993). In floriculture, *rolC* genes have been used to obtain dwarf plants (Winefield et al. 1999). Reduced apical dominance, more compact growth with more shoots, and early flowering, all typical of *rolC* plants, are useful characteristics in horticulture (Scorza et al. 1994; Casanova et al. 2004, 2005; Gardner et al. 2006). In tomato, *rolC* results in smaller plants and fruits, but productivity is not improved. These plants show a reduction in apical IAA and ABA levels (Bettini et al. 2010). ABA reduction was also noted in tobacco and in hybrid aspen (Nilsson et al. 1993a; Fladung and Ahuja 1997).

rolC genes have also been used to stimulate the production of useful secondary metabolites in suspension cultures of different plant species, with both stimulatory and inhibitory effects (Bulgakov et al. 2005). In *Panax ginseng* cells, *rolC* induces defense genes (Kiselev et al. 2006).

3.4 Gene 6b

The *6b* gene is one of the most studied *plast* genes and has been investigated for over 30 years by several research groups. Summaries of *6b* studies have been published recently (Ishibashi et al. 2014; Ito and Machida 2015). However, as stated by Takahashi et al. (2013), with regard to its mechanism "…an unequivocal conclusion has not been reached." This is all the more frustrating as the *6b* gene shows some very remarkable growth induction and modification properties.

In 1981, an *A. tumefaciens* A6 T-DNA mutant was described which leads to very large tumors on *K. daigremontiana* stems and leaves (Garfinkel et al. 1981). This "tumor morphology large" or *tml* mutant also shows reduced shoot induction from lateral buds (Ream et al. 1983). No effects were seen on tobacco. It was proposed that the *tml* gene stimulates shoot development through a cytokinin-like effect. At that time T-DNA maps were not sufficiently precise to distinguish the *6a* and *6b* genes, but it is likely that *tml* corresponds to the *6b* gene. Several groups initially speculated that *6b* influences hormone responses induced by the *iaa* and *ipt* T-DNA genes. Various studies were therefore aimed at comparing *6b*-induced growth

modifications with auxin and cytokinin responses. *6b* reduces shoot induction in tobacco (Leemans et al. 1983; Spanier et al. 1989; Bonnard et al. 1989), indicating an auxin-like effect, antagonistic to cytokinin reactions. Stimulation of *iaa* gene-induced tumor formation on *Nicotiana rustica* stems by *6b* confirmed the auxin-like effect (Bonnard et al. 1989). Although a T-DNA fragment from C58 with the *6a* and *6b* genes does not induce tumors (Inzé et al. 1984), Hooykaas et al. (1988) later found that Ach5-*6b* induces tumors on *N. glauca* and *Kalanchoe tubiflora* and can thus be considered as an oncogene. Neither *iaa* nor *ipt* genes induce tumors on *K. tubiflora*, indicating that *6b* does not act like auxin or cytokinin genes but has a new, unknown growth-stimulating function. *A. vitis* T-*6b* stimulates *rolABC*-induced root formation like the *A. rhizogenes* TR *iaa* genes. However, it also results in thick roots and callusing, which is atypical of auxin reactions, and does not enhance sensitivity to exogenous auxins (Tinland et al. 1990). In *N. rustica*, a heat shock-inducible T-*6b* gene leads to strong growth changes in seedlings, with formation of pin-like leaves, ectopic meristems along the hypocotyl, and stunted growth. Leaves of mature plants punctured with a needle show exaggerated wound-induced cell division, with the formation of small calli. Leaky expression increases seedling cotyledon and leaf size (Tinland et al. 1992). Tobacco plants transformed with the native AK-*6b* gene are fairly normal, but leaf fragments form shoots on hormone-free medium, contrary to normal tobacco (Wabiko and Minemura 1996), resembling cytokinin effects. *6b* responses therefore seem to mimick effects of both exogenous auxins and cytokinins, depending on the type of growth assay, and could possibly increase endogenous hormone levels. However, studies on C-*6b* tobacco plants showed no changes in cytokinin content compared to normal tobacco, although *6b* tissues tolerate normally inhibitory cytokinin concentrations (Gális et al. 1999). Decapitated AK-*6b* tobacco plants form thin shoots from buds and small leaf-like structures form along the veins on the adaxial surface. Green shoots are formed from wounded roots. Agrobacteria that transfer either *iaa* or *ipt* genes induce tumors on AK-*6b* tobacco plants, but not on normal tobacco (Wabiko and Minemura 1996). Tobacco seedlings expressing the native C-*6b* gene grow faster than do controls (Gális et al. 2002). All these studies clearly indicate that there is no simple relationship between *6b* and auxin or cytokinin, as initially proposed.

A study on 35S-AB-*6b* tobacco plants (Helfer et al. 2003) revealed a remarkable and highly specific phenotype, completely different from plants that overexpress auxin or cytokinin synthesis genes. This phenotype comprises (among other features) double, mirrored leaves (enations), thin tube-like shoots, and double flowers (catacorollas). Some spontaneous tobacco mutants, described at the beginning of the 20th century, also produce enations and catacorollas, but the corresponding mutations have not been identified. Catacorollas in tobacco plants transformed with *Agrobacterium* were already reported in 1990 (Komari 1990), but the T-DNA gene leading to this phenotype was not identified. Most probably it was *6b*. In some 35S-AB-*6b* regenerants, partial leaf blade doubling generates small dark green spots on the leaves, with large amounts of starch; these include a small additional and isolated vascular system with inverted polarity. Expression of cell

cycle-specific genes is not notably increased in *6b* leaves. The enation phenotype is graft transmissible, showing that some mobile factor (called enation factor) is produced and trafficks in *6b* plants, leading to the local accumulation of an active principle that could initiate or prolong cell division (Helfer et al. 2003). Movement of 6b protein (Grémillon et al. 2004) and *6b* mRNA (Chen and Otten 2015) have been reported, but the actual enation factor has not yet been identified. *Arabidopsis 6b* plants produce small tubes (ectopic meristems) growing out from the abaxial leaf surface, but no enations. Thus, the *6b*-induced phenotype varies with the plant species. A follow-up study with a dexamethasone-inducible dex-T-*6b* gene in tobacco (Grémillon et al. 2004) showed complex growth changes upon *6b* induction. Roots placed on inducing medium increase in diameter due to cell expansion and abnormal cell division in the pericycle and vascular parts, and shoots become tubular. AK-*6b* induces ectopic cell division in the abaxial side of the tobacco leaf, with adventitious vascular cells in the petiole (Terakura et al. 2006). These authors found strong expression of cell cycle-related genes in *6b* tissues, which could be a secondary effect of stimulation of cell division. They used the special property of the glucocorticoid receptor (GR) to enter the nucleus in the presence of dexamethasone to show that an AK-6b::GR fusion protein only produces its effects when it moves into the nucleus.

An exhaustive study investigated morphological changes induced by dex-T-*6b* (Chen and Otten 2015). These authors described 70 morphological and anatomical modifications resulting from four basic modifications: leaf chlorosis, cell expansion, ectopic shoots, and ectopic vascular bundles. The authors proposed that all changes may be caused by abnormal sucrose uptake and accumulation (see below).

Several *6b* genes have been described. A-*6b*, T-*6b,* and C-*6b* have different biological activities (Tinland et al. 1989). Six *6b* genes separately placed in the same 35S expression cassette show different levels of oncogenicity on *N. glauca*, *N. tabacum*, and *K. daigremontiana*. A-6b/T-6b hybrid proteins identified a short region that confers strong oncogenicity (Helfer et al. 2002). Recently, three *6b* genes have been found in a natural transformant, *N. otophora*. One was placed under control of a 2x35S promoter. It does not induce tumors on *N. rustica* stems, and most of the phenotypic aspects of the enation syndrome are missing. However, this gene increases vein formation in tobacco, causes leaf wrinkling, outgrowth of leaflets at the leaf margins, and most remarkably, early germination of embryos in the seed capsules (vivipary; Chen 2016; Chen et al. 2018).

Several *6b* gene promoters have been studied. Bo542-*6b* is induced by auxins and wounding, but not by cytokinins. Expression is high in roots and low in leaves (Bagyan et al. 1994, 1995). A Chry5-*6b*-*GUS* reporter gene in tobacco (Reddy et al. 2003) is also inducible by auxins, and to a lesser extent by cytokinins. C-*6b* transcription on the contrary is increased on cytokinin-containing medium and counteracted by NAA (Gális et al. 1999, 2002). GUS expression is mainly found in the vascular system and in the shoot meristem. AK-*6b* expression is high in shoot-forming calli and low in mature leaves (Wabiko and Minemura 1996). Further studies are needed to establish whether the specificities of these various promoters are really different, or result from differences in experimental factors

such as the size of the selected promoter region, position effects in different reporter lines, plant age, or growth conditions.

Different models exist for the activity of the *6b* genes. These are: 1. the PAL-auxin model, 2. the nuclear factor model, 3. the sucrose model, 4. the silencing model, 5. the miR319/TCP model.

1. The *6b*-PAL-auxin model. Resistance of C-*6b* tobacco seedlings to high cytokinin is accompanied by high IAA levels and activation of the phenylpropanoid pathway (Gális et al. 2002). It was hypothesized that *6b* gene increases PAL, C4H, and 4CL expression (similar to the action of cytokinins), leading to higher chlorogenic acid (CGA) and scopoletin levels. These metabolites are known to inhibit IAA degradation, which would lead to higher IAA levels (Gális et al. 2002). A follow-up study detected changes in phenolic compounds in AK-*6b* tobacco (Gális et al. 2004). Polar auxin transport is reduced in dex-AK-*6b* tobacco (Kakiuchi et al. 2006), and it was proposed that the increased levels of phenolic compounds in *6b* tissues interfere with IAA transport, thus leading to the *6b* phenotype. Venation patterns were studied in young AK-*6b* seedlings and shown to be modified, thereby resembling auxin mutant phenotypes and effects of auxin transport inhibitors (Kakiuchi et al. 2006, 2007). Dex-AK-*6b* plants were used to study IAA and cytokinin accumulation (Takahashi et al. 2013), using IAA and cytokinin antibodies and an auxin reporter gene (DR5::GUS). This study showed that IAA accumulates in *6b* tumors and in new vascular tissues. It was proposed that the accumulation of phenolic compounds changes hormone transport, thereby leading to the *6b* phenotype. Studies on dex-T-*6b* tobacco plants (Clément et al. 2006) showed that T-*6b* induction causes a strong increase in hexose concentrations and leaf cell expansion. This expansion is auxin independent, and T-*6b* does not lead to IAA increase or induction of IAA-responsive genes, contradicting other studies. 35S-AB-*6b* tobacco plants accumulate starch (Helfer et al. 2003), and a further study with dex-T-*6b* tobacco plants revealed a link between *6b* and sucrose accumulation in leaf disks (Clément et al. 2006, see below). dex-T-*6b* tobacco roots accumulate very high levels of phenolics upon induction (Clément et al. 2007), confirming earlier studies (Gális et al. 2004). Addition of a PAL inhibitor in the medium abolishes the accumulation of the phenolic compounds, but does not change the typical *6b* phenotype, clearly showing that changes in phenolic compounds are a secondary effect and do not cause the *6b* phenotype (Clément et al. 2007).
2. The *6b*-nuclear factor model. AK-6b protein interaction studies using yeast two-hybrid technology (Y2H) were reported in 2002. Three interacting proteins were found. The tobacco nuclear protein NtSIP1 (for 6b-interacting protein) is a transcription factor-like protein (Kitakura et al. 2002) and also interacts with AK-6b in vivo with purified proteins. An artificial transcription system using an AK-6b::GAL4 fusion protein and a luciferase reporter protein with a GAL4-binding site and a minimal 35S promoter allowed induction of the reporter. Nuclear localization of a GFP::6b fusion protein is enhanced by NtSIP1, as shown by transient expression assays in tobacco protoplasts.

This could be due to active transport or to trapping in the nucleus after passive diffusion. A C-terminal stretch of 21 acidic residues is essential for nuclear localization, transactivation, and hormone-independent growth. Thus, 6b might have a function in transcriptional regulation. In another study (Grémillon et al. 2004), a T-6b::GFP protein was detected in the cytoplasm and the nucleus of BY2 cells, but the fusion protein was biologically inactive.

In the Kitakura study, two other Y2H-based interactants were reported (NtSIP2 and NtSIP3). NtSIP2 is a nucleolar protein homologous to the TNP1 protein from a transposable element (Kitakura et al. 2008). NtSIP3 is histone H3, and 6b binds to it in vivo. The 6b protein was therefore proposed to act as an H3 chaperone and to influence the transcription of unknown target genes. A 6b mutant lacking the C-terminal region lost H3 binding activity (Terakura et al. 2007; Wang et al. 2011).

3. The *6b*-sucrose model. Studies with the dex-T-*6b* construct in tobacco (Clément et al. 2006, 2007) showed that its induction in leaves and roots causes a rapid and strong increase in uptake and accumulation of sucrose. Accumulation at the site of uptake leads to cell expansion. The early effects on root growth were described in detail by 3D analysis of root anatomy and cell division patterns, using dex-T-*6b* tobacco seedlings (Pasternak et al. 2017). Chlorogenic acid accumulates to very high levels in such roots and leads to browning (see above). These are secondary *6b* effects, because they are abolished with a PAL inhibitor without affecting the *6b* phenotype. They probably result from osmotic stress caused by excessive hexose accumulation. How *6b* leads to uptake and retention of sucrose remains unknown.
4. A major *6b* paper appeared in 2011 (Wang et al. 2011). This work started from the assumption that some T-DNA genes may have anti-silencing activity, in order to prevent the plant from shutting off T-DNA gene expression. The phenotypes of the *Arabidopsis* RNA-silencing mutants *ago1-27* and *se-1* are somewhat similar to *6b* phenotypes. In *6b* tissues, miR162, miR164, miR319, and miR165/166 levels are decreased. The mRNA targets of these miRNAs accumulate accordingly.

   AK-6b was crystallized and found to resemble the 3D structures of exotoxin A and cholera toxin protein, although no protein sequence homology was found. The 6b crystal contains an NAD+ molecule. Exotoxin A and cholera toxin have ADP-ribosylation activity. The purified AK-6b protein does not show ADP-ribosylation activity in vitro, but an AtARF protein could be isolated from *Arabidopsis* that resembles the human AFR ADP-ribosylation cofactor and enables this activity in vitro. Point mutations in the putative AK-6b catalytic site cause loss of activity. AtARF and AK-6b could be co-precipitated from plant extracts. AK-6b ADP-ribosylates AGO1 and SE proteins. These results are very interesting and important, but the data need independent confirmation by other groups, and further analysis. Do all 6b proteins show this activity? Can the ADP-ribosylated AGO and SE proteins be identified, both in vitro and in vivo, and be shown to accumulate upon induction of a *6b* gene? Do plants with

mutant AGO or SE proteins that are insensitive to ADP-ribosylation become resistant to 6b activity? How do the proposed target protein modifications lead to the enation phenotype? Do other Plast proteins show similar activity; in other words, is ADP-ribosylation an ancestral property of the Plast proteins, or was it derived from a different, earlier function? The ADP-ribosylation model also raises the question of its relation to the three other *6b* models. Can ADP-ribosylation explain increased sucrose uptake? How does it fit with binding to the three NtSIP proteins? It seems unlikely that the small 6b proteins are multifunctional. It will also be important to establish crystal structures for other Plast proteins and compare their structures with that of 6b, to identify active sites and modes of action.

5. The miR319/TCP model. Recently, the TE-*6b* gene from *N. otophora* was shown to induce a very unusual phenotype in tobacco, with leaflets growing out from the leaf rim (Chen et al. 2018). This phenotype was unexpected as it was completely different from the earlier observed AB-*6b*, T-*6b,* and AKE-*6b* enation phenotypes (see above). The dentate leaves of TE-1-6b-L tobacco plants strongly resemble those of mutants of the miR319/jaw-TCP module in *Arabidopsis*, *Antirrhinum majus,* and tomato. Indeed, overexpression of *MIR319A* in *Arabidopsis* (jaw-D mutants) or mutation of the miR319 target genes (the class II TCP transcription factor genes *TCP2*, *TCP3*, *TCP4*, *TCP10,* and *TCP*24) leads to folded and crinkled leaves with outgrowth of the leaf margins, due to abnormal maintenance of the leaf marginal meristem (Nath et al. 2003; Palatnik et al. 2003; Alvarez et al. 2016; Bresso et al. 2017). Leaf vasculature is also increased as in TE-6b plants. Possibly, TE-*6b* genes interfere with the miR319/jaw-TCP module. Intriguingly, the RolC protein interacts strongly with TCP13 (Mohajjel-Shoja 2010). TCP13 belongs to the class II TCP proteins, but is not targeted by miR319. This could indicate that Plast proteins affect the class II TCP protein function by binding to them. This hypothesis will require further study.

A recent study has described the first practical application for a *6b* gene (Jin et al. 2017). Interestingly, T-*6b* increases oil production in *Arabidopsis* seeds and might therefore become very useful in biodiesel plants such as rapeseed.

3.5 orf13

The *orf13* gene is part of the *A. rhizogenes* T-DNA and is also found in natural transformants. Several studies showed that it acts synergistically with *rolB* and may replace auxin required for HR induction on carrots, suggesting it has auxin-like activity (Capone et al. 1989; Hansen et al. 1993; Aoki and Syono 1999b). The first *orf13* tobacco plants (Hansen et al. 1993) were 35S-8196-*orf13* plants. They grew slowly, were short with reduced apical dominance, had small, wrinkled, asymmetric, dark green and rounded leaves, and reduced root systems. Flowers were also

affected. The phenotype does not resemble an auxin phenotype. Remarkably, it was reported that the *orf13* phenotype is transmissible by grafting. Although 35S-A4-*orf13* tobacco plants (Lemcke and Schmülling 1998a) have a similar phenotype, no increase was found in cytokinin levels, and the phenotype could not be transmitted by grafting. Tet-inducible tet-A4-*orf13* lines (Lemcke and Schmülling 1998a) showed severe dwarfing, lanceolate leaves, and fasciation on inflorescences. Roots were curly, but had a normal gravitropic response. Root elongation was reduced by 30%, with a reduction in cell number and cell length. The phenotype resembles neither auxin nor cytokinin phenotypes. 8196-*orf13* and 35S-t-*orf13-1* induce dark green calli on carrots (Fründt et al. 1998). These two genes do not stimulate root induction by *rolABC* genes, similar to 1855-*orf13*. Expression of the 1724-*orf13* gene in tobacco (Satuti et al. 2007) causes dwarfing, wrinkled dark green leaves, spikes on the adaxial and abaxial leaf side and on petals, and flowers with a short corolla. In *Arabidopsis*, 35S-A4-*orf13* leads to extreme dwarfing with leaf and flower size reduction (Kodahl et al. 2016).

Several natural transformants carry *orf13* genes on their cT-DNAs. A 35S-t-*orf13* construct from *N. tabacum* induces callus growth on carrot discs (Fründt et al. 1998). 35S-Ng-*orf13* from *N. glauca* promotes HRI-*rolB*-mediated root induction on tobacco leaf fragments, similar to HRI-*orf13* (Aoki and Syono 1999a). Expression of 35S-Ng-*orf13* in tobacco (Aoki and Syono 1999a; Aoki 2004) led to stunted seedling roots, which later normalized. These plants have rounded dark green leaves with slightly wrinkled edges, and short flowers, stems, and corollas.

In 35S-8196-*orf13* tobacco plants (Meyer et al. 2000), endoreduplication is reduced, indicating interaction of 8196-*orf13* with the cell cycle. Spikes occur on leaves and petals, resembling KNOX- and D-type cyclin-overexpression phenotypes. The authors found an LxCxE Retinoblastoma (RB) binding motif present in all Orf13 proteins.

These initial observations led to an important paper on the 8196-Orf13 protein (Stieger et al. 2004), which reported binding between 8196-Orf13 and plant Retinoblastoma (RB) protein in vitro. Mutants of 8196-Orf13 and RB abolish binding. 35S-8196-*orf13* and 35S-8196-*orf13** (the latter with an LxCxE to LxAxK mutation in the RB-binding domain) were introduced into tomato. 35S-8196-*orf13* plants show two groups of modifications. The first group consists of small and fewer leaves, advanced flowering, increased numbers of petals and sepals, modifications in phyllotaxis, and changes in the arrangement of lateral and intercalary leaflets. These modifications are absent in 35S-8196-*orf13** plants. The second group consists of reduced length, loss of apical dominance, spikes on leaves, sepals, petals, and fruits, fasciation of stems and petals, and petals fused to the stem. These modifications are also found in 35S-8196-*orf13** plants. No effects are seen on roots. It was concluded that some, but not all Orf13 effects depend on RB binding. Although the overall shoot apical meristem structure is not changed, 8196-*orf13* increases the rate of cell division and leaf formation. A model was proposed in which Orf13 binds to RB and enhances KNOX transcription factor expression, thereby leading to cell cycle changes. These results and the role of the LxCxE motif

require further analysis. None of the other Plast proteins contains this motif, which suggests that RB binding is specific to Orf13 proteins.

Several *orf13* promoters have been studied. Expression of 8196-*orf13* is high in roots, in the vascular system, and in wounded tissues (Hansen et al. 1997). Expression of A4-*orf13* is low in roots. Ng-*orf13* is expressed in genetic tumors, not in leaves (Aoki et al. 1994). A later paper found that Ng-*orf13* is expressed in normal and tumor parts of genetic hybrids but not in meristems, is low in roots, and induced by methyl jasmonate (Udagawa et al. 2004). t-*orf13-1* and t-*orf13-2* from tobacco are expressed in sepals, petals, shoot tips, and young leaves, with low expression in older leaves, roots, and stems. Growth of tobacco leaf discs on auxin and cytokinin decreases t-*orf13* expression (Fründt et al. 1998).

3.6 orf14

orf14 is closely linked to *orf13*. Overexpression of *orf14* did not induce a phenotype on tobacco and did not modify the *orf13*-induced phenotype (Lemcke and Schmülling 1998a; Aoki 2004). Ng-*orf14* weakly promoted *rolB*-induced rooting of tobacco leaf discs (Aoki and Syono 1999a, b) but gave no phenotype in tobacco plants (Aoki 2004). Two *Nicotiana orf14* cT-DNA genes, TA-*orf14* and TD-*orf14* (Chen et al. 2014), and a duplicated TE-*orf14* gene from *N. otophora* (Chen et al. 2018) are highly divergent and may have different properties. Strong expression of a large range of *orf14* genes in different plants may reveal morphogenetic properties, as is the case for many *plast* genes. It is also possible that the *orf14* gene has no activity on its own, but modulates the effects of other Plast proteins. Ng-*orf14* is expressed in genetic tumors, but not in leaves of *N. glauca* (Aoki et al. 1994).

3.7 Gene 5

Gene *5* has been found in T-DNAs of *A. tumefaciens* and *A. vitis*. An A6 gene *5* mutant is fully virulent (Garfinkel et al. 1981), but a C58 gene *5* mutant is attenuated (Joos et al. 1983). C58-*5* is non-oncogenic in different assays (Inzé et al. 1984; Otten et al. 1999). The p5 protein from *A. tumefaciens* Ach5 was reported to convert tryptophan to the IAA competitor indole-3-lactate (ILA) in vitro (Körber et al. 1991). Transgenic tobacco plants expressing gene *5* contain more ILA and are more tolerant to inhibitory levels of auxin. They have a normal phenotype, although young seedlings grow somewhat slower. AB-*5* does not modify AB-*iaa*-induced roots (Otten and De Ruffray 1994). Unfortunately, the reported enzymatic activity of the p5 protein was not further investigated, and the enzymatic parameters or substrate range of the p5 protein are still unknown. Enzymatic activities of p5 proteins from other *Agrobacterium* strains, and from related Plast proteins like p3' and d, remain to be explored.

An Ach5-*5* promoter construct could be induced by auxin in tobacco shoots and weakly in tumors (Koncz and Schell 1986), with a slow response to auxin in protoplasts, similar to *rolB*. ILA counteracts this induction. Expression of gene *5* is confined to the vascular protophloem in tobacco (Körber et al. 1991).

3.8 lso

The *lso* (*Lippia canescens* strain oncogene) gene was found in an unusual *Agrobacterium* strain, the AB2/73 or Lippia strain (Otten and Schmidt 1998). The AB2/73 T-DNA only contains two genes: *lso* and, an nopaline synthase-like gene, *lsn*. The *lso* gene causes tumors on *N. rustica* leaf fragments in vitro and on *K. daigremontiana* and *K. tubiflora* stems, similar to *6b*. The central part of the protein is most similar to Plast proteins e, Orf13, and RolB. Tobacco plants were transformed with the native *lso* gene, and a 35S-*lso* construct (Schmidt 1999). *lso* tobacco plants appear normal, whereas 35S-*lso* plants are stunted with wrinkled leaves. This phenotype is not transmissible by grafting.

3.9 *Gene* 6a

A single but intriguing study has appeared on the *6a* gene (Messens et al. 1985). Analysis of opine secretion by tobacco tumors in liquid medium showed that tumors induced with *6a* mutants of *Agrobacterium* octopine and nopaline strains do not secrete nopaline or octopine. Secretion of agrocinopine is not affected. Tumors induced by *6b*, *3'*, or *4' plast* gene mutants do not show such an effect (Salomon et al. 1984). The *6a* gene was therefore called *ons* gene (for octopine/nopaline secretion). These results indicate a possible link between *plast* genes and opine production. They have not yet been confirmed and analyzed in more detail. It would be interesting to study plants with inducible *6a* gene expression and to explore the specificity of the *6a* system with respect to different opines. Octopine synthase (Ocs) not only produces octopine, but a large range of related structures such as octopinic acid, lysopine, histopine, and methiopine (Otten et al. 1977; Hack and Kemp 1980). In addition, *ocs* genes coding for enzymes with different substrate specificities have been found in *A. vitis* (Otten and Szegedi 1985), and possible adaptations of the corresponding *6a* genes may be investigated. Metabolome analysis of root exudates from plants with inducible *6a* genes might show whether *6a* can also modify the secretion of normal plant compounds. It is not known where the *6a* gene is expressed, nor whether it would cause a special phenotype. In some strains, *6a* and *6b* are replaced by the *3'* gene and another type of *6b*, suggesting that *3'* can replace *6a* (Drevet et al. 1994; Otten and De Ruffray 1994). This could be an adaptation to allow secretion of other opine types.

3.10 Genes 3', 4', 7', b, c', d, *and* e

These genes seem to have little influence on growth. Some of them could be inactive variants or remnants of *plast* genes. B6S3-*3'* does not seem to play a role in tumor formation (Salomon et al. 1984). Neither AB-*3'* (Otten and De Ruffray 1994) nor AK-*3'* (Wabiko and Minemura 1996) induces growth, and AB-*3'* does not modify the effects of the nearby genes *6b* or *6b-ipt*. Only Ach5-*3'* is weakly oncogenic on *K. tubiflora* (Otten et al. 1999). Proteins p4' and p7 show little similarity to other Plast proteins and are distantly related to each other.

Gene *4'* (also known as gene *780*) has been studied because of its promoter, used in biotechnology (Bruce and Gurley, 1987, 1988; O'Grady and Gurley 1995). A protein from cauliflower binds to a *4'* promoter element (Adams and Gurley 1994). Nothing is known about the biological activity of this gene. Genes Ach5-*7* and *A. larrymoorei-7* potentially code for unusually short proteins (126 AA); the Bo542-7 protein is longer (180 AA).

C58-*b*, C58-*c'*, and C58-*d* do not seem to be biologically active (Otten et al. 1999). The C58-e protein has a large, 50 AA extension at its C-terminal part, compared to other Plast proteins. Gene *e* mutants of strain C58 do not affect tumor formation (Joos et al. 1983). A later study found that this mutant induces smaller tumors on *K. daigremontiana* and tobacco and is non-oncogenic on tomato stems (Broer et al. 1995), but this could not be confirmed (Otten et al. 1999).

4 What Could Be the Ancestral *Plast* Gene Mechanism?

Norf8, *orf8*, *rolC*, *orf13*, *rolB*, *rolBTR*, *lso*, and *6b* each induce different growth patterns. Other *plast* genes, like *orf14*, gene *3'*, *4'*, *5*, *6a*, *7*, *b*, *c'*, *d*, and *e*, and the FBP genes, remain to be tested. The *plast* gene studies have led to many different hypotheses concerning their modes of action. These are for RolB: indoxyl-β-glucosidase, tyrosine phosphatase, stimulation of auxin binding to membranes, interaction with a 14-3-3 protein leading to localization in the nucleus; for protein 6b: stimulation of sucrose uptake by an unknown mechanism, interference with the miR319/TCP module, enzymatic activity (ADP-ribosylation) leading to changes in silencing, a chaperone function by binding to histone H3, a modulation of transcriptional activity by binding to transcription factor NtSIP1, and binding to a nucleolar protein (NtSIP2) with unknown consequences; for protein p5: indole-3-lactate synthase; for RolC: cytokinin β-glucosidase; and for protein 6a an unknown mechanism to enhance opine secretion. It is hard to imagine that all these activities could be derived from one original, ancestral Plast protein activity. It is likely that such an ancestral Plast activity improved one of the two essential characteristics of *Agrobacterium*-induced tumors and roots: opine production and growth. In view of the many strong *plast*-induced growth changes, the second possibility seems more likely. Several *plast* genes seem to be able to induce or

stimulate cell division. In *6b* plants, the appearance of additional photosynthetic cells (at the base of glandular trichomes, along veins, and along the abaxial leaf surface) and vascular cells (in cotyledons, leaves, and petioles) is particularly striking. Tumor induction by this gene (as in *N. rustica*) seems to result from prolongation of wound-induced cell division, as intact tissues of dex-T-*6b* tobacco do not form tumors upon induction. The *lso* gene induces tumors as does *6b*, although the sequences of these genes are very different. The *rolB* gene induces roots, but also floral meristems, or at least favors their outgrowth. The *orf13* gene induces callus on carrot roots. It will be important to study whether *6b*, *lso*, *rolB*, and *orf13* affect the cell cycle. Can these *plast* genes prevent the arrest of the cell cycle (e.g., after wound-induced division or after a series of normal divisions within local meristems such as those at the base of glandular trichomes)? A role for *plast* genes in cell division does not necessarily imply a role in hormone metabolism or sensitivity. Neither is it necessary to postulate that cell division genes are direct targets of Plast proteins. Modification of the synthesis, transport, and accumulation of essential metabolites such as sucrose may also be involved in abnormal prolongation of cell division patterns. Such a mechanism could link effects on cell division to opine transport (as proposed for the *6a* gene). Dark green leaves or leaves with high starch levels seem to be common to plants expressing *Norf8*, *orf8*, *rolB*, *6b*, *orf13*, and *rolC*. Leaf wrinkling was reported for *rolB*, *rolBTR*, *6b*, *lso*, and *orf13*. What could be the connection with cell division? Does greening precede or follow cell division? Can one be obtained without the other? If one assumes that bacterial FBP genes are functional in plant-associated bacteria other than *Agrobacterium* (this has not yet been investigated), this would mean that *plast* genes do not necessarily target eucaryotic cell division genes or silencing systems, or in the case of fungi, do not act on plant-specific mechanisms such as photosynthesis. In that case, it would also become unlikely that Plast proteins need specific plant protein partners. Instead, they might function autonomously and affect cells in a very general way.

However, it cannot be excluded that FBP proteins leave fungal or bacterial cells and enter plant cells by active transport (like effector proteins), in which case they could still target plant-specific functions and require plant protein partners.

5 Future Research

5.1 Search for the Fundamental Mechanism of Plast *Genes*

The present *plast* models should be investigated in more detail, using different approaches. The 6b ADP-ribosylation model can be tested by isolating the proposed ADP-ribosylated target proteins AGO and SE. By rendering the targets resistant to ADP-ribosylation, *6b* effects should be abolished. The sucrose uptake model can be further investigated by interfering with normal sucrose transporters (e.g., by using

inducible gene silencing of sucrose transporter genes) in order to detect possible interactions between normal and *6b*-induced sucrose uptake. Sucrose uptake assays using simpler organisms such as yeast might also be of interest. *plast* gene-expressing plants show modifications in different plant organs. These observations suggest that Plast proteins target essential growth mechanisms. The graft transmission of the *6b* (and to a lesser extent *orf13*) phenotypes might allow the isolation and characterization of the moving enation factor, but if this is the *6b* mRNA (Chen and Otten, 2015) or the 6b protein (Grémillon et al. 2004), this will not be of much help. The protein and mRNA enation factor hypotheses can be tested using controlled destruction or capture of the 6b protein and *6b* mRNA in the receiving tissues.

Many *plast* genes remain to be studied. Some of these, such as the very short A-*7* gene or the unusually long *e* gene from C58, could be abnormal and inactive; others might reveal new phenotypes. Recent sequencing data suggest a large number of variants for a given Plast type. By the end of 2017, nineteen 6b proteins were reported (Chen et al. 2018) and studies on six of them showed clear differences in tumor induction. This should be extended with studies on other transgenic 6b plants including tobacco, *Arabidopsis* and other species outside the Angiosperms. Such studies could provide interesting data on common, and therefore more fundamental, growth effects.

In order to exclude effects of one modified plant organ on another, it is preferable to study isolated organs from plants with inducible *plast* genes. These might be small tobacco root fragments maintained in vitro, or simple tissues such as the *N. rustica* stem surface, which forms tumors on infection with Agrobacteria carrying *6b* (Chen and Otten 2015). Similarly, induction of *N. tabacum* leaf necrosis by infiltration with 2x35S-*rolB*-carrying bacteria (Mohajjel-Shoja 2010) may be simplified by using epidermal cell layers, mesophyll protoplasts, or tobacco BY-2 cells.

5.2 *Role of* Plast *Genes in Tumors, Hairy Roots, and Natural Transformants*

The role of *plast* genes in the initiation and maintenance of tumors, hairy roots, and natural transformants requires further investigation. This includes regulation of their expression in different tissue types and at different stages of development. Reporter genes should be used in a normal T-DNA context in order to create the natural conditions for their expression. This may reveal complex spatiotemporal patterns of T-DNA gene expression during tumor and hairy root induction. We also need a better view of the initial transformation steps: how much variability in T-DNA structure and expression exists at the start, and what kind of selection occurs for cells with particular combinations of T-DNA genes and expression levels? Inducible interfering RNA (RNAi) constructs would be useful to remove *plast* gene function for any given gene at any given time. Finally, the activity of different

variants of the same *plast* gene should be compared within a given T-DNA context and include the properties of their promoters. These data, although largely descriptive, will be of great use as more becomes known about the basic activities of *plast* genes.

5.3 Possible Applications

Several *plast* genes have been used to increase the production of secondary plant metabolites in vitro using hairy root cultures. They also have a considerable potential for use in agriculture because they modify essential growth parameters such as leaf and root growth and could improve seed properties as shown for T-*6b* (Jin et al. 2017). They might also be used to create apomictic plants. A few such attempts have already been made, for example with potato and rice, but these studies will require much more effort, using different tissue-specific promoters and different expression levels. Applications might be easier to achieve in horticulture, especially for *rolC*. A large-scale approach (e.g., by using promoter-less *plast* genes and random insertion in the plant genome), followed by field selection of thousands of different transformants, could conceivably yield agronomically useful varieties. However, it is likely that once the precise mechanism of action of given *plast* genes is known, special constructs can be designed to obtain predictable morphologies, and it will become much more attractive to use them for crop improvement, especially if such a mechanism would be expected to improve the plant's productivity. An interesting possibility might be the use of natural *A. rhizogenes* strains and selection of modified plants with certain T-DNA structures (intact or truncated) regenerated from hairy roots. It has been argued that horticultural plants obtained in this way should not be considered as genetically modified organisms as they are obtained by a natural infection process followed by spontaneous regeneration (Lütken et al. 2012). In view of the high variability among *A. rhizogenes* strains and the very large amount of combinations of T-DNA genes and expression levels, this strategy might also apply to crop plants. Indeed, nature has already produced such a modified crop plant with *Agrobacterium*-derived cT-DNA sequences, the sweet potato (Kyndt et al. 2015), but it is not yet known whether and how the cT-DNA has changed the growth of this species.

6 Conclusions

In spite of their often dramatic effects on plant growth and morphology, no general mechanisms have been found for the activity of the *plast* genes. The last 30 years have produced a large body of data, both at the phenotypical and molecular level, and several molecular models have been proposed for genes *5*, *6a*, *6b*, *rolB*, *rolC*, and *orf13*. However, there are few, if any, logical connections between these

models, and it is unlikely that all are correct. Existing proposals for *plast* gene activities need to be tested systematically and further developed by other research groups so that they can be either confirmed or rejected. If a model for one *plast* gene is agreed on, it should be systematically tested for other *plast* genes, starting with the most similar ones. Careful observation of phenotypes and growth patterns (although often criticized as being "only descriptive") will continue to be important. With new and more powerful techniques at hand, and much more detailed knowledge about basic plant functions, it is now time for a large-scale comprehensive effort to identify unambiguously the molecular basis of *plast* gene activities, and to answer some long-standing questions in *Agrobacterium* biology.

References

Adams EC, Gurley WB (1994) Nuclear protein 780BP from cauliflower binds an element in the *780* gene promoter of T-DNA. Plant Mol Biol 26:377–392

Altabella T, Angel E, Biondi S et al (1995) Effect of the *rol* genes from *Agrobacterium rhizogenes* on polyamine metabolism in tobacco roots. Physiol Plant 95:479–485

Altamura MM, Archilletti T, Capone I et al (1991) Histological analysis of the expression of *Agrobacterium rhizogenes rolB*-GUS gene fusion in transgenic tobacco. New Phytol 118:69–78

Altamura MM, Capitani F, Gazza L et al (1994) The plant oncogene *rolB* stimulates the formation of flower and root meristemoids in tobacco thin cell layers. New Phytol 126:283–293

Alvarez JP, Furumizu C, Efroni I et al (2016) Active suppression of a leaf meristem orchestrates determinate leaf growth. *eLife* 5:e15023

Aoki S (2004) Resurrection of an ancestral gene: functional and evolutionary analyses of the *Ngrol* genes transferred from *Agrobacterium* to *Nicotiana*. J Plant Res 117:329–337

Aoki S, Syono K (1999a) Function of *Ngrol* genes in the evolution of *Nicotiana glauca*: conservation of the function of *NgORF13* and *NgORF14* after ancient infection by an *Agrobacterium rhizogenes*-like ancestor. Plant Cell Physiol 40:222–230

Aoki S, Syono K (1999b) Synergistic function of *rolB*, *rolC*, *ORF13*, and *ORF14* of TL-DNA of *Agrobacterium rhizogenes* in hairy root induction in *Nicotiana tabacum*. Plant Cell Physiol 40:252–256

Aoki S, Syono K (1999c) Horizontal gene transfer and mutation: *Ngrol* genes in the genome of *Nicotiana glauca*. Proc Natl Acad Sci USA 96:13229–13234

Aoki S, Syono K (2000) The roles of *Rirol* and *Ngrol* genes in hairy root induction in *Nicotiana debneyi*. Plant Sci 159:183–189

Aoki S, Kawaoka A, Sekine M et al (1994) Sequence of the cellular T-DNA in the untransformed genome of *Nicotiana glauca* that is homologous to ORFs *13* and *14* of the Ri plasmid and analysis of its expression in genetic tumors of *N. glauca* x *N. langsdorffii*. Mol Gen Genet 243:706–710

Arshad W, Haq I, Waheed MT et al (2014) *Agrobacterium*-mediated transformation of tomato with *rolB* gene results in enhancement of fruit quality and foliar resistance against fungal pathogens. PLoS ONE 9:e96979

Bagyan IL, Revenkova EV, Kraev AS et al (1994) Functional analysis of the 5'-flanking region of gene *6b* from TL-DNA pTiBo542 in transgenic tobacco. Mol Biol 28:487–492

Bagyan IL, Revenkova EV, Pozmogova GE et al (1995) 5′-regulatory region of *Agrobacterium tumefaciens* T-DNA gene *6b* directs organ-specific, wound-inducible and auxin-inducible expression in transgenic tobacco. Plant Mol Biol 29:1299–1304

Barbier-Brygoo H, Maurel C, Shen WH et al (1990) Use of mutants and transformed plants to study the action of auxins. Symp Soc Exp Biol 44:67–77

Baumann K, De Paolis A, Costantino P et al (1999) The DNA binding site of the Dof protein NtBBF1 is essential for tissue-specific and auxin-regulated expression of the rolB oncogene in plants. Plant Cell 11:323–334

Bellincampi D, Cardarelli M, Zaghi D et al (1996) Oligogalacturonides prevent rhizogenesis in *rolB*-transformed tobacco explants by inhibiting auxin-induced expression of the *rolB* gene. Plant Cell 8:477–487

Bellincampi D, Dipierro N, Salvi G et al (2000) Extracellular H_2O_2 induced by oligogalacturonides is not involved in the inhibition of the auxin-regulated *rolB* gene expression in tobacco leaf explants. Plant Physiol 122:1379–1385

Bettini P, Baraldi R, Rapparini F et al (2010) The insertion of the *Agrobacterium rhizogenes rolC* gene in tomato (*Solanum lycopersicum* L.) affects plant architecture and endogenous auxin and abscisic acid levels. Sci Hortic 123:323–328

Bettini P, Marvasi M, Fabiola F et al (2016) *Agrobacterium rhizogenes rolB* gene affects photosynthesis and chlorophyll content in transgenic tomato (*Solanum lycopersicum* L.) plants. J Plant Physiol 204:27–35

Bonnard G, Tinland B, Paulus F et al (1989) Nucleotide sequence, evolutionary origin and biological role of a rearranged cytokinin gene isolated from a wide host range biotype III *Agrobacterium* strain. Mol Gen Genet 216:428–438

Bouchez D, Camilleri C (1990) Identification of a putative *rolB* gene on the TR-DNA of the *Agrobacterium rhizogenes* A4 Ri plasmid. Plant Mol Biol 14:617–619

Bouzar H, Jones JB (2001) *Agrobacterium larrymoorei* sp. nov., a pathogen isolated from aerial tumours of *Ficus benjamina*. Int J Syst Evol Microbiol 51:1023–1026

Bresso EG, Chorostecki U, Rodriguez RE et al (2017) Spatial control of gene expression by miR319-regulated TCP transcription factors in leaf development. Plant Physiol. Epub ahead of print

Broer I, Dröge-Laser W, Barker RF et al (1995) Identification of the *Agrobacterium tumefaciens* C58 T-DNA genes *e* and *f* and their impact on crown gall tumour formation. Plant Mol Biol 27:41–57

Bruce W, Gurley WB (1987) Functional domains of a T-DNA promoter active in Crown Gall tumors. Mol Cell Biol 7:59–67

Bruce W, Gurley WB (1988) An enhancer-like element present in the promoter of a T-DNA gene from the Ti plasmid of *Agrobacterium tumefaciens*. Proc Natl Acad Sci USA 85:4310–4314

Bulgakov VP, Khodakovskaya MV, Labetskaya NV et al (1998) The impact of *rolC* oncogene on ginsenoside production by ginseng hairy root cultures. Phytochem 49:1929–1934

Bulgakov VP, Tchernoded GK, Mischenko NP et al (2002) Effect of salicylic acid, methyl jasmonate, ethephon and cantharidin on anthraquinone production by *Rubia cordifolia* callus cultures transformed with the *rolB* and *rolC* genes. J Biotech 97:213–221

Bulgakov VP, Veselova MV, Tchernoded GK et al (2005) Inhibitory effect of the *Agrobacterium rhizogenes rolC* gene on rabdosiin and rosmarinic acid production in *Eritrichium sericeum* and *Lithospermum erythrorhizon* transformed cell cultures. Planta 221:471–478

Bulgakov VP, Kiselev KV, Yakovlev KV et al (2006) *Agrobacterium* mediated transformation of sea urchin embryos. Biotech J 1:454–461

Bulgakov VP, Shkryl YN, Veremeichik GN et al (2013) Recent advances in the understanding of *Agrobacterium rhizogenes*-derived genes and their effects on stress resistance and plant metabolism. Adv Biochem Eng Biotech 134:1–22

Bulgakov VP, Veremeichik G, Grigorchuk VP et al (2016) The *rolB* gene activates secondary metabolism in Arabidopsis calli via selective activation of genes encoding MYB and bHLH transcription factors. Plant Physiol Biochem 102:70–79

Capone I, Spano L, Cardarelli M et al (1989) Induction and growth properties of carrot roots with different complements of *Agrobacterium rhizogenes* T-DNA. Plant Mol Biol 13:43–52

Capone I, Cardarelli M, Mariotti D et al (1991) Different promoter regions control level and tissue specificity of expression of *Agrobacterium rhizogenes rolB* gene in plants. Plant Mol Biol 16:427–436

Capone I, Frugis G, Costantino P et al (1994) Expression in different populations of cells of the root meristem is controlled by different domains of the *rolB* promoter. Plant Mol Biol 25:681–691

Cardarelli M, Mariotti D, Pomponi M et al (1987) *Agrobacterium rhizogenes* T-DNA genes capable of inducing hairy root phenotype. Mol Gen Genet 210:111–115

Carmi N, Salts Y, Dedicova D et al (2003) Induction of parthenocarpy in tomato via specific expression of the *rolB* gene in the ovary. Planta 217:726–735

Casanova E, Valdés AE, Zuker A et al (2004) *rolC*-transgenic carnation plants: adventitious organogenesis and levels of endogenous auxin and cytokinins. Plant Sci 167:551–560

Casanova E, Trillas MI, Moysset L et al (2005) Influence of *rol* genes in floriculture. Biotech Adv 23:3–39

Cecchetti V, Pomponi M, Altamura MM et al (2004) Expression of *rolB* in tobacco flowers affects the coordinated processes of anther dehiscence and style elongation. Plant J 38:512–525

Chen K (2016) Sequencing and functional analysis of cT-DNAs in *Nicotiana*. Thesis, University of Strasbourg

Chen K, Otten L (2015) Morphological analysis of the *6b* oncogene-induced enation syndrome. Planta 243:131–148

Chen K, Otten L (2017) Natural *Agrobacterium* transformants: recent results and some theoretical considerations. Front Plant Sci 8:1600

Chen K, Dorlhac de Borne F, Szegedi E et al (2014) Deep sequencing of the ancestral tobacco species *Nicotiana tomentosiformis* reveals multiple T-DNA inserts and a complex evolutionary history of natural transformation in the genus *Nicotiana*. Plant J 80:669–682

Chen K, Dorlhac de Borne F, Julio E et al (2016) Root-specific expression of opine genes and opine accumulation in some cultivars of the naturally occurring GMO *Nicotiana tabacum*. Plant J 87:258–269

Chen K, Dorlhac de Borne F, Sierro N et al (2018) Organization of the TC and TE cT-DNA regions in *Nicotiana otophora* and functional analysis of three diverged TE-*6b* genes. Plant J (in press)

Chichiriccò G, Costantino P, Spanò L (1992) Expression of the *rolB* oncogene from *Agrobacterium rhizogenes* during zygotic embryogenesis in tobacco. Plant Cell Physiol 33:827–832

Clément B, Pollmann S, Weiler E et al (2006) The *Agrobacterium vitis* T-6b oncoprotein induces auxin-independent cell expansion in tobacco. Plant J 45:1017–1027

Clément B, Perot J, Geoffroy P et al (2007) Abnormal accumulation of sugars and phenolics in tobacco roots expressing the *Agrobacterium* T-*6b* oncogene and the role of these compounds in *6b*-induced growth. Mol Plant-Microbe Interact 20:53–62

Comai L, Kosuge T (1982) Cloning and characterization of *iaaM*, a virulence determinant of *Pseudomonas savastanoi*. J Bacteriol 143:950–957

de Groot MJA, Bundock P, Hooykaas PJJ et al (1998) *Agrobacterium tumefaciens* mediated transformation of filamentous fungi. Nat Biotechnol 16:839–842

De Paolis A, Sabatini S, De Pascalis L et al (1996) A *rolB* regulatory factor belongs to a new class of single zinc finger plant proteins. Plant J 10:215–223

Dehio C, Schell J (1994) Identification of plant genetic loci involved in a posttranscriptional mechanism for meiotically reversible transgene silencing. Proc Natl Acad Sci USA 91:5538–5542

Delbarre A, Muller P, Imhoff V et al (1994) The *rolB* gene of *Agrobacterium rhizogenes* does not increase the auxin sensitivity of tobacco protoplasts by modifying the intracellular auxin concentration. Plant Physiol 105:563–569

Di Cola A, Poma A, Spano L (1997) *rolB* expression pattern in the early stages of carrot somatic embryogenesis. Cell Biol Int 21:595–600

Drevet C, Brasileiro AC, Jouanin L (1994) Oncogene arrangement in a shooty strain of *Agrobacterium tumefaciens*. Plant Mol Biol 25:83–90

Estruch JJ, Parets-Soler A, Schmülling T et al (1991a) Cytosolic localization in transgenic plants of the *rolC* peptide from *Agrobacterium rhizogenes*. Plant Mol Biol 17:547–550

Estruch JJ, Chriqui D, Grossmann K et al (1991b) The plant oncogene *rolC* is responsible for the release of cytokinins from glucoside conjugates. EMBO J 10:2889–2895

Estruch JJ, Schell J, Spena A (1991c) The protein encoded by the *rolB* plant oncogene hydrolyses indole glucosides. EMBO J 10:3125–3128

Faiss M, Strnad M, Redig P et al (1996) Chemically induced expression of the *rolC* encoded β-glucosidase in transgenic tobacco plants and analysis of cytokinin metabolism: *rolC* does not hydrolyze endogenous cytokinin glucosides *in planta*. Plant J 10:33–46

Filippini F, Lo Schiavo F, Terzi M (1994) The plant oncogene *rolB* alters binding of auxin to plant cell membranes. Plant Cell Physiol 35:767–771

Filippini F, Rossi V, Marin O et al (1996) A plant oncogene as a phosphatase. Nature 379:499–500

Fladung M (1990) Transformation of diploid and tetraploid potato clones with the *rolC* gene of *Agrobacterium rhizogenes* and characterization of transgenic plants. Plant Breeding 104:295–304

Fladung M, Ahuja M (1997) Excision of the maize transposable element Ac in periclinal chimeric leaves of 35S-Ac-*rolC* transgenic aspen-*Populus*. Plant Mol Biol 33:1097–2003

Fladung M, Gieffers W (1993) Resistance reactions of leaves and tubers of *rolC* transgenic tetraploid potato to bacterial and fungal pathogenes. Correlations with sugar, starch and chlorophyll content. Phys Mol Plant Physiol 42:123–132

Fladung M, Ballvora A, Schmülling T (1993) Constitutive or light-regulated expression of the *rolC* gene in transgenic potato plants has different effects on yield attributes and tuber carbohydrate composition. Plant Mol Biol 23:749–757

Fründt C, Meyer AD, Ichikawa T et al (1998) A tobacco homologue of the Ri-plasmid *orf13* gene causes cell proliferation in carrot root discs. Mol Gen Genet 259:559–568

Fujii N (1997) Pattern of DNA binding of nuclear proteins to the proximal *Agrobacterium rhizogenes rolC* promoter is altered during somatic embryogenesis of carrot. Gene 201:55–62

Fujii N, Uchimiya H (1991) Conditions favorable for the somatic embryogenesis in carrot cell culture enhance expression of the *rolC* promoter-GUS fusion gene. Plant Physiol 95:238–241

Fujii N, Yokoyama R, Uchimiya H (1994) Analysis of the *rolC* promoter region involved in somatic embryogenesis-related activation in carrot cell cultures. Plant Physiol 104:1151–1157

Gális I, Simek P, Macas J et al (1999) The *Agrobacterium tumefaciens* C58-*6b* gene confers resistance to N(6)-benzyladenine without modifying cytokinin metabolism in tobacco seedlings. Planta 209:453–461

Gális I, Simek P, Van Onckelen HA et al (2002) Resistance of transgenic tobacco seedlings expressing the *Agrobacterium tumefaciens* C58-*6b* gene to growth-inhibitory levels of cytokinin is associated with elevated IAA levels and activation of phenylpropanoid metabolism. Plant Cell Physiol 43:939–950

Gális I, Kakiuchi Y, Simek P et al (2004) *Agrobacterium tumefaciens* AK-*6b* gene modulates phenolic compound metabolism in tobacco. Phytochem 65:169–179

Gardner N, Melberg T, George M et al (2006) Differential expression of *rolC* results in unique plant phenotypes. J Am Soc Hortic Sci 131:82–88

Garfinkel DJ, Simpson RB, Ream LW et al (1981) Genetic analysis of crown gall: fine structure map of the T-DNA by site-directed mutagenesis. Cell 27:143–153

Gelvin SB (2012) Traversing the cell: *Agrobacterium* T-DNA's journey to the host genome. Front Plant Sci 26:3–52

Gidoni D, Bar M, Gilboa N (2001) FLP/FRT-mediated restoration of normal phenotypes and clonal sectors formation in *rolC* transgenic tobacco. Transgenic Res 10:317–328

Gorpenchenko TY, Kiselev KV, Bulgakov VP et al (2006) The *Agrobacterium rhizogenes rolC*-gene-induced somatic embryogenesis and shoot organogenesis in *Panax ginseng* transformed calluses. Planta 223:457–467

Graham MW, Craig S, Waterhouse PM (1997) Expression patterns of vascular-specific promoters RolC and Sh in transgenic potatoes and their use in engineering PLRV-resistant plants. Plant Mol Biol 33:729–735

Grémillon L, Helfer A, Clément B et al (2004) New plant growth-modifying properties of the *Agrobacterium* T-*6b* oncogene revealed by the use of a dexamethasone-inducible promoter. Plant J 37:218–228

Guivarc'h A, Carneiro M, Vilaine F et al (1996) Tissue-specific expression of the *rolA* gene mediates morphological changes in transgenic tobacco. Plant Mol Biol 30:125–134

Hack E, Kemp JD (1980) Purification and characterization of the crown gall-specific enzyme, octopine synthase. Plant Physiol 65:949–955

Hansen G, Larribe M, Vaubert D et al (1991) *Agrobacterium rhizogenes* pRi8196 T-DNA: mapping and DNA sequence of functions involved in mannopine synthesis and hairy root differentiation. Proc Natl Acad Sci USA 88:7763–7767

Hansen G, Vaubert D, Heron JH et al (1993) Phenotypic effects of overexpression of *Agrobacterium rhizogenes* T-DNA *ORF13* in transgenic tobacco plants are mediated by diffusible factor(s). Plant J 4:581–585

Hansen G, Vaubert D, Clérot D et al (1997) Wound-inducible and organ-specific expression of *ORF13* from *Agrobacterium rhizogenes* 8196 T-DNA in transgenic tobacco plants. Mol Gen Genet 254:337–343

Helfer A, Pien S, Otten L (2002) Functional diversity and mutational analysis of *Agrobacterium* 6B oncoproteins. Mol Gen Genom 267:577–586

Helfer A, Clément B, Michler P et al (2003) The *Agrobacterium oncogene* AB-*6b* causes a graft-transmissible enation syndrome in tobacco. Plant Mol Biol 52:483–493

Hildebrand EM (1940) Cane gall of brambles caused by *Phytomonas rubi* n. sp. J Agric Res 61:685–696

Hooykaas PJJ, Den Dulk-Ras H, Schilperoort RA (1988) The *Agrobacterium tumefaciens* T-DNA gene *6b* is an oncogene. Plant Mol Biol 11:791–794

Hu Y, Chen B, Ni T et al (2003) Promoter of the *rolC* gene of *Agrobacterium rhizogenes* can be strongly regulated in glandular cell of transgenic tobacco. Mol Biotech 24:121–126

Ichikawa T, Ozeki Y, Syono K (1990) Evidence for the expression of *rol* genes of *Nicotiana glauca* in genetic tumors of *N. glauca* X *N. langsdorfii*. Mol Gen Genet 220:177–180

Intrieri MC, Buiatti M (2001) The horizontal transfer of *Agrobacterium rhizogenes* genes and the evolution of the genus *Nicotiana*. Mol Phylogenet Evol 20:100–110

Inzé D, Follin A, Van Lijsebettens M et al (1984) Genetic analysis of the individual T-DNA genes of *Agrobacterium tumefaciens*: further evidence that two genes are involved in indole-3-acetic acid synthesis. Mol Gen Genet 194:265–274

Ishibashi N, Kitakura S, Terakura S et al (2014) Protein encoded by oncogene *6b* from *Agrobacterium tumefaciens* has a reprogramming potential and histone chaperone-like activity. Front Plant Sci 5:1–7

Ito M, Machida Y (2015) Reprogramming of plant cells induced by 6b oncoproteins from the plant pathogen *Agrobacterium*. J Plant Res 128:423–435

Jin Y, Hu J, Liu X et al (2017) T-*6b* allocates more assimilation product for oil synthesis and less for polysaccharide synthesis during the seed development of *Arabidopsis thaliana*. Biotechnol Biofuels 10:19

Joos H, Inzé D, Caplan A et al (1983) Genetic analysis of T-DNA transcripts in nopaline crown galls. Cell 32:1057–1067

Kado CI (2014) Historical account on gaining insights on the mechanism of crown gall tumorigenesis induced by *Agrobacterium tumefaciens*. Front Microbiol 5:340

Kakiuchi Y, Gális I, Tamogami S et al (2006) Reduction of polar auxin transport in tobacco by the tumorigenic *Agrobacterium tumefaciens* AK-*6b* gene. Planta 223:237–247

Kakiuchi Y, Takahashi S, Wabiko H (2007) Modulation of the venation pattern of cotyledons of transgenic tobacco for the tumorigenic *6b* gene of *Agrobacterium tumefaciens* AKE10. J Plant Res 120:259–268

Kares C, Prinsen E, Van Onckelen H et al (1990) IAA synthesis and root induction with *iaa* genes under heat shock promoter control. Plant Mol Biol 15:225–236

Kerr A, Panagopoulos CG (1977) Biotypes of *Agrobacterium radiobacter* var. *tumefaciens* and their biological control. Phytopath Z 90:172–179

Kiselev KV, Kusaykin MI, Dubrovina AS et al (2006) The *rolC* gene induces expression of a pathogenesis-related beta-1,3-glucanase in transformed ginseng cells. Phytochemistry 67:2225–2231

Kiselev KV, Dubrovina AS, Veselova MV et al (2007) The *rolB* gene-induced overproduction of resveratrol in *Vitis amurensis* transformed cells. J Biotechnol 128:681–692

Kitakura S, Fujita T, Ueno Y et al (2002) The protein encoded by oncogene *6b* from *Agrobacterium tumefaciens* interacts with a nuclear protein of tobacco. Plant Cell 14:451–463

Kitakura S, Terakura S, Yoshioka Y et al (2008) Interaction between *Agrobacterium tumefaciens* oncoprotein 6b and a tobacco nucleolar protein that is homologous to TNP1 encoded by a transposable element of *Antirrhinum majus*. J Plant Res 121:425–433

Kiyokawa S, Kobayashi K, Kikuchi Y (1994) Root-inducing region of mikimopine type Ri plasmid pRi1724. Plant Physiol 104:801–802

Klee H, Horsch R, Hinchee MA et al (1987) The effects of overproduction of two *Agrobacterium tumefaciens* T-DNA auxin biosynthetic gene products in transgenic *Petunia* plants. Genes Dev 1:86–96

Kodahl N, Müller R, Lütken H (2016) The *Agrobacterium rhizogenes* oncogenes *rolB* and *ORF13* increase formation of generative shoots and induce dwarfism in *Arabidopsis thaliana* (L.) Heynh. Plant Sci 252:22–29

Koltunow AM, Johnson SD, Lynch M et al (2001) Expression of *rolB* in apomictic *Hieracium piloselloides* Vill. causes ectopic meristems in planta and changes in ovule formation, where apomixis initiates at higher frequency. Planta 214:196–205

Komari T (1990) Genetic characterization of a double-flowered tobacco plant obtained by a transformation experiment. Theor Appl Genet 80:167–171

Koncz C, Schell J (1986) The promoter of TL-DNA gene *5* controls the tissue-specific expression of chimaeric genes carried by a novel type of *Agrobacterium* binary vector. Mol Gen Genet 204:383–396

Körber H, Strizhov H, Staiger D et al (1991) T-DNA gene *5* of *Agrobacterium* modulates auxin response by autoregulated synthesis of a growth hormone antagonist in plants. EMBO J 10:3983–3991

Kyndt T, Quispe D, Zhai H et al (2015) The genome of cultivated sweet potato contains *Agrobacterium* T-DNAs with expressed genes: an example of a naturally transgenic food crop. Proc Natl Acad Sci USA 112:5844–5849

Leach F (1991) Promoter analysis of the highly expressed rolC and rolD root-inducing genes of *Agrobacterium rhizogenes*: enhancer and tissue-specific DNA determinants are dissociated. Plant Sc 79:69–76

Leemans J, Hernalsteens JP, Deblaere R et al (1983) Genetic analysis of T-DNA and regeneration of transformed plants. In: Molecular genetics of the bacteria-plant interaction. Proc Life Sci 322–330

Legué V, Driss-Ecole D, Maldiney R et al (1996) The response to auxin of rapeseed (*Brassica napus* L.) roots displaying reduced gravitropism due to transformation by *Agrobacterium rhizogenes*. Planta 200:119–124

Lemcke K, Schmülling T (1998a) Gain of function assays identify non-rol genes from *Agrobacterium rhizogenes* TL-DNA that alter plant morphogenesis or hormone sensitivity. Plant J 15:423–433

Lemcke K, Schmülling T (1998b) A putative *rolB* gene homologue of the *Agrobacterium rhizogenes* TR-DNA has different morphogenetic activity in tobacco than *rolB*. Plant Mol Biol 36:803–808

Lemcke K, Prinsen E, Van Onckelen H et al (2000) The *ORF8* gene product of *Agrobacterium rhizogenes* TL-DNA has tryptophan 2-monooxygenase activity. Mol Plant-Microbe Interact 13:787–790

Levesque H, Delepelaire P, Rouzé P et al (1988) Common evolutionary origin of the central portion of the Ri TL-DNA of *Agrobacterium rhizogenes* and the Ti T-DNAs of *Agrobacterium tumefaciens*. Plant Mol Biol 11:731–744

Lütken H, Clarke JL, Müller R (2012) Genetic engineering and sustainable production of ornamentals: current status and future directions. Plant Cell Rep 31:1141–1157

Marchler-Bauer A, Bo Y, Han L et al (2017) CDD/SPARCLE: functional classification of proteins via subfamily domain architectures. NAR 45:D200–D203

Matsuki R, Uchimiya H (1994) A 43-kDa nuclear tobacco protein interacts with a specific single-stranded DNA sequence from the 5′-upstream region of the *Agrobacterium rhizogenes rolC* gene. Gene 140:201–205

Matsuki R, Onodera H, Yamaguchi T (1989) Tissue-specific expression of the *rolC* promoter of the Ri plasmid in transgenic rice plants. Mol Gen Genet 220:12–16

Matveeva TV, Bogomaz DI, Pavlova OA et al (2012) Horizontal gene transfer from genus *Agrobacterium* to the plant *Linaria* in nature. Mol Plant-Microbe Interact 25:1542–1551

Maurel C, Brevet J, Barbier-Brygoo H et al (1990) Auxin regulates the promoter of the root-inducing *rolB* gene of *Agrobacterium rhizogenes* in transgenic tobacco. Mol Gen Genet 223:58–64

Maurel C, Barbier-Brygoo H, Spena A et al (1991) Single *rol* genes from the *Agrobacterium rhizogenes* T(L)-DNA alter some of the cellular responses to auxin in *Nicotiana tabacum*. Plant Physiol 97:212–216

Maurel C, Leblanc N, Barbier-Brygoo H et al (1994) Alterations of auxin perception in *rolB*-transformed tobacco protoplasts. Time course of *rolB* mRNA expression and increase in auxin sensitivity reveal multiple control by auxin. Plant Physiol 105:1209–1215

Messens E, Lenaerts A, Van Montagu M et al (1985) Genetic basis for opine secretion from crown gall tumor cells. Mol Gen Genet 199:344–348

Meyer AD, Ichikawa T, Meins F (1995) Horizontal gene transfer: regulated expression of tobacco homologue of the *Agrobacterium rhizogenes rolC* gene. Mol Gen Genet 249:265–273

Meyer A, Tempé J, Costantino P (2000) Hairy root: a molecular overview. Functional analysis of *Agrobacterium rhizogenes* T-DNA genes. In: Stacey G, Keen NT (eds) Plant-Microbe Interact, vol 5. American Phytopathological Society Press, St Paul, pp 93–139

Mohajjel-Shoja H (2010) Contribution to the study of the *Agrobacterium rhizogenes plast* genes *rolB* and *rolC*, and their homologs in *Nicotiana tabacum*. Thesis, University of Strasbourg

Mohajjel-Shoja H, Clément B, Perot J et al (2011) Biological activity of the *Agrobacterium rhizogenes*-derived *trolC* gene of *Nicotiana tabacum* and its functional relationship to other *plast* genes. Mol Plant-Microbe Interact 24:44–53

Moriuchi H, Okamoto C, Nishihama R et al (2004) Nuclear localization and interaction of RolB with plant 14-3-3 proteins correlates with induction of adventitious roots by the oncogene *rolB*. Plant J 38:260–275

Nagata N, Kosono S, Sekine M et al (1995) The regulatory functions of the *rolB* and *rolC* genes of *Agrobacterium rhizogenes* are conserved in the homologous genes (*Ngrol*) of *Nicotiana glauca* in tobacco genetic tumors. Plant Cell Physiol 36:1003–1012

Nagata N, Kosono S, Sekine M et al (1996) Different expression patterns of the promoters of the *NgrolB* and *NgrolC* genes during the development of tobacco genetic tumors. Plant Cell Physiol 37:489–498

Nath U, Crawford BCW, Carpenter R et al (2003) Genetic control of surface curvature. Science 299:1404–1407

Nester EW (2015) *Agrobacterium*: nature's genetic engineer. Front Plant Sci 5, article 730

Nilsson O, Olsson O (1997) Getting to the root: the role of the *Agrobacterium rhizogenes rol* genes in the formation of hairy roots. Physiol Plant 100:463–473

Nilsson O, Moritz T, Imbault N et al (1993a) Hormonal characterization of transgenic tobacco plants expressing the *rolC* gene of *Agrobacterium rhizogenes* TL-DNA. Plant Physiol 102:363–371

Nilsson O, Crozier A, Schmülling T et al (1993b) Indole-3-acetic acid homeostasis in transgenic tobacco plants expressing the *Agrobacterium rhizogenes rolB* gene. Plant J 3:681–689

Nilsson O, Moritz T, Sundberg B et al (1996a) Expression of the *Agrobacterium rhizogenes rolC* gene in a deciduous forest tree alters growth and development and leads to stem fasciation. Plant Physiol 112:493–502

Nilsson O, Little CH, Sandberg G et al (1996b) Expression of two heterologous promoters, *Agrobacterium rhizogenes rolC* and cauliflower mosaic virus 35S, in the stem of transgenic hybrid aspen plants during the annual cycle of growth and dormancy. Plant Mol Biol 31:887–895

Nilsson O, Tuominen H, Sundberg B et al (1997) The *Agrobacterium rhizogenes rolB* and *rolC* promoters are expressed in pericycle cells competent to serve as root initials in transgenic hybrid aspen. Physiol Plant 100:456–462

O'Grady K, Gurley WB (1995) Site-directed mutagenesis of the enhancer region of the *780* gene promoter of T-DNA. Plant Mol Biol 29:99–108

Oono Y, Handa T, Kanaya K et al (1987) The TL-DNA of Ri plasmids responsable for dwarfness of tobacco plants. Jpn J Genet 62:501–505

Oono Y, Kanaya K, Uchimiya H (1990) Early flowering in transgenic tobacco plants possessing the *rolC* gene of *Agrobacterium rhizogenes* Ri plasmid. Jpn J Genet 65:7–16

Oono Y, Satomi T, Uchimiya H (1991) *Agrobacterium rhizogenes lacZ-rolC* gene expression in *Escherichia coli*: detection of the product in transgenic plants using RolC-specific antibodies. Gene 104:95–98

Oono Y, Suzuki T, Toki S et al (1993) Effects of the over-expression of the *rolC* gene on leaf development in transgenic periclinal chimeric plants. Plant Cell Physiol 34:745–752

Ophel K, Kerr A (1990) *Agrobacterium vitis*-new species for strains of *Agrobacterium* biovar 3 from grapevine. Int J Syst Bacteriol 40:236–241

Otten L, De Ruffray P (1994) *Agrobacterium vitis* nopaline Ti plasmid pTiAB4: relationship to other Ti plasmids and T-DNA structure. Mol Gen Genet 245:493–505

Otten L, Helfer A (2001) Biological activity of the *rolB*-like 5′ end of the A4-*orf8* gene from the *Agrobacterium rhizogenes* TL-DNA. Mol Plant-Microbe Interact 14:405–411

Otten L, Schmidt JA (1998) T-DNA from the *Agrobacterium* limited-host range strain AB2/73 contains a single oncogene. Mol Plant-Microbe Interact 11:335–342

Otten L, Szegedi E (1985) Crown galls induced by octopine-degrading biotype 3 strains of *Agrobacterium tumefaciens* contain a new form of lysopine dehydrogenase. Plant Sci 40:81–85

Otten L, Vreugdenhil D, Schilperoort RA (1977) Properties of D(+)-lysopine dehydrogenase from crown gall tumour tissue. Biochim Biophys Acta 485:268–277

Otten L, Salomone JY, Helfer A et al (1999) Sequence and functional analysis of the left-hand part of the T-region from the nopaline-type Ti plasmid, pTiC58. Plant Mol Biol 41:765–776

Palatnik JF, Allen E, Wu X, Schommer C et al (2003) Control of leaf morphogenesis by microRNAs. Nature 425:257–263

Palazon J, Cusido RM, Roig C et al (1998) Expression of the *rolC* gene and nicotine production in transgenic roots and their regenerated plants. Plant Cell Rep 17:384–390

Pasternak T, Haser T, Falk T et al (2017) A 3D digital atlas of the *Nicotiana tabacum* root tip and its use to investigate changes in the root apical meristem induced by the *Agrobacterium 6b* oncogene. Plant J 92:31–42

Ream LW, Gordon MP, Nester EW (1983) Multiple mutations in the T region of the *Agrobacterium tumefaciens* tumor-inducing plasmid. Proc Natl Acad Sci USA 80:1660–1664

Reddy S, Dasgupta S, Rymarquis L et al (2003) Analysis of the *Agrobacterium tumefaciens* pTiChry5 *6b* promoter. J Plant Biochem Biotech 12:87–91

Riker AJ (1930) Studies on infectious hairy root of nursery apple trees. J Agric Res 41:507–540

Röder FT, Schmülling T, Gatz C (1994) Efficiency of the tetracycline-dependent gene expression system: complete suppression and efficient induction of the *rolB* phenotype in transgenic plants. Mol Gen Genet 243:32–38

Saha P, Chakraborti D, Sarkar A et al (2007) Characterization of vascular-specific RSs1 and *rolC* promoters for their utilization in engineering plants to develop resistance against hemipteran insect pests. Planta 226:429–442

Salomon F, Deblaere R, Leemans J et al (1984) Genetic identification of functions of TR-DNA transcripts in octopine crown galls. EMBO J 3:141–146

Satuti NSN, Moriuchi H, Yamakawa M et al (2005) Characterization of the *rolB* promoter on mikimopine-type pRi1724 T-DNA. Plant Sci 108:1353–1364

Satuti NSN, Tanaka N, Yoshida K et al (2007) Phenotype of transgenic tobacco plants (*Nicotiana tabacum* cv. Petit Havana SR-1) expressing 1724*orf13* gene of *Agrobacterium rhizogenes* strain MAFF301724. Indonesian J Biotechnol 12:980–987

Schell J, Van Montagu M, De Beuckeleer M et al (1979) Interactions and DNA transfer between *Agrobacterium tumefaciens*, the Ti-plasmid and the plant host. Proc R Soc Lond B Biol Sci 204:251–266

Schmidt J (1999) Etude d'un oncogène de la souche AB2/73 d'*Agrobacterium tumefaciens*. Thesis, Louis Pasteur University of Strasbourg

Schmülling T, Schell J (1993) Transgenic tobacco plants regenerated from leaf disks can be periclinal chimeras. Plant Mol Biol 21:705–708

Schmülling T, Schell J, Spena A (1988) Single genes from *Agrobacterium rhizogenes* influence plant development. EMBO J 7:2621–2629

Schmülling T, Schell J, Spena A (1989) Promoters of the *rolA*, *B*, and *C* genes of *Agrobacterium rhizogenes* are differentially regulated in transgenic plants. Plant Cell 1:665–670

Schmülling T, Fladung M, Grossmann K et al (1993) Hormonal content and sensitivity of transgenic tobacco and potato plants expressing single *rol* genes of *Agrobacterium rhizogenes* T-DNA. Plant J 3:371–382

Schröder G, Waffenschmidt S, Weiler W et al (1984) The T-region of Ti plasmid codes for an enzyme synthesizing indole-3-acetic acid. Eur J Biochem 138:387–391

Scorza R, Zimmerman TW, Cordts JM et al (1994) Horticultural characteristics of transgenic tobacco expressing the *rolC* gene from *Agrobacterium rhizogenes*. J Am Soc Hortic Sci 119:1091–1098

Serino G, Clérot D, Brevet J et al (1994) *rol* genes of *Agrobacterium rhizogenes* cucumopine strain: sequence, effects and pattern of expression. Plant Mol Biol 26:415–422

Shabtai S, Salts Y, Kaluzky G et al (2007) Improved yielding and reduced puffiness under extreme temperatures induced by fruit-specific expression of *rolB* in processing tomatoes. Theor Appl Genet 114:1203–1209

Shen WH, Petit A, Guern J et al (1988) Hairy roots are more sensitive to auxin than normal roots. Proc Natl Acad Sci USA 85:3417–3421

Shen WH, Davioud E, David C et al (1990) High sensitivity to auxin is a common feature of hairy root. Plant Physiol 94:554–560

Shkryl YN, Veremeichik GN, Bulgakov VP et al (2008) Individual and combined effects of the *rolA*, *B*, and *C* genes on anthraquinone production in *Rubia cordifolia* transformed calli. Biotechnol Bioeng 100:118–125

Sinkar VP, Pythoud F, White FF et al (1988) *rolA* locus of the Ri plasmid directs developmental abnormalities in transgenic tobacco plants. Genes Dev 2:688–697

Sitbon F, Hennion S, Sundberg B et al (1992) Transgenic tobacco plants coexpressing the *Agrobacterium tumefaciens iaaM* and *iaaH* genes display altered growth and indoleacetic acid metabolism. Plant Physiol 99:1062–1069

Spanier K, Schell J, Schreier PH (1989) A functional analysis of T-DNA gene *6b*: the fine-tuning of cytokinin effects on shoot development. Mol Gen Genet 219:209–216

Spano L, Mariotti D, Cardarelli M et al (1988) Morphogenesis and auxin sensitivity of transgenic tobacco with different complements of Ri T-DNA. Plant Physiol 87:479–483

Spena A, Schmülling T, Koncz C et al (1987) Independent and synergistic activity of *rolA*, *B* and *C* loci in stimulating abnormal growth in plants. EMBO J 6:3891–3899

Spena A, Aalen RB, Schulze SC (1989) Cell-autonomous behavior of the *rolC* gene of *Agrobacterium rhizogenes* during leaf development: a visual assay for transposon excision in transgenic plants. Plant Cell 1:1157–1164

Stieger PA, Meyer AD, Kathmann P et al (2004) The *orf13* T-DNA gene of *Agrobacterium rhizogenes* confers meristematic competence to differentiated cells. Plant Physiol 135:1798–1808

Studholme DJ, Downie JA, Preston GM (2005) Protein domains and architectural innovation in plant-associated Proteobacteria. BMC Genomics 16:6–17

Sugaya S, Uchimiya H (1992) Deletion analysis of the 5′-upstream region of the *Agrobacterium rhizogenes* Ri plasmid *rolC* gene required for tissue-specific expression. Plant Physiol 99:464–467

Sugaya S, Hayakawa K, Handa T et al (1989) Cell-specific expression of the *rolC* gene of the TL-DNA of Ri plasmid in transgenic tobacco plants. Plant Cell Physiol 30:649–653

Suzuki A, Kato A, Uchimiya H (1992) Single-stranded DNA of 5′-upstream region of the *rolC* gene interacts with nuclear proteins of carrot cell cultures. Biochem Biophys Res Commun 188:727–733

Suzuki K, Yamashita I, Tanaka N (2002) Tobacco plants were transformed by *Agrobacterium rhizogenes* infection during their evolution. Plant J 32:775–787

Takahashi S, Sato R, Takahashi M et al (2013) Ectopic localization of auxin and cytokinin in tobacco seedlings by the plant-oncogenic AK-*6b* gene of *Agrobacterium tumefaciens* AKE10. Planta 238:753–770

Tanaka N, Ikeda T, Oka A (1994) Nucleotide sequence of the *rol* region of the mikimopine-type root-inducing plasmid pRi1724. Biosci Biotechnol Biochem 58:548–551

Terakura S, Kitakura S, Ishikawa M et al (2006) Oncogene *6b* from *Agrobacterium tumefaciens* induces abaxial cell division at late stages of leaf development and modifies vascular development in petioles. Plant Cell Physiol 47:664–672

Terakura S, Ueno Y, Tagami H et al (2007) An oncoprotein from the plant pathogen *Agrobacterium* has histone chaperone-like activity. Plant Cell 19:2855–2865

Thomashow LS, Reeves S, Thomashow MF (1984) Crown gall oncogenesis: evidence that a T-DNA gene from the Agrobacterium Ti plasmid pTiA6 encodes an enzyme that catalyzes synthesis of indoleacetic acid. Proc Nat Acad Sci USA 81:5071–5075

Thomashow MF, Hugly S, Buchholz WG et al (1986) Molecular basis for the auxin-independent phenotype of crown gall tumor tissue. Science 231:616–618

Tinland B, Huss B, Paulus F et al (1989) *Agrobacterium tumefaciens 6b* genes are strain-specific and affect the activity of auxin as well as cytokinin genes. Mol Gen Genet 219:217–224

Tinland B, Rohfritsch O, Michler P et al (1990) *Agrobacterium tumefaciens* T-DNA gene *6b* stimulates *rol*-induced root formation, permits growth at high auxin concentrations and increases root size. Mol Gen Genet 223:1–10

Tinland B, Fournier P, Heckel T et al (1992) Expression of a chimaeric heat-shock-inducible *Agrobacterium 6b* oncogene in *Nicotiana rustica*. Plant Mol Biol 18:921–930

Udagawa M, Aoki S, Syono K (2004) Expression analysis of the *NgORF13* promoter during the development of tobacco genetic tumors. Plant Cell Physiol 45:1023–1031

Umber M, Voll L, Weber A et al (2002) The *rolB*-like part of the *Agrobacterium rhizogenes orf8* gene inhibits sucrose export in tobacco. Mol Plant-Microbe Interact 15:956–962

Umber M, Clément B, Otten L (2005) The T-DNA oncogene A4-*orf8* from *Agrobacterium rhizogenes* A4 induces abnormal growth in tobacco. Mol Plant-Microbe Interact 18:205–211

van Altvorst AC, Bino RJ, van Dijk AJ et al (1992) Effects of the introduction of *Agrobacterium rhizogenes rol* genes on tomato plant and flower development. Plant Sci 83:77–85

Van Onckelen H, Rüdelsheim P, Inzé D et al (1985) Tobacco plants transformed with the *Agrobacterium* T-DNA gene *1* contain high amounts of indole-3-acetamide. FEBS Lett 181:373–376

Van Onckelen H, Prinsen E, Inzé D et al (1986) *Agrobacterium* T-DNA gene *1* codes for tryptophan monooxygenase activity in tobacco crown gall cells. FEBS Lett 198:357–360

Veremeichik GN, Shkryl YN, Bulgakov VP et al (2012) Molecular cloning and characterization of seven class III peroxidases induced by overexpression of the agrobacterial *rolB* gene in *Rubia cordifolia* transgenic callus cultures. Plant Cell Rep 31:1009–1019

Wabiko H, Minemura M (1996) Exogenous phytohormone-independent growth and regeneration of tobacco plants transgenic for the *6b* gene of *Agrobacterium tumefaciens* AKE10. Plant Physiol 112:939–951

Wang M, Soyano T, Machida S et al (2011) Molecular insights into plant cell proliferation disturbance by *Agrobacterium* protein 6b. Genes Dev 25:64–76

Wasserman LA, Sergeev AI, Vasil'ev VG et al (2015) Thermodynamic and structural properties of tuber starches from transgenic potato plants grown *in vitro* and *in vivo*. Carbohydr Polym 125:214–223

White FF, Garfinkel DJ, Huffman GA et al (1983) Sequence homologous to *Agrobacterium rhizogenes* T-DNA in the genomes of uninfected plants. Nature 301:348–350

White FF, Taylor BH, Huffman GA et al (1985) Molecular and genetic analysis of the transferred DNA regions of the root-inducing plasmid of *Agrobacterium rhizogenes*. J Bacteriol 164:33–44

Winefield C, Lewis D, Arathoon S et al (1999) Alteration in *Petunia* plant form through the introduction of the *rolC* gene from *Agrobacterium rhizogenes*. Mol Breeding 5:543–551

Yokoyama R, Hirose T, Fujii N et al (1994) The *rolC* promoter of *Agrobacterium rhizogenes* Ri plasmid is activated by sucrose in transgenic tobacco plants. Mol Gen Genet 244:15–22

Zhang M, Pereira e Silva Mde C, Chaib De Mares M et al (2014) The mycosphere constitutes an arena for horizontal gene transfer with strong evolutionary implications for bacterial-fungal interactions. FEMS Microbiol Ecol 89:516–526

Zhu L, Holefors A, Ahlman A et al (2001) Transformation of the apple rootstock M.9/29 with the *rolB* gene and its influence on rooting and growth. Plant Sc 160:433–439

Agrobacterium-Mediated Transformation in the Evolution of Plants

Tatiana V. Matveeva

Contents

1 Introduction 422
2 T-DNA-like Sequences in Plant Species 423
2.1 Nicotiana 423
2.2 Linaria 426
2.3 Ipomoea 427
3 Common Features of cT-DNA Structures in Different Species 428
4 Methods Used to Search for cT-DNAs 428
5 Functional Characterization of cT-DNA Genes in Different Naturally Transgenic Species 431
6 Possible Functions of cT-DNA 431
6.1 Plant Adaptation to Arid Conditions 432
6.2 Regeneration Ability 432
6.3 Life Cycles 435
6.4 Immunity Hypothesis 436
6.5 Genes of Opine Synthesis and the Microbiome 436
6.6 cT-DNA and Secondary Metabolism 437
7 Conclusion 438
References 438

Abstract In most cases, the genetic engineering of plants uses *Agrobacterium*-mediated transformation to introduce novel genes. In nature, insertion of T-DNA into the plant genome and its subsequent transfer via sexual reproduction have been shown for several species in the genera *Nicotiana, Ipomoea*, and *Linaria*. A sequence homologous to T-DNA of the Ri plasmid of *Agrobacterium rhizogenes* was found in the genome of wild-type *Nicotiana glauca* (section Noctiflorae) more than 30 years ago and was named "cellular T-DNA" (cT-DNA). It comprises an imperfect inverted repeat and contains homologs of several T-DNA oncogenes (Ng*rolB*, Ng*rol*C, Ng*orf13*, Ng*orf14*) and an opine synthesis gene (Ng*mis)*.

T. V. Matveeva (✉)
St. Petersburg State University, 7/9 University emb, St. Petersburg 199034, Russia
e-mail: radishlet@gmail.com

Current Topics in Microbiology and Immunology (2018) 418:421–441
DOI 10.1007/82_2018_80

Published Online: 3 March 2018

Multiple cT-DNAs have also been found in species of the sections *Tomentosae* and *Nicotiana* of the genus *Nicotiana*. These ancient cT-DNA genes are still expressed, indicating that they may play a role in the evolution of these plants. In 2012–2013, cT-DNA was detected and characterized in *Linaria vulgaris* and *L. genistifolia* ssp. *dalmatica*. Their cT-DNA is present in two copies and organized as an imperfect direct tandem repeat, containing Lv*ORF2*, Lv*ORF3*, Lv*ORF8*, Lv*rolA*, Lv*rolB*, Lv*rolC*, Lv*ORF13*, Lv*ORF14*, and the Lv*mis* genes. In 2015, cT-DNA was found in *Ipomoea*. Two types of T-DNA-like sequences were described within this genera, and their distribution varied among cultured hexaploid, tetraploid, and wild diploid forms. Thus, several independent T-DNA integration events occurred in the genomes of these three plant genera. We propose that the events of T-DNA insertion in the plant genome might have affected their evolution, resulting in the creation of new plant species. In this chapter, we focus on the structure and functions of cT-DNA in *Linaria*, *Nicotiana,* and *Ipomoea* and discuss their possible evolutionary role.

1 Introduction

Horizontal gene transfer (HGT) takes place widely in prokaryotes, where its ecological and evolutionary effects are well-studied (Koonin et al. 2001). Prokaryotes have acquired a large number of important traits, including antibiotic and pathogen resistance, new metabolic pathways, and better adaptability to some environmental factors, via HGT. These horizontally transferred genes enable bacteria and archaea to explore new habitats and hence facilitate their rapid evolution (Koonin et al. 2001; Gogarten et al. 2002). At the same time, a number of recent discoveries indicate the possible contribution of HGT in the evolution of eukaryotes, as well as prokaryotes. Comparative and phylogenetic analyses of eukaryotic genomes show that considerable numbers of genes have been acquired by HGT. However, mechanisms of HGT in eukaryotic organisms are poorly understood in comparison with gene transfer among the prokaryota. The persistence of horizontally transferred genes in eukaryotic organisms may confer selective advantages that can be classified into two groups:

- improving existing functions;
- providing the recipient with new functions (e.g., altered host nutrition, protection and adaptation to environmental factors) (Koonin et al. 2001; Richardson and Palmer 2007; Husnik and McCutcheon 2017).

In multi-cellular eukaryotes, early developmental stages that are exposed to their environment (spores, zygotes, or embryos) and totipotent plant cells are likely to be recipient cells for foreign DNA that can be passed to offspring (Huang 2013). In higher plants chloroplast and mitochondrial DNA are often involved in HGT and have been the subject of numerous reviews (Dong et al. 1998; Richardson and Palmer 2007). Evidence of gene transfer from bacteria to the nuclei of multi-cellular

eukaryotes is rare (Richards et al. 2006; Acuna et al. 2012). The close contact which frequently occurs in parasitism, symbiosis, pathogen, epiphyte, entophyte, and grafting interactions could promote HGT between species. Besides these direct transfer methods, genes can be exchanged with such "vectors" as pollen, fungi, bacteria, viruses, viroids, plasmids, transposons, and insects (Gao et al. 2014). One of the best studied examples of natural HGT from bacteria to plants is HGT between *Agrobacterium sp.* and species of the genera *Nicotiana*, *Ipomoea*, and *Linaria* (White et al. 1983; Intrieri and Buiatti 2001; Matveeva et al. 2012; Pavlova et al. 2013; Kyndt et al. 2015).

Agrobacterium rhizogenes and *A. tumefaciens* are soil bacteria that are able to transfer a segment, called T-DNA, of a large tumor (or root)-inducing plasmid into plant host cells (White et al. 1982; Otten et al. 1992; Vain 2007). T-DNA is integrated into the plant genome and is expressed there. Expression of T-DNA genes leads to the formation of hairy roots or crown galls, transgenic tissues, formed on a non-transgenic plant. This process is beneficial for bacteria because the transgenic tissue synthesizes substances useful for the nutrition of the inciting bacterium. This example of HGT may in some cases be beneficial for plants because there are footprints of HGT from *Agrobacterium* to plants in the genomes of several present-day plant species. These sequences are homologous to *Agrobacterium* T-DNA sequences and have been transmitted through numerous sexual generations. Such sequences are called cellular T-DNAs, or cT-DNAs.

2 T-DNA-like Sequences in Plant Species

2.1 Nicotiana

cT-DNA was first discovered accidentally in *Nicotiana glauca* in 1982. The sequence had more than 80% similarity with *A. rhizogenes* T-DNA and was organized as an imperfect inverted repeat (White et al. 1982; Fürner et al. 1986). Later, it was shown that this cT-DNA included sequences similar to that of the following pRi genes: *orf11* (*rolB*), *orf12* (*rolC*), *orf13*, *orf14*, *orf15* (*rolD*) and homologs of the mikimopine synthase genes (*mis*). The *N. glauca* cT-DNA genes were called Ng*rolB*, Ng*rolC*, Ng*orf13*, Ng*orf14*, Ng*rolD*, and Ng*mis*. Each gene, except *rolB*, comes in two slightly different copies. This points to the origin of cT-DNA from an *A. rhizogenes* Ri plasmid of the mikimopine type (Fürner et al. 1986; Aoki et al. 1994; Suzuki et al. 2002). The lengths of the repeats, referred to as left and right arms of cT-DNA, are 7968 and 5778 bp, respectively (Suzuki et al. 2002). Fürner et al. (1986) investigated *Nicotiana glauca* plants collected in several geographical locations and showed the presence of cT-DNA in all studied varieties of *N. glauca*, with slight structural variation. Because cT-DNA has been identified in all studied varieties of *N. glauca*, it is reasonable to suggest that the transformation event occurred before the formation of this species. These results suggest that other related species may contain cT-DNA.

Goodspeed divided *Nicotiana* into three subgenera *Rustica, Tabacum*, and *Petunioides*, and into 14 sections (Goodspeed 1954). Since then, the number of subgenera of *Nicotiana* has remained constant, whereas the number and composition of the sections have been revised (Clarkson et al. 2005). *N. glauca* now belongs to the subgenus *Petunioides*, section *Noctiflorae*. It would be interesting to analyze other representatives of the *Noctiflorae* section for the presence of T-DNA.

Several research groups searched for cT-DNA in various species of *Nicotiana* from all three subgenera (Fürner et al. 1986; Intrieri and Buiatti 2001; Chen et al. 2014). More than 40 species were screened for T-DNA-like sequences. These data are summarized in Table 1. T-DNA-like sequences were found in each subgenus. It is important to note that there are some inconsistencies among the data of Fürner et al. (1986) and Intrieri and Buiatti (2001). For example, Intrieri and Buiatti (2001) showed that T-DNA is present in *N. debneyi* and *N. cordifolia*. Fürner et al. (1986) found no T-DNA in these species. This contradiction requires additional studies. One of the possible explanations could be variability among cultivars of the controversial species.

Comprehensive data were obtained by Chen et al. (2014) based on deep sequencing of genomes of several *Nicotiana* species of the subgenus *Tabacum*. They showed that the genome of *N. tomentosiformis* (section *Tomentosae*) carries four cT-DNA inserts (designated as TA, TB, TC, and TD) transferred to *Nicotiana* by several successive infections with different *Agrobacterium* strains. These cT-DNAs are different from the *N. glauca* cT-DNA which is referred to as gT. Each insert shows an incomplete inverted repeat structure. A fifth cT-DNA (TE) was discovered in *N. otophora* (section *Tomentosae*); its structure has not yet been assembled. However, TE contains the following types of sequences: vitopine synthase (*vis*)-like sequences, *6b*, *rolC*, *orf13*, and *orf14* (Chen et al. 2014). Other *Nicotiana* species in the same section as *N. tomentosiformis* contain the following sequences: TC in *N. otophora*; TC, TB, and TD in *N. tomentosa*; TC, TB, TD, and TA in *N. kawakamii*. *N. tabacum* (section *Nicotiana)* has three cT-DNAs: TA, TB, and TD and has lost TC (Chen et al. 2014). The TA insert resembles part of the *Agrobacterium rhizogenes* 1724 mikimopine-type T-DNA but has unusual *orf14* and *mis* genes. TB carries a fragment *orf14–mis*, similar to the T-DNA part of the 1724 strain, and a fragment with mannopine and agropine synthesis genes (*mas2′–mas1′–ags*). The TC insert is similar to that of the left part of the *A. rhizogenes* A4 T-DNA but also carries octopine synthase-like (*ocl*) sequences on the left and *c*-like genes previously found in *A. tumefaciens* on the right. TD has a T-DNA fragment similar to that of the right end of the TL-DNA of the *A. rhizogenes* strain A4 and includes an *orf14*-like gene, a gene *orf511* with unknown function, and remnants of the *orf18* gene (Chen et al. 2014). The presence in *Nicotiana* cT-DNA of some genes previously found in *A. tumefaciens* and *A. vitis* does not yet prove their origin from one or the other species. We know too little about the diversity of Ri plasmids of agrobacteria. However, it is likely that an ancient *Nicotiana* species underwent transformation by previously unknown types of plasmids. This idea is also supported by the following observations: In *N. tomentosiformis*, previously unknown T-DNA genes were found in each of

Table 1 Distribution of T-DNA-like sequences among *Nicotiana* species

Subgenus	Section	Species	T-DNA	SequenceAcc#	Ref
Rustica	*Paniculatae*	*N. paniculata*	–		1, 2[a]
		N. knightiana	–		1, 2
		N. solanifolia	–		2
		N. benavidesii	+(*rolC*)	n/a[b]	1, 2
	Rusticae	*N. cordifolia*	+(*rolB-orf14*)	AF281252.1 AF281248.1 AF281244.1	1, 2
		N. raimondi	–		2
		N. rustica	–		1, 2
Tabacum	*Tomentosae*	*N. tomentosiformis*	TA, TB, TC, TD	KJ599826 KJ599827 KJ599828 KJ599829	3
		N. tomentosa	TB, TC, TD		3
		N. kawakamii	TB, TC, TD		3
		N. otophora	TC, TE		3
		N. setchelli	+(*rolC*)	n/a	2
	Nicotiana	*N. tabacum*	TA, TB, TD		3
Petunioides	*Undulatae*	*N. undulata*	–		2
		N. glutinosa	–		1, 2
		N. arentsii	+(*rolC*)	n/a	2
	Trigonophyllae	*N. trigonophylla*	–		2
	Sylvestris	*N. sylvestris*	–		2
	Alatae	*N. alata*	–		2
		N. langsdorffi	–		2
		N. longiflora	–		2
		N. forgetiana	–		2
		N. sanderae	–		2
		N. plumbaginifolia	–		2
	Repandae	*N. nesophila*	–		2
		N. stocktonii	–		2
		N. repanda	–		2
		N. nudicaulis	–		2
	Noctiflorae	*N. glauca*	gT	X03432.1; D16559.1 AB071334.1; AB071335.1	1, 4–6
	Petunioides	*N. petunioides*	–		2
		N. acuminata	+(*rolC*)	n/a	2
		N. pauciflora	–		2
		N. attenuata	–		2
		N. miersii	+(*rolB*)	n/a	2

(continued)

Table 1 (continued)

Subgenus	Section	Species	T-DNA	SequenceAcc#	Ref
	Bigelovianae	*N. bigelovii*	+(*rolB*)	n/a	2
	Polydiclae	*N. clevelandi*	–		2
	Suaveolentes	*N. umbratica*	–		2
		N. debneyi	+(*rolC*)	AF281251.1	2
		N. gossei	+(*rolC*)	n/a	2
		N. rotundifolia	–		2
		N. suaveolens	+(*rolC*)	n/a	2
		N. exigua	+(*rolC*)	n/a	2
		N. goodspeedii	–		2

[a]1—Fürner et al. (1986); 2—Intrieri and Buiatti (2001); 3—Chen et al. (2014); 4—White et al. (1983); 5—Aoki et al. (1994); 6—Suzuki et al. (2002); [b]n/a, not available

the cT-DNA inserts. A gene distantly related to *orf14* was found in TA; a gene coding for a protein with weak similarity to agropine synthase (*ags*) was found in TB; one gene coding for a protein with weak similarity to octopine synthase (*ocs*) and one for a C-like protein (*c-like* gene) were found in TC; another gene related to *orf14* and a gene for a large, completely unknown protein (*orf511*) were found in TD (Chen et al. 2014; Chen and Otten 2017). In the *N. otophora* TE insert there is an unusual combination of genes: vitopine synthase (*vis*)-like sequences, *6b*, *rolC*, *orf13*, and *orf14*.

Because the NGS data allowed the detection of previously unknown multiple cT-DNAs in *Nicotiana* from the subgenus *Tabacum* (Chen et al. 2014), it raises the question whether there are also multiple cT-DNA insertions in the *N. glauca* genome. Deep sequencing this genome confirmed the structure of the previously described gT insertion and demonstrated the absence of other types of cT-DNA (Khafizova et al. 2018).

Thus, to date, we have well-documented outcomes of several independent acts of horizontal gene transfer from agrobacteria to species of *Nicotiana*. It is also interesting to note that each of the *Nicotiana* cT-DNAs has an incomplete repeat structure.

2.2 *Linaria*

The evolution of the genome of the toadflax plant (*Linaria*) of the Plantaginaceae family also involved agrobacterial transformations. A cT-DNA was found and characterized in detail in the genomes of *Linaria vulgaris* and *L. genistifolia* ssp. *dalmatica* (Matveeva et al. 2012; Pavlova et al. 2013). In *Linaria*, cT-DNA is organized as an imperfect direct repeat. The left arm of cT-DNA contains sequences homologous to the genes *acs*, *orf2*, *orf3*, *orf8*, *rolA*, *rolB*, *rolC*, *orf 13*, *orf14*, and *mis*. The right arm is shorter and does not carry the *acs* gene homolog.

Certain genes from the cT-DNAs of *L. vulgaris* ssp. *acutiloba* and *L. creticola* have also been sequenced (Matveeva and Kosachev 2013; Matveeva et al. 2018). Analysis of the sequence of the *rolC* of different toadflax species shows that all homologs of this gene can encode a full-length peptide. This gene is the most conserved cT-DNA gene in *Linaria* (Matveeva and Lutova 2014; Matveeva et al. 2018)

The *Linaria* cT-DNA is integrated into a DNA sequence similar to that of a retrotransposon. Analysis of the site of localization of cT-DNA indicates a single transformation of *Linaria* with this sequence and a monophyletic origin of *L. vulgaris*, *L. acutiloba*, *L. creticola*, and *L. genistifolia*, belonging, respectively, to sections *Linaria* and *Speciosae*.

2.3 *Ipomoea*

There are two types of cT-DNA in the *Ipomoea batatas* genome. The first type was named *Ib*TDNA1. It is polymorphic among sweet potato cultivars. In the Huachano cultivar, it contains the genes *acs*, *C-prot*, *iaaH*, and *iaaM* from *Agrobacterium* spp. as well as a defective copy of the *iaaM* gene in an inverted orientation. In the Xu781 cultivar, it has two copies of each of the following genes: *acs*, *C-prot*, *iaaH*, and *iaaM*. In the latter cultivar, a transposon inserted into the *iaaM* gene of the right arm. In the second type of cT-DNA, named *Ib*TDNA2, there are at least five intact ORF homologs of the genes *orf14*, *orf17n*, *rolB/rolC*, *orf13*, and *orf18/orf17n* from *A. rhizogenes*. Both T-DNA types are arranged as inverted repeats (Kyndt et al. 2015). *Ib*TDNA1 and 2 are present at different loci and segregate independently. *Ib*T-DNA1 is present in all domesticated varieties investigated, whereas *Ib*T-DNA2 is restricted to only some accessions (Quispe-Huamanquispe et al. 2017).

Whether *Ib*T-DNA1 and *Ib*T-DNA2 were introduced by one or two transformation events is not clear, because both could be derived from a single *Agrobacterium* strain and divergence of the sequences of their repeats does not contradict with both hypotheses. *orf13* sequences from *Ib*T-DNA2 were also detected in *I. trifida* (Kyndt et al. 2015). This result suggests that as in *Nicotiana*, cT-DNAs were introduced in an ancestral species and transmitted across speciation events. However, *Ib*T-DNA2 could also have been transferred by interspecific hybridization, known to occur between *I. batatas* and *I. trifida* (Chen and Otten 2017).

Thus at present, naturally transgenic plants are known in three genera—*Nicotiana*, *Ipomoea*, and *Linaria*—belonging to three families: Solanaceae, Convolvulaceae, and Plantaginaceae, respectively. Solanaceae and Convolvulaceae are members of the order Solanales, while Plantaginaceae belongs to the order of Lamiales. Both orders belong to of the superorder Lamiids of the subclass Asterids (Angiosperm Phylogeny Group 2016). Thus, the orders in which T-DNA-containing species were found are relatively close to each other in terms of phylogeny.

3 Common Features of cT-DNA Structures in Different Species

The common features of cellular T-DNAs of *Nicotiana, Linaria*, and *Ipomoea* are the prevalence of genes characteristic for *A. rhizogenes* and the repeat structure of the inserts (Fig. 1). DNA repeats, especially inverted ones, are prone to formation of secondary structures; this may decrease gene expression. Oncogenes of agrobacteria, actively expressed in plant cells, cause deviations in the development of plants. The silencing of these genes likely leads to milder effects on plant growth and development. Possibly, cT-DNAs arranged in the form of repeats were conserved in the genomes of ancestral forms of toadflax, sweet potato, and tobacco, suppressing expression of the oncogenes and permitting the regeneration of plants.

T-DNA insertion events provide interesting clues to reconstruct plant evolution. All species with a cT-DNA at the same insertion site are likely derived from a common ancestor. Based on this idea, we can conclude that within the genus *Linaria*, the sections of *Linaria* and *Speciosae* are of monophyletic origin, and the transformation of the ancestral form of the toadflaxes occurred before the separation of the sections. Based on the calculations of Blanco-Pastor et al. (2012), the separation of these sections occurred about 1.5 million years ago.

In the case of the *Tomentosae* section of *Nicotiana*, it may be possible to date the different insertion events. The most diverged *Nicotiana* cT-DNA (TC) shows 5.8% divergence between the repeats, which leads to an estimated age of 1 million years (Chen and Otten 2017). The sequence and timing of insertional events in *Nicotiana* are shown in Fig. 2.

Thus, the fact that all known cT-DNAs have a repeat structure not only could be an important property for preserving them in plant genomes, but also provides a welcome opportunity for evolutionary reconstruction of the transformation events.

4 Methods Used to Search for cT-DNAs

In early investigations of *Agrobacterium*-mediated transformation of plants, most researchers assumed that there was no significant homology between the T-DNA and non-transformed plant genomes. White et al. (1982) attempted to detect pRiA4b T-DNA sequences in the genome of *Nicotiana glauca*, transformed in laboratory conditions by *Agrobacterium rhizogenes* strain A4. DNA (southern) blot analysis detected a fragment of pRiA4 in the transgenic tissue. Surprisingly, a hybridization signal was also detected in uninfected tissues of *N. glauca*. Further analysis confirmed the presence of DNA homologous to T-DNA in the *N. glauca* genome. Thus, the first detection of T-DNA occurred accidentally.

Subsequent screening for T-DNA in tobacco was carried out by DNA blot hybridization or PCR followed by DNA blot hybridization of the products (Fürner et al. 1986; Intrieri and Buiatti 2001). PCR made it possible to detect individual

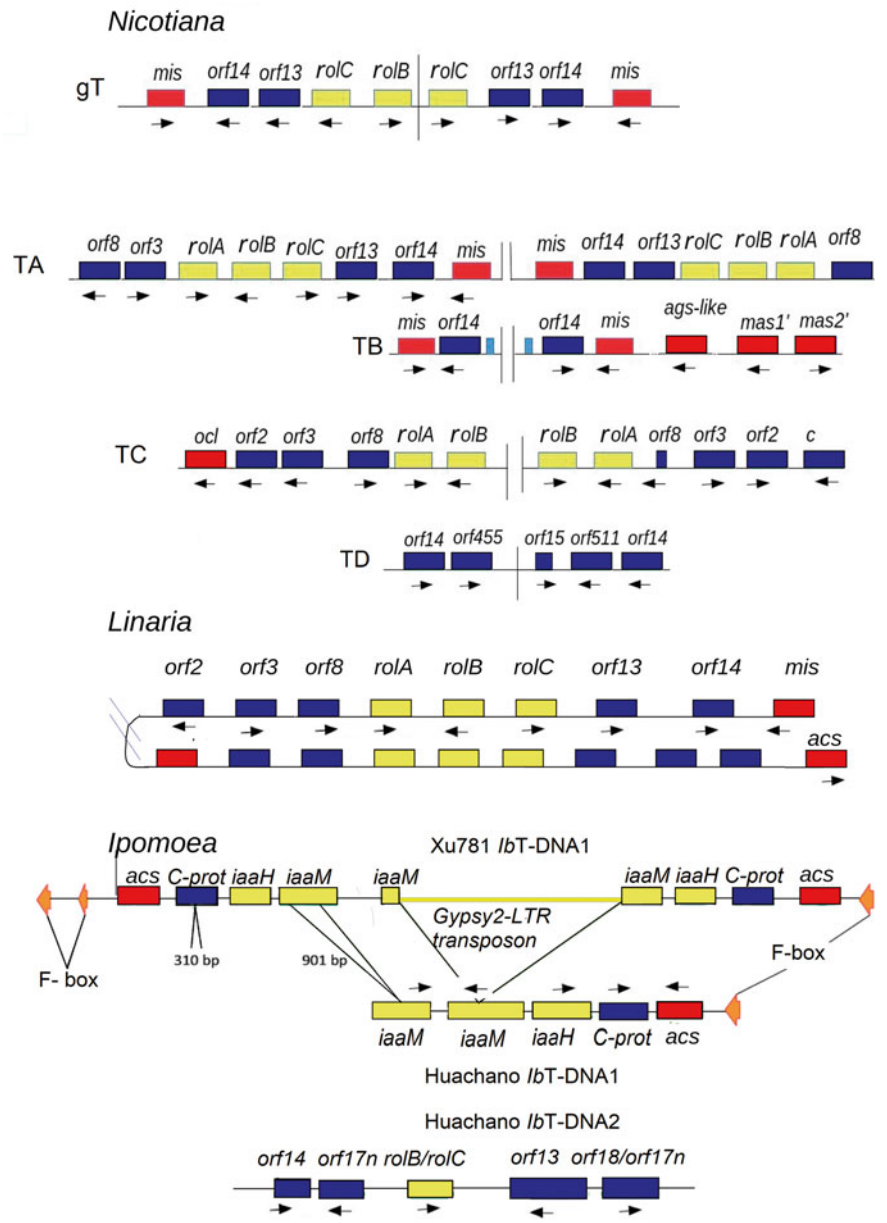


Fig. 1 Structural organization of cT-DNA in naturally transgenic plants. Oncogenes are indicated in yellow, opine genes are red, and genes with unknown function are blue. Modified from Matveeva and Sokornova (2017)

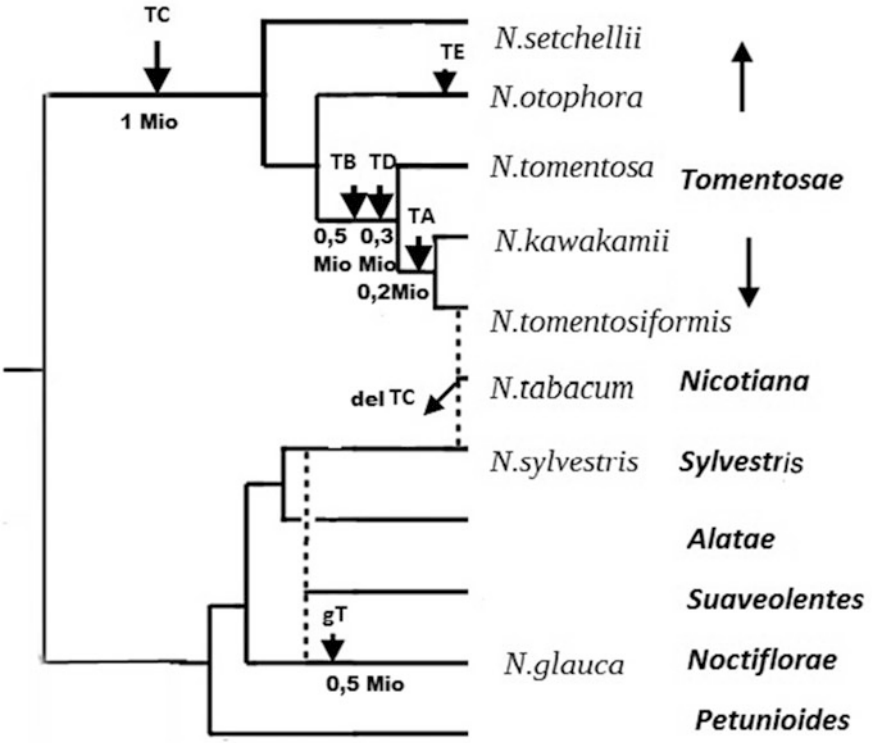


Fig. 2 Fragment of phylogenetic tree of *Nicotiana*. Sections names are in bold font, and species names are in standard font. Arrows show T-DNA integration events. Dotted lines show hybridization events. Modified from Matveeva and Sokornova (2017)

genes but did not indicate their conformation or relative positions. T-DNA in toadflax was identified through a targeted search for homologs of *Agrobacterium* oncogenes among more than a hundred species of plants (Matveeva et al. 2012). Real-time PCR was used for this research work. T-DNA in sweet potato was found by chance during a study of its miRNAs (Kyndt et al. 2015).

Real-time PCR is a fast, inexpensive, and reliable method to search for T-DNA homologs in genomes. However, for more detailed studies of the structure of T-DNA, additional experiments are required. These include DNA blot hybridization followed by cloning and sequencing the fragment, or long-range PCR and sequencing of its product. Using these methods, it is possible to describe the composition of a specific T-DNA insert. However, one cannot say for sure whether there are other T-DNAs in the genome. The analysis of genomic sequencing data can answer this question, as was demonstrated by Chen et al. (2014) and Khafizova et al. (2018).

Genome sequencing of *Nicotiana* species from the Tabacum subgenus has revealed the presence of homologous genes in different cT-DNAs (three *orf14* gene

homologs on TA, TB, and TD; two *orf3*, *orf8*, *rolA,* and *rolB* homologs on TA and TC; two *mis* homologs on TA and TB of *N. tomentosiformis*). Under these circumstances, the phylogenetic analysis of partial cT-DNA sequences from different species can only be carried out after it has been shown whether they belong to the same cT-DNA or not.

Thus, PCR-based methods are optimal for rapid cT-DNA sequence searching, whereas genomic sequencing is the most reliable method for detailed analysis of T-DNA in the genomes. Using this set of methods, one can look for and characterize new examples of HGT.

5 Functional Characterization of cT-DNA Genes in Different Naturally Transgenic Species

One of the most interesting problems is the elucidation of cT-DNA function. The first step in answering this question is to find out which of the cT-DNA genes are still expressed. It is also important to analyze which sequences potentially encode full-length proteins. Currently available information is summarized in Table 2.

Thus, *Nicotiana* and *Ipomoea* show structural integrity and expression of the following genes for the synthesis of opines: *mis, mas2′, acs.* Moreover, it was shown that, as a result of the functioning of the enzymes encoded by these genes, opines are synthesized (mikimopine in the heterologous system of *E. coli* and desoxyfructosylglutamine (DFG) in tobacco; Suzuki et al. 2002; Chen et al. 2016).

The *rolC* oncogene is expressed in tobacco and toadflaxes, and *rolB* and *orf13* are expressed in tobacco and sweet potato.

orf14 is expressed in *Nicotiana*, *geneC*, *iaaH*, *iaaM* in *Ipomoea*.

In order to understand the evolutionary role of the cT-DNAs in the species that harbor them, it will be essential to learn more about the functions of these genes.

6 Possible Functions of cT-DNA

The existence of several independent *Agrobacterium*-mediated transformation events and the maintenance of intact forms of cT-DNA genes in plant genomes during the process of evolution suggest that T-DNA-like sequences may confer some selective advantage upon the transformed plants (Ichikawa et al. 1990; Matveeva et al. 2012).

Several possible functions of cT-DNA genes are stated in the literature:

(1) increase in root mass as a plant adaptation to arid conditions;
(2) immunity to repeated agrobacterial infection;
(3) increased regeneration ability;

(4) passage to earlier flowering and, consequently, to an annual life cycle;
(5) influence on microbial communities within the plant rhizosphere and phyllosphere (Suzuki et al. 2002; Matveeva and Lutova 2014).

The available data allow us to identify some functional patterns in relation to naturally transgenic plants and help outline processes for further investigations.

6.1 Plant Adaptation to Arid Conditions

The hypothesis about the impact of cT-DNA on root mass was first described by Suzuki. Increasing root mass would seem beneficial for tolerance to arid conditions. Hence, ancient transformed plants with increased root mass might have demonstrated increased tolerance to dry environments, thus surviving in arid conditions (Tanaka 2008). However, no hairy root phenotype is observed in *Nicotiana* and *L. vulgaris* plants. In contrast, *L. vulgaris* and *L. creticola* exhibit a shooty phenotype. In the case of sweet potato, the picture appears more complicated. Crop roots of sweet potato represent overgrown side roots. Therefore, it is difficult to compare correctly the biomass of the roots of cultured naturally transgenic varieties of sweet potato with that of related non-transgenic wild species.

6.2 Regeneration Ability

A high regeneration capacity of the transformed plants may be one of the prerequisites for natural emergence of species containing agrobacterial DNA in their genomes. In this case, the progeny of transformed plants is expected to retain this capability. Another hypothesis proposes the influence of cT-DNA on an endogenous hormonal metabolism that may modify features of regeneration processes (Ichikawa et al. 1990).

Of all naturally transgenic plants, tobacco is the most studied plant in vitro. *N. tabacum* is a model species for genetic engineering (Draper et al. 1988). It exhibits typical reactions to exogenous hormones, such as callus formation in response to a combination of auxins and cytokinins, root formation in response to auxins, and shoot formation in response to cytokinins (Bogani et al. 1985).

Many *Nicotiana* species are easily regenerated in vitro, but different morphogenetic reactions predominate in different species. Some tobacco hybrids give rise to spontaneous tumors. To describe and model the tumor-forming mechanisms in tobacco hybrids, Naf (1958) proposed dividing all species of this genus into "+" and "−" groups. Species *N. langsdorffii*, *N. alata*, *N. longiflora*, *N. plumbaginifolia*, *N. sanderae*, *N. forgetiana*, *N. bonariensis*, and *N. noctiflora* were assigned to the "+" group. Species *N. glauca*, *N. tabacum*, *N. suaveolens*, *N. debneyi*, *N. rustica*, *N. paniculata*, *N. miersii*, and *N. bigelovii* were assigned to the "−"

Table 2 cT-DNA genes in different natural transformants

Species	cT-DNA	Genes	Intact	Expressed	Reference
Nicotiana glauca	gT	*rolB*	–	+	Aoki and Syono (1999a)
		rolC	+	+	
		orf13	+	+	Aoki and Syono (1999b)
		orf14	+	+	
		mis	+	+	Suzuki et al. (2002)
N. tomentosiformis and *N. tabacum* (N.tab)	TA	*orf8*	–	nt	Chen et al. (2014)
		rolA	–	nt	
		rolB	–	nt	
		rolC	+	+	Mohajjel-Shoja et al. (2011)
		orf13	+	+	Fründt et al. (1998)
		orf14-like	–	+	Chen et al. (2014)
		mis	–	nt	
	TB	*orf14*	+	nt	
		mis	–	nt	
		ags	–	nt	
		mas1'	–	nt	
		mas2'	+	+	Chen et al. (2016)
	TC	*ocs-like*	+	nt	Chen et al. (2014)
		orf2	–	nt	
		orf 3	–	nt	
		orf 8	–	nt	
		rolA	–	nt	
		rolB	–	nt	
		geneC	–	nt	
	TD	*orf18*	–	nt	
		orf14-like	+	nt	
		orf15	+	nt	
		orf511	?	nt	
	TE	*vis*	?	nt	
		6b	?	nt	
		mas1'	?	nt	
		mas2'	?	nt	
		rolB	?	nt	
		rolC	?	nt	
		orf13	?	nt	
		orf14	?	nt	
		iaaH	?	nt	
		iaaM	?	nt	
		acs	?	nt	

(continued)

Table 2 (continued)

Species	cT-DNA	Genes	Intact	Expressed	Reference
Linaria vulgaris	*Lvc*T-DNA	*acs*		nt	Matveeva et al. (2012)
		orf2	−	nt	
		orf3	−	nt	
		orf8	−	nt	
		rolA	−	nt	
		rolB	−	−	
		rolC	+	+	Matveeva et al. (2018)
		orf13	−	−	Matveeva et al. (2012)
		orf14	−	−	
		mis	−	nt	
L. vulgaris ssp. *acutiloba*	*Lac*T-DNA	*rolC*	+	nt	Matveeva and Kosachev (2013)
L. genistifolia ssp. *dalmatica*	*Lgc*T-DNA	*rolC*	+	nt	Pavlova et al. (2013)
L. genistifolia ssp. *genistifolia*	*Lgc*T-DNA	*rolC*	+	−	Matveeva et al. (2018)
		mis	−	−	Kovacova et al. (2014)
L. creticola	*Lcc*T-DNA	*rolC*	+	+	Matveeva et al. (2018)
Ipomoea batatas	*Ib*T-DNA1	*acs*	+	+	Kyndt et al. (2015)
		geneC	+	+	
		iaaH	+	+	
		iaaM	+	+	
	*Ib*T-DNA2	*orf14*	−	nt	
		orf17n	−	nt	
		rolB-like	+	+	
		orf13	+	+	
		orf18/0rf17n	+	nt	
Ipomoea trifida	*Ib*T-DNA2	*orf13*	+	nt	

nt—Not tested

group. Hybridization between representatives of different groups led to tumor-forming hybrids, but crosses within one group did not result in tumors (Kostoff 1930; Kehr and Smith 1954). Ichikawa and Syono (1991) advanced the hypothesis that the "−" group comprises tobacco species which contain cT-DNA and are therefore prone to root formation, whereas the species of the "+" group do not contain cT-DNA; they have an elevated cytokinin status and are thus prone to shoot formation. However, comparison of this list of species with data on T-DNA in genomes (Table 1) identifies inconsistencies. In the "−" group, at least *N. rustica* does not contain cT-DNA. Experiments in our laboratory (Matveeva et al. 2009)

showed that *N. suaveolens* does not form roots, but tends to form shoots in vitro, demonstrating intraspecific polymorphism on this trait. These facts do not support the hypothesis of Ichikawa and Syono (1991).

In the genus *Linaria*, regeneration in vitro was studied by our group using *L. vulgaris*, *L. creticola*, *L. genistifolia*, *L. maroccana*, *L. purpurea*, and *L. aeruginea* (Matveeva 2013; Matveeva et al. 2018). Internode explants of all species demonstrated a high ability to form calli on a hormone-free medium. The effectiveness of the process varied from 80 to 100% in different species. Furthermore, *L. vulgaris*, *L. creticola*, and *L. aeruginea* displayed intense shoot formation on a medium without hormones. This feature was most pronounced in naturally transgenic *L. creticola* and *L. vulgaris*, although other species were also highly regenerative. It is interesting to note the ability of *L. vulgaris* to regenerate shoots from roots (through a callus stage as well as directly) with a concomitant increase in the expression level of *rolC* (Matveeva et al. 2018).

In the genus *Ipomoea*, regeneration capability has also been actively investigated in vitro (Dessai et al. 1995; Kumar et al. 2013). Sweet potato can regenerate shoots on culture medium supplied with hormones. However, more research is needed to obtain a more precise characterization of the regenerative capacity of *Ipomoea* species with or without T-DNA.

Because all examined species possess high regenerative potential, this presumably enabled the transformed tissues (hairy roots or other types of growth) to spontaneously regenerate and to establish themselves as new species.

In conclusion, naturally transgenic *Nicotiana*, *Linaria*, and *Ipomoea* do not express the "hairy root" phenotype in adult plants. Moreover, toadflax and some *Nicotiana* species show a tendency for shoot formation in vitro, but other naturally transgenic *Nicotiana* species are prone to root formation, and the regeneration of sweet potato may follow an embryogenesis pathway (Gama et al. 1996). Additional studies of regenerative processes are required, accompanied by an analysis of the structure and expression of cT-DNA genes.

6.3 Life Cycles

Early experiments on hairy root plants suggested that cT-DNA genes may convert biennial plants to annual plants. This possibility was experimentally shown for *Cichorium intybus* and *Daucus carota,* transformed by *rolC* or wild-type pRi (Limami et al. 1998). Such peculiarities of the life cycle might bring evolutionary benefits through a wider spread of the transgenic plants as compared with their non-transformed relatives. In the absence of a cold impact, early flowering might allow new ecotypes to expand their range. However, biennial species were not found in the genera *Nicotiana, Ipomoea*, and *Linaria* (Matveeva and Sokornova 2017). In addition, the cT-DNA-containing *Linaria* species, of the sections *Linaria* and *Speciosae*, are perennial and spread in Eurasia (and in America as invasive

species), whereas representatives of other sections in which no cT-DNA was found could be annual and grow only in the regions of the Mediterranean and Pyrenean peninsulas (Sutton 1988). However, the ecological flexibility of the species from the sections *Linaria* and *Speciosae* could be related to the function of cT-DNA. The naturally transgenic species of *Nicotiana* are mainly represented by perennial forms. This is also true for sweet potato and, presumably, is due to the plant's ability to develop shoots from dormant buds on underground organs. In turn, this ability may be caused by the cT-DNA oncogene activity (Matveeva and Sokornova 2017).

6.4 Immunity Hypothesis

Apart from changing plant growth, cT-DNA gene expression may confer immunity to *Agrobacterium* by silencing incoming T-DNA (Escobar et al. 2001). However, in the *Tomentosae* section different *Agrobacterium* strains were able to re-infect already transformed species, arguing against this possibility (Chen and Otten 2017).

6.5 Genes of Opine Synthesis and the Microbiome

T-DNAs contain gene coding for enzymes which catalyze the biosynthesis of opines. Opines are amino acid derivatives and can be consumed by *Agrobacterium* as sources of nitrogen, carbon, and energy (Hong et al. 1997). Some other bacterial species have also acquired the capacity to degrade opines (Oger et al. 1997, 2000). Therefore, if natural transgenic plants produce opines, one can expect that the presence of such compounds will affect the composition of the bacterial populations in the rhizosphere. As a rule, a given T-DNA carries one or two types of opine synthesis genes; the genes encoding the breakdown of opines are located on the same Ti or Ri plasmid, but outside the T-DNA. Consequently, opines secreted by the transgenic cell are metabolized only by agrobacteria possessing a Ti/Ri plasmid (or a derived plasmid) that is similar to the one that has transformed the plant tissue. This confers a selective advantage for such bacteria and creates a kind of ecological niche for pathogenic strains of *Agrobacterium* (Oger et al. 1997, 2000).

More than twenty Ti and Ri plasmids are known that carry different types of opine synthesis genes. Probably many more remain to be discovered. Nevertheless, mikimopine and agrocinopine synthesis genes are present in most naturally transgenic plants. Mikimopine and mannopine synthase genes are expressed in *Nicotiana* (Table 2). *Nicotiana tabacum* cultivars highly expressing *mas2′* do indeed produce the expected product of the gene—desoxyfructosylglutamine (DFG)—and are the only known cases so far of natural transformants which synthesize opines (Chen et al. 2016).

6.6 cT-DNA and Secondary Metabolism

It was initially thought that T-DNA functions are related to regulation of the hormonal status of transgenic cells. As early as 1987, it was reported that cultures of "hairy roots" are notable for their increased content of secondary metabolites (Flores et al. 1988; Hamill et al. 1987; Rhodes et al. 1987). Since those observations, it was hypothesized that T-DNA genes have additional functions and this led to detailed investigations of effects of T-DNA genes on plant secondary metabolism. The effects of particular T-DNA genes (*rolA*, *rolB*, and *rolC*) and their combinations on secondary metabolism have been discussed. More details are provided in Matveeva et al. (2015).

The most common biologically active secondary metabolites of plants comprise glycosides, alkaloids, polyphenols (flavonoids, terpenoids, coumarins, and saponins), and essential oils. All of these chemical groups may be efficiently synthesized in "hairy root" cultures (Matveeva and Sokornova 2016). In natural conditions, these compounds are of great significance for plant protection from different environmental factors as well as for plant interactions with different organisms.

Plants of the genera *Linaria*, *Nicotiana*, and *Ipomea* are widely used in folk medicine because of high levels of secondary metabolites in their tissues. Bioactive compounds produced by naturally transgenic plants belong to different chemical groups. They are not unique for particular species and are synthesized in appreciable amounts by different representatives of the same genera and families. However, it seems quite possible that T-DNA genes may increase the levels of major fractions of secondary metabolites of plants. As an example, we can consider interesting studies that have been performed on tobacco grown in tissue culture. Palazon et al. (1998) have shown that transformation of *Nicotiana tabacum* with a *rolC* gene leads to increased synthesis of nicotine. Growth capacities and nicotine production were also greatly increased in lines transformed with *rolA*, *rolB*, and *rolC*. Unfortunately, the mechanism of the effect of T-DNA genes on secondary metabolism is poorly studied. This effect can be a direct or indirect. Secondary metabolites could protect the plant, and this may play a role in the selection and survival of natural transformants.

Alkaloids play an essential role in protecting *Nicotiana* plants from pests and diseases. However, alkaloid synthesis is only a part of an induced integrated defense response which also involves phenolics, proteinase inhibitors, PR-proteins, and sesquiterpenoid phytoalexins. Although the effects of the *rol* genes on the accumulation of tobacco antimicrobial alkaloids have been demonstrated for man-made transformants (Palazon et al. 1998), published data imply a more complex regulation of levels of these compounds along with the participation of various metabolites in the control of plant resistance to pathogens and pests (Bush et al. 1999). At present, there is no direct evidence that cT-DNA genes in natural transformants change plant secondary metabolism.

7 Conclusion

To date, naturally transgenic species are reported for three genera of dicotyledonous plants. In the most cases, the most probable source of cT-DNA is strains of *Agrobacterium rhizogenes*. Currently characterized T-DNA inserts are organized as imperfect repeats. Some of the cT-DNA genes are expressed, suggesting an evolutionary role for cT-DNAs.

The following potential cT-DNA functions have been discussed in the literature:

(1) increased root mass as an adaptation to life under arid conditions;
(2) increased resistance to agrobacterial infection;
(3) elevated regenerative capacity;
(4) passage to earlier flowering and, as a consequence, passage to annual life cycle;
(5) influence on the composition of microbial communities of the plant's rhizosphere and phyllosphere.

Based on the currently available data, the hypotheses of the effect of T-DNA on the regenerative capacity and the interaction with microorganism communities look the most probable. However, additional research is needed to clarify the true role of cT-DNA. Genome editing technologies open up new perspectives in this field, and the constantly updated list of sequenced genomes will provide new material for searching for naturally transgenic plants.

Acknowledgements This publication was supported by a grant to Tatiana Matveeva from the Russian Science Foundation 16-16-10010. The author expresses her deep gratitude to Professor Leon Otten (Institut de biologie moléculaire des plantes, Strasbourg, France) for his critical reading of the manuscript.

References

Acuna R, Padilla BE, Florez-Ramos CP, Rubio JD, Herrera JC, Benavides P et al (2012) Adaptive horizontal transfer of a bacterial gene to an invasive insect pest of coffee. Proc Natl Acad Sci U S A 109:4197–4202

Angiosperm Phylogeny Group (2016) An update of the Angiosperm Phylogeny Group classification for the orders and families of flowering plants: APG IV. Bot J Linn Soc 181:1–20

Aoki S, Syono K (1999a) Function of Ng*rol* genes in the evolution of *Nicotiana glauca*: Conservation of the function of NgORF13and NgORF14 after ancient infection by an *Agrobacterium rhizogenes*-like ancestor. Plant Cell Physiol 40:222–230

Aoki S, Syono K (1999b) Horizontal gene transfer and mutation: Ng*rol* genes in the genome of *Nicotiana glauca*. Proc Natl Acad Sci U S A 96:13229–13234

Aoki S, Kawaoka A, Sekine M, Ichikawa T, Fujita T, Shinmyo A et al (1994) Sequence of the cellular T-DNA in the untransformed genome of *Nicotiana glauca* that is homologous to ORFs 13 and 14 of the Ri plasmid and analysis of its expression in genetic tumors of *N. glauca* × *N. langsdorffii*. Mol Gen Genet 243:706–710

Blanco-Pastor JL, Vargas P, Pfeilm BE (2012) Coalescent simulations reveal hybridization and incomplete lineage sorting in Mediterranean *Linaria*. PLoS ONE 7:e39089

Bogani P, Buiatti M, Tegli S, Pellegrini MG, Bettini P, Scala A (1985) Interspecific differences in differentiation and dedifferentiation patterns in the genus *Nicotiana*. Plant Syst Evol 151:19–29

Bush L, Hempflin WP, Burton H (1999) Biosynthesis of nicotine and related compounds. In: Gorrod JW, Jacob P III (eds) Analytical determination of nicotine and related compounds and their metabolites. Amsterdam, Netherlands: Elsevier Science, pp 13–44

Chen K, Otten L (2017) Natural *Agrobacterium* transformants, recent results and some theoretical considerations. Front Plant Sci. https://doi.org/10.3389/fpls.2017.01600

Chen K, Dorlhac de Borne F, Julio E, Obszynski J, Pale P, Otten L (2016) Root-specific expression of opine genes and opine accumulation in some cultivars of the naturally occurring GMO *Nicotiana tabacum*. Plant J 87:258–269

Chen K, Dorlhac de Borne F, Szegedi E, Otten L (2014) Deep sequencing of the ancestral tobacco species *Nicotiana tomentosiformis* reveals multiple T-DNA inserts and a complex evolutionary history of natural transformation in the genus Nicotiana. Plant J 80:669–682

Clarkson JJ, Yoong Lim K, Kovarik A, Chase MW, Knapp S, Leich AR (2005) Long-term genome diploidization in allopolyploid *Nicotiana* section Repandae (Solanaceae). New Phytol 168:241–252

Dessai AP, Gosukonda RM, Blay E, Dumenyo CK, Medina-Bolivar F, Prakash CS (1995) Plant regeneration of sweetpotato (*Ipomoea batatas* L.) from leaf explants *in vitro* using a two-stage protocol. Sci Hortic 62:217–224

Dong F, WilsonKG, Makaroff CA (1998) Analysis of the four *cox2* genes found in turnip (*Brassica campestris* Brassicaceae) mitochondria. Am J Bot 85:153–161

Draper J, Scott R, Armitage P, Walden R (eds) (1988) Plant genetic engineering laboratory manual. Blackwell Scientific Ltd, London

Escobar MA, Civerolo EL, Summerfelt KR, Dandekar AM (2001) RNAi-mediated oncogene silencing confers resistance to crown gall tumorigenesis. Proc Natl Acad Sci U S A 98: 13437–13442

Flores H, Pickard J, Hoy M (1988) Production of polyacetylenes and thiophenes in heterotrophic and photosynthetic root cultures of Asteraceae In: Lam J Breheler H Arnason T Hansen L (eds) Chemistry and biology of naturally occurring acetylenes and related compounds (NOARC) bioactive molecules, vol 7. Elsevier, Amsterdam, pp 233–254

Fründt C, Meyer AD, Ichikawa T, Meins F (1998) A tobacco homologue of the Ri-plasmid *orf13* gene causes cell proliferation in carrot root discs. Mol Gen Genet 259:559–568

Fürner IJ, Huffman GA, Amasino RM, Garfinkel DJ, Gordon MP, Nester EW (1986) An *Agrobacterium* transformation in the evolution of the genus *Nicotiana*. Nature 329:422–427

Gama MIC, Leite JRP, Cordeiro AR, Cantliffe DJ (1996) Transgenic sweet potato plants obtained by *Agrobacterium tumefaciens* mediated transformation. Plant Cell Tissue Organ Cult 46: 237–244

Gao C, Reno X, Mason AS, Liu H, Xiao M, Li J, Fu D (2014) Horizontal gene transfer in plants. Funct Integr Genomics 14:23–29

Gogarten JP, Doolittle WF, Lawrence JG (2002) Prokaryotic evolution in light of gene transfer. Mol Biol Evol 19:2226–2238

Goodspeed TH (1954) The genus nicotiana. Chronica Botanica, Waltham

Hamill JD, Parr AJ, Rhodes MJC, Robins RJ, Walton NJ (1987) New routes to plant secondary products. Nat Biotechnol 5:800–804

Hong SB, Hwang I, Dessaux Y, Guyon P, Kim KS, Farrand SK (1997) A T-DNA gene required for agropine biosynthesis by transformed plants is functionally and evolutionary related to a Ti plasmid gene required for catabolism by *Agrobacterium* strains. J Bacteriol 189:4831–4840

Huang J (2013) Horizontal gene transfer in eukaryotes: the weak-link model. BioEssays 35: 868–875

Husnik F, McCutcheon JP (2017) Functional horizontal gene transfer from bacteria to eukaryotes. Nature Rev Microbiol. https://doi.org/10.1038/nrmicro.2017.137

Ichikawa T and Syōno K (1991) Tobacco genetic tumors. Plant Cell Physiol 32:1123–1128

Ichikawa T, Ozeki Y, Syono K (1990) Evidence for the expression of the rol genes of *Nicotiana glauca* in genetic tumors of *N. glauca* × *N. langsdorfii*. Mol Gen Genet 220:177–180

Intrieri MC, Buiatti M (2001) The horizontal transfer of *Agrobacterium rhizogenes* genes and the evolution of the genus *Nicotiana*. Mol Phylogenet Evol 20:100–110

Khafizova G, Dobrynin P, Polev D, Matveeva T (2018) *Nicotiana glauca* whole-genome investigation for cT-DNA study. BMC Res Notes 11:18. https://doi.org/10.1186/s13104-018-3127-x

Kehr AE, Smith HH (1954) Genetic tumors in *Nicotiana* hybrids. Brookhaven Symp Biol 6:55–78

Koonin EV, Makarova KS, Aravind L (2001) Horizontal gene transfer in prokaryotes: quantification and classification. Annu Rev Microbiol 55:709–742

Kostoff D (1930) Tumors and other malformations on certain nicotiana hybrids. Zentral bl. f.Bakt. I.1. 81:244–260

Kovacova V, Zluvova J, Janousek B, Talianova M, Vyskot B (2014) The evolution of the horizontally transferred agrobacterial mikimopinesynthase gene in the genera *Nicotiana* and *Linaria*. PLoS ONE 9:e113872

Kumar AHG, Ku AT, Agarwal P, Mahesh Sand Yeh KW (2013) Protuberance-mediated high frequency of adventitious shoot regeneration from sweet potato leaf explants [*Ipomoea batatas* (L) Lam]. Biotechnol 94:445–450

Kyndt T, Quispe D, Zhai H, Jarret R, Ghislain M, Liu Q et al (2015) The genome of cultivated sweet potato contains *Agrobacterium* T-DNAs with expressed genes: An example of a naturally transgenic food crop. Proc Natl Acad Sci U S A 112:5844–5849

Limami MA, Sun LY, Douatm C, Helgeson J, Tepfer D (1998) Natural genetic transformation by *Agrobacterium rhizogenes*: annual flowering in two biennials belgianendive and carrot. Plant Physiol 118:543–550

Matveeva TV (2013) Horizontal gene transfer from *Agrobacterium* spp. to the higher plants. Extended abstract of doctoral (biol) dissertation. St Petersburg Gos Univ, St Petersburg

Matveeva TV, Kosachev PA (2013) Sequences homologous to *Agrobacterium rhizogenes rolC* in the genome of *Linaria acutiloba.* In: Zheng D (ed) International conference on frontiers of environment energy and bioscience (ICFEEB 2013) Lancaster, DES Tech Publ Inc, pp 541–546

Matveeva TV, Lutova LA (2014) Horizontal gene transfer from *Agrobacterium* to plants. Front Plant Sci 5:326

Matveeva TV, Bogomaz DI, Pavlova OA, Nester EW, Lutova LA (2012) Horizontal gene transfer from genus *Agrobacterium* to the plant *Linaria* in nature. Mol Plant-Microbe Interact 25:1542–1551

Matveeva TV, Bogomaz OD, Golovanova LA, Li YuS, Dimitrov D (2018) Homologs of the *rolC* gene of naturally transgenic toadflaxes *Linaria vulgaris* and *Linaria creticola* are expressed *in vitro*. Vavilov J Genet Breeding (in press)

Matveeva TV, Pavlova OA, Ivanitskii KI, Lutova LA (2009) Horizontal gene transfer from *Agrobacterium* to *Nicotiana* genus plants: evolutionary prerequisites and consequences. Vestn St. Petersburg Gos Univ Ser 3(4):58–65

Matveeva T, Sokornova S (2017) Biological traits of naturally transgenic plants and their evolutional roles. Russ J Plant Physiol (2017) 64(5):635–64

Matveeva T, Sokornova S (2016) *Agrobacterium rhizogenes* mediated transformation of plants for improvement of yields of secondary metabolites. In: Pavlov A, Bley T (ed) Reference series in phytochemistry. Bioprocessing of plant in vitro systems. Springer, Berlin, pp 1–42

Matveeva TV, Sokornova SV, Lutova LA (2015) Influence of *Agrobacterium* oncogenes on secondary metabolism of plants. Phytochem Rev 14:541–556

Mohajjel-Shoja H, Clément B, Perot J, Alioua M, Otten L (2011) Biological activity of the *Agrobacterium rhizogenes*-derived *trolC* gene of *Nicotiana tabacum* and its functional relationship to other plast genes. Mol Plant Microbe Interact 24:44–53

Naf U (1958) Studies on tumor formation in *Nicotiana hybrids* I. The classification of the parents into two etiologically significant groups. Growth 22:167–180

Oger P, Mansouri H, Dessaux Y (2000) Effect of crop rotation and soil cover on alteration of the soil microflora generated by the culture of transgenic plants producing opines. Mol Ecol 9:881–890

Oger P, Petit A, Dessaux Y (1997) Genetically engineered plants producing opines alter their biological environment. Nat Biotechnol 15:369–372

Otten L, Canaday J, Gerard JC, Fournier P, Crouzet P, Paulus F (1992) Evolution of agrobacteria and their Ti plasmids. Mol Plant Microbe Interact 5:279–287

Palazon J, Cusido RM, Roig C, Pino MT (1998) Expression of the *rolC* gene and nicotine production in transgenic roots and their regenerated plants. Plant Cell Rep 17:384–390

Pavlova OA, Matveeva TV, Lutova LA (2013) *Linaria dalmatica* genome contains a homologue of *rolC* gene of *Agrobacterium rhizogenes*. Ecol Genet 11:10–15

Quispe-Huamanquispe DG, Gheysen G, Kreuze JF (2017) Horizontal gene transfer contributes to plant evolution: the case of *Agrobacterium* T-DNAs. Front Plant Sci. https://doi.org/10.3389/fpls.2017.02015

Rhodes MJC, Robins RJ, Hamill JD, Parr AJ, Walton NJ (1987) Secondary product formation using *Agrobacterium rhizogenes*-transformed "hairy root" cultures. News Lett 53:2–15

Richards TA, Dacks JB, Campbell SA, Blanchard JL, Foster PG, McLeod R et al (2006) Evolutionary origins of the eukaryotic shikimate pathway: gene fusions horizontal gene transfer and endosymbiotic replacements. Eukaryot Cell 5:1517–1531

Richardson AO, Palmer JD (2007) Horizontal gene transfer in plants. J Exp Bot 58:1–9

Sutton DA (1988) A revision of the tribe antirrhineae. Oxford University Press, London

Suzuki K, Yamashita I, Tanaka N (2002) Tobacco plants were transformed by *Agrobacterium rhizogenes* infection during their evolution. Plant J 32:775–787

Tanaka N (2008) Horizontal gene transfer. In Tzfira T, Citovsky V (eds) *Agrobacterium*: from biology to biotechnology. Springer, New York, pp 623–647

Vain P (2007) Thirty years of plant transformation technology development. Plant Biotechnol J 5:221–229

White FF, Ghidossi G, Gordon MP, Nester EW (1982) Tumor induction by *Agrobacterium rhizogenes* involves the transfer of plasmid DNA to the plant genome. Proc Natl Acad Sci USA 79:3193–3319

White FF, Garfinkel DJ, Huffman GA, Gordon MP, Nester EW (1983) Sequence homologous to *Agrobacterium rhizogenes* T-DNA in the genomes of uninfected plants. Nature 301: 348–350

Beyond *Agrobacterium*-Mediated Transformation: Horizontal Gene Transfer from Bacteria to Eukaryotes

Benoît Lacroix and Vitaly Citovsky

Contents

1 Introduction 444
2 Pathways for Bacterium-to-Eukaryote HGT 445
2.1 Transformation 445
2.2 Transduction 446
2.3 Conjugation 446
3 Stable and Transient HGT 451
4 HGT Signatures in Eukaryotic Genomes 452
5 Bacterium-to-Eukaryote HGT as Tool for Research and Biotechnology 456
6 Conclusions 457
References 458

Abstract Besides the massive gene transfer from organelles to the nuclear genomes, which occurred during the early evolution of eukaryote lineages, the importance of horizontal gene transfer (HGT) in eukaryotes remains controversial. Yet, increasing amounts of genomic data reveal many cases of bacterium-to-eukaryote HGT that likely represent a significant force in adaptive evolution of eukaryotic species. However, DNA transfer involved in genetic transformation of plants by *Agrobacterium* species has traditionally been considered as the unique example of natural DNA transfer and integration into eukaryotic genomes. Recent discoveries indicate that the repertoire of donor bacterial species and of recipient eukaryotic hosts potentially are much wider than previously thought, including donor bacterial species, such as plant symbiotic nitrogen-fixing bacteria (e.g., *Rhizobium etli*) and animal bacterial pathogens (e.g., *Bartonella henselae*, *Helicobacter pylori*), and recipient species from virtually all eukaryotic clades. Here, we review the molecular

B. Lacroix (✉) · V. Citovsky
Department of Biochemistry and Cell Biology, State University of New York, Stony Brook, NY 11794-5215, USA
e-mail: benoit.lacroix@stonybrook.edu

Current Topics in Microbiology and Immunology (2018) 418:443–462
DOI 10.1007/82_2018_82

Published Online: 3 March 2018

pathways and potential mechanisms of these trans-kingdom HGT events and discuss their utilization in biotechnology and research.

Keywords Horizontal gene transfer · Agrobacterium · Bacterium-to-eukaryote HGT · Type IV secretion system

1 Introduction

Horizontal gene transfer (HGT) among bacterial species (Beiko et al. 2005; Gogarten and Townsend 2005; Koonin et al. 2001) represents a mechanism that dominates evolution of prokaryotic life and is essential for the survival of microbial populations (Koonin 2016). In contrast to the genetic promiscuity observed among prokaryotes, the importance of HGT from bacteria to eukaryotes in evolution is still debated. The transfer of many genes from the genomes of bacterial endosymbiont-derived organelles, i.e., mitochondria or plastids, to their eukaryotic hosts is well-documented (Archibald 2015). These episodes of massive gene transfer took place in the early evolution of eukaryotic cells and are usually referred to as endosymbiotic gene transfer (EGT) or organelle gene transfer (OGT; Huang 2013). Understanding the relative importance of continued HGT from non-organelle bacteria during the evolution of eukaryotes still requires more investigation. For example, whereas a recent study suggests that events of bacterium-to-eukaryote HGT did not lead to an accumulation of transferred genes in modern eukaryotes (Ku et al. 2015), an increasing number of reported cases of HGT argues for an important role in adaptive evolution (Fitzpatrick 2012; Husnik and McCutcheon 2018; Lacroix and Citovsky 2016; Schönknecht et al. 2014; Sieber et al. 2017). In the early stages of eukaryote evolution, the pool of genes of prokaryotes was likely much larger than that of eukaryotes. Thus, acquisition of genes from organelles might reflect the role of endosymbiosis as a way to acquire new genes (Fournier et al. 2009). Besides the HGT events revealed by analyses of genomic sequences of eukaryotes, DNA transfer from *Agrobacterium* spp. to their host plant cells represents a rare example of HGT frequently occurring in the natural world at present day. Accumulated knowledge of the mechanisms involved in this naturally occurring bacterium-to-eukaryote HGT is essential to understand general molecular pathways of bacterial DNA transfer and integration into the eukaryotic host cell genome. In this chapter, we will review the pathways and the genomic signatures of trans-kingdom HGT, with a focus on its likely mechanisms and their utilization in biotechnology and research.

2 Pathways for Bacterium-to-Eukaryote HGT

Three different pathways for lateral DNA exchange between bacteria have been identified and well characterized: transformation, transduction, and conjugation (Arber 2014; Johnsborg et al. 2007). Transformation relies on the uptake of free DNA segments present in the environment. Transduction, mediated by bacteriophages, may occur as restricted transduction (i.e., only the bacteriophage sequence and small adjacent sequences are transferred) or as generalized transduction (i.e., larger sequences from the bacterial genome are co-transferred with the phage). Conjugation, or plasmid-mediated transfer, usually requires close cell-to-cell contact between the donor and recipient cells and relies on the activity of complex secretion machinery. Theoretically, bacterium-to-eukaryote HGT may occur via any of these three pathways. In eukaryotic cells, however, the nuclear envelope, which acts as a physical barrier to invading nucleic acid molecules, makes the acquisition of foreign DNA more complex than in prokaryotic cells. Indeed, once the transferred DNA segment enters the cell cytoplasm, it must be transported into the nucleus and then find its way through the highly structured and packaged host chromatin toward a potential site of integration. It is unlikely that a macromolecule the size of a natural DNA segment encoding protein functions could diffuse freely in the crowded environment of a eukaryotic cell cytoplasm. Moreover, DNA molecules usually are too large to traffic passively through the nuclear pore complex (NPC); the molecular size for such diffusion is thought to be about 9 nm (Forbes 1992), and the passage of larger molecules requires active nuclear import. Thus, nuclear import of foreign DNA most likely depends on the host import machinery. Similarly, after the entry of the foreign DNA into the nucleus, the incoming DNA molecule must be guided to a potential site of integration in the host chromatin, unless it functions as a self-replicating element, such as a plasmid. In some instances, gene delivery to the nucleus may follow the breakdown of the nuclear membrane during mitosis, but this strategy is not effective in quiescent, non-dividing cells.

2.1 Transformation

Natural transformation, i.e., the uptake of free DNA segments, is observed in many different bacterial species (Johnsborg et al. 2007). In eukaryotic cells, several laboratory techniques are available to introduce free DNA into the cells via mechanical or chemical treatments. Such transformation protocols may lead to transient expression of the transgenes and to generation of stably transformed cells/organisms, thus demonstrating the possibility of acquisition and integration of free DNA segments by eukaryotic cells. However, it is generally considered that this pathway of DNA transfer to eukaryotic cells does not occur naturally. Although recent studies suggest that natural competence for exogenous DNA uptake does exist in yeast (Mitrikeski 2013), additional research is needed to assess the extent of this DNA acquisition pathway.

2.2 Transduction

Another possible pathway of DNA transfer from bacteria to eukaryotes is transfer via DNA replicating elements such as transposons, phages, and viruses. Exchange of genetic information via bacteriophages is well-documented between different bacterial species that are closely related as well as evolutionarily distant. Whereas phages or viruses able to cross-kingdom boundaries have not been identified so far, the analysis of sequences of bacterial origin in eukaryotic genomes has shown potential signatures of bacteriophage-mediated DNA transfer, for example, in the case of HGT from the intracellular pathogen *Wolbachia* spp. to its host *Aedes aegypti* (Klasson et al. 2009). Specifically, it has been suggested that giant viruses might mediate the transfer of genetic information from bacteria to eukaryotic cells (Schönknecht et al. 2013).

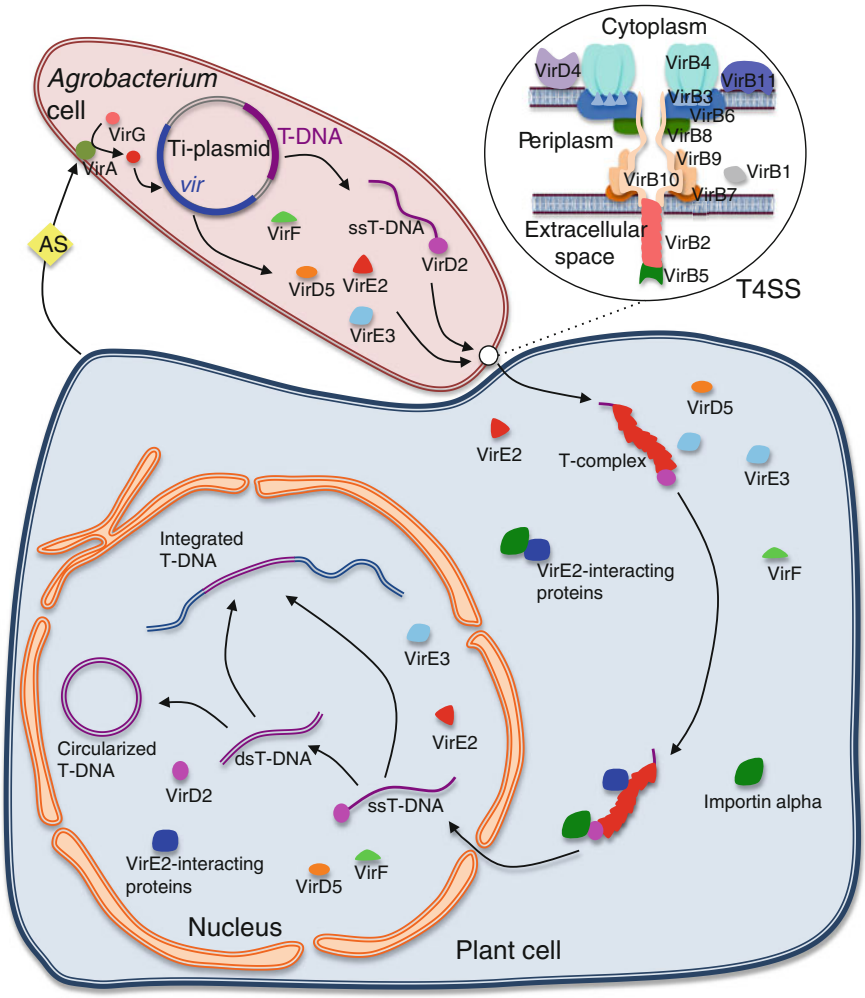
2.3 Conjugation

To date, trans-kingdom conjugation via a type IV secretion system (T4SS) remains the only demonstrated pathway of DNA transfer from bacteria to eukaryotic cells (Table 1). T4SSs are molecular machines specialized in the transport of macromolecules, proteins, and DNA, between bacteria and from bacteria to a variety of eukaryotic hosts. T4SSs are widespread among eubacteria (Alvarez-Martinez and Christie 2009; Maindola et al. 2014); originally, they were described in gram-negative bacterial species, although similar systems also exist in gram-positive bacteria (Goessweiner-Mohr et al. 2013). Generally, bacterial T4SSs are involved in the intercellular transport of macromolecules that fulfill many different functions in the host cell. Conjugation allows the exchange of genetic information between bacterial cells from the same or a related species via the transfer of plasmid or conjugative transposon DNA in the form of a nucleoprotein. Besides bacterial conjugation, T4SSs mediate the transport of macromolecules, proteins, or nucleoprotein complexes, from bacterial to eukaryotic cells. In many cases, T4SS-mediated transport plays a crucial role in the natural, pathogenic, or symbiotic interactions between bacterial cells and their eukaryotic hosts. For instance, effector proteins are transported from several animal prokaryotic pathogens to the host cells (Backert and Meyer 2006). This transport is exemplified by the facultative intracellular pathogen *Brucella* spp., which translocates several proteins that interfere with host functions, e.g., apoptosis inhibition or F-actin modulation, and facilitate infection (Siamer and Dehio 2015). Similarly, two human pathogens, *Helicobacter pylori* (Backert and Selbach 2008) and *Legionella pneumonia* (Hubber and Roy 2010), rely on their T4SS to translocate effector proteins into their hosts. Symbiotic plant-associated *Mesorhizobium loti* might also transfer proteins to host cells in a T4SS-dependent manner (Hubber et al. 2004). Finally, several species of the *Agrobacterium* genus naturally transfer DNA and proteins to many different species of plant hosts,

Table 1 DNA transfer from bacteria to eukaryotes in natural and artificial systems

Bacterial donor	Eukaryotic recipient	Reference
Agrobacterium tumefaciens	Most plant species	Lacroix et al. (2006)
	Saccharomyces cerevisiae	Bundock et al. (1995), Piers et al. (1996)
	Penicillium	de Groot et al. (1998)
	Other fungi	Lacroix et al. (2006), Soltani et al. (2008)
	Insect cells	Machado-Ferreira et al. (2015)
	Sea urchin	Bulgakov et al. (2006)
	Human cells in culture	Kunik et al. (2001)
Rhizobium trifolii Phyllobacterium myrsinacearum	Tobacco	Hooykaas et al. (1977)
Rhizobium sp. NGR234 *Sinorhizobium meliloti Rhizobium loti*	Tobacco	Broothaerts et al. (2005), Wendt et al. (2011)
Ensifer adhaerens	Tobacco, rice	Wendt et al. (2012), Zuniga-Soto et al. (2015)
Rhizobium etli	Tobacco	Lacroix and Citovsky (2016), Wang et al. (2017)
Escherichia coli	*Saccharomyces*	Heinemann and Sprague (1989)
	Other fungi	Hayman and Bolen (1993), Inomata et al. (1994)
	Human cells in culture	Waters (2001)
Bartonella henselae	Endothelial human cells in culture	Fernández-González et al. (2011), Schröder et al. (2011)
Helicobacter pylori	Human cells in culture	Varga et al. (2016)

resulting in the genetic transformation of the host cell. In nature, most potential *Agrobacterium* host species belong to several families of dicotyledonous plants (Gelvin 2010; Lacroix and Citovsky 2013), and the transferred DNA leads to uncontrolled cell division (tumors), and production of opines (small molecules used as carbon and nitrogen sources by *Agrobacterium* cells; Escobar and Dandekar 2003). *Agrobacterium*-mediated T-DNA transfer represents the most extensively studied example of HGT; a current model of its molecular mechanism is described in Fig. 1. Under laboratory conditions, using various plant tissue culture techniques and exogenously added enhancers of *Agrobacterium* virulence, virtually all plant species, including monocotyledonous plants, are amenable to genetic transformation by *Agrobacterium*, albeit often with low efficiency. Moreover, DNA transfer from *Agrobacterium* was also demonstrated toward non-plant species. Indeed, yeast (Bundock et al. 1995; Piers et al. 1996), other fungi (Bundock et al. 1999; de Groot et al. 1998; Gouka et al. 1999), sea urchin (Bulgakov et al. 2006), and cultured animal cells (Kunik et al. 2001; Machado-Ferreira et al. 2015) were shown to be



recipients of *Agrobacterium*-mediated DNA transfer under laboratory conditions (Lacroix et al. 2006), with variable efficiencies.

On the bacterial side, the repertoire of potential donor species for DNA transfer to eukaryotic cells can also be expanded beyond *Agrobacterium* spp. Indeed, several species belonging to the Rhizobiales order (which includes *Agrobacterium* spp. and many species involved in symbiotic interactions with plants that result in fixation of nitrogen) can mediate DNA transfer to plant cells when they are provided with plasmid(s) containing the machinery for DNA transfer (virulence region) and the sequence to be transferred (T-DNA) from a virulent species of *Agrobacterium*. For example, *Rhizobium trifolii* became virulent, i.e., able to induce tumors on several plant species, after conjugative transfer of the Ti plasmid from

◂**Fig. 1** Summary of the molecular mechanism of T-DNA transfer from *Agrobacterium* to plant cells. More details and most references may be found in several extensive reviews (Gelvin 2003; Lacroix and Citovsky 2013). Upon wounding of plant tissues, plant-produced small phenolic signals (e.g., acetosyringone, AS) activate the VirA sensor, which in turn activates the VirG transcriptional inducer by phosphorylation. Phosphorylated VirG recognizes regulatory elements (*vir* boxes) in different *vir* gene promoters on the Ti (tumor-inducing) plasmid and induces their expression. The Vir proteins then initiate the transfer of DNA. VirD2 and VirD1 generate the single-stranded (ss) T-DNA (corresponding to the sequence element between the two 25-bp left and right T-DNA borders) from the Ti plasmid via a strand-replacement mechanism, and the VirD2 endonuclease remains covalently attached to the 5' end of the T-DNA molecule. T4SS composed of the protein products of the *vir*B operon and of VirD4 assembles at the membrane and mediates the export of the nucleoprotein complex VirD2-T-DNA as well as *vir* gene-encoded effector proteins VirD5, VirE2, VirE3, and VirF out of the bacterial cell and into the host cell cytoplasm. Noteworthy, the mechanism of passage of these macromolecules through the host cell barriers is not understood. In the host cell cytoplasm, molecules of the ssDNA-binding protein VirE2 likely associate with the ssT-DNA, forming the T-complex. Intracellular transport and nuclear import of the T-complex rely on its interactions with bacterial effectors and host factors such as cytoskeletal elements, VirE2-interacting proteins, and importin alpha proteins. Once the T-DNA enters the cell nucleus, it is uncoated from its associated proteins, most likely with the help of the bacterial F-box protein effector VirF and the host ubiquitin/proteasome system. Following uncoating, several scenarios are possible. The T-DNA may ligate into a genomic double-stranded DNA break (DSB) as a single-stranded molecule presumably via the action of a DNA polymerase theta-like enzyme (van Kregten et al. 2016), or be converted into a double-stranded form before integration into DSBs via one of the host's DSB repair pathways. In some hosts, such as yeast cells, double-stranded T-DNA may also circularize and form a plasmid able to replicate

Agrobacterium (Hooykaas et al. 1977). Then, several other *Rhizobiaceae* species (i.e., *R. leguminosarum*, *R. trifolii*, and *Phyllobacterium myrsinacearum*) displayed the ability to transfer T-DNA to *Arabidopsis*, tobacco, and rice after being transformed with two plasmids: a helper plasmid that carries the virulence (*vir*) region and a binary plasmid that carries the T-DNA (Broothaerts et al. 2005). Introduction of a similar set of plasmids into *Sinorhizobium meliloti*, *Rhizobium* sp. strain NGR234, and *Mesorhizobium loti* conferred onto these bacteria the ability to genetically transform potato plants (Wendt et al. 2011). More recently, *Ensifer adhaerens* (also known as *Sinorhizobium adhearens*) was used to transfer DNA to potato and rice plants after introduction of a plasmid containing both *vir* and T-DNA regions of *Agrobacterium* (Wendt et al. 2012; Zuniga-Soto et al. 2015).

Lacroix and Citovsky (2016) recently reported that *R. etli* harbors a plasmid with a *vir* region highly similar to that encoded by *Agrobacterium* and that this bacterium was able to mediate plant genetic transformation when a plasmid carrying a T-DNA sequence was provided to the bacterial cell. Unlike the examples described above, where the *vir* region had to be supplied to the different Rhizobiales strains to render them virulent, *R. etli* harbors a complete and functional plant genetic transformation machinery in its p42a plasmid. Moreover, the transcriptional regulation of *vir* genes in *R. etli* overall is similar to that in *Agrobacterium* (Wang et al. 2017), with a few minor differences. For example, the *R. etli virB2* gene, located outside of the *vir*B operon, showed almost no activation upon acetosyringone treatment, which may partially explain the lower level of virulence of *R. etli* compared to *A. tumefaciens*.

Among several other Rhizobiales strains that possess genes sharing homology with the *Agrobacterium vir* genes, none contain close homologs of all the *vir* genes essential for T-DNA transfer (i.e., *virA*, *virB1* to *virB1*1, *virC*, *virD1*, *virD2*, *virD*4, *virE1*, *vir*E2, and *virG*). However, the presence of *vir* gene homologs in many bacterial species might represent the remnants of more widely spread systems of DNA transfer. Although no sequences similar to *Agrobacterium* T-DNA were detected in the *R. etli* genome, it is clear that HGT to plant cells can be mediated by a non-*Agrobacterium* species carrying its own virulence system.

In addition, *E. coli* was reported to mediate the transfer of plasmid DNA to eukaryotic host cells via a T4SS-dependent mechanism. Initially, plasmid transfer to *Saccharomyces cerevisiae* was demonstrated via a mechanism that shares essential features with bacterial conjugation (Heinemann and Sprague 1989): the requirement for close cell-to-cell contact and for the *mob* and *ori*T functions of the conjugative DNA transfer. Other yeast species, such as *Kluyveromyces lactis* and *Pichia angusta* or *S. kluyveri*, can be recipients of DNA transfer from *E. coli* (Hayman and Bolen 1993; Inomata et al. 1994). Furthermore, *E. coli* was able to transfer DNA to other types of eukaryotic cells, such as human cells in culture (Waters 2001) and diatoms, a type of unicellular algae (Karas et al. 2015). Another bacterial species, *Bartonella henselae*, could mediate the transfer of plasmid DNA to endothelial human cells in vitro via a mechanism related to conjugation (Fernández-González et al. 2011; Schröder et al. 2011). *B. henselae* is a facultative intracellular human pathogen, known to transfer effector proteins into its host cells via a T4SS-dependent mechanism during the infection process (Siamer and Dehio 2015). In these two studies (Fernández-González et al. 2011; Schröder et al. 2011), different plasmids, a modified cryptic plasmid or R388 plasmid derivatives, were transferred from *B. henselae* to their host cells, resulting in stable transgenic human cell lines, which signifies that the transferred DNA was integrated in the host cell genome. A functional T4SS was essential to the transfer, as *B. henselae* strains mutated in the *virB* region that encodes T4SS were unable to mediate DNA transfer. Unlike genetic transformation by *Agrobacterium* and *R. etli*, host cell division was required for *B. henselae* transgene expression, suggesting that transport of transferred DNA from *B. henselae* into the host cell nucleus relies on the disruption of the host nuclear envelope rather than on the host nuclear import machinery. Recently, it was shown that another human bacterial pathogen, *Helicobacter pylori*, known as a risk factor for gastric cancer, may translocate to its host cells not only its effector proteins but also DNA molecules that are recognized by an intracellular receptor TLR9 (Varga et al. 2016). The nature of the transported *H. pylori* DNA and its fate in the host cell, however, remain unknown. Another indication of bacterium-to-human cell DNA transfer comes from analyses of the genome sequences of human tumors, in which integrated bacterial DNA was detected (Riley et al. 2013; Sieber et al. 2016). These analyses could not be performed in non-tumor cells because it is next to impossible to detect HGT reliably in a single cell genome, whereas tumors represent clones of a single cell, allowing such studies.

The data summarized here indicate that HGT mediated by a conjugation-like mechanism is not restricted to *Agrobacterium* spp. and their host plants, and it

likely occurs in other instances. Interestingly, however, except for *Agrobacterium*, no demonstrated roles for DNA transfer in infection, symbiosis, or cancer induction have been identified so far for the eukaryote-associated bacterial species that may transfer DNA to their host cells.

3 Stable and Transient HGT

Regardless of the molecular mechanisms of HGT, one approach to assess its importance and relevance for evolution is to investigate the presence of sequences from bacterial origin in eukaryotic genomes (see next section). It is important to mention, however, that formation of these signatures of the bacterium-to-eukaryote HGT requires several additional molecular reactions following the initial transport of DNA from the bacterial cell to the host cell cytoplasm. Indeed, fixation of bacterial sequences in a eukaryotic genome and their vertical transmission do not depend merely on the ability of the bacterial donor species to transfer a DNA segment to a eukaryotic cell. The accumulated knowledge of the molecular pathways of *Agrobacterium*-mediated genetic transformation (Gelvin 2003, 2017; Lacroix and Citovsky 2013) provides us with precious clues on the steps required for expression of the transferred DNA and its integration into the host genome, which represent two distinct processes. In fact, DNA may be imported into the host nucleus, and the genes it carries may be expressed without actual integration in the genomic DNA. This phenomenon is well known during plant genetic transformation and is termed transient expression of the transgene. Only a fraction of transiently expressed DNA undergoes integration, and this DNA mediates what is termed stable expression of the transgene. Integration of the transferred DNA is necessary, but not sufficient for production of a transgenic organism. First, even if the transferred DNA was integrated into the genome of the target cell, it still must be conserved through rearrangements of the genome during subsequent cell divisions. Second, the transformed cell must be of a type that will regenerate into a fully functional and fertile organism. In the case of unicellular eukaryotes and most fungi, for which at least a part of the cell cycle is in the unicellular form, genes acquired by cells in this unicellular stage may be transmitted to the progeny by simple cell division. For plants, cells from many tissues can dedifferentiate and regenerate into a functional organism under appropriate conditions; a transformed cell may thus regenerate into an organism carrying the transforming DNA, i.e., a genetically modified organism. For most animals, the transgene may be transmitted to the progeny only if the target cell is a germline cell, with the exception of sessile organisms that do not possess dedicated germline cells, e.g., some marine organisms (Degnan 2014). Finally, the transgene must be transmitted through generations and conserved during evolution of the surviving transgenic line. Obviously, the conservation of a sequence of bacterial origin is more likely if it confers a selective advantage or is at least neutral for the recipient, or confers an advantage to the DNA segment itself in the case of a selfish DNA element. The consequence of these

multiple steps is that the HGT event signatures identified by analyzing genome sequences likely resulted from a much higher frequency of initial gene transfer from bacteria to eukaryotes that did not lead to integration and fixation of the transferred sequences in the progeny of the original recipient. The possibility of gene transfer and transient expression without integration in the host genome also raises the question of a potential role of this phenomenon as a strategy to express bacterium-encoded effector proteins in the host cell, providing an alternative to the direct introduction of the effector proteins via T3SS or T4SS pathways and further facilitating infection. Indeed, in the case of *Agrobacterium*, expression of T-DNA genes very early after inoculation, presumably before integration, leads to changes in the host cell that facilitate subsequent development of a tumor (Lee et al. 2009). In this scenario, the transferred genes, similarly to the *Agrobacterium* T-DNA genes, would harbor regulatory sequences compatible with expression in eukaryotic cells.

4 HGT Signatures in Eukaryotic Genomes

Based on the examples of DNA transfer in natural and experimental systems described above, it appears that HGT is possible from many bacterial species to cells of organisms belonging to virtually all clades of eukaryotes. Indeed, genomic signatures of bacterium-to-eukaryote HGT have been found in many organisms (Table 2). Based on these examples, it is possible to outline general conditions required for the bacterium-to-eukaryote gene transfer. Most important is the close interaction between the bacterial donor and eukaryotic recipient cells: at the minimum a shared habitat, and sometimes cell-surface attachment or even intracellular location of the donor bacteria. The close association between donor and recipient cells must be concomitant with the HGT occurrence. Thus, when considering the signature of ancient HGT, one must also consider the interactions that may have taken place between ancestors of the involved species. Interestingly, among the donor bacterial species at the origin of HGT into eukaryotes, a higher proportion of proteobacteria and cyanobacteria is found (Le et al. 2014); consistently, the most common endosymbionts and closely eukaryote-associated bacteria belong to these families (Huang 2013).

Usually, the first hint for detecting an HGT event is the unexpected phyletic distribution of a sequence that differs from the known species phylogenetic tree. Indeed, a topological discrepancy between the known species tree and the predicted gene tree represents a good indication that a sequence was acquired via HGT (Fitzpatrick 2012; Koonin et al. 2001; Syvanen 2012). Nevertheless, it cannot be excluded that the presence of a sequence only in one species of a clade is due to differential gene loss in all other species. Thus, to confirm that the presence of an unexpected sequence is due to HGT, several secondary clues must be taken into account. Such auxiliary evidence includes base composition, codon usage, presence or absence of introns, synteny analysis, and a certain level of ecological association

Table 2 Genomic signatures of bacterium-to-eukaryote HGT

Bacterial donor	Eukaryotic recipient	Transferred sequences	Reference
Unknown	*Colletotrichum* spp.	11 genes involved in interaction with plants	Jaramillo et al. (2015)
Proteobacteria	*Verticillium* spp.	Glucan glycosyltransferase	Klosterman et al. (2011)
Thermoacidophilic bacteria	*Galdieria sulphuraria*	Several genes facilitating ecological adaptation	Schönknecht et al. (2013)
Unknown	*Dictyostelium discoideum*	18 genes, some encoding functional proteins	Eichinger et al. (2005)
Unknown	*Blastocystis* spp.	Up to 2.5% of the genes	Eme et al. (2017)
Agrobacterium rhizogenes	*Nicotiana glauca*	T-DNA genes	Aoki et al. (1994, White et al. (1983)
	Nicotiana spp.	T-DNA genes	Furner et al. (1986), Intrieri and Buiatti (2001)
	Linaria spp.	T-DNA genes	Matveeva et al. (2012)
	Ipomea batatas	T-DNA genes	Kyndt et al. (2015)
Unknown	*Arabidopsis thaliana*	Genes for xylem formation, defense, growth regulation	Yue et al. (2012)
Actinobacteria	Land plants	Transaldolase	Yang et al. (2015)
Unknown	Land plants	Auxin biosynthesis pathway	Yue et al. (2014)
Unknown	Most land plants	Glycerol transporter	Zardoya et al. (2002)
Alpha-proteobacteria	Most land plants	Gamma-glutamylcysteine ligase, glutathione synthesis	Copley and Dhillon (2002)
Unknown	Most land plants	DNA-3-methyladenine glycosylase (DNA repair)	Fang et al. (2017)
Unknown	*Hydra magnipapillata*	Unknown	Chapman et al. (2010)
Unknown	Bdelloid rotifers	Unknown	Gladyshev et al. (2008)
Wolbachia spp.	Arthropods (8 species)	Large sequences (up to 30% of the bacterial genome)	Dunning Hotopp et al. (2007)
Rhizobiales	Plant-parasitic nematodes	Multiple genes (e.g., invertase)	Danchin et al. (2016)
Bacillus spp.	*Hypothenemus hampei*	Mannanase (HhMAN1)	Chapman et al. (2010)

between the species involved, e.g., shared niche or habitat or closer physical association. Claims of HGT events from bacteria to eukaryotes must be examined carefully. Indeed, several publications that reported large numbers of genes acquired from bacterial donors via HGT have been later contested because of insufficiently stringent criteria for identification of HGT-derived genes or because of possible contaminations of the sequenced DNA. For example, the human genome was reported to contain many HGT-derived genes (Lander et al. 2001), but subsequent phylogenetic studies including a larger number of eukaryotic species contested this initial finding (Crisp et al. 2015; Stanhope et al. 2001). Similarly, the discovery of a large fraction of HGT-acquired genes in a tardigrade *Hypsibius dujardini* genome (Boothby et al. 2015) was likely the result of contamination by bacterial DNA (Koutsovoulos et al. 2016). However, the number of HGT signatures in eukaryotic genomes may also be underestimated for different reasons, such as an extinct donor bacterium or a large number of mutations in the transferred gene, which may prevent the identification of the bacterial origin of these genes (Degnan 2014; Sieber et al. 2017).

The early evolution of eukaryote lineages was marked by events of vast transfer of genetic information from intracellular organelles of bacterial origin, i.e., plastids and mitochondria, to their host cells. Indeed, acquisition of genes from these organelles by the nuclear genome represents major and most evolutionary significant transfer of genetic information from bacterial to eukaryotic genomes (Archibald 2015). The mechanism of these HGT events is unknown; however, gene transfer from chloroplasts to nuclear genome was measured experimentally at significant rates (Huang et al. 2003; Stegemann et al. 2003). Note that, in some cases, genes had been acquired by eukaryotic cells from different bacterial sources to compensate for the loss of the corresponding functions in their organelles (e.g., Nowack et al. 2016), which has been described as maintenance HGT (Husnik and McCutcheon 2018). Besides the gene transfer that originated from permanent organelles, the acquisition of bacterial genes was demonstrated in several instances (see Table 2) and played an important role in the adaptive evolution of eukaryotes.

Eukaryotic species displaying a predominant unicellular stage in their life cycle, which include unicellular eukaryotes and most fungi, represent a favored target for HGT from bacteria because they do not need to dedifferentiate and regenerate as do multicellular organisms. The acquisition of genes from bacteria may have played an important role in evolution of many fungal species, particularly in the rhizosphere, where plant-associated bacteria and fungi live in close proximity to each other (Gardiner et al. 2013). Indeed, the adaptive evolution of plant-associated fungi is likely to be facilitated by the acquisition of bacterial genes, many of which encode factors involved in pathogenicity, niche specification, and adaptation to different metabolic requirement, via HGT (Fitzpatrick 2012). For example, sequencing of the genomes of three species of *Colletotrichum*, plant-pathogenic fungi responsible for a crop-destructive anthracnose disease, revealed at least 11 independent events of HGT from bacteria (Jaramillo et al. 2015). These transferred genes likely are important for niche adaptation because most of them encode proteins involved in interactions with the host plant or virulence. In two species of the *Verticillum* genus,

the causative agent of vascular wilt in more than two hundred plant species, a gene encoding a glucan glycosyltransferase and playing a role in virulence via the synthesis of extracellular glucans, was also acquired by HGT from proteobacteria (Klosterman et al. 2011).

Genome sequencing of several unicellular eukaryotes also revealed genes from bacterial origin. For example, genes obtained from bacteria and archaea were found in the genome of *Galdieria sulphuraria*, a red alga living in hot, acidic, and heavy-metal-rich extreme environment. Taking into account duplication and diversification, these genes of suspected bacterial origin may represent up to 5% of all *G. sulphuraria* protein-encoding genes and encode proteins that may have helped the evolution of this species toward adaptation to extreme environments. Among these proteins is an arsenic membrane protein pump similar to those found in thermoacidophilic bacteria (Schönknecht et al. 2013). Similarly, 18 genes likely resulting from HGT from bacteria were found in the soil-living amoeba *Dictyostelium discoideum* genome. Some of these genes encoded proteins conferring a new function, such as a dipeptidase that may degrade bacterial cell walls (Eichinger et al. 2005). In genomes of several species of *Blastocystis*, parasites found notably in the human gut, up to 2.5% of the genes have been acquired through HGT mostly, i.e., 80%, from bacterial donors. Many of these genes appear to be functional and important for adaptation to the environment (Eme et al. 2017).

The first series of HGT events in plants was discovered in the genomes of several species as a result of T-DNA transfer from a donor bacterial species related to *Agrobacterium rhizogenes*. Based on the known biology of *Agrobacterium*-plant interactions, it is likely that the first step after T-DNA integration was the regeneration of a functional organism from dedifferentiated transformed cells induced by T-DNA gene expression, followed by vertical transmission via sexual reproduction (Matveeva and Lutova 2014). The presence of T-DNA genes acquired via HGT was first discovered in *Nicotiana glauca* (Aoki et al. 1994; White et al. 1983), and further studies showed that it could detected in many species of the *Nicotiana* genus (Furner et al. 1986; Intrieri and Buiatti 2001). Analysis of more than 100 dicotyledonous plant species detected the presence of T-DNA in the genomes of *Linaria vulgaris* and *Linaria dalmatica* (Matveeva et al. 2012; Matveeva and Lutova 2014). Recently, genome sequencing of several varieties of cultivated sweet potato *Ipomea batatas* revealed the presence of T-DNA sequences derived from HGT (Kyndt et al. 2015). Unlike most of the HGT events described in this section, signatures of *Agrobacterium*-to-plant HGT correspond to the rare case for which the transfer pathway is known and the source of transferred genes is clearly identified. Indeed, it was determined that all T-DNA sequences found in *Nicotiana* and *Linaria* originated from a mikimopine strain of *A. rhizogenes*, whereas in the case of *I. batatas* the donor bacterium was an ancestral form of *A. rhizogenes*. It is still unknown whether these T-DNA genes, which have been preserved during evolution and some of which are expressed at detectable levels, play a role in the plant biology. However, a recent analysis suggested that T-DNA genes acquired by HGT likely play a role in the recipient plant evolution, for example, by affecting root development (Quispe-Huamanquispe et al. 2017). A comparative genomic study of

the early land plant *Physcomitrella patens* and of *Arabidopsis thaliana* suggested that several families of nuclear genes potentially originated from bacterium-to-plant HGT events (Yue et al. 2012). Because some of these genes were specific for land plant activities, such as growth regulation and xylem formation, these HGT events were suggested to be important for transition from aquatic to terrestrial environments. For example, the genes encoding a transaldolase enzyme found in many land plant species may have derived from an ancient event of HGT from an *Actinobacterium* to an ancestor of land plants (Yang et al. 2015). In addition, genes involved in several other pathways may have been acquired by HGT from bacteria. These include essential genes for auxin biosynthesis (Yue et al. 2014), a glycerol transporter (Zardoya et al. 2002), a gamma-glutamylcysteine ligase that catalyzes glutathione synthesis (Copley and Dhillon 2002), and a DNA-3-methyladenine glycosylase involved in base excision repair (Fang et al. 2017).

HGT from bacteria to animals has been confirmed in a limited number of cases, mostly invertebrates. These HGT events often originate from DNA transfer from unknown bacteria to asexual animals, i.e., sessile organisms able to regenerate into a functional organism by asexual reproduction, or from DNA transfer from endosymbiotic bacteria to their host germline cells (Dunning Hotopp 2011). Indeed, several freshwater asexual animals were recipients for HGT from bacteria, as in the case of the freshwater cnidarian *Hydra magnipapillata* (Chapman et al. 2010) and bdelloid rotifers (Gladyshev et al. 2008). *Wolbachia*, identified as donor of heritable HGT to their host arthropods or nematodes, are maternally inherited endosymbiotic bacteria transmitted through the egg cytoplasm. Eight out of 11 completely sequenced genomes of arthropods display *Wolbachia* sequences acquired via HGT, which may represent up to 30% of the recipient genome in some cases (Dunning Hotopp et al. 2007). Plant-parasitic nematodes have acquired several genes by HGT from bacteria, including a gene encoding an invertase, which is important for metabolism of host plant carbohydrates and likely originated from bacteria of the Rhizobiales order (Danchin et al. 2016). A major pest of coffee plants, *Hypothenemus hampei*, the coffee berry borer beetle, harbors a gene encoding a mannanase HhMAN1 protein that hydrolyzes the major coffee storage polysaccharide galactomannan and that most likely originated from a *Bacillus* gene (Acuña et al. 2012).

5 Bacterium-to-Eukaryote HGT as Tool for Research and Biotechnology

In plant research and biotechnology, *Agrobacterium* represents the major vector used for gene transfer. Since the first successful transformation and regeneration of tobacco transgenic plants in the early 1980s (Horsch et al. 1984), *Agrobacterium* has been used to genetically transform virtually all plant species for generation of transgenic organisms (Banta and Montenegro 2008). Whereas most plant species

are susceptible to *Agrobacterium*-mediated genetic transformation (Lacroix et al. 2006), the efficiency of transgenic plant generation remains low for many of them. *Agrobacterium*-mediated DNA transfer is also used as an inexpensive and convenient method for transient gene expression in plant tissues allowing, for example, rapid evaluation of protein subcellular localization or promoter regulation (Krenek et al. 2015). Both transient expression and stable transformation are widely used in plant research for studies of gene function. The random nature of T-DNA insertion into the host plant genome has allowed generation of numerous collections of insertional mutants of *Arabidopsis thaliana* plants (Alonso et al. 2003), which represented a transformative advance in plant genomics research. New molecular tools allowing targeted genomic modifications, such as the CRISPR/Cas9 system, were also introduced into plants via *Agrobacterium*-mediated transformation (Char et al. 2017). The many cases of bacterium-to-eukaryote HGT described in this chapter show that, at least under experimental conditions, the repertoire of bacterial donor cells and of eukaryote recipient cells is much wider than the mere cases of plant genetic transformation mediated by *Agrobacterium* spp. Since the discovery of genetic transformation of yeast and other fungi by *Agrobacterium*, *Agrobacterium* has also become the major tool for genetic modification and insertional mutagenesis of these organisms (Frandsen 2011). Potentially, protocols based on several other systems of gene transfer from bacterial to eukaryotic cells will be developed in the future. For example, the ability of other bacterial species, such as *R. etli*, to transform plants could be exploited with host species that are recalcitrant to *Agrobacterium* (Broothaerts et al. 2005; Lacroix and Citovsky 2016; Wendt et al. 2012). Moreover, it has been suggested that *E. coli* could be used as tool to transform yeast or other fungal cells (Moriguchi et al. 2016), whereas *B. henselae* could be used as tool to transform human cells (Llosa et al. 2012). As a potential application, *B. henselae*—or similar bacterial species such as *H. pylori*, engineered to transfer DNA to human cells—could be used as a vector for gene therapy (Elmer et al. 2013; Walker et al. 2017), which might be useful when other types of vectors are not usable or to avoid viral vectors that may inherently represent health risks.

6 Conclusions

HGT is known to occur from *Agrobacterium* spp. to their host plant cells during infection of plants by *Agrobacterium*, resulting in plant diseases such as tumors (termed crown galls) and hairy roots. Moreover, it is also known that sequences of bacterial origin and resulting from HGT are found in many eukaryotic genomes. These sequences most likely were derived from endosymbionts that produced permanent organelles as well as from non-organelle bacteria, often resulting in the acquisition of genes important for adaptive evolution. Recent studies have also shown that DNA transfer from various bacterial species to different eukaryotic cells may be performed under laboratory conditions. Consequently, it is likely that

bacterium-to-eukaryote HGT occurs or has occurred among a wide variety of combinations of donor bacterial species and recipient eukaryotic species. Several compelling reasons exist for continuing research of these gene transfer systems: understanding the potential evolutionary and ecological significance of HGT, deciphering the pathways of transport and integration of the incoming bacterial DNA, and developing new tools for the use of HGT in fundamental and applied research.

Extensive studies of DNA transfer from *Agrobacterium* to its host cells have provided many invaluable insights in the mechanisms by which a segment of bacterial DNA can be transported into a eukaryotic host cell nucleus and integrated in its genome. Yet, numerous unanswered questions remain, many of which center on the largely unknown responses of the recipient cell to the donor DNA. For example, introduction of foreign DNA into a cell genome represents an aggressive act against the cell, and eukaryotic organisms may have evolved adapted responses to maintain the integrity of their genomes. Even when transferred DNA is not integrated, the products of expression of transiently transferred bacterial genes may act similarly to bacterial or viral effector proteins, also eliciting host defense responses.

Acknowledgements The work in the VC laboratory is supported by grants from USDA/NIFA, NIH, NSF, and BARD to VC.

References

Acuña R, Padilla BE, Flórez-Ramos CP et al (2012) Adaptive horizontal transfer of a bacterial gene to an invasive insect pest of coffee. Proc Natl Acad Sci U S A 109:4197–4202

Alonso JM, Stepanova AN, Leisse TJ et al (2003) Genome-wide insertional mutagenesis of *Arabidopsis thaliana*. Science 301:653–657

Alvarez-Martinez CE, Christie PJ (2009) Biological diversity of prokaryotic type IV secretion systems. Microbiol Mol Biol Rev 73:775–808

Aoki S, Kawaoka A, Sekine M et al (1994) Sequence of the cellular T-DNA in the untransformed genome of *Nicotiana glauca* that is homologous to ORFs 13 and 14 of the Ri plasmid and analysis of its expression in genetic tumours of *N. glauca* x *N. langsdorffii*. Mol Gen Genet 243:706–710

Arber W (2014) Horizontal gene transfer among bacteria and its role in biological evolution. Life 4:217–224

Archibald JM (2015) Endosymbiosis and eukaryotic cell evolution. Curr Biol 25:R911–R921

Backert S, Meyer TF (2006) Type IV secretion systems and their effectors in bacterial pathogenesis. Curr Opin Microbiol 9:207–217

Backert S, Selbach M (2008) Role of type IV secretion in *Helicobacter pylori* pathogenesis. Cell Microbiol 10:1573–1581

Banta L, Montenegro M (2008) *Agrobacterium* and plant biotechnology. In: Tzfira T, Citovsky V (eds) *Agrobacterium*: from biology to biotechnology. Springer, New York, pp 72–147

Beiko RG, Harlow TJ, Ragan MA (2005) Highways of gene sharing in prokaryotes. Proc Natl Acad Sci U S A 102:14332–14337

Boothby TC, Tenlen JR, Smith FW et al (2015) Evidence for extensive horizontal gene transfer from the draft genome of a tardigrade. Proc Natl Acad Sci U S A 112:15976–15981

Broothaerts W, Mitchell HJ, Weir B et al (2005) Gene transfer to plants by diverse species of bacteria. Nature 433:629–633

Bulgakov VP, Kiselev KV, Yakovlev KV et al (2006) *Agrobacterium*-mediated transformation of sea urchin embryos. Biotechnol J 1:454–461

Bundock P, den Dulk-Ras A, Beijersbergen A et al (1995) *Trans*-kingdom T-DNA transfer from *Agrobacterium tumefaciens* to *Saccharomyces cerevisiae*. EMBO J 14:3206–3214

Bundock P, Mroczek K, Winkler AA et al (1999) T-DNA from *Agrobacterium tumefaciens* as an efficient tool for gene targeting in *Kluyveromyces lactis*. Mol Gen Genet 261:115–121

Chapman JA, Kirkness EF, Simakov O et al (2010) The dynamic genome of *Hydra*. Nature 464:592–596

Char SN, Neelakandan AK, Nahampun H et al (2017) An *Agrobacterium*-delivered CRISPR/Cas9 system for high-frequency targeted mutagenesis in maize. Plant Biotechnol J 15:257–268

Copley SD, Dhillon JK (2002) Lateral gene transfer and parallel evolution in the history of glutathione biosynthesis genes. Genome Biol 3: research0025

Crisp A, Boschetti C, Perry M et al (2015) Expression of multiple horizontally acquired genes is a hallmark of both vertebrate and invertebrate genomes. Genome Biol 16:50

Danchin EG, Guzeeva EA, Mantelin S et al (2016) Horizontal gene transfer from bacteria has enabled the plant-parasitic nematode *Globodera pallida* to feed on host-derived sucrose. Mol Biol Evol 33:1571–1579

de Groot MJ, Bundock P, Hooykaas PJJ et al (1998) *Agrobacterium tumefaciens*-mediated transformation of filamentous fungi [published erratum appears in Nat. Biotechnol. 16, 1074 (1998)]. Nat Biotechnol 16:839–842

Degnan SM (2014) Think laterally: Horizontal gene transfer from symbiotic microbes may extend the phenotype of marine sessile hosts. Front Microbiol 5:638

Dunning Hotopp JC (2011) Horizontal gene transfer between bacteria and animals. Trends Genet 27:157–163

Dunning Hotopp JC, Clark ME, Oliveira DC et al (2007) Widespread lateral gene transfer from intracellular bacteria to multicellular eukaryotes. Science 317:1753–1756

Eichinger L, Pachebat JA, Glöckner G et al (2005) The genome of the social amoeba *Dictyostelium discoideum*. Nature 435:43–57

Elmer JJ, Christensen MD, Rege K (2013) Applying horizontal gene transfer phenomena to enhance non-viral gene therapy. J Control Release 172:246–257

Eme L, Gentekaki E, Curtis B et al (2017) Lateral gene transfer in the adaptation of the anaerobic parasite blastocystis to the gut. Curr Biol 27:807–820

Escobar MA, Dandekar AM (2003) *Agrobacterium tumefaciens* as an agent of disease. Trends Plant Sci 8:380–386

Fang H, Huangfu L, Chen R et al (2017) Ancestor of land plants acquired the DNA-3-methyladenine glycosylase (*MAG*) gene from bacteria through horizontal gene transfer. Sci Rep 7:9324

Fernández-González E, de Paz HD, Alperi A et al (2011) Transfer of R388 derivatives by a pathogenesis-associated type IV secretion system into both bacteria and human cells. J Bacteriol 193:6257–6265

Fitzpatrick DA (2012) Horizontal gene transfer in fungi. FEMS Microbiol Lett 329:1–8

Forbes DJ (1992) Structure and function of the nuclear pore complex. Annu Rev Cell Biol 8: 495–527

Fournier GP, Huang J, Gogarten JP (2009) Horizontal gene transfer from extinct and extant lineages: biological innovation and the coral of life. Philos Trans R Soc Lond B 364: 2229–2239

Frandsen RJ (2011) A guide to binary vectors and strategies for targeted genome modification in fungi using *Agrobacterium tumefaciens*-mediated transformation. J Microbiol Methods 87:247–262

Furner IJ, Huffman GA, Amasino RM et al (1986) An *Agrobacterium* transformation in the evolution of the genus *Nicotiana*. Nature 319:422–427

Gardiner DM, Kazan K, Manners JM (2013) Cross-kingdom gene transfer facilitates the evolution of virulence in fungal pathogens. Plant Sci 210:151–158
Gelvin SB (2003) *Agrobacterium*-mediated plant transformation: the biology behind the "gene-jockeying" tool. Microbiol Mol Biol Rev 67:16–37
Gelvin SB (2010) Plant proteins involved in *Agrobacterium*-mediated genetic transformation. Annu Rev Phytopathol 48:45–68
Gelvin SB (2017) Integration of *Agrobacterium* T-DNA into the plant genome. Annu Rev Genet 51:195–217
Gladyshev EA, Meselson M, Arkhipova IR (2008) Massive horizontal gene transfer in bdelloid rotifers. Science 320:1210–1213
Goessweiner-Mohr N, Arends K, Keller W et al (2013) Conjugative type IV secretion systems in gram-positive bacteria. Plasmid 70:289–302
Gogarten JP, Townsend JP (2005) Horizontal gene transfer, genome innovation and evolution. Nat Rev Microbiol 3:679–687
Gouka RJ, Gerk C, Hooykaas PJJ et al (1999) Transformation of *Aspergillus awamori* by *Agrobacterium tumefaciens*-mediated homologous recombination. Nat Biotechnol 17:598–601
Hayman GT, Bolen PL (1993) Movement of shuttle plasmids from *Escherichia coli* into yeasts other than *Saccharomyces cerevisiae* using trans-kingdom conjugation. Plasmid 30:251–257
Heinemann JA, Sprague JF Jr (1989) Bacterial conjugative plasmids mobilize DNA transfer between bacteria and yeast. Nature 340:205–209
Hooykaas PJJ, Klapwijk PM, Nuti MP et al (1977) Transfer of the *Agrobacterium tumefaciens* Ti plasmid to avirulent Agrobacteria and to *Rhizobium ex planta*. J Gen Microbiol 98:477–484
Horsch RB, Fraley RT, Rogers SG et al (1984) Inheritance of functional foreign genes in plants. Science 223:496–498
Huang CY, Ayliffe MA, Timmis JN (2003) Direct measurement of the transfer rate of chloroplast DNA into the nucleus. Nature 422:72–76
Huang J (2013) Horizontal gene transfer in eukaryotes: the weak-link model. BioEssays 35: 868–875
Hubber A, Roy CR (2010) Modulation of host cell function by *Legionella pneumophila* type IV effectors. Annu Rev Cell Dev Biol 26:261–283
Hubber A, Vergunst AC, Sullivan JT et al (2004) Symbiotic phenotypes and translocated effector proteins of the *Mesorhizobium loti* strain R7A VirB/D4 type IV secretion system. Mol Microbiol 54:561–574
Husnik F, McCutcheon JP (2018) Functional horizontal gene transfer from bacteria to eukaryotes. Nat Rev Microbiol 16:67–79
Inomata K, Nishikawa M, Yoshida K (1994) The yeast *Saccharomyces kluyveri* as a recipient eukaryote in transkingdom conjugation: Behavior of transmitted plasmids in transconjugants. J Bacteriol 176:4770–4773
Intrieri MC, Buiatti M (2001) The horizontal transfer of *Agrobacterium rhizogenes* genes and the evolution of the genus *Nicotiana*. Mol Phylogenet Evol 20:100–110
Jaramillo VD, Sukno SA, Thon MR (2015) Identification of horizontally transferred genes in the genus *Colletotrichum* reveals a steady tempo of bacterial to fungal gene transfer. BMC Genom 16:2
Johnsborg O, Eldholm V, Havarstein LS (2007) Natural genetic transformation: prevalence, mechanisms and function. Res Microbiol 158:767–778
Karas BJ, Diner RE, Lefebvre SC et al (2015) Designer diatom episomes delivered by bacterial conjugation. Nat Commun 6:6925
Klasson L, Kambris Z, Cook PE et al (2009) Horizontal gene transfer between *Wolbachia* and the mosquito *Aedes aegypti*. BMC Genom 10:33
Klosterman SJ, Subbarao KV, Kang S et al (2011) Comparative genomics yields insights into niche adaptation of plant vascular wilt pathogens. PLoS Pathog 7:e1002137
Koonin EV (2016) Horizontal gene transfer: essentiality and evolvability in prokaryotes, and roles in evolutionary transition. F1000 Res 5 1805

Koonin EV, Makarova KS, Aravind L (2001) Horizontal gene transfer in prokaryotes: Quantification and classification. Annu Rev Microbiol 55:709–742

Koutsovoulos G, Kumar S, Laetsch DR et al (2016) No evidence for extensive horizontal gene transfer in the genome of the tardigrade *Hypsibius dujardini*. Proc Natl Acad Sci U S A 113:5053–5058

Krenek P, Samajova O, Luptovciak I et al (2015) Transient plant transformation mediated by *Agrobacterium tumefaciens*: principles, methods and applications. Biotechnol Adv 33: 1024–1042

Ku C, Nelson-Sathi S, Roettger M et al (2015) Endosymbiotic origin and differential loss of eukaryotic genes. Nature 524:427–432

Kunik T, Tzfira T, Kapulnik Y et al (2001) Genetic transformation of HeLa cells by *Agrobacterium*. Proc Natl Acad Sci U S A 98:1871–1876

Kyndt T, Quispe D, Zhai H et al (2015) The genome of cultivated sweet potato contains *Agrobacterium* T-DNAs with expressed genes: An example of a naturally transgenic food crop. Proc Natl Acad Sci U S A 112:5844–5849

Lacroix B, Citovsky V (2013) The roles of bacterial and host plant factors in *Agrobacterium*-mediated genetic transformation. Int J Dev Biol 57:467–481

Lacroix B, Citovsky V (2016) A functional bacterium-to-plant DNA transfer machinery of *Rhizobium etli*. PLoS Pathog 12:e1005502

Lacroix B, Tzfira T, Vainstein A et al (2006) A case of promiscuity: *Agrobacterium*'s endless hunt for new partners. Trends Genet 22:29–37

Lander ES, Consortium IHGS, Linton LM et al (2001) Initial sequencing and analysis of the human genome. Nature 409:860–921

Le PT, Pontarotti P, Raoult D (2014) *Alphaproteobacteria* species as a source and target of lateral sequence transfers. Trends Microbiol 22:147–156

Lee CW, Efetova M, Engelmann JC et al (2009) *Agrobacterium tumefaciens* promotes tumor induction by modulating pathogen defense in *Arabidopsis thaliana*. Plant Cell 21:2948–2962

Llosa M, Schröder G, Dehio C (2012) New perspectives into bacterial DNA transfer to human cells. Trends Microbiol 20:355–359

Machado-Ferreira E, Balsemão-Pires E, Dietrich G et al (2015) Transgene expression in tick cells using *Agrobacterium tumefaciens*. Exp Appl Acarol 67:269–287

Maindola P, Raina R, Goyal P et al (2014) Multiple enzymatic activities of ParB/Srx superfamily mediate sexual conflict among conjugative plasmids. Nat Commun 5:5322

Matveeva TV, Bogomaz DI, Pavlova OA et al (2012) Horizontal gene transfer from genus *Agrobacterium* to the plant *Linaria* in nature. Mol Plant-Microbe Interact 25:1542–1551

Matveeva TV, Lutova LA (2014) Horizontal gene transfer from *Agrobacterium* to plants. Front Plant Sci 5:326

Mitrikeski PT (2013) Yeast competence for exogenous DNA uptake: towards understanding its genetic component. Antonie Van Leeuwenhoek 103:1181–1207

Moriguchi K, Yamamoto S, Ohmine Y et al (2016) A fast and practical yeast transformation method mediated by *Escherichia coli* based on a trans-kingdom conjugal transfer system: just mix two cultures and wait one hour. PLoS ONE 11:e0148989

Nowack EC, Price DC, Bhattacharya D et al (2016) Gene transfers from diverse bacteria compensate for reductive genome evolution in the chromatophore of *Paulinella chromatophora*. Proc Natl Acad Sci U S A 113:12214–12219

Piers KL, Heath JD, Liang X et al (1996) *Agrobacterium tumefaciens*-mediated transformation of yeast. Proc Natl Acad Sci U S A 93:1613–1618

Quispe-Huamanquispe DG, Gheysen G, Kreuze JF (2017) Horizontal gene transfer contributes to plant evolution: the case of *Agrobacterium* T-DNAs. Front Plant Sci 8:2015

Riley DR, Sieber KB, Robinson KM et al (2013) Bacteria-human somatic cell lateral gene transfer is enriched in cancer samples. PLoS Comput Biol 9:e1003107

Schönknecht G, Chen WH, Ternes CM et al (2013) Gene transfer from bacteria and archaea facilitated evolution of an extremophilic eukaryote. Science 339:1207–1210

Schönknecht G, Weber AP, Lercher MJ (2014) Horizontal gene acquisitions by eukaryotes as drivers of adaptive evolution. BioEssays 36:9–20

Schröder G, Schuelein R, Quebatte M et al (2011) Conjugative DNA transfer into human cells by the VirB/VirD4 type IV secretion system of the bacterial pathogen *Bartonella henselae*. Proc Natl Acad Sci U S A 108:14643–14648

Siamer S, Dehio C (2015) New insights into the role of *Bartonella* effector proteins in pathogenesis. Curr Opin Microbiol 23:80–85

Sieber KB, Bromley RE, Dunning Hotopp JC (2017) Lateral gene transfer between prokaryotes and eukaryotes. Exp Cell Res 358:421–426

Sieber KB, Gajer P, Dunning Hotopp JC (2016) Modeling the integration of bacterial rRNA fragments into the human cancer genome. BMC Bioinformatics 17:134

Soltani J, van Heusden GP, Hooykaas PJJ (2008) *Agrobacterium*-mediated ransformation of non-plant organisms. In: Tzfira T, Citovsky V (eds) *Agrobacterium*: from biology to biotechnology. Springer, New York, pp 649–675

Stanhope MJ, Lupas A, Italia MJ et al (2001) Phylogenetic analyses do not support horizontal gene transfers from bacteria to vertebrates. Nature 411:940–944

Stegemann S, Hartmann S, Ruf S et al (2003) High-frequency gene transfer from the chloroplast genome to the nucleus. Proc Natl Acad Sci U S A 100:8828–8833

Syvanen M (2012) Evolutionary implications of horizontal gene transfer. Annu Rev Genet 46:341–358

van Kregten M, de Pater S, Romeijn R et al (2016) T-DNA integration in plants results from polymerase-theta-mediated DNA repair. Nat Plants 2:16164

Varga MG, Shaffer CL, Sierra JC et al (2016) Pathogenic *Helicobacter pylori* strains translocate DNA and activate TLR9 via the cancer-associated cag type IV secretion system. Oncogene 35:6262–6269

Walker BJ, Stan GV, Polizzi KM (2017) Intracellular delivery of biologic therapeutics by bacterial secretion systems. Expert Rev Mol Med 19:e6

Wang L, Lacroix B, Guo J et al (2017) Transcriptional activation of virulence genes of *Rhizobium etli*. J Bacteriol 199:e00841–00816

Waters VL (2001) Conjugation between bacterial and mammalian cells. Nat Genet 29:375–376

Wendt T, Doohan F, Mullins E (2012) Production of *Phytophthora infestans*-resistant potato (*Solanum tuberosum*) utilising *Ensifer adhaerens* OV14. Transgenic Res 21:567–578

Wendt T, Doohan F, Winckelmann D et al (2011) Gene transfer into *Solanum tuberosum* via *Rhizobium* spp. Transgenic Res 20:377–386

White FF, Garfinkel DJ, Huffman GA et al (1983) Sequences homologous to *Agrobacterium rhizogenes* T-DNA in the genomes of uninfected plants. Nature 301:348–350

Yang Z, Zhou Y, Huang J et al (2015) Ancient horizontal transfer of transaldolase-like protein gene and its role in plant vascular development. New Phytol 206:807–816

Yue J, Hu X, Huang J (2014) Origin of plant auxin biosynthesis. Trends Plant Sci 19:764–770

Yue J, Hu X, Sun H et al (2012) Widespread impact of horizontal gene transfer on plant colonization of land. Nat Commun 3:1152

Zardoya R, Ding X, Kitagawa Y et al (2002) Origin of plant glycerol transporters by horizontal gene transfer and functional recruitment. Proc Natl Acad Sci U S A 99:14893–14896

Zuniga-Soto E, Mullins E, Dedicova B (2015) *Ensifer*-mediated transformation: an efficient non-*Agrobacterium* protocol for the genetic modification of rice. SpringerPlus 4:600

Agrobacterium: A Genome-Editing Tool-Delivery System

Nagesh Sardesai and Subhashree Subramanyam

Contents

1 Introduction 464
2 *Agrobacterium* in Plant Biotechnology 465
3 History of Genome Editing and Gene Targeting 467
3.1 Methods of Editing and Targeting 468
3.2 Meganucleases and Homing Endonucleases 470
3.3 Zinc-Finger Nucleases 470
3.4 Transcription Activator-like Effector Nucleases 471
3.5 Clustered Regularly Interspersed Short Palindromic Repeats (CRISPR)/Cas9 Nuclease 473
4 *Agrobacterium* in Genome Editing and Gene Targeting 474
4.1 *Agrobacterium tumefaciens*-Mediated Genome Editing 474
4.2 *Agrobacterium rhizogenes*-Mediated Genome Editing 476
5 *Agrobacterium* Engineering 477
5.1 Modifications in *Agrobacterium* to Facilitate Genome Editing 477
5.2 Improvement of Transformation Efficiency Using Additional Copies of *Vir* Genes 480
6 Future of *Agrobacterium* in Gene Targeting 481
References 482

The original version of this chapter was revised: The name of one of the authors has been corrected. The erratum to this chapter is available at DOI: 10.1007/82_2018_136

N. Sardesai (✉)
Corteva Agriscience™, Agriculture Division of DowDuPont,
8305 NW 62nd Avenue, Johnston, IA, USA
e-mail: nsardesai@dow.com

S. Subramanyam
Department of Agronomy, Purdue University, 915 W State Street,
West Lafayette, IN 47907, USA
e-mail: shubha@purdue.edu

Current Topics in Microbiology and Immunology (2018) 418:463–488
DOI 10.1007/82_2018_101

Published Online: 25 July 2018

Abstract With the rapidly increasing global population, it will be extremely challenging to provide food to the world without increasing food production by at least 70% over the next 30 years. As we reach the limits of expanding arable land, the responsibility of meeting this production goal will rely on increasing yields. Traditional plant breeding practices will not be able to realistically meet these expectations, thrusting plant biotechnology into the limelight to fulfill these needs. Better varieties will need to be developed faster and with the least amount of regulatory hurdles. With the need to add, delete, and substitute genes into existing genomes, the field of genome editing and gene targeting is now rapidly developing with numerous new technologies coming to the forefront. *Agrobacterium*-mediated crop transformation has been the most utilized method to generate transgenic varieties that are better yielding, have new traits, and are disease and pathogen resistant. Genome-editing technologies rely on the creation of double-strand breaks (DSBs) in the genomic DNA of target species to facilitate gene disruption, addition, or replacement through either non-homologous end joining or homology-dependent repair mechanisms. DSBs can be introduced through the use of zinc-finger nucleases (ZFNs), transcription activator-like effector nucleases (TALENs), or clustered regularly interspersed short palindromic repeats (CRISPR)/Cas nucleases, among others. *Agrobacterium* strains have been employed to deliver the reagents for genome editing to the specific target cells. Understanding the biology of transformation from the perspective not only of *Agrobacterium*, but also of the host, from processing of T-DNA to its integration in the host genome, has resulted in a wealth of information that has been used to engineer *Agrobacterium* strains having increased virulence. As more technologies are being developed, that will help overcome issues of *Agrobacterium* host range and random integration of DNA, combined with highly sequence-specific nucleases, a robust crop genome-editing toolkit finally seems attainable.

1 Introduction

One of the major challenges of the twenty-first century is the rapid increase in global population. Global population is projected to reach 9.7 billion by 2050 and 11.2 billion by 2100 (UN WPP Data Booklet 2015). According to the Food and Agricultural Organization (FAO), feeding a world population of this magnitude will necessitate increasing overall food production by ~70% between now and 2050. In developing countries, 80% of the necessary production increase would come from increasing yields and cropping intensity, whereas only 20% would come from expansion of arable land. Farming practices and crops cultivated today have developed over a relatively short span of time. Varieties planted today are a result of improvements that have been made possible through traditional plant breeding techniques. These traditional practices have been responsible for making farming predictable and productive. Despite the successes of these techniques, the time taken to produce new varieties is a factor worth considering. Whereas some yield

increases are possible through traditional plant breeding practices, the bulk of yield increases will have to come from the effective use of plant biotechnology.

Over the past few decades, a number of technologies have emerged that have advanced our understanding of all aspects of plant sciences, leading to the development of new and better varieties of crops. There has been a significant advancement in the area of plant tissue culture that has bolstered the application of traditional plant breeding techniques. Manipulation of plant growth media compositions has allowed scientists not only to propagate plants from tissues and organs (Murashige 1974; Hussey 1978; Morrison and Evans 1988), but also to regenerate them from just a few cells, and in some cases, even protoplasts (Cocking 1972; Rhodes et al. 1988). Another area that has expanded exponentially is the field of genetic engineering or recombinant DNA technology. The ability to identify and purify genetic material from one organism and tailor it for insertion into a different one opened up avenues for the generation of plants containing new traits that would have taken several years to achieve through traditional plant breeding techniques. With continuously improving recombinant DNA technology, uptake of DNA by plants was demonstrated in numerous systems throughout the 1970s, though at the time the methods did not definitively prove integration of the DNA into the host genome (Kleinhofs and Behki 1977). As the plant cell wall was seen to be the major barrier to uptake of DNA in plant cells, protoplasts were prepared and shown to integrate DNA (Krens et al. 1982). Somatic hybrid plants of potato and tomato were regenerated from fused protoplasts (Melchers et al. 1978). Development of techniques in genetic engineering led to the idea of a suitable vector for transferring desirable genetic information into plant cells. Some prominent potential vehicles included Ti plasmids of *Agrobacterium tumefaciens*, DNA plant viruses such as cauliflower mosaic virus, and the chloroplast and mitochondrial genomes of higher plants (Levings and Pring 1979). Together, the rapid progress in these two areas led to the discoveries that initiated the age of plant biotechnology and the complete transformation of traditional agriculture.

2 *Agrobacterium* in Plant Biotechnology

The discovery, over 30 years ago, that *Agrobacterium tumefaciens* could be used to generate transgenic plants (Barton et al. 1983; Caplan et al. 1983) heralded the beginning of the era of plant biotechnology. This discovery of plant transformation revolutionized agriculture. In the 1980s, scientists learned to delete from T-DNA (the region of the *Agrobacterium* Ti plasmid that is transferred to plants) the oncogenes and the opine synthase genes, effectively disarming the virulent strains so that plant tissues infected with *Agrobacterium* containing this engineered T-DNA could regenerate into normal plants (Fraley et al. 1983; Herrera-Estrella et al. 1983). Recombinant *Agrobacterium* strains, in which the native T-DNA is replaced with genes of interest, were developed as the most efficient vehicles for the introduction of foreign genes into plants. Because of the complexity of introducing foreign genes

directly into the T-region of the very large native Ti-plasmid, a number of different strategies were developed by scientists to either clone these genes using indirect means into the Ti plasmid, or clone the genes into a separate *Vir* gene-independent replicon, within a T-region (Gelvin 2003). This second method resulted in the development of T-DNA binary vector systems (Hoekema et al. 1983).

T-DNA binary vector systems revolutionized the use of *Agrobacterium* to transform plants. The plasmid containing the T-region constitutes the binary vector, whereas the replicon containing the *Vir* genes constitutes the *Vir* helper. The *Vir* helper plasmid generally contains a partial or complete deletion of the native T-region, resulting in the production of a 'disarmed' strain that is incapable of inducing tumors. Several nononcogenic *Agrobacterium* strains were developed that are currently being used in plant biotechnology, including LBA4404 (Ooms et al. 1981), GV3101::pMP90 (Koncz and Schell 1986), AGL0 (Lazo et al. 1991), EHA101 and its derivative strain EHA105 (Hood et al. 1986, 1993), and NT1 (pKPSF2) (Palanichelvam et al. 2000). The binary vectors are small and capable of being propagated in both *Escherichia coli* and *Agrobacterium* and can be manipulated to have (i) different plant selectable markers, (ii) regulatory elements driving genes of interest, (iii) translational enhancers to increase expression of transgenes, and (iv) protein-targeting signals that direct the transgene-encoded protein to specific locations in the plant cell (Hellens et al. 2000).

Agrobacterium-mediated transformation is considered to be the more preferred method of genetic transformation over other artificial approaches such as electroporation, microinjection, polyethylene glycol (PEG)-mediated transformation, or biolistic bombardment of cells with highly accelerated naked DNA molecules (gene gun) not only because of the ease and low cost of the method, but also because of the relatively low complexity of intact transgenes integrated into the plant genome (Kohli et al. 2003; Olhoft et al. 2004; Anderson and Birch 2012; Jackson et al. 2013). However, many economically important plant species, or cultivars of some species, are recalcitrant to *Agrobacterium*-mediated transformation. Although there has been significant progress in improving the efficiency of genetic transformation of corn and soybeans, many other cereal grains, legumes, and trees of horticultural and industrial importance remain problematic (Shrawat and Lorz 2006). A few different strategies have been adopted by scientists to increase plant transformation efficiency, including the identification of highly virulent *Agrobacterium* strains, manipulation of strains to make them highly virulent (Hansen et al. 1994), or optimizing plant culture conditions (Newell 2000). Strain manipulations such as addition of copies of various *Vir* genes to create 'supervirulent' strains and 'superbinary vectors' have increased transformation frequencies of many plants, including cereals (Hiei et al. 1994; Ishida et al. 1996). However, limits to improvement of genetic transformation might have been reached using such approaches (Gelvin 2000). This paved the way for developing alternative approaches to increase plant transformation by manipulation of the plants themselves. A prerequisite to this approach was the identification of plant genes that are involved in the plant transformation process.

One approach to identify plant genes involved in plant transformation was to identify plant proteins that interact with specific transferred (effector) or

surface-localized *Agrobacterium* Vir proteins. As a number of virulence effector proteins are transferred from *Agrobacterium* to plant cells, their interaction with plant proteins is likely to be important in transformation. Protein–protein two-hybrid interaction traps in yeast identified several proteins from a plant cDNA library (Schrammeijer et al. 2001; Tzfira et al. 2001; Bakó et al. 2003; Hwang and Gelvin 2004) that were verified to interact *in planta* with *Agrobacterium* Vir proteins using bimolecular fluorescence complementation (Citovsky et al. 2006; Lee et al. 2008). Another approach for identification of plant genes important in transformation is by comparing transcript profiles between uninfected plants and ones infected with virulent (or avirulent) *A. tumefaciens* strains. Several such studies have been carried out in *Arabidopsis* (Deeken et al. 2006; Ditt et al. 2006; Kim et al. 2007; Lee et al. 2009) and tobacco (Veena et al. 2003) that identified plant defense response genes, hormone signaling pathways, and chromatin-associated proteins such as histones. However, a forward genetic screening of mutant organisms to identify altered phenotypic characteristics remains the classical way of identifying genes and gene function. Using *Arabidopsis* as the model system for generating and screening mutants, plants that are resistant to *Agrobacterium*-mediated transformation (*rat* mutants) (Nam et al. 1999; Zhu et al. 2003) or hypersusceptible to *Agrobacterium*-mediated transformation (*hat* mutants) (Sardesai et al. 2013; Sardesai and Gelvin unpublished) have been identified. Manipulation of such genes can lead to increased transformation.

3 History of Genome Editing and Gene Targeting

Genome editing in the plant world is as old as the existence of plants themselves, arising in the form of the natural process of random mutagenesis. This non-directed genome editing via spontaneous and induced mutations has been responsible for the generation of new and altered traits for mankind, as crop plants were domesticated approximately 10,000 years ago and selected for higher seed set and seedling vigor, reduced seed dormancy and dispersal, and better architectural characteristics such as compact growth or reduced tillering (Gepts 2002; Meyer et al. 2012). As natural selection moved into the more deliberate realms of plant breeding and improvement through intra- and interspecific, and intergeneric crosses, the increase in genetic variability continued to rely on natural and spontaneous mutations (Sleper and Poehlman 2006). The next logical step was the introduction of mutations via the use of different mutagens such as X-rays, gamma rays, beta and ultraviolet irradiation, and neutrons (FAO/IAEA 1977). In addition, chemical mutagens, such as ethyl methanesulfonate (EMS), were identified that caused base substitutions (Greene et al. 2003; Caldwell et al. 2004). Another source of mutations within species is transposable elements that have been used for induced mutagenesis and gene tagging (Geiser et al. 1982; Federoff et al. 1983; Voytas and Ausubel 1988; Grandbastien et al. 1989; Chuck et al. 1993).

3.1 Methods of Editing and Targeting

A large number of reports in recent years have been published describing precision editing of genomes using nucleases and recognition sequences. These technologies can be summarized as: zinc-finger nucleases (ZFN), transcription activator-like effector nucleases (TALENs), clustered regularly interspersed short palindromic repeats (CRISPR)/Cas 9 nuclease, or Cpf1 nuclease. Genome editing relies on DNA repair. Typically, natural damage occurring in the cell to DNA, including a double-strand break (DSB), because of free radicals, UV radiation, or any other type of metabolic by-product, may be repaired by one of two major DNA repair pathways: non-homologous end joining (NHEJ) and homology-directed repair (HDR) (Kanaar et al. 1998; Pastwa and Blasiak 2003; Hartlerode and Scully 2009; Steinert et al. 2016). In the case of a DSB, NHEJ, an error-prone pathway, may be used by the cell to repair the break in two ways. In classical NHEJ, different proteins, such as Ku70 and Ku80, are recruited to the DSB that bind to the broken ends of the DNA, followed by ligation of the ends by a ligase, which may result in errors by the insertion or deletion (indel) of nucleotides. In the alternative microhomology-based NHEJ pathway, 5' ends are resected and the single-strand regions of DNA bind to complementary sequences; non-homologous regions of DNA are consequently excised. This reaction results in deletions of DNA sequences that are flanked by the homologous sequences. The errors caused by NHEJ may lead to premature stop codons or nonsense codons, making the genes non-functional.

The second DNA repair pathway, HDR, relies on homologous recombination that occurs in somatic cells to repair DSBs, and in meiotic cells to exchange genetic material between parental chromosomes. In plants, almost all DSBs in somatic cells are repaired through synthesis-dependent strand annealing (SDSA), which is the most common conserved HDR mechanism (Puchta 2005; Steinert et al. 2016). During HDR, 3′ overhangs are extended from a DSB site, following which a 5′ end invades the homologous strand forming a D-loop. The homologous DNA is used as the template for filling in the gaps, while the 3′ end reanneals with the second 3′ end, without crossover (Fig. 1). This process results in the precise integration of the template DNA strand into the DSB. In nature, a sister chromatid or homologous chromosome could supply the template or 'donor' DNA, whereas a synthetic template DNA could also be provided exogenously for gene insertion/replacement. One of the first demonstrations that homologous recombination could occur between T-DNA and plant chromosomal sequences came from the work of Lee et al. (1990). They showed that a deleted, non-functional fragment from the acetolactate synthase (*ALS*) gene carrying a single amino acid mutation could be delivered into tobacco protoplasts using *A. tumefaciens* LBA4404, and they could recover chlorsulfuron-resistant colonies from which plants could be regenerated. Using *A. tumefaciens* EHA105, a highly efficient gene targeting system was developed for rice utilizing a positive/negative selection system (Ozawa et al. 2012). The system consisted of a targeting vector that harbored a hygromycin resistance gene (*HPT*) for positive selection of targeted events, a *Waxy* or *Xyl* gene

providing sequence homology to the endogenous targeted *Waxy* or *Xyl* locus, and two highly expressed diphtheria toxin genes flanking the *Waxy* (or *Xyl*) sequence/*HPT* cassette, outside of the homology regions, for providing a strong negative selection by eliminating random insertions. Targeting of the endogenous *Waxy* (or *Xyl*) locus with the HPT cassette would result in hygromycin-resistant targeted plants, whereas any random insertion would result in the expression of the diphtheria toxin genes, eliminating such events.

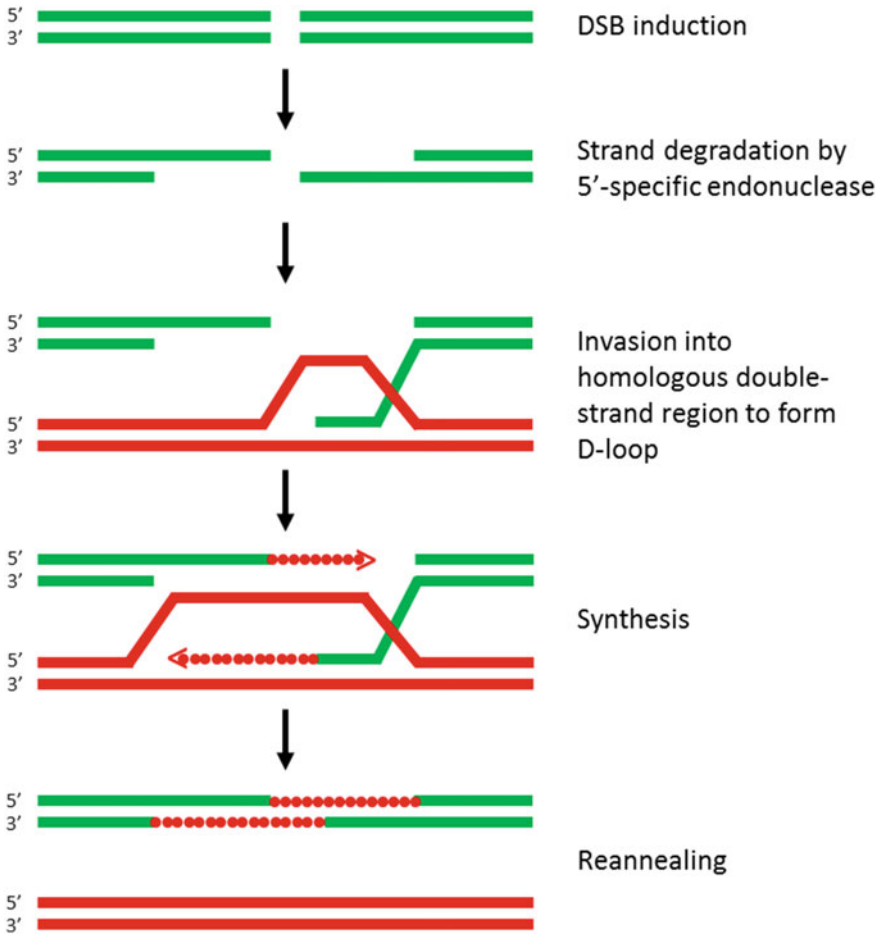


Fig. 1 Synthesis-dependent strand annealing (SDSA) pathway of homologous recombination. After a double-strand break (DSB) occurs, both ends are resected to expose 3' ends of the single-stranded DNA. One 3' end invades a homologous sequence by displacing one of the strands, producing a displacement loop (D-loop). The invading strand elongates by copying the sequence information from the intact donor DNA (red). The elongated single strand is then released from the D-loop and reanneals with the homologous single-stranded DNA on the opposite side of the break site, producing a non-crossover event

3.2 Meganucleases and Homing Endonucleases

It became increasingly apparent that if genome editing were to occur at DSB sites, then precise targeting of the genome would necessitate the ability to create DSBs at the desired locations that would allow either NHEJ or HDR to operate and edit the genome. This model resulted in the rapid evolution of sequence-specific nucleases (SSNs) for plant genome editing. One such family of nucleases is the meganucleases, or homing endonucleases. Meganucleases are site-specific endonucleases that recognize DNA sequences over 12 bp long. They are found in bacteria, archaea, and eukaryotes (Paques and Duchateau 2007). Several hundred meganucleases have been discovered so far and categorized into five groups based on sequence and structural motifs: LAGLIDADG, His-Cys box, GIY-YIG, HNH, and PD-(D/E)XK (Orlowski et al. 2007; Paques and Duchateau 2007). The commonly used homing endonucleases I-*Sce*I and I-*Cre*I belong to the LAGLIDADG group. As only a few amino acid residues of the meganuclease makes direct contact with the DNA, binding specificity can be altered for targeting endogenous genes, making them one of the first useful tools for targeted induction of DSBs. Naturally occurring meganucleases, such as I-*Sce*I, have been used in a series of experiments for targeted mutagenesis and gene targeting *in planta*. Cotransfection of *Nicotiana plumbaginifolia* with an I-*Sce*I-carrying plasmid and recombination substrates carrying homologous sequences flanked by a I-*Sce*I site produced targeted events that were rate-limited by the induction of DSBs (Puchta et al. 1993). Tzfira et al. (2003) used *A. tumefaciens* first to generate *Nicotiana tabacum* plants that carried an I-*Sce*I recognition site, and then demonstrated targeting to that site by retransforming these plants with an *Agrobacterium* strain allowing transient expression of I-*Sce*I for creating the DSB and another carrying a promoterless integration marker on a T-DNA with an I-*Sce*I site. Crystal structure analysis of the I-*Cre*I homing endonuclease and two mutants revealed molecular interactions responsible for their DNA target specificities, leading to development of homing endonucleases with novel target specificities (Rosen et al. 2006). Development of a re-engineered meganuclease for cleavage of a specific endogenous cotton target gene sequence allowed the targeted insertion of herbicide tolerance trait genes in addition to a transgenic insect control locus (D'Halluin et al. 2013).

3.3 Zinc-Finger Nucleases

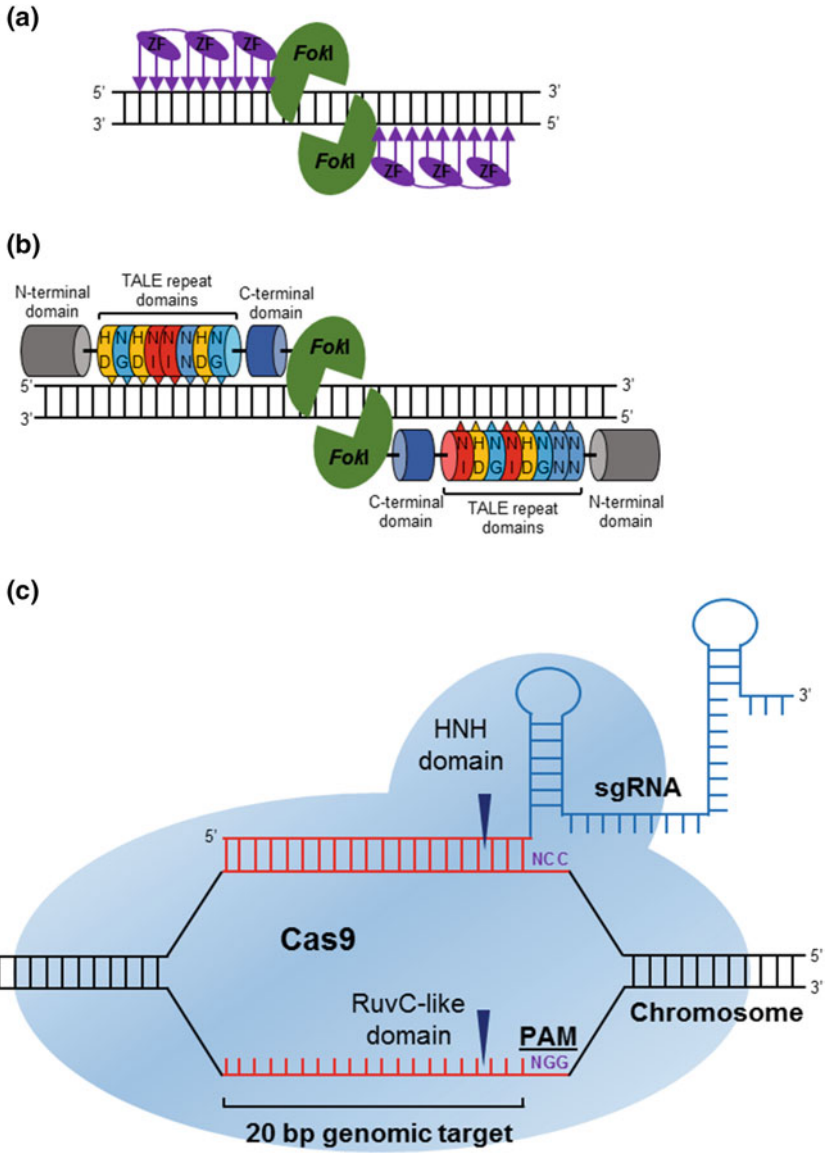
Zinc-finger nucleases (ZFNs) consist of nuclease and DNA-binding domains that can be designed to recognize specific DNA sequences (Urnov et al. 2010). The DNA-binding region consists of zinc-finger protein domains that can bind to specific DNA sequences and is fused to the catalytic domain of the type II restriction enzyme *Fok*I that serves as the nuclease for DNA cleavage (Kim et al. 1996). DNA binding occurs as a result of a tethered array of 4–6 zinc-finger protein

domains, each recognizing 3 bp of DNA (Fig. 2a). Although this allows ZFNs to be built that have binding specificity, the interactions between neighboring zinc fingers can pose significant challenges in design (Urnov et al. 2010). For the *Fok*I nuclease to be functional, it must dimerize, allowing the two ZFN pairs to orient in a way that increases the level of specificity of the ZFN complex. Variants of *Fok*I have been developed that require heterodimerization, enhancing the specificity and reducing off-target cleavage (Miller et al. 2007). Because of their specificity, ZFNs can be designed to bind and cleave virtually any stretch of DNA sequence, creating double-strand breaks at defined loci and consequently making genome editing a very controlled process.

In a proof-of-concept study for targeted mutagenesis, a pre-integrated transgene was successfully targeted and mutated by a ZFN pair in *Arabidopsis thaliana* using *A. tumefaciens* LBA4404 in a floral dip transformation strategy (Lloyd et al. 2005). Zhang et al. (2010) described an efficient method for targeted mutagenesis of the *Arabidopsis* genes *ADH1* and *TT4* employing estrogen-inducible ZFNs. The constructs targeting these genes were delivered by a floral dip protocol using *A. tumefaciens* GV3101. This strain was also used successfully to deliver a heat-inducible ZFN targeting the *Arabidopsis ABA-INSENSITIVE4* (*ABI4*) gene (Osakabe et al. 2010). Designed ZFNs targeting a pre-integrated reporter construct containing dual partial, non-functional reporter genes flanked by large stretches of non-homologous sequence in *N. tabacum* BY2 suspension cells were used to enable site-specific cleavage and integration of a donor DNA containing a full-length promoter and a 5' partial *pat* herbicide resistance gene flanked by appropriate homologous sequences (Cai et al. 2009). *A. tumefaciens* strain LBA4404-mediated delivery of the plasmid harboring the ZFNs and the donor DNA construct flanked by short stretches of homology to the target locus yielded up to 10% targeted, homology-directed transgene integration into the ZFN cleavage site (Cai et al. 2009). NHEJ-mediated replacement of a *gfp* gene with *hpt* was demonstrated in tobacco and *Arabidopsis* when constructs for acceptor and donor DNA sequences flanked by ZFN sites were transformed using *A. tumefaciens* strain EHA105 (Weinthal et al. 2013). This work highlighted the success of NHEJ-induced gene exchange and the use of this method for transgene replacement and gene stacking in plants. ZFN-mediated gene targeting (GT) at a pre-integrated transgenic locus in *Arabidopsis* using *A. tumefaciens* AGL1 carrying the ZFN genes and a T-DNA GT construct showed site-specific mutagenesis and a 0.1% GT frequency (de Pater et al. 2009) when targeting an endogenous PPO gene with a ZFN by introducing two mutations, resulting in herbicide-resistant plants at a GT frequency of 0.31% (de Pater et al. 2013).

3.4 Transcription Activator-like Effector Nucleases

Transcription activator-like effector nucleases (TALENs) have emerged as an alternative to ZFNs for genome editing and introducing DSBs. TALEs were derived



from the type III effector protein AvrBs3 of the plant pathogen *Xanthomonas campestris* (Kay et al. 2007; Römer et al. 2007). Most have a variable number of nearly identical tandem 34 amino acid repeats comprising the DNA-binding domain with specificity conferred by two hypervariable residues, termed repeat variable diresidues (RVDs). These residues are typically found at positions 12 and 13 of the repeat and recognize one base pair per TALE motif (Lahaye and Bonas 2001). TALENs are similar to ZFNs and comprise a nonspecific *Fok*I nuclease domain

◂**Fig. 2** Genome modification methods via double-strand break (DSB) repair. **a** A ZFN pair recognizes the target sequence via 4–6 zinc-finger (ZF) protein domains that each recognize 3 bp of DNA. The *Fok*I cleavage domain dimerizes to make a DSB. **b** TALENs bind and cleave as dimers on a target DNA site. TALE repeats are shown as colored disks. Each repeat region contains 33–35 amino acids of which the ones on position 12 and 13 are hypervariable residues specifically binding to a single base of DNA (NN, NI, HD and NG recognize G, A, C, and T, respectively). The N- and C-terminal domains that are required for DNA-binding are indicated. *Fok*I dimerizes to create a DSB. **c** In the CRISPR/Cas9 system, the single-guide RNA (sgRNA) with a 20-bp genomic DNA-binding sequence forms a complex with the Cas9 nuclease and guides it to a DNA target sequence adjacent to the protospacer adjacent motif (PAM) having the sequence NGG. Cas9 contains two nuclease domains homologous to RuvC and HNH nucleases. HNH nuclease domain cleaves the complementary DNA strand, whereas RuvC-like domain cleaves the non-complementary strand, creating a blunt DSB (indicated by the blue triangle) in the target DNA, 3 bp upstream of the PAM

fused to a customizable DNA-binding domain (Fig. 2b). Individual TALE repeats in an array specifically bind to a single base of DNA, and there is a simple correlation between the RVD and the base bound by each repeat, making the engineering of TALE repeat arrays with novel specificities possible (Boch et al. 2009). Nearly, all engineered TALE repeat arrays use four domains that contain the RVDs NN, NI, HD, and NG for the recognition of guanine, adenine, cytosine, and thymine, respectively. TALENs have been used to introduce knockout mutations in *Arabidopsis* using PEG-mediated transfection (Cermak et al. 2011) and confer resistance to infection by *Xanthomonas* in rice via *A. tumefaciens* strain EHA105-mediated delivery of TALENs targeting the rice bacterial blight susceptibility gene *OsSWEET14* (Li et al. 2012). De novo-engineered TALENS were demonstrated to target a 12-bp effector binding element in tobacco when delivered via *A. tumefaciens* strain GV3101 (Mahfouz et al. 2011). TALEN-mediated editing in tomato of the *PROCERA* gene, a negative regulator of gibberellic acid, was demonstrated using DNA delivered by *A. tumefaciens* AGL1 (Lor et al. 2014). This strain was also used for the delivery of TALENs targeting the intron of a *Ubiquitin7* gene in potato to generate a targeted insertion of a promoterless herbicide resistance gene through one-sided HR-mediated integration (Forsyth et al. 2016).

3.5 Clustered Regularly Interspersed Short Palindromic Repeats (CRISPR)/Cas9 Nuclease

CRISPR/Cas systems are important components of adaptive immunity in bacteria and archaea that are involved in the elimination of viral and plasmid DNA from invading parasites. They are based on a RNA-guiding system (Bhaya et al. 2011). CRISPR is a genomic locus of tandem direct repeat sequences and protospacers, the sequences between the repeats that are derived from invading elements. The CRISPR locus contains a combination of Cas9 genes, non-coding RNA

elements called CRISPR RNA (crRNA), and sequences for small trans-activating crRNA (tracrRNA). The crRNA and tracrRNA form a complex called the guide RNA (gRNA). The Cas9 endonuclease forms a complex with the guide RNA that recognizes the specific sequences on the target site in the presence of a downstream protospacer adjacent motif (PAM) with the sequence 5′-NGG-3′ (Fig. 2c). The tracrRNA:crRNA heteroduplex can be replaced by an engineered single guide RNA (sgRNA), thus reprogramming the Cas9/sgRNA system to a specific complementary target site (Cong et al. 2013; Jinek et al. 2012; Mali et al. 2013). The sgRNA, in the form of the DNA sequence N20-NGG, can be targeted by altering the first 20 nucleotides of the sgRNA for novel genome-editing applications (Sander and Joung 2014). With a very high occurrence of PAMs in genomes, almost any gene can be targeted by a Cas9/sgRNA complex.

Cas9-mediated plant genome editing or targeting can be achieved by the use of Cas9 and sgRNA expression cassettes. These reagents could be part of a single plasmid and delivered to the plant cells or delivered separately, using conventional transformation methods such as *Agrobacterium* or biolistics (Xing et al. 2014; Svitashev et al. 2015). In either of these cases, the delivered gRNA, Cas9, and selectable marker genes can integrate into the plant genome leading to gene disruption, chimerism, and potential off-target cleavage (Svitashev et al. 2015; Kanchiswamy 2016). In addition, the integration of these components into the DSB at the target sites reduced the efficiency of gene insertion and editing (Kim et al. 2014; Svitashev et al. 2015). Transgene-free genome editing has been demonstrated in wheat by transiently expressing CRISPR/Cas9 DNA and delivery of Cas9 and gRNA as in vitro transcribed RNA molecules (Zhang et al. 2016), and in maize by the biolistic delivery of pre-assembled Cas9-gRNA ribonucleoproteins into embryo cells (Svitashev et al. 2016). Despite these advances, *A. tumefaciens* remains one of the preferred methods of delivery of CRISPR/Cas genome-editing reagents due to the stability of the DNA vectors used in this system.

4 *Agrobacterium* in Genome Editing and Gene Targeting

4.1 Agrobacterium tumefaciens-*Mediated Genome Editing*

One of the first examples of the use of *A. tumefaciens* in site-specific integration (SSI) mediated by Cre recombinase in *Arabidopsis thaliana* was reported by Vergunst et al. (1998). A chromosomally introduced *loxP* site was targeted with a frequency of 1.2–2.3%. Target plants were generated using *A. tumefaciens* MOG101 containing the plasmid p35S-*lox-cre,* and root explants of the transgenic plants were cocultivated with an *A. tumefaciens* strain containing the plasmid p*lox-npt-lox*, followed by selection for site-specific recombinants with kanamycin. Eighty-nine percent of the recombinants resulted from precise site-specific

integration (Vergunst et al. 1998). Homologous recombination (HR)-dependent gene targeting was demonstrated in rice where herbicide-tolerant plants were generated by the introduction of two-point mutations into the endogenous *ALS* gene via *A. tumefaciens* strain EHA101-mediated transfer of T-DNA carrying the modified piece of DNA (Endo et al. 2007). Two-thirds of the generated plants were shown to only carry the two mutations without the presence of foreign DNA such as border sequences of the T-DNA.

A site-directed integration (SDI) system was designed for an *Agrobacterium*-mediated recombinase-mediated cassette exchange (RMCE) strategy in tobacco. This system precisely integrated a single copy of a desired gene into a predefined target locus by gene replacement (Nanto et al. 2005). In this study, the R-*RS* system, derived from the pSR1 plasmid of *Zygosaccharomyces rouxii*, and consisting of the recombinase (R) and its recognition sites (*RS*) was used (Matsuzaki et al. 1990) and the T-DNA delivered using *A. tumefaciens* LBA4404. A Cre-based targeting strategy using heterospecific *lox* sites (*loxP* and *lox5171*) was developed for *Arabidopsis* using RMCE of T-DNA delivered by *A. tumefaciens* LBA1100, a C58C1 strain with a disarmed octopine-type pTiB6 plasmid (Louwerse et al. 2007). To avoid instability between the *lox* sites of the exchange cassette in the presence of Cre in *A. tumefaciens*, the exchange cassette T-DNA and the Cre T-DNA were mobilized separately into *A. tumefaciens* LBA1100, and the two strains were mixed in a 1:1 ratio before being used for transformation of transgenic root explants.

Using *Agrobacterium*-mediated transformation, 100% mutational efficiency was demonstrated for two lignin and flavonoid biosynthesis pathway genes, *4CL1* and *4CL2*, in *Populus* (Zhou et al. 2015). The group designed gRNAs for the *4CL* genes after careful analysis of in-house generated RNA-sequencing data to avoid single nucleotide polymorphisms (SNPs), as the CRISPR/Cas9 system is highly sensitive to SNPs. For their experiments, the gRNA/Cas9 cassette-containing binary vector was delivered by *A. tumefaciens* C58(pMP90). This system generated over 30 events per construct with no off-target cleavage detected in the mutants. A procedure called *in planta* GT employed a pre-inserted CRISPR/Cas nuclease gene for double-strand break induction in the *PPO* gene in *Arabidopsis*, as well as a pre-inserted repair template with a 5' truncated *PPO* gene containing two mutations for butafenacil resistance. These constructs were delivered via *A. tumefaciens* AGL1 by floral dip transformation with the two strains mixed together (de Pater et al. 2018). *In planta* GT resulted in plants that were resistant to butafenacil, showing that a DSB in a target locus can be repaired by a homologous sequence present elsewhere in the genome (de Pater et al. 2018). Experiments in maize using *A. tumefaciens* LBA4404 for delivery of the donor DNA and the nuclease to obtain pre-integrated sequences showed similar results (Ayar et al. 2013). Other experiments in maize targeted the *Zmzb7* gene using *A. tumefaciens* EHA101, resulting in the production of albino phenotypes (Feng et al. 2016), and the *PSY1* gene using *A. tumefaciens* EHA105 (Zhu et al. 2016).

A high-frequency targeted mutagenesis system in maize was described by Char et al. (2017) using *Agrobacterium*-delivered CRISPR/Cas9. The team of researchers demonstrated the flexibility of their system by cloning up to four guide RNAs

(gRNAs) for single or multiplex gene targeting. They constructed gRNAs that targeted two closely related but polymorphic *Argonaute* (*Ago*) genes, *ZmAgo18a* and *ZmAgo18b*, and the dihydroflavonol 4-reductase gene *a1* (*anthocyaninless 1*) and its homolog *a4*. The gRNAs were tested as specifically targeting a single gene at a time or as having both gRNAs in one construct to target both genes simultaneously. These gRNA constructs were moved into a T-DNA vector containing *Cas9* regulated by the maize ubiquitin promoter. *Agrobacterium* strain EHA101 was transformed with the gRNA/*Cas9* constructs and was used to infect immature Hi-II maize embryos. Single-gene targeting constructs showed similar transformation and mutagenesis frequencies of 70–74%, whereas the duplex targeting experiments showed 58% mutations in both the genes. The group also investigated the feasibility of mutating two genes (or groups of genes) through a single procedure by mixing *Agrobacterium* strains harboring different gRNA constructs and carrying out a co-transformation of the explants. They used a mixture of *A. tumefaciens* EHA101 strains containing the two independent *Ago* gRNAs to infect maize B104 immature embryos. Their results indicated that mixing two *Agrobacterium* strains generated mutation frequencies in individual target genes similar to those of single strain infections. Their turn-around time for the *Agrobacterium*-mediated process, called ISU Maize CRISPR, was reported as seven months for maize genome editing.

More recently, a group from DuPont Pioneer has developed an *Agrobacterium*-mediated sorghum transformation system that has been used to also develop stable CRISPR/Cas9-mediated gene knockouts (Che et al. 2018). Here, they used the *A. tumefaciens* strain LBA4404 which harbored a ternary vector system comprising a T-DNA-less helper plasmid, pVIR, carrying an optimal set of *Agrobacterium* virulence genes (Anand et al. 2018) along with a T-DNA transfer-competent binary plasmid containing the Cas9 and gRNA gene editing machinery recognizing a region in the *Sb-CENH3* gene. With three designs for gRNAs targeting *Sb-CENH3*, the researchers achieved an editing efficiency of 37–40% based on the analyses of T0 plants. The high-targeted editing efficiencies are in great part due to increased transformation frequencies using a ternary transformation system that yielded transformation efficiencies ranging from 25 to 29% for 17–18 kb T-DNA, compared to 15% with a co-integrate system carrying a 4.6 kb T-DNA, in which a superbinary vector, pSB1, carrying a DNA fragment with extra *vir* genes (B, G, part of C and D) recombined with a T-DNA carrying vector in *Agrobacterium*, giving rise to a large co-integrate vector (Komari et al. 1996).

4.2 Agrobacterium rhizogenes-*Mediated Genome Editing*

With the widespread adoption and large-scale use of ZFN technology, various robust and publicly available methods for engineering zinc-finger arrays are now available. A publicly available platform of reagents and software, context-dependent assembly (CoDA), has been developed for generating active zinc-finger arrays.

CoDA-generated ZFNs induced targeted insertion or deletion mutations in a target site present in two duplicated soybean genes with mutation frequencies of 10–18%. Cotyledons of soybean were transformed with the ZFN constructs using *A. rhizogenes* K599 to generate hairy roots where the ZFN transgenes were induced by beta-estradiol (Sander et al. 2011). A similar method was employed for the targeted mutagenesis of a transgene and nine endogenous genes in soybean using ZFNs by generating hairy roots from soybean cotyledons with *A. rhizogenes* K599 (Curtin et al. 2011). In addition to this, whole plant transformation was carried out with the disarmed *A. rhizogenes* strain K599 variant 18r12, which lacks the root-inducing genes, carrying the ZFN construct targeting *DCL4* paralogs. The ZFN transgene was induced during the co-cultivated, shoot induction, and shoot elongation steps of transformation by supplying estrogen to the tissue culture media (Curtin et al. 2011).

The CRISPR/Cas system was first tested in tomato using hairy root transformation by *A. rhizogenes* ATCC 15834 (Ron et al. 2014). The potential of the Cas9 system to induce knockouts in a pre-integrated *mGFP5* gene was tested first, followed by the ability of the system to target the endogenous transcription factor genes SHORTROOT (*SHR*) and SCARECROW (*SCR*). The presence of the shortened root phenotype in hairy roots generated using the *SHR*-specific sgRNAs demonstrated that *A. rhizogenes* and the CRISPR/Cas9 system provides a facile means to test gene function in root development (Ron et al. 2014).

A. rhizogenes strain K599 was also used for delivering the genome-editing components to soybean hypocotyls (Cai et al. 2015). The Cas9 cassette and sgRNAs targeting different sites of two endogenous genes (*GmFEI2* and *GmSHR*) or one sgRNA targeting two homologous genes (*GmFEI1* and *GmFEI2*) were assembled on one vector and delivered to generate soybean hairy roots that showed the presence of mutations in all the targets. Thus, CRISPR/Cas-mediated genome editing via *A. rhizogenes* provides a powerful tool for root-specific functional genomics studies in soybean. In addition to delivery of CRISPR/Cas genome-editing components, *A. rhizogenes* strain K599 has been used for delivery of TALENs to generate a mutation in the soybean *Dicer-like2* gene (Curtin et al. 2018), and *phytoene desaturase* (*PDS*) genes (Du et al. 2016).

5 *Agrobacterium* Engineering

5.1 *Modifications in* Agrobacterium *to Facilitate Genome Editing*

The machinery that operates within *Agrobacterium* cells, resulting in the transfer of T-DNA to plant cells, is a good candidate for manipulation to use this system as an efficient genome-editing tool. The transfer of T-DNA occurs as a single-stranded DNA-protein complex mediated by a set of virulence (Vir) proteins that are

encoded by the Ti plasmid. This transfer occurs through the Type IV secretion system (T4SS) comprising 11 VirB (VirB1 to VirB11) and VirD4 membrane-bound proteins that span the inner and outer bacterial membrane (Christie 2001; Lai and Kado 2000; Zupan et al. 1998). VirD2, along with VirD1, nicks the left and right borders of the T-DNA region to generate a single-strand T-DNA molecule called the T-strand. VirD2 attaches covalently to the 5' end of the T-strand at the right border (Ward and Barnes 1988), and the VirD2/T-strand complex is exported through the T4SS channel along with several other effector proteins VirF, VirE2, VirE3, and VirD5 (Schrammeijer et al. 2003; Vergunst et al. 2000, 2005). The carboxy termini of the effector proteins contain the T4SS translocation signal (Atmakuri et al. 2003; Simone et al. 2001; Vergunst et al. 2000). The C-terminal transport signal for recruitment and translocation of effector proteins is predicted to have a net positive charge and a consensus motif of R-X(7)-R-X-R-X-R-X-X(n) (Vergunst et al. 2005) and has the potential of translocating various genome-editing reagents as a fusion protein from *A. tumefaciens* to the target plant cell. Seminal work was done by fusing Cre recombinase to VirE2 or VirF to demonstrate protein translocation into plant cells (Vergunst et al. 2000). Transgenic *Arabidopsis thaliana* plants containing a *loxP*-flanked DNA segment that prevented expression of an *nptII* marker gene were used for transformation with *A. tumefaciens* strains LBA1149 (*virE2*::Tn3HoHo1) and LBA2561 (ΔvirF) containing the Cre::VirE2 fusion protein and nuclear localization signal (NLS)::Cre::VirFΔ42 N fusion protein, respectively. The efficient transfer of the Cre::Vir fusion proteins was visualized as kanamycin-resistant calli produced by the root explants as a result of deletion of the *loxP*-flanked DNA and fusion of the 35S promoter region to the *nptII* gene (Vergunst et al. 2000). Using GFP fluorescence as a readout for Cre-Vir fusion protein transfer from *A. tumefaciens* to the plant, the translocation of VirD2 and VirD5 was demonstrated (Vergunst et al. 2005). Thus, the ability of *A. tumefaciens* effector proteins to translocate to plant cells as fusion proteins could be leveraged to transfer ZFNs, Cas9, or other genome-editing reagents along with T-DNA containing donor DNA sequences to cells containing target DNA sequences.

C-terminal fusions of zinc-finger moieties combined with the nuclease domain of *Fok*I (to create ZFN-type domains) with VirD2 followed by the C-terminus of VirF (VirD2-ZFN-VirF^{CT}) could be translocated through the T4SS into *Arabidopsis* root explants using the *A. tumefaciens* strain LBA2585 (Δ*VirD2*, ΔT-DNA) at the same level as their counterparts that did not contain a *Fok*I nuclease domain (van Kregten et al. 2011a). In addition, a fusion protein comprising VirD2 with the homing endonuclease I-*Sce*I followed by VirF^{CT} (VirD2-I-SceI-VirF^{CT}) was translocated into an *Arabidopsis* line carrying the I-*Sce*I target site, and showed nuclease activity at this site (van Kregten et al. 2011b). Targeted integration mediated by translocated I-*Sce*I was demonstrated during the transformation of yeast by *A. tumefaciens* (Rolloos et al. 2015). Two approaches were tested: one was based on translocation of T-strands piloted by a fusion NLS-VirD2-I-SceI-VirF^{CT} protein; the other was based on T-strand delivery via an isogenic *A. tumefaciens* strain containing a wild-type *VirD2* locus (LBA1100) concomitant with transfer of NLS-I-SceI-VirF^{CT}.

The second approach with the wild-type VirD2-directed translocation of the T-strand showed higher transformation efficiency when I-*Sce*I restriction sites were present at the target locus.

A unique method of bringing donor DNA to target DNA using the *A. tumefaciens* VirD2 protein is described by Hommelsheim et al. (2016). An important consideration during genome editing is the ability of the donor and genomic target nucleic acid sequences to be in close proximity for the duration of the editing carried out by the host repair machinery. The authors addressed this issue by designing a nucleic acid carrier molecule comprising of the general formula: M-S_1-L-W-S_2, where M and W are polypeptides that bind to the donor and target DNA, respectively, L is a linker that allows M and W flexibility, and S_1 and S_2 are signal peptides fused to M and W proteins either N- or C-terminal (Fig. 3a). They made a fusion construct containing VirD2 fused to a TAL-effector specific for GFP followed by D2TS, the VirD2 type IV translocation signal (Fig. 3b) and transformed this into a *virD2* deletion *Agrobacterium* GV3101 pM6000. The construct also contained a T-DNA region harboring a 35S promoter-driven mCherry expression cassette. Transgenic tobacco plants containing GFP in a heterozygous state were infiltrated with the *virD2* deletion strain, and protoplasts isolated after 9 days that showed strong expression of mCherry but not GFP. They also demonstrated that using the same strategy for large fusion proteins such as nucleases, most of the genome editing and programmable proteins could be reliably expressed in *Agrobacterium* and exported into the plant cells.

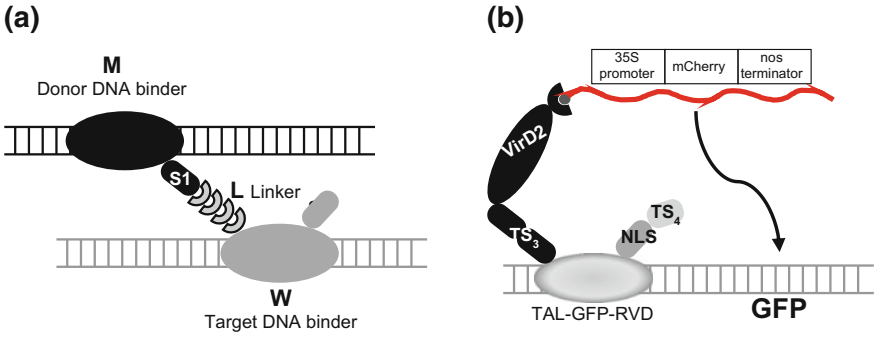


Fig. 3 Delivery of donor DNA to the target DNA using *Agrobacterium* VirD2. **a** A nucleic acid carrier molecule comprising of the general formula: M-S_1-L-W-S_2, where M and W are polypeptides that bind to the donor and target DNA, respectively, L is a linker that allows M and W flexibility, and S_1 and S_2 are signal peptides fused to M and W proteins either N- or C-terminal. **b** A fusion construct containing VirD2 fused to a Type III translocation signal (TS_3) followed by a TAL-effector specific for GFP (TAL-GFP-RVD) fused with a nuclear localization signal (NLS) followed by the VirD2 type IV translocation signal (TS_4). Upon transformation into tobacco using a VirD2 mutant *Agrobacterium* strain, the complex consisting of the VirD2/TS_3/TAL-GFP-RVD/NLS/TS_4 fusion protein covalently bound to a T-DNA harboring a 35S promoter-driven mCherry expression cassette is brought in close proximity to the GFP gene being targeted in the genome. The TAL-GFP-RVD induces a DSB disrupting GFP, and the mCherry cassette is integrated in the DSB

VirD2 is the key *A. tumefaciens* protein that is involved in processing of the T-strand and facilitating its translocation through the T4SS into the plant cell. The N-terminal 228 amino acids comprise the region that bestows the T-DNA border-specific endonuclease activity, whereas the C-terminal region contains a bipartite nuclear localization signal (NLS) and the omega (ω) region (Shurvinton et al. 1992; Narasimhulu et al. 1996; Mysore et al. 1998). Mutation of the C-terminal of VirD2 resulting in the loss of NLS and the ω region rendered the bacterium avirulent (Narasimhulu et al. 1996). However, this mutation did not abolish the translocation of the T-DNA from *Agrobacterium* to the plant cell as evident from the transient accumulation of *gusA* transcripts in tobacco suspension cells infected with *A. tumefaciens* At829 (*VirD2*Δω) lacking the omega sequence. In addition, the integration of T-DNA into high-molecular-weight plant DNA after infection with the *A. tumefaciens* wild-type strain was much higher than with the VirD2 ω mutant strain, demonstrating that the ω domain is involved in T-DNA integration (Mysore et al. 1998). Strikingly, although stable transformation decreased ∼50-fold using the VirD2 ω mutant strain, transient transformation decreased by only 4–5-fold. This observation could be the basis for the development of *A. tumefaciens* strains that will translocate genome-editing reagents on the T-DNA into the target plant cell where transient expression of the proteins could achieve the desired effect, such as induction of DSBs, without integration of the T-DNA into the plant genome.

5.2 *Improvement of Transformation Efficiency Using Additional Copies of* Vir *Genes*

As increased transformation efficiency can result in increased genome editing efficiency using *Agrobacterium*, several laboratories have improved *A. tumefaciens* strains by supplementing existing strains with extra copies of *Vir* genes from 'supervirulent' strains. A superbinary vector was developed for rice and maize transformation (Hiei et al. 1994; Ishida et al. 1996) that was an improved version of a binary vector and carried *VirB*/*VirC* operons, and the *VirG* gene from pTiBo542, which is responsible for the supervirulence phenotype of *A. tumefaciens* A281 (Jin et al. 1987). Cloning genes of interest into the superbinary vector, however, was challenging due to the relatively large vector size, and hence led to the strategy of cointegration through homologous recombination using intermediate vectors such as pSB11 and an acceptor vector such as pSB1 (Komari et al. 1996). This strategy resulted in an *A. tumefaciens* strain carrying the cointegrated superbinary vector in addition to an intrinsic disarmed plasmid (pTiAch5) containing a full set of virulence genes. However, the cointegration event generated a pair of large directly repeated sequences due to recombination between the homologous regions of pSB11 and pSB1 that could be preferred targets for intramolecular recombination, leading to deletion and other rearrangements of the T-DNA region. In addition, the

cointegrant plasmid has two distinct ColE1-type origins of replication (ori), as well as a third one, IncP-type, for replication in *Agrobacterium*. Whereas the ColE1 ori is normally non-functional in *Agrobacterium*, genomic mutations may allow for its stable maintenance in *Agrobacterium*, making it potentially highly unstable. Merlo et al. (2017) described a simplified method for development of an *Agrobacterium* strain containing additional *Vir* genes. They cloned the *VirB*/*VirC* operon, *VirG*, *VirD1* fragment of pTiBo542 on a plasmid that had an ori which was compatible with a binary plasmid containing T-DNA borders and an IncP-type ori. Further, to reduce intramolecular recombinations that could be caused by the presence of repeated sequence elements within plasmids residing in *Agrobacterium*, the authors generated RecA-deficient strains in the C58 genetic background, and introduced, separately, disarmed Ti plasmids from EHA105 (disarmed pTiBo542) and GV3101 (disarmed pTiC58). The EHA105 RecA-deficient ternary strain comprising the disarmed pTiBo542, helper plasmid with *VirB*/*VirC* operon, *VirG*, *VirD1* from pTiBo542, and a binary vector containing a gene of interest resulted in a maize transformation frequency that was six times higher than that of strain EHA105 alone. The group also developed a supervirulent strain of LBA4404 by integrating the superbinary *Vir* genes into the LBA4404 chromosome, creating a SUPERCHROME strain that was twice as efficient in transformation frequency as was the binary LBA4404 system. This RecA-deficient strain of LBA4404 was capable of maintaining 100% stability of the resident plasmid as compared to the LBA4404 wild-type strain (Gupta et al. 2017). Further refinements have been made to the ternary system by reducing the size of the vector compared to the superbinary vector, enhancing stability, improving the bacterial selection marker, and amending the *Vir* genes by adding the *VirC*, *VirD*, and *VirE* operons for improved T-DNA delivery (Anand et al. 2018).

6 Future of *Agrobacterium* in Gene Targeting

Improvement of *Agrobacterium*-mediated transformation through modifications in the *Vir* gene system has played a major role in expanding the host range of this tool. However, there are still several crops and genotypes within the crops of economic importance that are recalcitrant to *Agrobacterium*-mediated transformation. Addressing these issues of increasing the host range will likely improve the chances of genome editing and targeting in these species/cultivars. An important question to address for the future is how can we harness the ability of *Agrobacterium* to deliver T-DNA to specific target sites without random integration? One way could be to further optimize the C-terminal mutations in VirD2 to be able to deliver the T-DNA without random integration. As we run out of options to improve the *Agrobacterium* strains themselves, we can avail the treasure trove of knowledge that we have assimilated over the last couple of decades about plant genes that play a role in transformation and T-DNA integration. A recent report has identified the mechanism of T-DNA integration in plants as being dependent on

polymerase-θ-mediated DNA repair (van Kregten et al. 2016). The authors found that Pol θ mutants in *Arabidopsis* were resistant to T-DNA integration demonstrating a way to potentially disrupt the random integration of T-DNA and making the incoming T-DNA amenable to HR-mediated recombination into specific target sites. This, in combination with site-specific integration mediated through *A. tumefaciens* and the use of sequence-specific nucleases such as ZFNs, TALENs, and CRISPR/Cas to create DSBs, will herald the way toward an efficient, robust, and rapid crop genome-editing toolkit.

Acknowledgements The authors thank Prof. Stanton Gelvin for permission to include some unpublished data from his laboratory. Some of the work described here was supported by the Applied Science and Technology function of Corteva Agriscience™, Agriculture Division of DowDuPont. The authors apologize to their colleagues whose work was not cited because of space limitations.

References

Anand A, Bass SH, Wu E et al (2018) An improved ternary vector system for *Agrobacterium*-mediated rapid maize transformation. Plant Mol Biol 97:187–200

Anderson DJ, Birch RG (2012) Minimal handling and superbinary vectors facilitate efficient, *Agrobacterium*-mediated, transformation of sugarcane (*Saccharum* spp. hybrid). Trop Plant Biol 5:183–192

Atmakuri K, Ding Z, Christie PJ (2003) VirE2, a Type IV secretion substrate, interacts with the VirD4 transfer protein at cell poles of *Agrobacterium tumefaciens*. Mol Microbiol 49:1699–1713

Ayar A, Wehrkamp-Richter S, Laffaire J-B et al (2013) Gene targeting in maize by somatic ectopic recombination. Plant Biotech J 11:305–314

Bakó L, Umeda M, Tiburcio AF et al (2003) The VirD2 pilot protein of *Agrobacterium*-transferred DNA interacts with the TATA box-binding protein and a nuclear protein kinase in plants. Proc Natl Acad Sci USA 100:10108–10113

Barton KA, Binns AN, Matzke AJ et al (1983) Regeneration of intact tobacco plants containing full length copies of genetically engineered T-DNA, and transmission of T-DNA to R1 progeny. Cell 32:1033–1043

Bhaya D, Davison M, Barrangou R (2011) CRISPR-Cas systems in bacteria and archaea: versatile small RNAs for adaptive defense and regulation. Annu Rev Genet 45:273–297

Boch J, Scholze H, Schornack S et al (2009) Breaking the code of DNA binding specificity of TAL-type III effectors. Science 326:1509–1512

Cai CQ, Doyon Y, Ainley WM et al (2009) Targeted transgene integration in plant cells using designed zinc finger nucleases. Plant Mol Biol 69:699–709

Cai Y, Chen L, Liu X et al (2015) CRISPR/Cas9-mediated genome editing in soybean hairy roots. PLoS ONE 10(8):e0136064

Caldwell DG, McCallum N, Shaw P et al (2004) A structured mutant population for forward and reverse genetics in barley (*Hordeum vulgare* L.). Plant J 40:143–150

Caplan A, Herrera-Estrella L, Inze D et al (1983) Introduction of genetic material into plant cells. Science 222:815–821

Cermak T, Doyle EL, Christian M et al (2011) Efficient design and assembly of custom TALEN and other TAL effector-based constructs for DNA targeting. Nucleic Acids Res 39:e82

Char SN, Neelakandan AK, Nahampun H et al (2017) An *Agrobacterium*-delivered CRISPR/Cas9 system for high-frequency targeted mutagenesis in maize. Plant Biotech J 15:257–268

Che P, Anand A, Wu E et al (2018) Developing a flexible, high-efficiency *Agrobacterium*-mediated sorghum transformation system with broad application. Plant Biotech J 16:1388–1395

Christie PJ (2001) Type IV secretion: intercellular transfer of macromolecules by systems ancestrally related to conjugation machines. Mol Microbiol 40:294–305

Chuck G, Robbins T, Nijjar C et al (1993) Tagging and cloning of a petunia flower color gene with the maize transposable element Activator. Plant Cell 5:371–378

Citovsky V, Lee L-Y, Vyas S et al (2006) Subcellular localization of interacting proteins by bimolecular fluorescence complementation *in planta*. J Mol Biol 362:1120–1131

Cocking EC (1972) Plant cell protoplasts—isolation and development. Ann Rev Plant Physiol 23:29–50

Cong L, Ran FA, Cox D et al (2013) Multiplex genome engineering using CRISPR/Cas systems. Science 339:819–823

Curtin SJ, Xiong Y, Michno J-M et al (2018) CRISPR/Cas9 and TALENs generate heritable mutations for genes involved in small RNA processing of *Glycine max* and *Medicago truncatula*. Plant Biotechnol J 16:1125–1137

Curtin SJ, Zhang F, Sander JD et al (2011) Targeted mutagenesis of duplicated genes in soybean with zinc-finger nucleases. Plant Physiol 156:466–473

Deeken R, Engelmann JC, Efetova M et al (2006) An integrated view of gene expression and solute profiles of *Arabidopsis* tumors: a genome-wide approach. Plant Cell 18:3617–3634

de Pater S, Klemann BJPM, Hooykaas PJJ (2018) True gene-targeting events by CRISPR/Cas-induced DSB repair of the PPO locus with an ectopically integrated repair template. Sci Rep 8:3338

de Pater S, Neuteboom LW, Pinas JE et al (2009) ZFN-induced mutagenesis and gene-targeting in *Arabidopsis* through *Agrobacterium*-mediated floral dip transformation. Plant Biotechnol J 7:821–835

de Pater S, Pinas JE, Hooykaas PJJ et al (2013) ZFN-mediated gene targeting of the *Arabidopsis* protoporphyrinogen oxidase gene through *Agrobacterium*-mediated floral dip transformation. Plant Biotechnol J 11:510–515

D'Halluin K, Vanderstraeten C, van Hulle J et al (2013) Targeted molecular trait stacking in cotton through targeted double-strand break induction. Plant Biotechnol J 11:933–941

Ditt RF, Kerr KF, de Figueiredo P et al (2006) The *Arabidopsis thaliana* transcriptome in response to *Agrobacterium tumefaciens*. Mol Plant-Microbe Interact 19:665–681

Du H, Zeng X, Zhao M et al (2016) Efficient targeted mutagenesis in soybean by TALENs and CRISPR/Cas9. J Biotech 217:90–97

Endo M, Osakabe K, Ono K et al (2007) Molecular breeding of a novel herbicide-tolerant rice by gene targeting. Plant J 52:157–166

FAO/IAEA (1977) Manual on mutation breeding, 2nd edn, Technical Report Series, No. 119. Food and Agricultural Organization, International Atomic Energy Agency, Vienna, Austria

Federoff N, Wessler S, Shure M (1983) Isolation of the transposable maize controlling elements Ac and Ds. Cell 35:235–242

Feng C, Yuan J, Wang R et al (2016) Efficient targeted genome modification in maize using CRISPR/Cas9 system. J Genet Genomics 43:37–43

Forsyth A, Weeks T, Richael C et al (2016) Transcription activator-like effector nucleases (TALEN)-mediated targeted DNA insertion in potato plants. Front Plant Sci 7:1–12

Fraley RT, Rogers SG, Horsch RB et al (1983) Expression of bacterial genes in plant cells. Proc Natl Acad Sci USA 80:4803–4807

Geiser M, Weck E, Döring HP et al (1982) Genomic clones of a wild-type allele and a transposable element-induced mutant allele of the sucrose synthase gene of *Zea mays* L. EMBO J 11:1455–1460

Gelvin SB (2000) *Agrobacterium* and plant genes involved in T-DNA transfer and integration. Annu Rev Plant Physiol Plant Mol Biol 51:223–256

Gelvin SB (2003) *Agrobacterium*-Mediated Plant Transformation: the Biology behind the "Gene-Jockeying" Tool. Microbiol Mol Biol Rev 67:16–37

Gepts P (2002) A comparison between crop domestication, classical plant breeding, and genetic engineering. Crop Sci 42:1780–1790

Grandbastien MA, Spielmann A, Caboche M (1989) Tnt1, a mobile retroviral-like transposable element of tobacco isolated by plant cell genetics. Nature 337:376–380

Greene EA, Codomo CA, Taylor NE et al (2003) Spectrum of chemically induced mutations from a large-scale reverse-genetic screen in *Arabidopsis*. Genetics 164:731–740

Gupta M, Bennett S, Kumar S et al (2017) Methods and compositions for recombination a gene-deficient strains of *Agrobacterium tumefaciens*. United States Patent 9(765):350

Hansen G, Das A, Chilton M-D (1994) Constitutive expression of the virulence genes improves the efficiency of plant transformation by *Agrobacterium*. Proc Natl Acad Sci USA 91:7603–7607

Hartlerode AJ, Scully R (2009) Mechanisms of double-strand break repair in somatic mammalian cells. Biochem J 423:157–168

Hellens R, Mullineaux P, Klee H (2000) A guide to *Agrobacterium* binary Ti vectors. Trends Plant Sci 5:446–451

Herrera-Estrella L, DeBlock M, Messens E et al (1983) Chimeric genes as dominant selectable markers in plant cells. EMBO J 2:987–996

Hiei Y, Ohta S, Komari T et al (1994) Efficient transformation of rice (*Oryza sativa* L.) mediated by *Agrobacterium* and sequence analysis of the boundaries of the T-DNA. Plant J 6:271–282

Hoekema A, Hirsh PR, Hooykaas PJJ et al (1983) A binary plant vector strategy based on separation of *vir*-and T-region of the *Agrobacterium tumefaciens* Ti-plasmid. Nature 303:179–180

Hommelsheim C, Frantzeskakis L, Uelker B (2016) Improved gene targeting and nucleic acid carrier molecule, in particular for use in plants. United States Patent Application 2016/0060637

Hood EE, Gelvin SB, Melchers LS et al (1993) New *Agrobacterium* helper plasmids for gene transfer to plants. Transgenic Res 2:208–218

Hood EE, Helmer GL, Fraley RT et al (1986) The hypervirulence of *Agrobacterium tumefaciens* A281 is encoded in a region of pTiBo542 outside of T-DNA. J Bacteriol 168:1291–1301

Hussey G (1978) The application of tissue culture to the vegetative propagation of plants. Sci Prog 65:185–208

Hwang H-H, Gelvin SB (2004) Plant proteins that interact with VirB2, the *Agrobacterium tumefaciens* pilin protein, mediate plant transformation. Plant Cell 16:3148–3167

Ishida Y, Saito H, Ohta S et al (1996) High efficiency transformation of maize (*Zea mays* L.) mediated by *Agrobacterium tumefaciens*. Nat Biotechnol 14:745–750

Jackson MA, Anderson DJ, Birch RG (2013) Comparison of *Agrobacterium* and particle bombardment using whole plasmid or minimal cassette for production of high-expressing, low-copy transgenic plants. Transgenic Res 22:143–151

Jin S, Komari T, Gordon MP et al (1987) Genes responsible for the supervirulence phenotype of *Agrobacterium tumefaciens* A281. J Bacteriol 169:4417–4425

Jinek M, Chylinski K, Fonfara I et al (2012) A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. Science 337:816–821

Kanaar R, Hoeijmakers JH, van Gent DC (1998) Molecular mechanisms of DNA double strand break repair. Trends Cell Biol 8:483–489

Kanchiswamy CN (2016) DNA-free genome editing methods for targeted crop improvement. Plant Cell Rep 35:1469–1474

Kay S, Hahn S, Marois E et al (2007) A bacterial effector acts as a plant transcription factor and induces a cell size regulator. Science 318:648–651

Kim S, Kim D, Cho SW et al (2014) Highly efficient RNA-guided genome editing in human cells via delivery of purified Cas9 ribonucleoproteins. Genome Res 24:1012–1019

Kim S-I, Veena, Gelvin SB (2007) Genome-wide analysis of *Agrobacterium* T-DNA integration sites in the *Arabidopsis* genome generated under nonselective conditions. Plant J 51:779–791

Kim YG, Cha J, Chandrasegaran S (1996) Hybrid restriction enzymes: Zinc finger fusions to *Fok* I cleavage domain. Proc Natl Acad Sci USA 93:1156–1160

Kleinhofs A, Behki R (1977) Prospects for plant genome modification by nonconventional methods. Ann Rev Genet 11:79–101

Kohli A, Twyman RM, Abranches R et al (2003) Transgene integration, organization and interaction in plants. Plant Mol Biol 52:247–258

Komari T, Hiei Y, Saito Y et al (1996) Vectors carrying two separate T-DNAs for co-transformation of higher plants mediated by *Agrobacterium tumefaciens* and segregation of transformants free from selection markers. Plant J 10:165–174

Koncz C, Schell J (1986) The promoter of TL-DNA gene 5 controls the tissue-specific expression of chimaeric genes carried by a novel type of *Agrobacterium* binary vector. Mol Gen Genet 204:383–396

Krens FA, Molendijk L, Wullems GJ et al (1982) In vitro transformation of plant protoplasts with Ti-plasmid DNA. Nature 296:72–74

Lahaye T, Bonas U (2001) Molecular secrets of bacterial type III effector proteins. Trends Plant Sci 6:479–485

Lai E-M, Kado CI (2000) The T-pilus of *Agrobacterium tumefaciens*. Trends Microbiol 8:361–369

Lazo GR, Stein PA, Ludwig RA (1991) A DNA transformation competent *Arabidopsis* genomic library in *Agrobacterium*. Bio/Technology 9:963–967

Lee C-W, Efetova M, Engelmann JC et al (2009) *Agrobacterium tumefaciens* promotes tumor induction by modulating pathogen defense in *Arabidopsis thaliana*. Plant Cell 21:2948–2962

Lee KY, Lund P, Lowe K et al (1990) Homologous recombination in plant cells after *Agrobacterium*-mediated transformation. Plant Cell 2:415–425

Lee L-Y, Fang M-J, Kuang L-Y et al (2008) Vectors for multi-color bimolecular fluorescence complementation to investigate protein-protein interactions in living plant cells. Plant Methods 4:24

Levings CS III, Pring DR (1979) Mitochondrial DNA of higher plants and genetic engineering. In: Setlow JK et al (eds) Genetic engineering. Plenum Press, New York, pp 205–222

Li T, Liu B, Spalding MH et al (2012) High-efficiency TALEN-based gene editing produces disease-resistant rice. Nat Biotech 30:390–392

Lloyd A, Plaisier CL, Carroll D et al (2005) Targeted mutagenesis using zinc-finger nucleases in *Arabidopsis*. Proc Natl Acad Sci USA 102:2232–2237

Lor VS, Starker CG, Voytas DF et al (2014) Targeted mutagenesis of the tomato *PROCERA* gene using transcription activator-like effector nucleases. Plant Physiol 166:1288–1291

Louwerse JD, van Lier MCM, van der Steen DM et al (2007) Stable recombinase-mediated cassette exchange in Arabidopsis using *Agrobacterium tumefaciens*. Plant Physiol 145:1282–1293

Mahfouz MM, Li L, Shamimuzzaman M et al (2011) De novo-engineered transcription activator-like effector (TALE) hybrid nuclease with novel DNA binding specificity creates double-strand breaks. Proc Natl Acad Sci USA 108:2623–2628

Mali P, Yang L, Esvelt KM et al (2013) RNA-guided human genome engineering via Cas9. Science 339:823–826

Matsuzaki H, Nakajima R, Nishiyama J et al (1990) Chromosome engineering in *Saccharomyces cerevisiae* by using a site-specific recombination system of a yeast plasmid. J Bacteriol 172:610–618

Melchers G, Sacristan MD, Holder AA (1978) Somatic hybrid plants of potato and tomato regenerated from fused protoplasts. Carlsberg Res Commun 43:203–218

Merlo DJ, Russell SM, Retallack DN et al (2017) Method of increasing plant transformation frequency using modified strains of Agrobacteria. United States Patent 9(617):551

Meyer RS, DuVal AE, Jensen HR (2012) Patterns and processes in crop domestication: an historical review and quantitative analysis of 203 global food crops. New Phytol 196:29–48

Miller JC, Holmes MC, Wang J et al (2007) An improved zinc-finger nuclease architecture for highly specific genome editing. Nat Biotechnol 25:778–785

Morrison RA, Evans DA (1988) Haploid plants from tissue culture: new plant varieties in a shortened time frame. Bio/Technology 6:684–690

Murashige T (1974) Plant propagation through tissue cultures. Ann Rev Plant Physiol 25:135–166

Mysore KS, Bassuner B, Deng XB et al (1998) Role of the *Agrobacterium tumefaciens* VirD2 protein in T-DNA transfer and integration. Mol Plant-Microbe Interact 11:668–683

Nam J, Mysore KS, Zheng C et al (1999) Identification of T-DNA tagged *Arabidopsis* mutants that are resistant to transformation by *Agrobacterium*. Mol Gen Genet 261:429–438

Nanto K, Yamada-Watanabe K, Ebinuma H (2005) *Agrobacterium*-mediated RMCE approach for gene replacement. Plant Biotechnol J 3:203–214

Narasimhulu SB, Deng XB, Sarria R et al (1996) Early transcription of *Agrobacterium* T-DNA genes in tobacco and maize. Plant Cell 8:873–886

Newell CA (2000) Plant transformation technology: developments and applications. Mol Biotechnol 16:53–65

Olhoft PM, Flagel LE, Somers DA (2004) T-DNA locus structure in a large population of soybean plants transformed using the *Agrobacterium*-mediated cotyledonary-node method. Plant Biotechnol J 2:289–300

Ooms G, Hooykaas PJJ, Moolenaar G et al (1981) Crown gall plant tumors of abnormal morphology, induced by *Agrobacterium tumefaciens* carrying mutated octopine Ti plasmids; analysis of T-DNA functions. Gene 14:33–50

Orlowski J, Boniecki M, Bujnicki JM (2007) I-Ssp6803I: the first homing endonuclease from the PD-(D/E)XK superfamily exhibits an unusual mode of DNA recognition. Bioinformatics 23:527–530

Osakabe K, Osakabe Y, Toki S (2010) Site-directed mutagenesis in *Arabidopsis* using custom designed zinc finger nucleases. Proc Natl Acad Sci USA 107:12034–12039

Ozawa K, Wakasa Y, Ogo Y et al (2012) Development of an efficient *Agrobacterium*-mediated gene targeting system for rice and analysis of rice knockouts lacking granule-bound starch synthase (*Waxy*) and β1,2-Xylosyltransferase. Plant Cell Physiol 53:755–761

Palanichelvam K, Oger P, Clough SJ et al (2000) A second T-region of the soybean-supervirulent chrysopine type Ti plasmid pTiChry5, and construction of a fully disarmed *vir* helper plasmid. Mol Plant-Microbe Interact 13:1081–1091

Pastwa E, Blasiak J (2003) Non-homologous DNA end joining. Acta Biochim Pol 50:891–908

Paques F, Duchateau P (2007) Meganucleases and DNA double-strand break-induced recombination: perspectives for gene therapy. Curr Gene Ther 7:49–66

Puchta H (2005) The repair of double-strand breaks in plants: mechanisms and consequences for genome evolution. J Exp Bot 56:1–14

Puchta H, Dujon B, Hohn B (1993) Homologous recombination in plant cells is enhanced by in vivo induction of double strand breaks into DNA by a site-specific endonuclease. Nucleic Acids Res 21:5034–5040

Rhodes CA, Pierce DA, Mettler IJ, Mascarenhas D, Detmer JJ (1988) Genetically transformed maize plants from protoplasts. Science 240:204–207

Rolloos M, Hooykaas PJJ, van der Zaal BJ (2015) Enhanced targeted integration mediated by translocated I-SceI during the *Agrobacterium* mediated transformation of yeast. Sci Rep 5:8345

Römer P, Hahn S, Jordan T et al (2007) Plant pathogen recognition mediated by promoter activation of the pepper *Bs3* resistance gene. Science 318:645–648

Ron M, Kajala K, Pauluzzi G et al (2014) Hairy root transformation using *Agrobacterium rhizogenes* as a tool for exploring cell type-specific gene expression and function using tomato as a model. Plant Physiol 166:455–469

Rosen LE, Morrison HA, Masri S et al (2006) Homing endonuclease I-*Cre*I derivatives with novel DNA target specificities. Nucleic Acids Res 34:4791–4800

Sander JD, Dahlborg EJ, Goodwin MJ et al (2011) Selection-free zinc-finger-nuclease engineering by context-dependent assembly (CoDA). Nat Methods 8:67–69

Sander JD, Joung JK (2014) CRISPR-Cas systems for editing, regulating and targeting genomes. Nat Biotechnol 32:347–355

Sardesai N, Lee L-Y, Chen H et al (2013) Cytokinins secreted by *Agrobacterium* promote transformation by repressing a plant Myb transcription factor. Science Sig 6(302):ra100

Schrammeijer B, den Dulk-Ras A, Vergunst AC et al (2003) Analysis of Vir protein translocation from *Agrobacterium tumefaciens* using *Saccharomyces cerevisiae* as a model: evidence for transport of a novel effector protein VirE3. Nucleic Acids Res 31:860–868

Schrammeijer B, Risseeuw E, Pansegrau W et al (2001) Interaction of the virulence protein VirF of *Agrobacterium tumefaciens* with plant homologs of the yeast Skp1 protein. Curr Biol 11:258–262

Shrawat AK, Lorz H (2006) *Agrobacterium*-mediated transformation of cereals: a promising approach crossing barriers. Plant Biotechnol J 4:575–603

Shurvinton CE, Hodges L, Ream W (1992) A nuclear localization signal and the C-terminal omega sequence in the *Agrobacterium tumefaciens* VirD2 endonuclease are important for tumor formation. Proc Natl Acad Sci USA 89:11837–11841

Simone M, McCullen CA, Stahl LE et al (2001) The carboxy-terminus of VirE2 from *Agrobacterium tumefaciens* is required for its transport to host cells by the *virB*-encoded type IV transport system. Mol Microbiol 41:1283–1293

Sleper DA, Poehlman JM (2006) Breeding field crops, 5th edn. Blackwell Publishing, Ames, IA

Steinert J, Schiml S, Puchta H (2016) Homology-based double-strand break-induced genome engineering in plants. Plant Cell Rep 35:1429–1438

Svitashev S, Schwartz C, Lenderts B et al (2016) Genome editing in maize directed by CRISPR-Cas9 ribonucleoprotein complexes. Nat Commun 7:13274

Svitashev S, Young JK, Schwartz C et al (2015) Targeted mutagenesis, precise gene editing, and site-specific gene insertion in maize using Cas9 and guide RNA. Plant Physiol 169:931–945

Tzfira T, Vaidya M, Citovsky V (2001) VIP1, an *Arabidopsis* protein that interacts with *Agrobacterium* VirE2, is involved in VirE2 nuclear import and *Agrobacterium* infectivity. EMBO J 20:3596–3607

Tzfira T, Frankman LR, Vaidya M et al (2003) Site-specific integration of *Agrobacterium tumefaciens* T-DNA via double-stranded intermediates. Plant Physiol 133:1011–1023

UN WPP (2015) United Nations, Department of Economic and Social Affairs, Population Division. World population prospects 2015—Data booklet (ST/ESA/SER.A/377) (https://esa.un.org/unpd/wpp/Publications/Files/WPP2015_DataBooklet.pdf)

Urnov FD, Rebar EJ, Holmes MC et al (2010) Genome editing with engineered zinc finger nucleases. Nat Rev Genet 11:636–646

van Kregten M, de Boer P, Hooykaas PJJ et al (2011a) Translocation of novel recombinant effector proteins from *Agrobacterium tumefaciens* to *Arabidopsis thaliana*. In: VirD2 of *Agrobacterium tumefaciens*: functional domains and biotechnological applications. Ph.D. thesis, Leiden University

van Kregten M, de Pater S, Romeijn R et al (2016) T-DNA integration in plants results from polymerase-θ-mediated DNA repair. Nat Plants 2:16164

van Kregten M, Hooykaas PJ, van der Zaal BJ (2011b) *Agrobacterium*-mediated delivery of a meganuclease into target plant cells. In: VirD2 of *Agrobacterium tumefaciens*: functional domains and biotechnological applications. Ph.D. thesis, Leiden University

Veena, Jiang H, Doerge RW et al (2003) Transfer of T-DNA and Vir proteins to plant cells by *Agrobacterium tumefaciens* induces expression of host genes involved in mediating transformation and suppresses host defense gene expression. Plant J 35:219–236

Vergunst AC, Jansen LE, Hooykaas PJ (1998) Site-specific integration of *Agrobacterium* T-DNA in *Arabidopsis thaliana* mediated by Cre recombinase. Nucleic Acids Res 26:2729–2734

Vergunst AC, Schrammeijer B, den Dulk-Ras A et al (2000) VirB/D4-dependent protein translocation from *Agrobacterium* into plant cells. Science 290:979–982

Vergunst AC, van Lier MCM, den Dulk-Ras A et al (2005) Positive charges an important feature of the C-terminal transport signal of the VirB/D4-translocated proteins of *Agrobacterium*. Proc Natl Acad Sci USA 102:832–837

Voytas DF, Ausubel FM (1988) A copia-like transposable element family in *Arabidopsis thaliana*. Nature 336:242–244

Ward ER, Barnes WM (1988) VirD2 protein of *Agrobacterium tumefaciens* very tightly linked to the 5′ end of T-strand DNA. Science 242:927–930

Weinthal DM, Taylor RA, Tzfira T (2013) Nonhomologous end joining-mediated gene replacement in plant cells. Plant Physiol 162:390–400

Xing H-L, Dong L, Wang Z-P et al (2014) A CRISPR/Cas9 toolkit for multiplex genome editing in plants. BMC Plant Biol 14:327

Zhang F, Maeder ML, Unger-Wallace E et al (2010) High frequency targeted mutagenesis in *Arabidopsis thaliana* using zinc finger nucleases. Proc Natl Acad Sci USA 107:12028–12033

Zhang Y, Liang Z, Zong Y et al (2016) Efficient and transgene-free genome editing in wheat through transient expression of CRISPR/Cas9 DNA or RNA. Nat Commun 7:12617

Zhou X, Jacobs TB, Xue L-J et al (2015) Exploiting SNPs for biallelic CRISPR mutations in the outcrossing woody perennial *Populus* reveals 4-coumarate:CoA ligase specificity and redundancy. New Phytol 208:298–301

Zhu J, Song N, Sun S et al (2016) Efficiency and inheritance of targeted mutagenesis in maize using CRISPR-Cas9. J Genet Genomics 43:25–36

Zhu Y, Nam J, Humara JM et al (2003) Identification of *Arabidopsis rat* mutants. Plant Physiol 132:494–505

Zupan JR, Ward D, Zambryski PC (1998) Assembly of the VirB transport complex for DNA transfer from *Agrobacterium tumefaciens* to plant cells. Curr Opin Microbiol 1:649–655

Advancing *Agrobacterium*-Based Crop Transformation and Genome Modification Technology for Agricultural Biotechnology

Ajith Anand and Todd J. Jones

Contents

1 Introduction 490
2 Dawn of Genotype-Independent Plant Transformation Technology 490
2.1 Cell Proliferation Factors 491
2.2 Rapid *Agrobacterium*-Mediated Cereal Transformation 491
2.3 Improved Vector Systems 492
3 Transgenic Event Quality 494
3.1 Quality Events: Breakdown of Event Quality in Maize 495
3.2 Genomic Insertion Site and Transgene Expression 496
4 Gene Targeting and Genome Modification Using AMT 496
4.1 *Agrobacterium*-Mediated Site-Specific Integration in Maize 497
4.2 Application of Agro-mediated CRISPR/Cas Technology for Maize, Sorghum, and Soybean Genome Editing 499
5 Non-*Agrobacterium*-Based Plant Transformation 500
6 Looking Forward 501
References 502

Abstract The last decade has seen significant strides in *Agrobacterium*-mediated plant transformation technology. This has not only expanded the number of crop species that can be transformed by *Agrobacterium*, but has also made it possible to routinely transform several recalcitrant crop species including cereals (e.g., maize, sorghum, and wheat). However, the technology is limited by the random nature of DNA insertions, genotype dependency, low frequency of quality events, and variation in gene expression arising from genomic insertion sites. A majority of these deficiencies have now been addressed by improving the frequency of quality

A. Anand (✉) · T. J. Jones
Corteva Agriscience™, Agriculture Division of DowDuPont™,
8325 NW 62nd Avenue, Johnston, IA, USA
e-mail: ajith.anand@pioneer.com

Current Topics in Microbiology and Immunology (2018) 418:489–507
DOI 10.1007/82_2018_97

Published Online: 30 June 2018

events, developing genotype-independent transformation capability in maize, developing an *Agrobacterium*-based site-specific integration technology for precise gene targeting, and adopting *Agrobacterium*-delivered CRISPR-Cas genes for gene editing. These improved transformation technologies are discussed in detail in this chapter.

1 Introduction

Sustainable crop production faces challenges from an increasing population, the emerging societal and dietary changes in developing countries, climate change, and limited arable land around the world. Agricultural biotechnology has played a pivotal role in sustainable crop production with the introduction of crops that are insect- and/or herbicide-tolerant. A comprehensive meta-analysis of data on the use of crops with biotech traits indicates that worldwide pesticide use has been reduced by 37% while yields, on average, increased up to 25%, while lowering levels of mycotoxins (Pellegrino et al. 2018). Plant transformation is a core technology for agricultural biotechnology, making it possible to insert foreign DNA sequences into the plant genomes. The two most common methods for plant transformation are *Agrobacterium*-mediated transformation (AMT) and particle bombardment (biolistics). AMT is the preferred method for plant genetic engineering, mainly due to its simple operation, high reproducibility, capacity to transfer large (30–150 kb) DNA fragments, and proclivity to produce low copy events (see reviews Gelvin 2003, 2009; Komari et al. 2004). However, AMT still is limited in its application across a narrow range of genotypes within a species, and some crops remain recalcitrant to transformation by *Agrobacterium* (Altpeter et al. 2016). While AMT is an effective technology, the random nature of transgene integration, low frequency of quality events (QE = single copy, vector backbone negative), and event-to-event variation present major challenges for product development using AMT. In this chapter, we discuss approaches to address some of these challenges, including the use of gene-targeting technologies for precise genome modification and integration.

2 Dawn of Genotype-Independent Plant Transformation Technology

The major limitation of AMT for commercial crop genome engineering is the limited ability to transform elite commercial genotypes. This limitation can be due to a poor response to conventional tissue culture techniques and/or poor gene delivery and integration (Altpeter et al. 2016). There are several potential ways to improve cell and tissue culture response, including choosing the appropriate

explant, manipulating tissue culture nutrients, altering plant growth regulators, and incorporating media supplements. These approaches have all been extensively discussed elsewhere. In this chapter, we will review alternate approaches and recent advances for improving crop plant transformation efficiency.

2.1 Cell Proliferation Factors

One alternative to modifying tissue culture parameters is to express plant, or non-plant, growth-stimulating transgenes, transcription factors, stress factors, or signaling pathway genes. The heterologous expression of growth-stimulating genes such as *ipt*, *iaaM*, or *rol* improved dicot plant transformation (Ebinuma et al. 1997, 2005; Gordon-Kamm et al. 2002). The developmental reprogramming of somatic cells to form embryogenic cells, the very basis of cellular totipotency, is controlled by a complex molecular system involving somatic embryogenesis-related genes (Zeng et al. 2007). Somatic embryogenesis involves several characteristic events: cellular dedifferentiation, cell division, reprogramming of cell physiology, metabolism, and regulation of gene expression patterns (Yang and Zhang 2010). The ectopic expression of several transcription factors, such as LEAFY COTYLEDON1, (LEC1; (Lotan et al. 1998), LEAFY COTYLEDON2, (LEC2; (Stone et al. 2001), WUSCHEL (WUS; (Zuo et al. 2002), and BABY BOOM (BBM; (Boutilier et al. 2002), has been described to improve plant regeneration via stimulating somatic embryogenesis (Deng et al. 2009; Srinivasan et al. 2006). The role of stress factors has been implicated in the dedifferentiation of plant cells and subsequent callus induction from somatic cells (Fehér 2015; Florentin et al. 2013; Grafi and Barak 2015). Several signal transduction pathway genes involving auxin response factors (Fan et al. 2012); cytokinin type-B Arabidopsis response regulators (Tajima et al. 2004); receptor-like protein kinase 1 (Motte et al. 2014), and somatic embryogenesis receptor kinases (SERKs; (Singh and Khurana 2017) have been implicated in maintaining plant growth and development. With no direct evidence to suggest that signal transduction pathway genes improve tissue culture response or enhance plant regeneration, their role in plant transformation remains uncertain. The knowledge of factors controlling tissue response and improved plant regeneration has been limited in most crop species (Lowe et al. 2016) and continues to present a challenge.

2.2 *Rapid* Agrobacterium-*Mediated Cereal Transformation*

Even though many dicot species can be routinely transformed, genetic transformation of cereals requires multiple steps in tissue culture, predominantly with an intermediate callus stage followed by embryogenesis or shoot organogenesis and subsequent root formation (Ji et al. 2013; Que et al. 2014). The need to use

immature embryos (IE) or IE-derived embryogenic cultures as target explants for transformation is a further constraint and is often genotype dependent (Ji et al. 2013). Callus-based approaches for transformation, such as is common for rice (Toki 1997), are usually genotype dependent and prone to somaclonal variation (Ji et al. 2013; Wei et al. 2016). Meristem cells have also been investigated as explants for transformation. Although some early results were promising, meristem transformation has rarely been used due to low transformation frequency and the frequent production of chimeric plants (Ji et al. 2013).

In a recent publication, Lowe et al. (2016) reported ectopic expression of the morphogenic genes *Bbm* and *Wus2* to stimulate somatic embryogenesis in monocot species to overcome genotype and callus dependency. The use of morphogenic genes also allows the use of alternative target tissues, such as mature seeds and leaf segments, as explants for AMT with the potential for automation (Lowe et al. 2016). Another significant breakthrough came with the development of a rapid genotype-independent maize transformation protocol via direct somatic embryogenesis. By co-expressing *Bbm* driven by a maize phospholipid transferase promoter (Zm-PLTP Pro) and *Wus2* driven by a maize auxin-inducible promoter (Zm-AXIG1 Pro), Lowe and colleagues (Lowe et al. 2018) recovered stable transgenic maize plants from more than 22 Pioneer elite inbreds as well as the public lines B73 and Mo17. Transgenic plants were generated at very high frequencies via a callus-free transformation method within a month after initiation. The use of tissue specifically expressed morphogenic genes for plant transformation is a breakthrough in monocot transformation with the potential to boost transformation rates in a range of genotypes. This technology has enabled the development of a selectable marker-free transformation system (Mookkan et al. 2017), genome editing in maize with ribonucleoproteins (Svitashev et al. 2016), creation of new and improved waxy corn hybrids (Chilcoat et al. 2017), and creation of a complex trait locus (CTL) for gene insertion through the application of CRISPR-Cas genome editing (Chilcoat et al. 2017). The full potential of the morphogenic genes in plant transformation will soon be realized with its wider application for crop genome engineering and genome modification (e.g., soybean, canola, wheat, sorghum).

2.3 *Improved Vector Systems*

The efficiency of AMT in monocot species is dependent on technical factors and the genetic makeup of the explant used. Technical factors include explant preparation, *Agrobacterium* strain, infection conditions, culture regimes, media composition, timing of plant growth regulators, selectable markers employed, and the choice of specific binary vectors (reviewed in Cheng et al. 2004; Que et al. 2014). A breakthrough in AMT was achieved with the introduction of the super-binary plasmid pSB1 (Komari et al. 1996, 2006) that greatly improved cereal transformation and broadened the host range of plants amenable to transformation with *Agrobacterium tumefaciens* (Cheng et al. 1997; Cho et al. 2014; Hiei et al. 1994;

Ishida et al. 1996; Tingay et al. 1997; Wu et al. 2014; Zhi et al. 2015). However, the large size of the plasmid pSB1 (~37 kb) and the co-integration step required for generation of the T-DNA binary vector complicated vector construction and structural confirmation of plasmid integrity, restricting the application of the vector system for high-throughput vector construction (Anand et al. 2018). A ternary vector system harboring a disarmed Ti plasmid, an accessory plasmid with additional *Agrobacterium* virulence (vir) gene(s), and the T-DNA launched from an independent binary vector is a notable option to simplify vector construction. A similar design was demonstrated to improve AMT of dicots (Kessler and Baldwin 2002; van der Fits et al. 2000). However, there are limited descriptions of a similar system for monocot transformation. In an attempt to address the deficiencies in plasmid pSB1, a series of pVIR plasmids (Fig. 1a) were created featuring a smaller replicon, corrected *vir* genes (corrected a frame shift in the *virC* operon, and replaced a truncated *virD2* gene with a functional gene), a superior bacterial selectable marker, and compatibility with Gateway™ cloning technology (Invitrogen) (Anand et al. 2017a, b). Different pVIR plasmids have been designed that can serve both as a T-DNA vector and as an accessory plasmid in ternary vector systems (Anand et al. 2017a, b, 2018).

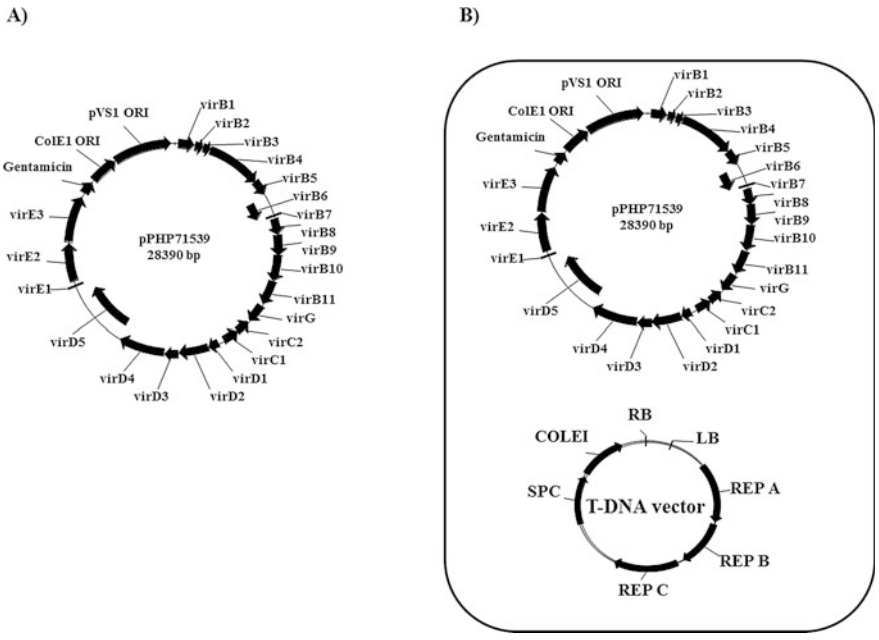


Fig. 1 Design of a pVIR plasmid (**a**) and the ternary vector design (**b**) containing a disarmed Ti plasmid, the pVIR as an accessory plasmid, and a T-DNA binary vector used for maize, sorghum, and wheat transformation

Table 1 Average stable transformation frequency with the ternary vectors containing the accessory pVIR plasmid pPHP71539 and three different T-DNA vectors with proprietary trait gene cassettes in the elite maize inbred PH2RT

Accessory plasmid in the ternary design	Total embryos	Total T0 events	T0 transformation frequency (%) ± Standard error
pSB1	1609	220	13.7 ± 1.6
pPHP71539	1691	526	31.1 ± 2.5

The use of pVIR vectors as accessory plasmids enabled a switch from large co-integrate super-binary vectors (often referred to as co-integrate vectors or CIVs) to simple T-DNA binary vectors for transforming multiple cereal crops including maize, sorghum, and wheat (Anand et al. 2017b). The design improved construct throughput and simplified the process of ensuring vector quality control. Using a ternary vector system with the accessory pVIR plasmids (Fig. 1b) for sorghum transformation, Che and colleagues demonstrated highly efficient transformation in the sorghum line TX430 (Che et al. 2018). Additionally, they were able to transform multiple recalcitrant African sorghum varieties and reported the first targeted genome editing in sorghum using CRIPSR-Cas9 (Che et al. 2018). In maize, ternary vectors with different pVIR accessory plasmids improved transformation efficiency in multiple maize inbreds (Table 1; Anand et al. 2018). The same ternary design in combination with the use of T-DNA vectors containing the morphogenic genes *Bbm* and *Wus2* enabled the development of a callus-free, rapid, genotype-independent maize transformation technology (Lowe et al. 2018). Furthermore, the pVIR accessory plasmids are now routinely used in ternary vector systems supporting AMT in both dicot and monocot species. The development of pVIR ternary vectors marks a significant improvement in crop transformation technology.

3 Transgenic Event Quality

For a commercial product, consistent, multi-generational expression of the transgenes is necessary. Depending on the transformation method employed, several factors are likely to impact transgene expression over time: transgene copy number, accuracy of the insertion or additional DNA (deletions or rearrangements), presence of vector backbone elements, site of integration, and optimization of the expression cassette. Due to the inherent variability, historically, transgenic trait testing has involved generating large numbers of sister transgenic events from the same construct, to identify a "best" or "representative" event (Strauss and Sax 2016). Eliminating multiple copy, truncated, or rearranged transgene events is a necessary step requiring robust molecular analysis.

3.1 Quality Events: Breakdown of Event Quality in Maize

Transgenic events produced by conventional random transformation methods often contain complex transgene integrations which contribute to variability in gene expression between sister events (Depicker et al. 2005; Srivastava et al. 2004). Consistent transgene expression has been highly correlated with transformation events of high molecular quality (De Buck et al. 2004). Therefore, molecular screening of transgenic events is critical and relies on robust analytical tools for identifying QE. For molecular event characterization, a combination of qPCR (transgene copy determination), multiplex PCR (detecting vector backbone integration; Zhi et al. 2015), followed by Southern-by-Sequencing™ (SbS™) (Zastrow-Hayes et al. 2015) is used to determine the exact, intact transgenic insert sequence, while simultaneously detecting the presence of exogenous DNA in the events. A combination of these assays provides a robust framework for characterization and selection of QE. Typically, ~70–80% of the T0 events are discarded early in the analysis for not meeting the stringent QE criteria (Anand et al. 2018). A breakdown of non-quality events in maize suggests three distinct categories, (1) multi-copy events (>50%), (2) rearranged or truncated events (40–50% of the single-copy events), and (3) single-copy events positive for plasmid backbone insertions (3–5%). In our analysis, it is common to detect a higher percent of transformants with large (100 bp to >2 kbp) RB deletions (20–28%) compared to large LB deletions (<2%), contradicting the T-DNA integration patterns described previously (reviewed in Gelvin 2003). Our T-DNA vectors are designed with the selectable marker closer to the LB region, likely favoring the recovery of events with an intact, fully functional marker gene. Transformed plants with multiple copies of transgenes, or rearranged or truncated copies of T-DNA, have variable levels of transgene expression (reviewed in Depicker et al. 2005). Therefore, it is common practice in the industry to advance only events that meet the analytical criteria of a QE for further phenotypic analysis and field performance. Beyond these requirements for consistent phenotype, regulatory agencies also prefer transgenic events that do not disrupt endogenous genes or have an insert within 1–2 kbp of an open reading frame (EFSA GMO Panel 2015). Incorporation of SbS™ analysis in molecular characterization allows identification of nearly fully intact transgene insertions along with the exact location of insertion site (Zastrow-Hayes et al. 2015). Extensive analysis of many QE transformants by SbS™ suggests 30–50% of the transformants fail to meet the regulatory requirement, resulting in a cumulative attrition rate of >90% of the transformation event. Comprehensive molecular analysis of maize transformants generated by AMT showed less than 10% of the transformants truly pass the QE criteria for eventual product development.

3.2 Genomic Insertion Site and Transgene Expression

Differences in gene expression have also been attributed to the genomic location of the insertion. The expression differences could be explained by interactions with flanking host DNA or chromosomal location (Eszterhas et al. 2002; Matzke et al. 2000), the chromatin state, integration into an intergenic or gene region, or into an exon or an intron (Gelvin and Kim 2007; Pröls and Meyer 1992; Schnell et al. 2015). The variability arising from the genomic insertion site often confounds functional analysis of genetic elements such as promoters, protein coding genes, and non-coding RNAs. Most of the differences in gene expression between single-copy events were attributed to genomic location (Chawla et al. 2006; Day et al. 2000), even when multi-copy or silenced transformants were excluded. In light of the different experimental methodologies applied, the choice of molecular assays, and differences between species, the conclusion drawn on the effect of genomic location on gene expression is inconclusive. However, the notion that "genomic insertion location is critical to transgenic product development" historically has driven "event sorting" for identifying superior transgenic events (Mumm and Walters 2001). An extensive multi-year analysis of a large set of multiple random quality transgenic events (single-copy intact T-DNA with no vector sequences) in maize and soybean events found that genomic insertion sites had minimal impact on transgene expression (Mutti et al. in preparation). The differences among transgene expression across all sites measured as recombinant protein concentration never varied greater than twofold to threefold, which is in contrast to numerous published reports. Interestingly, even though they found a minor impact of genomic insertion site on transgene expression, this study illustrated larger effects on transgene expression arising from the choice of promoter and the cis-regulatory elements. These results suggest that the construct design is more pivotal to variability in gene expression than is the insertion site.

4 Gene Targeting and Genome Modification Using AMT

From a transgenic product development perspective, randomly generated transgenic events require a laborious process of event sorting, trait introgression, and deregulation of a trait; this requires substantial time and resources. Therefore, it is highly desirable to develop strategies that precisely target transgenes to predefined target sites that are well characterized for controlled transgene integration, gene expression, and non-interference with endogenous genes. The next and current generation of trait products will require integration of multiple transgenes, necessitating a flexible and modular gene stacking approach. Adopting precise gene-targeting (GT) strategies for genome modification and genetic engineering not only facilitates improved recovery of quality events, but significantly simplifies event sorting due to the lessened regulatory consideration and trait introgression of stacked genes. By

targeting multiple insertions at a predefined genetic locus, traits can be genetically stacked, providing a means to deliver polygenic traits to the market (Akbudak et al. 2010; Cardi and Stewart 2016; De Buck et al. 2009; Li et al. 2009; Nandy and Srivastava 2011; Nanto et al. 2009; Rinaldo and Ayliffe 2015; Srivastava and Thomson 2016). Preferred methods for GT have often relied on either of two processes: recombinase-mediated site-specific integration (SSI) or homologous recombination (HR) mediated by nucleases, to insert template DNA in plants (Lyznik et al. 2003; Ow 2007; Terada et al. 2002; Tzfira and White 2005). The SSI approach relies on recombinase-mediated cassette exchange (RMCE) to locate and exchange a segment of donor DNA into a predetermined genomic locus. The predetermined genomic locus either relies on a promoter trap or is marked by a tag consisting of two heterologous recombination target sites (RT) flanking a selectable marker gene, allowing efficient gene swapping with a compatible donor DNA also flanked by RT (Schlake and Bode 1994; Turan et al. 2010, 2013). This approach has been applied to target genes at specific locations in plants (Akbudak et al. 2010; Ebinuma et al. 2015; Li et al. 2009; Nandy and Srivastava 2011; Nanto and Ebinuma 2008; Nanto et al. 2005, 2009), often with low efficiency. The low efficiency has limited broad application of the technology. Gene targeting via HR depends on nuclease-mediated generation of double strand breaks at a specific genomic location, and co-delivery of the template donor DNA (gene edited, insertions, or replacement) for HR at the locus. Recently, sequence-specific nucleases (ZFNs, TALENs, CRISPR/Cas) have been used to induce double strand breaks at specified genomic loci followed by site-specific mutagenesis, gene integration, or gene replacement (Cardi and Stewart 2016; Puchta and Fauser 2014; Sprink et al. 2015; Weeks et al. 2016). Even though sequence-specific nucleases are now routinely used for targeted mutagenesis in plants (for reviews, see Lyznik et al. 2003; Ow 2007; Terada et al. 2002), the applicability of these reagents for precise genome integration remains plagued by low efficiency.

4.1 Agrobacterium-*Mediated Site-Specific Integration in Maize*

Direct DNA delivery methods have been preferred over *Agrobacterium*-mediated site-specific integration (Agro-SSI) in plants because larger amounts of plasmid DNA and donor template can be delivered to the nucleus (Albert et al. 1995; Louwerse et al. 2007; Srivastava and Ow 2002). However, while effective for SSI, direct DNA delivery methods often result in complex random DNA integration patterns and multi-copy insertions at the integration site (Kohli et al. 2003), making it less attractive and inefficient for commercial application. *Agrobacterium*-based delivery systems for SSI have been reported, but at very low efficiency (Louwerse et al. 2007; Nanto et al. 2005; Vergunst et al. 1998), making agro-based SSI impractical. Therefore, biolistic delivery has remained the method of choice for SSI until recently.

An efficient agro-SSI technology using the FLP/*FRT* recombinase system was developed recently in maize. Improvements in agro-based SSI were achieved by evaluating several factors including *Agrobacterium* strain, the choice of heterologous *FRT* pairs, and the use of morphogenic genes (*Bbm* and *Wus2*; Anand et al. in preparation). In maize, we have developed an SSI design utilizing a chromosomal promoter trap coupled with a DNA donor construct containing a promoter-less selectable marker (Fig. 2) for improved RMCE efficiency. Added to this design, the use of the supervirulent *A. tumefaciens* strain AGL1 resulted in a fivefold improvement (from 0.05 to 0.27%) in SSI frequency over the previously preferred strain LBA4404 for maize transformation. Introduction of the morphogenic genes *Bbm* and *Wus2* into the vector design increased efficiency an additional tenfold (from 0.27 to 2.65%). Further improvements were achieved by replacing the highly cross-reactive heterologous *FRT* pairs, *FRT1*/*FRT87*, with *FRT* pairs that showed very low cross-reactivity, *FRT1*/*FRT6*, which further boosted the SSI frequency by twofold to threefold. By combining these improvements, along with the tissue culture optimizations in an elite inbred HC69, we have developed a reliable and highly efficient agro-SSI technology for maize which can consistently achieve a 19–22.5% T0 transformation frequency (expressed as percent number of infected embryos producing transformants) with 50–60% of the events recovered producing perfect RMCE events ("clean SSI"). These RMCE events are molecularly perfect (intact *FRT* junctions flanking the intended DNA sequence of the donor DNA), inserted at defined, well-characterized locations in the genome, and have remarkably consistent gene expression (Betts et al. in preparation). This efficient protocol

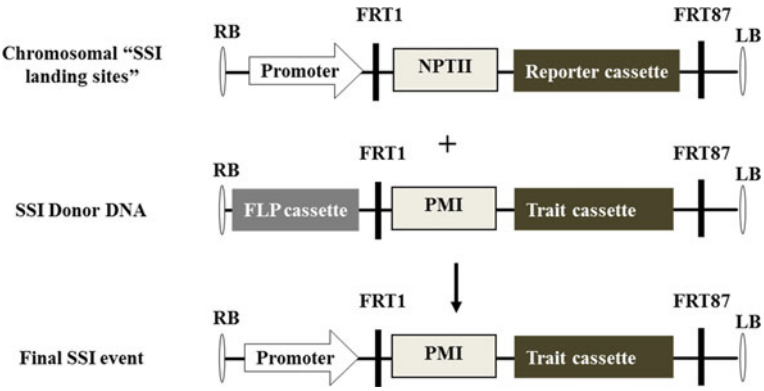


Fig. 2 Design of the SSI strategy in maize. An SSI landing site is generated containing heterologous FRT pairs either through random transformation or at target locations using CRISPR/Cas. This landing site contains an FRT-promoter trap flanking a selectable marker. Hemizygous transgenic events containing the landing sites are retransformed with a T-DNA binary vector harboring the donor DNA consisting of the FLP recombinase cassette and the template DNA containing a new selectable marker (phosphomannose isomerase, PMI) and the desired trait cassette. The FLP-mediated cassette exchange between the donor DNA and target site results in the final SSI product, containing the new selectable marker and the desired trait gene

is now routinely used in multiple maize inbreds for transgene evaluation and product development, replacing *Agrobacterium*-mediated transformation with random T-DNA integration.

4.2 Application of Agro-mediated CRISPR/Cas Technology for Maize, Sorghum, and Soybean Genome Editing

The application of CRISPR/Cas technology for genome editing and genome engineering in maize and soybean has predominantly utilized biolistics for introducing the reagents (Chilcoat et al. 2017). As with any transformation technology, biolistics has numerous problems, particularly for generating clean genome engineered events for phenotypic analysis or product development. An alternate method for plant transformation that can simplify the overall process and produce cleaner, less complicated events is desirable. Historically, AMT has long been preferred over biolistics for producing simple, clean transformants. A robust CRISPR/Cas gene-editing pipeline has been developed internally that either uses biolistics or *Agrobacterium* for CRISPR/Cas gene editing (mutagenesis or gene deletion). Gene edits have been verified at over 144 loci in multiple elite maize inbreds. The mutation rates with *Agrobacterium*-mediated CRISPR/Cas delivery are highly efficient with 90% of the guides generating plants with edits in at least 50% of the plants (Fig. 3). In the last few years, the *Agrobacterium*-mediated plant transformation frequency has significantly improved with the adoption of the rapid, morphogenic gene-enabled transformation method described earlier, and through improvements in the CRISPR/Cas reagents. A robust pipeline for generating indels or frameshifts and gene dropouts has been developed by strengthening existing bioinformatic platforms, developing genomic resources for designing gRNAs, and combining the rapid maize transformation protocol for *Agrobacterium*-delivered CRISPR/Cas9 genome engineering reagents to maize. Very recently, Che et al. (2018) used a maize-optimized Cas9 design to generate stable gene knockouts of centromere-specific histone H3 (Sb-CEN3) in the sorghum genotype TX430. They demonstrated editing efficiencies of 37–40% with monoallelic gene knockouts at 20–37%, based on the limited number of events analyzed. No biallelic knockouts were identified, consistent with the hypothesis that biallelic *Sb-CENH3* mutations would be lethal in plants. *Agrobacterium*-mediated delivery of CRISPR/Cas reagents has also been evaluated in soybean to generate gene knockouts of multiple *FAD2* and *FAD3* genes using the immature cotyledon as explants (unpublished). *Agrobacterium*-delivered CRISPR/Cas reagents were utilized for targeted mutagenesis of marker genes (Feng et al. 2016; Zhu et al. 2016) and gene families encoding Agronaute 18 and dihydroflavonol 4-reductase. Char and collaborators described developing a vector design that enables multiplexing of up to four

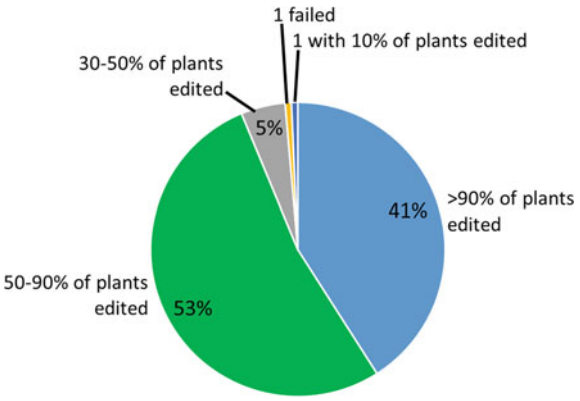


Fig. 3 Breakdown of mutation rates at the cut sites in maize using CRISPR/Cas nucleases. The data represent the percent of maize plants with targeted mutations at 144 locations in the maize genome. The data are a compilation of experiments conducted with *Agrobacterium* and biolistic-delivered Cas9, with over 80% of the events regenerated with *Agrobacterium*-delivered CRISPR/Cas in multiple maize genotypes

gRNAs, and a method for co-transformation using two different *Agrobacterium* strains harboring different gRNAs for mutating multiple genes (Char et al. 2017). The mutation rates based on indels or frameshifts ranged anywhere from 2.8 to 58% across multiple genes tested.

5 Non-*Agrobacterium*-Based Plant Transformation

While AMT is a widely used technique for gene transfer in plants, there has been significant interest in developing non-*Agrobacterium*-based gene delivery systems due to the complexity of the *Agrobacterium* patent landscape and the prohibitive costs of deregulating products engineered with *A. tumefaciens*, which is considered a plant pathogen. The use of non-AMT approaches for plant transformation was reported by Broothaerts and colleagues in 2005, where they demonstrated plant transformation using different *Rhizobia* species such as *Sinorhizobium meliloti*, *Mesorhizobium loti*, and NGR 234 (collectively called Transbacter; Broothaerts et al. 2005). The above *Rhizobia* species required the acquisition of a disarmed Ti plasmid (EHA105) and a binary vector to be competent for plant transformation. The use of Transbacter-mediated plant transformation technology for crop biotechnology has been limited. Two alternate soil-related, non-pathogenic bacteria, *Ensifer adhaerens* (OV14) and *Ochrobactrum haywardense* (H1, NRRL B-67078), can genetically modify plants and provide viable options to overcome the restrictions of *Agrobacterium* (Anand et al. 2017a, b; Wendt et al. 2012). *E. adhaerens* (OV14)-mediated transformation (EMT) utilizing pCAMBIA binary vectors was

demonstrated in dicot species (Rathore et al. 2016; Wendt et al. 2012) and as well as in two japonica rice varieties, Curinga and Nipponbare, and an indica variety, IR64 (Zuniga-Soto et al. 2015). However, the efficacy of EMT was lower compared to that of AMT. *Ochrobactrum haywardense* (H1)-mediated plant transformation (OMT) was reported to work in both dicot and monocot species including tobacco, *Arabidopsis*, soybean, and sorghum (Anand et al. 2017a, b). The introduction of a pVIR binary plasmid enabled development of a gene transfer competent strain of *O. haywardense*. Successful OMT has been reported using half-seeds and embryonic axes in soybean, tobacco leaf, *Arabidopsis* floral dip, and transient transformation of sorghum leaf discs (Anand et al. 2017a). These two new microbes have opened new avenues for crop genome engineering and crop genome modification, which may simplify the regulatory process and be more cost-effective compared to AMT-derived cultivars.

6 Looking Forward

The technical breakthroughs made in the last decade in utilizing *Agrobacterium*-based technologies for plant genetic engineering and plant genome editing have opened the doors for precision plant genome engineering. The development of robust and efficient genotype-independent AMT of maize and sorghum, supported by vigorous molecular event characterization, has allowed rapid progress in the genetic engineering of these crops. Even though random transform by AMT is an attractive technology for functional characterization of a gene or for rapid screening of candidate genes in a trait testing pipeline, random integration of events and screening of multiple sister events are a major limitation for this technology.

With the progress made in application of morphogenic genes and improvements in vector systems, we are close to realizing a germplasm-independent transformation system of maize. The morphogenic genes and orthologs currently being evaluated to facilitate transformation using alternate explants, such as maize leaves, may potentially be extended to most other crops, cultivars, and genotypes. However, there are limitations to AMT technology, including building the capacity for a crop-independent transformation technology, developing a robust and efficient transformation process, rapidly moving from random to precision genome integration, and advancing genome modification technology to economically important crops. A robust genotype-independent transformation technology may be achieved by broadening the application of morphogenic genes to diverse crops, which partially has been demonstrated through modulating gene delivery and expression by advancing cassette optimization and transient T-DNA delivery (Lowe et al. 2018). The limitations of random DNA integration are likely to be addressed by deployment of targeted DNA insertions into predefined sites mediated by SSI, and by application of sequence-specific nucleases (ZFNs, TALENs, CRISPR/Cas) to induce double strand breaks to support site-specific mutagenesis, gene integration, and gene replacement. The potential for SSI to facilitate stacking of multiple genes

into complex trait loci (CTL), or to replace genes at will in elite germplasm, has the potential to be a quantum leap in plant biotechnology.

A large question is how the technologies such as SSI, CTLs, and genome editing for introducing novel traits or genome modification of plant genomes will be viewed by the public and governmental organizations. Clearly, plants with transgenic traits, those with introduced DNA, and, presumably, template-based edits, will continue to be regulated by governmental agencies. However, what about new traits introduced into well-characterized and previously de-regulated loci in a CTL? Will they be subject to the same current extensive scrutiny? Is there an opportunity to lessen the regulatory burden using SSI? With respect to genome editing, the USDA has already stated that simple gene knockouts and deletions will not be subject to regulation. Recently, an advocate in the European Court of Justice issued an opinion that gene-edited plants may not be regulated by the same rules used for genetically modified organisms. While preliminary, these recent events provide some hope that genome editing with tools such as CRISPR-Cas may not face the same regulatory hurdles as those currently in place for product development of biotech crops.

Acknowledgements The work described here was supported by the Applied Science and Technology organization including Vector Construction, Crop Genome Engineering, Controlled Environments, Genomics and Nucleic Acid Analysis at Corteva Agriscience™, Agriculture Division of DowDuPont™. Individual projects were led by Keith Lowe, William Gordon-Kamm, Emily Wu, Ping Che, Jeffery Sander, and Scott Betts.

References

Akbudak MA, More AB, Nandy S, Srivastava V (2010) Dosage-dependent gene expression from direct repeat locus in rice developed by site-specific gene integration. Mol Biotechnol 45:15–23

Albert H, Dale EC, Lee E, Ow DW (1995) Site-specific integration of DNA into wild-type and mutant sites placed in the plant genome. Plant J 7:649–659

Altpeter F, Springer NM, Bartley LE, Blechl AE, Brutnell TP, Citovsky V, Conrad LJ, Gelvin SB, Jackson DP, Kausch AP, Lemaux PG, Medford JI, Orozco-Cárdenas ML, Tricoli DM, Van Eck J, Voytas DF, Walbot V, Wang K, Zhang ZJ, Stewart CN (2016) Advancing crop transformation in the era of genome editing. Plant Cell 28:1510

Anand A, Bass SH, Bertain SM, Cho HJ, Kinney AJ, Klein TM, Lassner M, McBride KE, Moy Y, Rosen BAM, Wei JZ (2017a) *Ochrobactrum*-mediated transformation of plants. In: WO/2017/040343

Anand A, Bass SH, Cho HJ, Klein TM, Lassner M, McBride KE (2017b) Methods and compositions of improved plant transformation. In: WO/2017/078836 A1

Anand A, Bass SH, Wu E, Wang N, McBride KE, Annaluru N, Miller M, Hua M, Jones TJ (2018) An improved ternary vector system for *Agrobacterium*-mediated rapid maize transformation. Plant Mol Biol. https://doi.org/10.1007/s11103-018-0732-y

Boutilier K, Offringa R, Sharma VK, Kieft H, Ouellet T, Zhang L, Hattori J, Liu C-M, van Lammeren AAM, Miki BLA, Custers JBM, van Lookeren Campagne MM (2002) Ectopic expression of baby boom triggers a conversion from vegetative to embryonic growth. Plant Cell 14:1737–1749

Broothaerts W, Mitchell HJ, Weir B, Kaines S, Smith LMA, Yang W, Mayer JE, Roa-Rodríguez C, Jefferson RA (2005) Gene transfer to plants by diverse species of bacteria. Nature 433:629–633

Cardi T, Neal Stewart C (2016) Progress of targeted genome modification approaches in higher plants. Plant Cell Rep 35:1401–1416

Char SN, Neelakandan AK, Nahampun H, Frame B, Main M, Spalding MH, Becraft PW, Meyers BC, Walbot V, Wang K, Yang B (2017) An *Agrobacterium*-delivered CRISPR/Cas9 system for high-frequency targeted mutagenesis in maize. Plant Biotechnol J 15:257–268

Chawla R, Ariza-Nieto M, Wilson AJ, Moore SK, Srivastava V (2006) Transgene expression produced by biolistic-mediated, site-specific gene integration is consistently inherited by the subsequent generations. Plant Biotechnol J 4:209–218

Che P, Anand A, Wu E, Sander JD, Simon MK, Zhu W, Sigmund AL, Zastrow-Hayes G, Miller M, Liu D, Lawit SJ, Zhao Z-Y, Albertsen MC, Jones TJ (2018) Developing a flexible, high-efficiency *Agrobacterium*-mediated sorghum transformation system with broad application. Plant Biotechnol J. https://doi.org/10.1111/pbi.12879

Cheng M, Fry JE, Pang S, Zhou H, Hironaka CM, Duncan DR, Conner TW, Wan Y (1997) Genetic transformation of wheat mediated by *Agrobacterium tumefaciens*. Plant Physiol 115:971–980

Cheng M, Lowe BA, Spencer TM, Ye X, Armstrong CL (2004) Factors influencing *Agrobacterium*-mediated transformation of monocotyledonous species. In Vitro Cell Devel Biol Plant 40:31–45

Chilcoat D, Liu Z-B, Sander J (2017) Chapter two—Use of CRISPR/Cas9 for crop improvement in maize and soybean. In: Weeks DP, Yang B (eds) Progress in Molecular Biology and Translational Science. Academic Press, Cambridge, pp 27–46

Cho M-J, Wu E, Kwan J, Yu M, Banh J, Linn W, Anand A, Li Z, TeRonde S, Register J III, Jones T, Zhao Z-Y (2014) *Agrobacterium*-mediated high-frequency transformation of an elite commercial maize (*Zea mays* L.) inbred line. Plant Cell Rep 33:1767–1777

Day CD, Lee E, Kobayashi J, Holappa LD, Albert H, Ow DW (2000) Transgene integration into the same chromosome location can produce alleles that express at a predictable level, or alleles that are differentially silenced. Genes Dev 14:2869–2880

De Buck S, Windels P, De Loose M, Depicker A (2004) Single-copy T-DNAs integrated at different positions in the *Arabidopsis* genome display uniform and comparable *β-glucuronidase* accumulation levels. Cell Mol Life Sci 61:2632–2645

De Buck S, Podevin N, Nolf J, Jacobs A, Depicker A (2009) The T-DNA integration pattern in *Arabidopsis* transformants is highly determined by the transformed target cell. Plant J 60:134–145

Deng W, Luo K, Li Z, Yang Y (2009) A novel method for induction of plant regeneration via somatic embryogenesis. Plant Sci 177:43–48

Depicker A, Sanders M, Meyer P (2005) Transgene silencing. In: Meyer P (ed) Plant epigenetics. Blackwell Publishing Ltd., Oxford, pp 1–31

Ebinuma H, Sugita K, Matsunaga E, Yamakado M (1997) Selection of marker-free transgenic plants using the isopentenyl transferase gene. Proc Natl Acad Sci 94:2117–2121

Ebinuma H, Sugita K, Endo S, Matsunaga E, Yamada K (2005) Elimination of marker genes from transgenic plants using MAT vector systems. In: Peña L (ed) Methods in molecular biology. Human Press, Totowa, NJ, pp 237–253

Ebinuma H, Nakahama K, Nanto K (2015) Enrichments of gene replacement events by *Agrobacterium*-mediated recombinase-mediated cassette exchange. Mol Breeding 35:82

EFSA Panel on Genetically Modified Organisms (2015) Guidance for renewal applications of genetically modified food and feed authorised under Regulation (EC) No. 1829/2003. EFSA J 13:4129

Eszterhas SK, Bouhassira EE, Martin DIK, Fiering S (2002) Transcriptional interference by independently regulated genes occurs in any relative arrangement of the genes and is influenced by chromosomal integration position. Mol Cell Biol 22:469–479

Fan M, Xu C, Xu K, Hu Y (2012) LATERAL ORGAN BOUNDARIES DOMAIN transcription factors direct callus formation in *Arabidopsis* regeneration. Cell Res 22:1169

Fehér A (2015) Somatic embryogenesis—stress-induced remodeling of plant cell fate. Biochim Biophys Acta—Gene Regul Mech 1849:385–402

Feng C, Yuan J, Wang R, Liu Y, Birchler JA, Han F (2016) Efficient targeted genome modification in maize using CRISPR/Cas9 System. J Genet Genomics 43:37–43

Florentin A, Damri M, Grafi G (2013) Stress induces plant somatic cells to acquire some features of stem cells accompanied by selective chromatin reorganization. Dev Dyn 242:1121–1133

Gelvin SB (2003) *Agrobacterium*-mediated plant transformation: the biology behind the "gene-jockeying" tool. Microbiol Mol Biol Rev 67:16–37

Gelvin SB (2009) *Agrobacterium* in the genomics age. Plant Physiol 150:1665–1676

Gelvin SB, Kim S-I (2007) Effect of chromatin upon *Agrobacterium* T-DNA integration and transgene expression. Biochim Biophys Acta—Gene Struct Expr 1769:410–421

Gordon-Kamm W, Dilkes BP, Lowe K, Hoerster G, Sun X, Ross M, Church L, Bunde C, Farrell J, Hill P, Maddock S, Snyder J, Sykes L, Li Z, Y-m Woo, Bidney D, Larkins BA (2002) Stimulation of the cell cycle and maize transformation by disruption of the plant retinoblastoma pathway. Proc Natl Acad Sci 99:11975–11980

Grafi G, Barak S (2015) Stress induces cell dedifferentiation in plants. Biochim Biophys Acta—Gene Regul Mech 1849:378–384

Hiei Y, Ohta S, Komari T, Kumashiro T (1994) Efficient transformation of rice (*Oryza sativa* L.) mediated by *Agrobacterium* and sequence analysis of the boundaries of the T-DNA. Plant J 6:271–282

Ishida Y, Saito H, Ohta S, Hiei Y, Komari T, Kumashiro T (1996) High efficiency transformation of maize (*Zea mays* L.) mediated by *Agrobacterium tumefaciens*. Nature Biotech 14:745–750

Ji Q, Xu X, Wang K (2013) Genetic transformation of major cereal crops. In Vitro Cell Dev Biol Plant 57:495–508

Kessler A, Baldwin IT (2002) Plant responses to insect herbivory: the emerging molecular analysis. Ann Rev Plant Biol 53:299–328

Kohli A, Twyman RM, Abranches R, Wegel E, Stoger E, Christou P (2003) Transgene integration, organization and interaction in plants. Plant Mol Biol 52:247–258

Komari T, Hiei Y, Saito Y, Murai N, Kumashiro T (1996) Vectors carrying two separate T-DNAs for co-transformation of higher plants mediated by *Agrobacterium tumefaciens* and segregation of transformants free from selection markers. Plant J 10:165–174

Komari T, Ishida Y, Hiei Y (2004) Plant transformation technology: *Agrobacterium*-mediated transformation. In: Handbook of plant biotechnology. Wiley, Hoboken

Komari T, Takakura Y, Ueki J, Kato N, Ishida Y, Hiei Y (2006) Binary Vectors and super-binary vectors. In: Wang K (ed) Agrobacterium protocols. Humana Press, Totowa, NJ, pp 15–42

Li Z, Xing A, Moon BP, McCardell RP, Mills K, Falco SC (2009) Site-specific integration of transgenes in soybean via recombinase-mediated DNA cassette exchange. Plant Physiol 151:1087–1095

Lotan T, M-a Ohto, Yee KM, West MAL, Lo R, Kwong RW, Yamagishi K, Fischer RL, Goldberg RB, Harada JJ (1998) *Arabidopsis* LEAFY COTYLEDON1 is sufficient to induce embryo development in vegetative cells. Cell 93:1195–1205

Louwerse JD, van Lier MCM, van der Steen DM, de Vlaam CMT, Hooykaas PJJ, Vergunst AC (2007) Stable recombinase-mediated cassette exchange in *Arabidopsis* using *Agrobacterium tumefaciens*. Plant Physiol 145:1282–1293

Lowe K, Wu E, Wang N, Hoerster G, Hastings C, Cho M-J, Scelonge C, Lenderts B, Chamberlin M, Cushatt J, Wang L, Ryan L, Khan T, Chow-Yiu J, Hua W, Yu M, Banh J, Bao Z, Brink K, Igo E, Rudrappa B, Shamseer PM, Bruce W, Newman L, Shen B, Zheng P, Bidney D, Falco SC, RegisterIII JC, Zhao Z-Y, Xu D, Jones TJ, Gordon-Kamm WJ (2016) Morphogenic regulators baby boom and wuschel improve monocot transformation. Plant Cell 28:1998–2015

Lowe K, Rota ML, Hoerster G, Hastings C, Wang N, Chamberlin M, Wu E, Jones T, Gordon-Kamm W (2018) Rapid genotype "independent" maize transformation via direct

somatic embryogenesis. In Vitro Cell Dev Biol Plant. https://doi.org/10.1007/s11627-018-9901-6
Lyznik LA, Gordon-Kamm WJ, Tao Y (2003) Site-specific recombination for genetic engineering in plants. Plant Cell Rep 21:925–932
Matzke MA, Mette MF, Matzke AJM (2000) Transgene silencing by the host genome defense: implications for the evolution of epigenetic control mechanisms in plants and vertebrates. Plant Mol Biol 43:401–415
Mookkan M, Nelson-Vasilchik K, Hague J, Zhang ZJ, Kausch AP (2017) Selectable marker independent transformation of recalcitrant maize inbred B73 and sorghum P898012 mediated by morphogenic regulators BABY BOOM and WUSCHEL2. Plant Cell Rep 36:1477–1491
Motte H, Vercauteren A, Depuydt S, Landschoot S, Geelen D, Werbrouck S, Goormachtig S, Vuylsteke M, Vereecke D (2014) Combining linkage and association mapping identifies RECEPTOR-LIKE PROTEIN KINASE1 as an essential *Arabidopsis* shoot regeneration gene. Proc Natl Acad Sci 111:8305–8310
Mumm RH, Walters DS (2001) Quality control in the development of transgenic crop seed Products. Crop Sci 41:1381–1389
Nandy S, Srivastava V (2011) Site-specific gene integration in rice genome mediated by the FLP–FRT recombination system. Plant Biotechnol J 9:713–721
Nanto K, Ebinuma H (2008) Marker-free site-specific integration plants. Trans Res 17:337–344
Nanto K, Yamada-Watanabe K, Ebinuma H (2005) *Agrobacterium*-mediated RMCE approach for gene replacement. Plant Biotechnol J 3:203–214
Nanto K, Sato K, Katayama Y, Ebinuma H (2009) Expression of a transgene exchanged by the recombinase-mediated cassette exchange (RMCE) method in plants. Plant Cell Rep 28:777–785
Ow DW (2007) GM maize from site-specific recombination technology, what next? Curr Opin Biotechnol 18:115–120
Pellegrino E, Bedini S, Nuti M, Ercoli L (2018) Impact of genetically engineered maize on agronomic, environmental and toxicological traits: a meta-analysis of 21 years of field data. Sci Rep 8:3113
Pröls F, Meyer P (1992) The methylation patterns of chromosomal integration regions influence gene activity of transferred DNA in Petunia hybrida. Plant J 2:465–475
Puchta H, Fauser F (2014) Synthetic nucleases for genome engineering in plants: prospects for a bright future. Plant J 78:727–741
Que Q, Elumalai S, Li X, Zhong H, Nalapalli S, Schweiner M, Fei X, Nuccio M, Kelliher T, Gu W, Chen Z, Chilton M-DM (2014) Maize transformation technology development for commercial event generation. Front Plant Sci 5:000379
Rathore DS, Doohan F, Mullins E (2016) Capability of the plant-associated bacterium, *Ensifer adhaerens* strain OV14, to genetically transform its original host *Brassica napus*. Plant Cell Tissue Organ Cult 127:85–94
Rinaldo AR, Ayliffe M (2015) Gene targeting and editing in crop plants: a new era of precision opportunities. Mol Breeding 35:40
Schlake T, Bode J (1994) Use of mutated FLP recognition target (FRT) Sites for the exchange of expression cassettes at defined chromosomal loci. Biochemistry 33:12746–12751
Schnell J, Steele M, Bean J, Neuspiel M, Girard C, Dormann N, Pearson C, Savoie A, Bourbonnière L, Macdonald P (2015) A comparative analysis of insertional effects in genetically engineered plants: considerations for pre-market assessments. Transgenic Res 24:1–17
Singh A, Khurana P (2017) Ectopic expression of *Triticum aestivum* SERK genes (TaSERKs) control plant growth and development in *Arabidopsis*. Sci Rep 7:12368
Sprink T, Metje J, Hartung F (2015) Plant genome editing by novel tools: TALEN and other sequence specific nucleases. Curr Opin Biotech 32:47–53
Srinivasan C, Liu Z, Heidmann I, Supena EDJ, Fukuoka H, Joosen R, Lambalk J, Angenent G, Scorza R, Custers JBM, Boutilier K (2006) Heterologous expression of the BABY BOOM

AP2/ERF transcription factor enhances the regeneration capacity of tobacco (*Nicotiana tabacum* L.). Planta 225:341–351

Srivastava V, Ow DW (2002) Biolistic mediated site-specific integration in rice. Mol Breed 8:345–349

Srivastava V, Thomson J (2016) Gene stacking by recombinases. Plant Biotechnol J 14:471–482

Srivastava V, Ariza-Nieto M, Wilson AJ (2004) Cre-mediated site-specific gene integration for consistent transgene expression in rice. Plant Biotechnol J 2:169–179

Stone SL, Kwong LW, Yee KM, Pelletier J, Lepiniec L, Fischer RL, Goldberg RB, Harada JJ (2001) LEAFY COTYLEDON2 encodes a B3 domain transcription factor that induces embryo development. Proc Natl Acad Sci 98:11806–11811

Strauss SH, Sax JK (2016) Ending event-based regulation of GMO crops. Nat Biotechnol 34:474–477

Svitashev S, Schwartz C, Lenderts B, Young JK, Mark Cigan A (2016) Genome editing in maize directed by CRISPR–Cas9 ribonucleoprotein complexes. Nat Commun 7:13274

Tajima Y, Imamura A, Kiba T, Amano Y, Yamashino T, Mizuno T (2004) Comparative studies on the type-b response regulators revealing their distinctive properties in the His-to-Asp phosphorelay signal transduction of *Arabidopsis thaliana*. Plant Cell Physiol 45:28–39

Terada R, Urawa H, Inagaki Y, Tsugane K, Iida S (2002) Efficient gene targeting by homologous recombination in rice. Nat Biotechnol 20:1030–1034

Tingay S, McElroy D, Kalla R, Fieg S, Wang M, Thornton S, Brettell R (1997) *Agrobacterium tumefaciens*-mediated barley transformation. Plant J 11:1369–1376

Toki S (1997) Rapid and efficient *Agrobacterium*-mediated transformation in rice. Plant Mol Biol Rep 15:16–21

Turan S, Kuehle J, Schambach A, Baum C, Bode J (2010) Multiplexing RMCE: versatile extensions of the Flp-recombinase-mediated cassette-exchange technology. J Mol Biol 402:52–69

Turan S, Zehe C, Kuehle J, Qiao J, Bode J (2013) Recombinase-mediated cassette exchange (RMCE)—a rapidly-expanding toolbox for targeted genomic modifications. Gene 515:1–27

Tzfira T, White C (2005) Towards targeted mutagenesis and gene replacement in plants. Trends Biotechnol 23:567–569

van der Fits L, Deakin EA, Hoge JHC, Memelink J (2000) The ternary transformation system: constitutive *virG* on a compatible plasmid dramatically increases *Agrobacterium*-mediated plant transformation. Plant Mol Biol 43:495–502

Vergunst AC, Jansen LE, Hooykaas PJ (1998) Site-specific integration of *Agrobacterium* T-DNA in *Arabidopsis thaliana* mediated by Cre recombinase. Nucl Acids Res 26:2729–2734

Weeks DP, Spalding MH, Yang B (2016) Use of designer nucleases for targeted gene and genome editing in plants. Plant Biotechnol J 14:483–495

Wei F-J, Kuang L-Y, Oung H-M, Cheng S-Y, Wu H-P, Huang L-T, Tseng Y-T, Chiou W-Y, Hsieh-Feng V, Chung C-H, Yu S-M, Lee L-Y, Gelvin SB, Hsing Y-IC (2016) Somaclonal variation does not preclude the use of rice transformants for genetic screening. Plant J 85:648–659

Wendt T, Doohan F, Mullins E (2012) Production of *Phytophthora infestans*-resistant potato (*Solanum tuberosum*) utilising *Ensifer adhaerens* OV14. Transgenic Res 21:567–578

Wu E, Lenderts B, Glassman K, Berezowska-Kaniewska M, Christensen H, Asmus T, Zhen S, Chu U, Cho M-J, Zhao Z-Y (2014) Optimized *Agrobacterium*-mediated sorghum transformation protocol and molecular data of transgenic sorghum plants. In Vitro Cell Dev Biol—Plant 50:9–18

Yang X, Zhang X (2010) Regulation of somatic embryogenesis in higher plants. Critical Rev Plant Sci 29:36–57

Zastrow-Hayes GM, Lin H, Sigmund AL, Hoffman JL, Alarcon CM, Hayes KR, Richmond TA, Jeddeloh JA, May GD, Beatty MK (2015) Southern-by-sequencing: a robust screening approach for molecular characterization of genetically modified crops. Plant Genome 8

Zeng F, Zhang X, Cheng L, Hu L, Zhu L, Cao J, Guo X (2007) A draft gene regulatory network for cellular totipotency reprogramming during plant somatic embryogenesis. Genomics 90:620–628

Zhi L, TeRonde S, Meyer S, Arling ML, Register JC III, Zhao Z-Y, Jones TJ, Anand A (2015) Effect of *Agrobacterium* strain and plasmid copy number on transformation frequency, event quality and usable event quality in an elite maize cultivar. Plant Cell Rep 34:745–754

Zhu J, Song N, Sun S, Yang W, Zhao H, Song W, Lai J (2016) Efficiency and inheritance of targeted mutagenesis in maize using CRISPR-Cas9. J Genet Genomics 43:25–36

Zuniga-Soto E, Mullins E, Dedicova B (2015) *Ensifer*-mediated transformation: an efficient non-*Agrobacterium* protocol for the genetic modification of rice. SpringerPlus 4:600

Zuo J, Niu Q-W, Frugis G, Chua N-H (2002) The *WUSCHEL* gene promotes vegetative-to-embryonic transition in *Arabidopsis*. Plant J 30:349–359

Correction to: *Agrobacterium*: A Genome-Editing Tool-Delivery System

Nagesh Sardesai and Subhashree Subramanyam

Correction to: Chapter "*Agrobacterium*: A Genome-Editing Tool-Delivery System" in: N. Sardesai and S. Subramanyam, Current Topics in Microbiology and Immunology, https://doi.org/10.1007/82_2018_136

By mistake the chapter was published with incorrect author name. The chapter has now been corrected.

The updated online version of this chapter can be found at https://doi.org/10.1007/82_2018_101

N. Sardesai (✉)
Corteva Agriscience™, Agriculture Division of DowDuPont, 8305 NW 62nd Avenue, Johnston, IA, USA
e-mail: nsardesai@dow.com

S. Subramanyam
Department of Agronomy, Purdue University, 915 W State Street, West Lafayette, IN 47907, USA
e-mail: shubha@purdue.edu

Current Topics in Microbiology and Immunology (2018) 418:E1
DOI 10.1007/82_2018_136

Published Online: 20 September 2018