

The Genomes of Endophytic Bacteria

A. Carolin Frank

Abstract Genome sequencing and comparative genomics has had major impact on our understanding of the genetic potential, ecology, and evolution of microorganisms. Analysis of a dozen bacterial endophyte genomes have recently contributed insights into the molecular mechanisms that enable bacterial exploration of the plant interior, including genes for motility, colonization and synergistic interactions with the host. Known host-interaction systems include type IV pili, flagella, diverse dedicated secretion systems, genes for phytohormone synthesis and inhibition, bacterial volatiles, and antimicrobials. Different endophytes use different sets of known host interaction systems, suggesting that there are multiple strategies to colonize and persist within plants, and that there are different ways in which endophytes can interact with their host. The majority of host-interaction systems are shared with other bacteria, including plant- and animal pathogens. Functional exploration of the large sets of endophyte genes encoding hypothetical proteins (especially those shared with other phytobacteria) promises to further elucidate bacterial adaptation to life in plant tissue, especially in regards to plant colonization, defense evasion and plant growth promotion.

Abbreviations

ACC	1-aminocyclopropane-1-carboxylate
GI	genomic islands
MGEs	mobile genetic elements
IS	insertion sequence

A.C. Frank (✉)
School of Natural Sciences, University of California Merced, Merced, CA, USA
e-mail: cfrank3@ucmerced.edu

TAM	tryptamine pathway
IAN	indole-3-acetonitrile pathway
ISR	induced systemic resistance
GABA	Gamma-aminobutyrate
IAM	indole-3-acetamide pathway
IPA	Indole-3-pyruvate pathway
TSO	the tryptophan side-chain oxidase pathway
IAA	indole acetic acid
LPS	lipopolysaccharide

1 Introduction

The knowledge of how bacterial endophytes interact with plants is essential both to those who wish to develop effective and environmentally sound endophyte biotechnology applications, and to those who wish to better understand the role of this ubiquitous and diverse set of organisms in natural ecosystems. In the last decade, genome-enabled studies have become a popular and effective means of deepening our understanding of life on earth. Genome sequencing of bacterial isolates began in the early 1990s and is now, thanks to recent development in sequence technologies and advances in bioinformatics, almost routine. Other genomics techniques include gene expression analysis, which can be performed using microarrays if a reference genome is available, or with next-generation sequencing technologies (Simon and Daniel 2011), an approach that does not require a genome sequence. Sequencing the metagenome, i.e. the collective genomes of all organisms in an environment, is an increasingly popular way to study the uncultured majority of bacteria (Wooley et al. 2010), as is single-cell genomics, genome sequencing that does not require bacterial isolation (Woyke et al. 2010). These genomics methods, including isolate sequencing, have been underused as a means of understanding bacterial endophyte biology. Only a dozen of the thousands of bacterial genomes sequenced so far are from endophytes. However, analysis of a handful of endophyte genomes sequenced in the last few years demonstrates the effectiveness of this approach for developing a better understanding of the endophytic niche.

This review summarizes the findings from endophyte isolate sequencing and analysis, with an emphasis on the correlation between life style and genome content/structure. Although these genome sequences only represent a snapshot of the rich diversity of endophytes that exists, they provide major insights into the ecology and evolution of bacterial endophytes. As most genomes are from endophytes of grasses, a review of forest tree-endophyte genomes is unfortunately premature. A rice endophyte metagenome has been sequenced, but is not yet published, and so will not be covered here.

1.1 Bacterial Genome Evolution

An understanding of the ecology, role and host-interaction of a particular bacterial isolate through analysis of its genome requires knowledge on how bacterial genomes evolve. Compared to eukaryotes (though there are exceptions) bacterial genomes are remarkably plastic, experiencing gene loss and gain at evolutionary – and even ecological – timescales. Horizontal gene transfer, the non-vertical transfer of genetic material between organisms, is an important source of novelty in bacterial evolution and adaptation to new environments (Ochman and Moran 2001), including the adaptation to hosts. Though the mechanisms of gene transfer (transduction, conjugation and transformation) have been known for decades, the era of genome sequencing brought some surprises as horizontal gene transfer turned out to be more rampant than previously thought. Initially a subject of heated debate, it is now generally accepted that horizontal gene transfer in bacteria is common. Still, there must also be ecological barriers to horizontal gene transfer, given that many high taxonomic levels (e.g. phyla) of bacteria share general life strategies or traits that distinguish them from members of other groups (Philippot et al. 2009, 2010).

Compared to eukaryotes, bacterial genomes are small. However, although a given bacterial cell will have a genome in the order of a few million nucleotides, the full functional spectrum of genes available to a given population – termed the ‘pan-genome’ – can be quite large. The genes in a bacterial population (for example a species or a genus) can usually be classified into the ‘core’ and ‘variable’ genome components, which together make up the pan-genome (Tettelin et al. 2008). The core genome is made up of the genes that are common to all but a few individuals of a species; a backbone of essential components. The ‘variable’, or accessory genome, is composed of genes not found in all strains, because they were differentially gained or lost. Variable genes contribute to the diversity within bacterial species, and may confer selective advantages in specific niches. It was recently demonstrated that in *Escherichia coli*, the variable genes make up more than 90% of the pan-genome and about 80% of a typical genome (Lukjancenko et al. 2010). Variable genes can further be divided into the categories ‘character’ and ‘accessory’, where the character genes represent genes essential for colonization and survival in particular environmental niches (e.g. symbiosis or photosynthesis), and the accessory genes are mostly unknown genes that can be used to distinguish strains and serotypes (Lapierre and Gogarten 2009). Generally, the function of most genes in the latter category is unknown.

Genome analysis goes beyond merely providing a ‘parts list’ of the capabilities of an organism. For example, comparative genomics – the process of identifying similarities and differences between strains or species – can improve our understanding of the evolutionary context of host–bacteria interactions. Powerful comparative analysis requires adequate datasets of several relatively closely related genomes, something that is not yet available for bacterial endophytes. However, analysis of the endophyte genomes sequenced so far has significantly advanced our understanding of life in the endosphere.

2 Significance of Sequenced Endophytes

To date, genomes from nine strains of bacterial endophytes have been sequenced and published (Table 1). Two are Alphaproteobacteria; the rice endophyte *Azospirillum* sp. B510 (Kaneko et al. 2010) and the sugarcane endophyte *Gluconacetobacter diazotrophicus* Pal5 (Bertalan et al. 2009). Five are Gammaproteobacteria; *Klebsiella pneumoniae* 342, isolated from maize stem (Fouts et al. 2008), and the poplar endophytes *Pseudomonas putida* W619, *Serratia proteamaculans* 568, *Stenotrophomonas maltophilia* R551-3 (Taghavi et al. 2009) and *Enterobacter* sp. strain 638 (Taghavi et al. 2009, 2010). Two betaproteobacterial endophytes have been sequenced and published; the rice endophyte, *Azoarcus* sp. BH72 (Krause et al. 2006), and *Variovorax paradoxus* S110, a strain that has been detected in the interior of various plants (Han et al. 2011). Finally, a cyanobacterial water fern endophyte, *Nostoc azollae* 0708 has been sequenced and published (Ran et al. 2010). In addition, several endophyte genomes are on the way, some of which are completed and present in the NCBI database, though not yet published. Table 1 compiles the 14 endophyte strains that have been sequenced so far, including the source of isolation (if indicated in the publication), along with their genome size and structure. Here, designation of replicons as chromosomes or plasmids follows that of the authors, although some of these replicons could probably be classified as so called ‘chromids’ a recently coined term for a replicon that is neither a chromosome, nor a plasmid (Harrison et al. 2010).

With the exception of the water fern endophyte, these organisms have all been chosen for genome sequencing because of their beneficial effects on plants and their potential in agrobiotechnological applications. More specifically, published endophyte genomes were sequenced for the following reasons: *Azospirillum* B510 increases rice stem number and seed yield, and enhances host resistance to rice blast fungus and to the bacterial pathogen *Xanthomonas oryzae* (Kaneko et al. 2010). *P. putida* W619, and *S. proteamaculans* 568 are both commonly isolated endophytes of poplar, and promote shoot and root development in their host (Taghavi et al. 2009). In the same study, *S. maltophilia* R551-3 was also isolated from poplar (although not as commonly) but it has no direct plant growth promoting effects on its host. *S. maltophilia* is of interest as an emergent opportunistic human pathogen, which means that though not necessarily adapted to humans, it can take advantage of immunocompromised individuals (Looney et al. 2009). *S. maltophilia* and is therefore a potential agent of phytonotic disease (of plant origin). *Enterobacter* sp. 638, also isolated from poplar, can increase poplar growth by up to 40%, and provides systemic drought resistance to poplar (Taghavi et al. 2009). *G. diazotrophicus* Pal5, *K. pneumoniae* 342, *Azoarcus* sp. BH72, and *N. azollae* 0708 are all diazotrophic endophytes (they fix nitrogen inside the plant). *G. diazotrophicus* Pal5 can, in addition to this, promote growth, particularly in roots, inhibit the growth of the sugarcane pathogen *Xanthomonas albilineans* (Blanco et al. 2005) protect against fungal pathogens (Mehnaz and Lazarovits 2006), and promote an increase in the solubility of phosphate and zinc (Saravanan et al. 2007). *K. pneumoniae* 342

Table 1 Endophytic bacterial strains with sequenced genomes

Species	Publication	Isolated from	Class; order	Chromosome size (bp)	Plasmid size (bp)
<i>Azoarcus</i> sp. BH72	(Krause et al. 2006)	Kallar grass roots	Betaproteobacteria; Rhodocyclales	4,376,040	
<i>Azospirillum</i> sp. B510	(Kaneko et al. 2010)	Rice stem	Alphaproteobacteria; Rhodospirillales	3,311,395	1,455,109 723,779 681,723 628,837 537,299 261,596
<i>B. phytofirmans</i> PsJN	Unpublished	Onion root	Betaproteobacteria; Burkholderiales	4,467,537 3,625,999	121,122
<i>Enterobacter</i> sp. strain 638	(Taghavi et al. 2010)	Poplar stem	Gammaproteobacteria; Enterobacteriales	4,518,712	157,749
<i>G. diazotrophicus</i> Pal5 A	(Bertalan et al. 2009)	Sugarcane root	Alphaproteobacteria; Rhodospirillales	3,944,163	38,818 16,610
<i>G. diazotrophicus</i> Pal5 B	Unpublished	Unknown	Alphaproteobacteria; Rhodospirillales	3,887,492	27,455
<i>H. seropedicae</i> SmR1	Unpublished	Unknown	Betaproteobacteria; Burkholderiales	5,513,887	
<i>K. pneumoniae</i> 342	(Fouts et al. 2008)	Maize stem	Gammaproteobacteria; Enterobacteriales	5,641,239	187,922 91,096

(continued)

Table 1 (continued)

Species	Publication	Isolated from	Class; order	Chromosome size (bp)	Plasmid size (bp)
<i>M. populi</i> B1001	Unpublished	Poplar plantlets and tissue cultures	Alphaproteobacteria; Rhizobiales	5,800,441	25,164 23,392
<i>N. azollae</i> 0708	(Ran et al. 2010)	The water fern <i>Azolla filiculoides</i>	Cyanobacteria; Nostocales;	5,354,700	109,570 21,875
<i>P. putida</i> W619	(Taghavi et al. 2009)	Poplar root and stem	Gammaaproteobacteria; Pseudomonadales	5,774,330	
<i>S. proteamaculans</i> 568	(Taghavi et al. 2009)	Poplar root	Gammaaproteobacteria; Enterobacteriales	5,448,853	46,804
<i>S. maltophilia</i> R551-3	(Taghavi et al. 2009) 90,435	Poplar rhizosphere, root and stem	Gammaaproteobacteria; Xanthomonadales	4,573,969	
<i>V. paradoxus</i> S110		Potato plant	Betaproteobacteria; Burkholderiales;	5,626,353 1,128,644	

is of interest because of its ability to colonize the interior of a wide range of host plants with a very small inoculum dose (Dong et al. 2003). Moreover, it is similar to *K. pneumoniae* strains that are opportunistic human pathogens. *Azoarcus* sp. BH72 colonizes not only its native host Kallar grass (Hurek et al. 1994), but also rice roots in high numbers and spreads systematically into rice shoots (Reinhold-Hurek et al. 2006). It is able to fix nitrogen in both plants. *V. paradoxus* S110 was isolated from potato plant leaves. Other *V. paradoxus* strains have been shown to stimulate root elongation in indian mustard (Belimov et al. 2005), biodegrade contaminants (Han et al. 2011), and engage in mutualistic beneficial interactions with both plants and other bacteria (Kanzler et al. 2005). Finally, *N. azollae* 0708 is a vertically transmitted extracellular cyanobacterial endophyte, which was sequenced to provide insight into the evolution of chloroplasts from a cyanobacterial ancestor (Ran et al. 2010).

3 Genome Architecture, Genome Content and Life Style

Broadly defined as bacteria inside healthy plant tissue, endophytes can display a range of different life styles, differing in the time spent free-living in the soil, as well as in their transmission dynamics, colonization strategies, and competence (Hardoim et al. 2008). Such differences should be reflected in the genomes of different endophytic bacteria. Genome size and structure usually correlates with bacterial life style. For example, host-restricted bacteria often have smaller genomes than free living relatives (Toft and Andersson 2010), whereas a large and complex genome with multiple replicons might indicate ability to survive in a variable environment (Ettema and Andersson 2009), as an increase in gene content increases robustness against environmental perturbation (Kitano 2007), promoting survival in multiple or variable niches. Large, versatile genomes are typical for e.g. nodule symbionts (Kaneko et al. 2000, 2002), which experience selection at two levels; as a result of adaptation to the stressful and variable environment in the soil, and as a result of adaptation to the plant host. Whereas nodule symbionts use soil as an alternative habitat in their life cycle, some endophytes of grasses are typically not isolated from soil (Reinhold-Hurek and Hurek 1998b). Poplar endophytes on the other hand, are believed to originate from soil and colonize the host via the roots (van der Lelie et al. 2009). Poplar is propagated by cuttings, and because these typically contain only a low number of endophytes, it is assumed that many species of poplar bacterial endophytes have an alternative life stage in soil (Taghavi et al. 2010).

Many of the endophyte genomes bear the same signatures of adaptation to a stressful and variable soil/rhizosphere environment as nodule symbiont genomes. The genomes of endophytes sequenced to date are relatively large and versatile, often comprising more than one chromosome and/or multiple plasmids or chromids. For example, *Azospirillum* sp. B510, which has the smallest chromosome of the endophyte genomes, has a remarkable set of six large plasmids. This genomic

versatility is consistent with the fact that *Azospirillum* is commonly isolated from both soil and plant interior (Hurek et al. 1994). The types of functions encoded by a genome can also indicate if an organism spends time in the soil or rhizosphere environment. For example, based on the number of transporters involved in carbohydrate, amino-acids and iron uptake, as well as some heavy metal resistance genes, Taghavi et al. (2010) suggest that *Enterobacter* sp. 638 is well adapted to survive in the plant rhizosphere.

Only three species; *Azoarcus* sp. BH72, *P. putida* W617 and *H. seropedicae* SmR1 have genomes that consist of a single chromosome without any chromids or plasmid (Table 1). Interestingly, both *Azoarcus* sp. BH72 and *H. seropedicae* SmR1 are candidates for tight association with the plant as attempts to isolate them from root-free soil have failed (Reinhold-Hurek and Hurek 1998a). Similarly, the sugarcane endophyte *G. diazotrophicus* is believed to survive poorly in soil (Kaneko et al. 2010). Though the published *G. diazotrophicus* strain does have two plasmids (of size 39 and 17 kb), they are not large enough to make a substantial contribution to the accessory genome and play an important role in adaptation to the stressful soil environment. Finally, *P. putida* W619 does not carry any plasmids, large or small. The paucity of plasmids in this strain may reflect its phylogeny rather than life style since most *Pseudomonas* genomes sequenced to date consist of a single chromosome. It is interesting to note however, that the *P. putida* endophyte strain has a smaller genome compared to non-endophytes *Pseudomonas*, possibly a result of host restriction.

4 The Plant-Associated Life Style

4.1 Transporters

Endophyte genomes are expected to encode a large diversity of transporters for the uptake of plant-produced nutrients. For example, *K. pneumoniae* 342 contains one of the highest percentage of transporters found in a bacterial genome, 15.4%, which is similar to plant/soil associated microbes like *Bradyrhizobium japonicum* and *Mesorhizobium loti* (Fouts et al. 2008). *Enterobacter* sp. 638 contains over 600 coding sequences for putative transporter proteins (Taghavi et al. 2010). The number of predicted transporters is however smaller for *Azospirillum* sp. B510 (~300) (Kaneko et al. 2010).

4.2 Motility and Colonization

Motility is an important feature of endophytes, used both to move towards the site of infection, and to systematically spread within the plant (Hardoim et al. 2008).

Endophytic host colonization depends on a variety of surface adhesion factors that allows attachment to the host outer or cell surface.

Type IV pili mediate twitching motility, which is essential e.g. for endophytic rice colonization by *Azoarcus* sp. BH72 (Bohm et al. 2007). The BH72 genome encodes 41 putative genes for pilus assembly and regulation (compared to 30 such genes in the closely related *Azoarcus* strain EbN1) (Krause et al. 2006). In fact, all endophyte genomes carry genes for type IV pili. Therefore, based on the limited set of bacterial endophyte genomes available, the presence of type IV pili genes appears to be a universal feature of endophytes.

Flagella and chemotaxis All the sequenced endophyte genomes (including the unpublished, as revealed by a simple Blast search), encode proteins for chemotaxis and almost all encode proteins for flagellar biosynthesis. The two exceptions are the water-fern endophyte *Nostoc azollae* 0708, which is not surprising given that cyanobacteria are known to lack flagella, and *K. pneumoniae* 342. Fouts et al. (2008) speculate that the lack of flagella in the *K. pneumoniae* 342 genome may contribute to its ability to colonize the host in high numbers, given that flagella are known to induce plant defense (Felix et al. 1999). Lack of flagella might indicate a higher level of adaptation to the endophytic niche, however it also raises questions about the organisms' ability to move towards the plant. *Azoarcus* sp. BH72 contains 48 genes for flagella and chemotaxis, whereas the soil-dwelling strain, *Azoarcus* sp. EbN1 does not contain any complete flagellar operons (Krause et al. 2006). The *P. putida* W619 genome contains a large cluster of 52 genes involved in flagellar biosynthesis (Wu et al. 2011). Gene expression results obtained with *P. putida* W619 showed that the transcription of some of those genes is induced in the presence of poplar roots, (Wu et al. 2011), supporting the importance of flagella in endophyte-host interactions. In the genome of *G. diazotrophicus* Pal5, the >40 motility genes are clustered in a region that is absent from the genomes of other *Gluconobacter* species (Bertalan et al. 2009). *Enterobacter* sp. 638 contains many genes for motility, including three flagellar biosynthesis operons (Taghavi et al. 2010). *Azospirillum* sp. B510 encodes an impressive number of around 100 putative chemotaxis/flagella genes, most of which are located on plasmids (Kaneko et al. 2010).

Curli fibers belong to a class of fibers known as amyloids. They are involved in surface adhesion, and promote colonization (Barnhart and Chapman 2006). Similar to so many other bacterial adhesion molecules, they are implicated in pathogenesis, but their presence in endophytes demonstrates that they likely are involved in commensal/mutualistic host colonization as well. The *P. putida* W619 genome is the only additional genome that harbors genes for curli fiber biogenesis, a feature that this strain shares with other sequenced *P. putida* strains (Wu et al. 2011).

Hemagglutinins Important for colonization in a number of plant and animal pathogens (Gottig et al. 2009; Balder et al. 2007), hemagglutinins are often described as 'pathogenicity factors', but their presence in *K. pneumoniae* 342 (Fouts et al. 2008) and *Enterobacter* sp. 638 (Taghavi et al. 2010) suggests that they could also be involved in endophyte colonization.

Cellulases Other factors involved in endophyte establishment within the plant include hydrolytic enzymes that macerate plant cell wall polymers. Whereas plant-pathogens are known to use e.g. glucoside hydrolases to degrade the host cell wall (Herron et al. 2000), the endophyte genomes carry few such genes. One exception is the *K. pneumoniae* 342 genome, with at least 38 genes encoding glycosyl hydrolases (Fouts et al. 2008). It is important to point out that a low production of macerating enzymes may be expected as these can assist in endophytic colonization, as has been shown for an *Azoarcus* sp. BH72 endoglucanase (Reinhold-Hurek et al. 2006).

Celluloses Bacteria can produce their own celluloses for attachment to the host surface. *P. putida* W619 encodes a protein involved in the production of beta-(1,2)-glucan, which in *Agrobacterium tumefaciens* is involved in attachment to plant cells (Rodriguez-Navarro et al. 2007). Interestingly, homologs are not present in two non-endophytic *P. putida* strains, but are present in other nonrelated endophytes (*Enterobacter* sp 638 and *S. proteamaculans*), pointing to a more general importance of this protein in endophyte colonization. *K. pneumoniae* 342 also encodes a beta-(1,2)-glucan, similar to *ndvB*, a gene involved nodule invasion in *Rhizobium meliloti* (Fouts et al. 2008).

4.3 Protection Against Plant Defense

To survive inside plants, endophytes need protection against non-specific plant defense compounds such as reactive oxygen species (ROS), nitric oxide, and phytoalexins (Zeidler et al. 2004). The endophyte genomes are well equipped with genes that provide protection against all three; superoxide dismutases, catalases, peroxidases, hydroperoxide reductases and glutathione-S-transferases.

To colonize the host, bacterial endophytes must overcome the strong plant defense system. Genes involved in plant defense evasion and suppression were mostly not identified in the endophyte genomes, however, genome sequences provide an excellent starting point for elucidating the factors that enable endophytes to escape plant rejection.

4.4 Quorum Sensing

Some bacteria engage in cooperative group behavior through signal molecules termed autoinducers, which can trigger specific functions in a cell-density dependent manner (Camilli and Bassler 2006). The name ‘quorum sensing’ comes from the fact that to regulate processes via autoinducers, bacteria need to reach a critical mass (Miller and Bassler 2001). Though in many bacteria quorum sensing regulates the expression of virulence genes (Antunes et al. 2010), it is not restricted to pathogenic interactions (Sanchez-Contreras et al. 2007). In rhizobia for

example, quorum sensing can play an important part in the regulation of transfer or plasmids and Integrated Conjugative Elements (ICE, a form of mobile genetic element that can carry genes important in symbiosis) (Ramsay et al. 2006). As pointed out by Rosenblueth and Martinez-Romero (2006), it would be interesting to determine if endophytes use quorum sensing to communicate inside plants. Despite (unpublished) evidence of quorum sensing in *Azoarcus* sp. BH72, the genes for the common autoinducer (N-acyl homoserine lactone, AHL), the autoinducer synthetase (LuxI) or the autoinducer receptor (LuxR) were not found in the genome (Krause et al. 2006), indicating that this strain uses a novel quorum sensing system.

Interestingly, some of the other endophyte genomes possess these well-studied quorum sensing genes. For example, *V. paradoxus* S110 encodes AHL synthase and its transcription regulator (Han et al. 2011). Three quorum sensing genes were found in the *G. diazotrophicus* Pal5 genome: one *luxI* autoinducer synthase gene and two *luxR*-type transcriptional regulator genes (Bertalan et al. 2009). The *Azospirillum* genome carries an AHL synthase gene and 22 *luxR* family transcriptional regulators (Kaneko et al. 2010). In the related free-living diazotroph *Azospirillum lipoferum*, quorum sensing regulates a range of activities important to plant-host interaction, such as pectinase activity, siderophore synthesis, and phytohormone production (Vial et al. 2006).

4.5 Plant Growth-Promoting Traits

Like rhizosphere bacteria, endophytic bacteria can promote plant growth in direct (e.g. through phytohormone synthesis) and indirect ways (e.g. through nitrogen fixation or defense). Many of the endophytes that have been sequenced are of interest because of their role in plant growth promotion, and consequently they encode proteins involved in such processes.

4.5.1 Phytohormone Production, Volatiles and ACC Deaminase

A wide variety of plant-associated bacteria – including both pathogens and mutualists – can affect plant physiology through the synthesis of compounds that alter plant hormone balance. These can be phytohormones, modulators of phytohormones, or compounds that mimic the action of phytohormones (Ping and Boland 2004).

IAA synthesis Auxin is produced by the plant to regulate various developmental processes. Some bacteria can interfere with plant growth and development by producing indole-3-acetic acid (IAA), the major naturally occurring auxin. IAA producers are found across the range of phytobacterial life styles and include phytopathogens (both gall-inducing- and other pathogens) (Spaepen and Vanderleyden 2010), nodule-forming bacteria (Perrine et al. 2005), free-living plant

growth-promoting bacteria in the rhizosphere (Ahmad et al. 2008), and endophytes (Madmony et al. 2005). In the context of endophytes, IAA production is generally considered a plant-beneficial characteristic (Hardoim et al. 2008).

Several different IAA pathways have been described for bacteria, including the indole-3-acetamide pathway (IAM), the Indole-3-pyruvate pathway (IPA), the tryptamine pathway (TAM), the tryptophan side-chain oxidase pathway (TSO), the Indole-3-acetonitrile pathway (IAN), and the tryptophan-independent pathway (Spaepen and Vanderleyden 2010). It has been suggested that the result of bacterial IAA production from the perspective of the plant may depend on the specific pathway used to synthesize IAA, as pathogens tend to use the IAM pathway, whereas IPA is observed in pathogenic as well as in nonpathogenic bacteria, and is the pathway used by most beneficial bacteria (Spaepen et al. 2007). Moreover, different IAA pathways may have different roles at different stages of colonization (Brandl et al. 2001; Manulis et al. 1998).

Why do bacteria synthesize IAA? Given that IAA can increase the efficiency of colonization (Suzuki et al. 2003), its production may be a colonization strategy; the stimulation of plant proliferation promotes the bacterial niche (whether plant surface or interior). However, IAA biosynthesis may also play an important role in colonization through circumvention of the host defense (Spaepen et al. 2007). Down-regulation of auxin signaling is part of a plant-induced immune response, and auxin promotes susceptibility to bacterial disease (Navarro et al. 2006). In either case, although bacterial IAA synthesis may be a desired quality of endophytes for use in agriculture and forestry, it is unclear whether bacterial phytohormone production actually promotes plant fitness.

In *Agrobacterium tumefaciens*, the genes *iaaM* and *iaaH* are involved in the IAM pathway. Two of the sequenced endophyte genomes, *Azospirillum* sp. B510 and *P. putida* W619, were reported to encode proteins involved in the IAM pathway (Kaneko et al. 2010; Wu et al. 2011), however the identified proteins do not appear to be orthologs of *A. tumefaciens* *iaaM* and *iaaH*, showing only relatively weak sequence similarity to those (data not shown). *P. putida* W619 is an efficient *in vitro* producer of IAA in comparison with other endophytic bacteria (Taghavi et al. 2009). One reason for *P. putida*'s efficient IAA production might be the presence of the IPA pathway in addition to the putative IAM pathway. Also consistent with the high level of IAA produced by *P. putida* W619 is the presence of three genes encoding putative auxin carriers (Wu et al. 2011).

IPA is common in bacteria, including those not associated with plants, probably because the genes required for the IPA pathway are also part of the Ehrlich pathway (Spaepen and Vanderleyden 2010). Though present in the close relative *Azospirillum brasilense*, no IPA genes were detected in *Azospirillum* B510. Some IPA genes were also detected in *G. diazotrophicus* Pal5, confirming the results of experiments that demonstrated IAA is mostly synthesized through this pathway (Lee et al. 2004). However, one of the required genes for IPA biosynthesis is lacking from the *G. diazotrophicus* Pal5 genome, suggesting more remains to be discovered regarding bacterial IAA production. The authors suggest that the biochemical activity could be executed by one of the many putative decarboxylases identified in the genome

(Bertalan et al. 2009). Also, the presence of genes encoding enzymes such as aromatic-L-amino-acid decarboxylase, amine oxidase and aldehyde dehydrogenases suggests that *G. diazotrophicus* Pal5 might synthesize IAA via the tryptamine pathway (TAM). Finally, the presence of two genes coding for putative nitrilases suggests that IAA might be produced by the indole-3-acetonitrile pathway (IAN), which is poorly characterized in bacteria.

Judging from the genomes sequenced so far, IAA production is not a universal feature of endophytes. The relative paucity of IAA genes in endophyte genomes might also reflect incomplete understanding of bacterial IAA pathways.

ACC deaminase Ethylene is a stress-induced plant hormone that can inhibit plant growth (Morgan and Drew 1997). Some bacteria can lower the level of ethylene in the plant by cleaving the plant-produced ethylene precursor 1-aminocyclopropane-1-carboxylate (ACC) (Glick et al. 1998). Inoculation of such bacteria can mitigate the effect of various stressors by sustaining plant growth in the face of ethylene (Belimov et al. 2009; Siddikee et al. 2011). This raises the question as to why bacteria produce ACC-deaminase. According to one hypothesis, plants have evolved to select bacteria with high ACC-deaminase activity to attenuate the stress caused by high levels of ethylene (Hardoim et al. 2008). If so, ACC-deaminase activity would be a highly competitive trait in the endosphere. On the other hand, if the growth-inhibiting action of ethylene is an adaptation to stress that ultimately enhance plant survival under adverse conditions (Morgan and Drew 1997), bacterial interference may not be beneficial or even desirable to the plant. Ethylene is also required for normal plant development, and is being produced in bursts to e.g. help break seed dormancy. ACC-deaminase producing bacteria may play a role in regulating ethylene levels after such bursts, ensuring that ethylene levels stay below the point where growth is impaired (Glick 1995; Hardoim et al. 2008). Finally, there is evidence suggesting that ethylene is a key regulator of the colonization of plant tissue by bacteria (Iniguez et al. 2005), which in turn suggests that the ethylene inhibiting effects of ACC-deaminase may be a bacterial colonization strategy. Regardless of why plant-associated bacteria produce ACC-deaminase, their application can clearly be a very useful strategy to mitigate the effects of various stressors on cultivated plants.

Only a few of the genomes carry *acdS*, the gene for ACC deaminase, including *V. paradoxus* S110, which can use ACC as a carbon and nitrogen source (Han et al. 2011), and *Azospirillum* sp. B510, although it is unknown whether this bacterium has ACC deaminase activity (Kaneko et al. 2010). Putative ACC deaminase genes were found in *P. putida* W619, *Enterobacter* sp. 638, and *S. proteamaculans* 568, but all three lack the conserved amino acid signature characteristic for a genuine ACC deaminase, and consequently do not grow on ACC as their sole nitrogen source (Taghavi et al. 2009). The unpublished but completed genome of *B. phytofirmans* PsJN carries an *acdS* homolog. Inactivation of this gene leads to lost ACC deaminase activity and lost ability to promote the elongation canola seedling roots (Sun et al. 2009). Finally, the unpublished genome of *H. seropedicae* SmR1 encodes an *acdS* gene, probably functional, given that other strains of this species

have ACC-deaminase activity (Rothballer et al. 2008). Thus, only three or four of the complete endophyte genomes encode a functional AcdS protein.

Bacterial volatiles Bacterial volatile compounds, especially acetoin and 2,3-butanediol, trigger plant defense responses, effectively stimulate plant growth, and mediate drought resistance (Ping and Boland 2004; Cho et al. 2008; Ryu et al. 2003, 2004). The genes involved in the pathway, *budA*, *budB* and *budC*, were discovered in *Klebsiella terrigena* and *Enterobacter aerogenes* (Blomqvist et al. 1993). BudB, the acetolactate synthase, converts pyruvate to acetolactate, which is subsequently converted by the acetoin decarboxylase BudA into acetoin. Acetoin is released by the bacteria or subsequently converted into 2,3-butanediol by the acetoin reductase BudC.

The *budB* gene is present in all endophyte genomes, which is not surprising, given that the protein product is involved in multiple metabolic pathways. The *budA* gene is present in *Azospirillum* sp. B510, *Enterobacter* sp. 638, *K. pneumoniae* 342, and *S. proteamaculans* 568. *Enterobacter* sp. 638 carries the *budC* gene, and so does *G. diazotrophicus* Pal5 despite lacking a *budA* homolog. Interestingly, both bacteria possessing all three *bud* genes – *Enterobacter* sp. 638 and *K. pneumoniae* 342 – each carry two more genes whose products contribute to the production of acetoin; an acetoin dehydrogenase, which can convert diacetyl (formed spontaneously from acetolactate under aerobic conditions) into acetoin, and PoxB, a pyruvate dehydrogenase (Fouts et al. 2008; Taghavi et al. 2010), which, although its main function is to convert pyruvate into acetaldehyde, converts some of the pyruvate to acetoin.

In *Enterobacter* sp. 638, the *budABC* operon is located in a genomic region next to the operon for sucrose uptake and metabolism, further indicating that acetoin production is important in the interaction between this bacterium and its poplar host. This region is discussed in more detail in Sect. 6.

The *P. putida* W619 genome also carries *poxB* (but not *budABC*). *P. putida* W619 also lacks the genes involved in the catabolic conversion of acetoin and 2,3-butanediol to central metabolites, whereas the non-endophytic strains *P. putida* F1 and *P. putida* KT2440 have them (Wu et al. 2011). Therefore, although the production of acetoin via PoxB in the *P. putida* endophyte may be very low, it can potentially be converted into 2,3-butanediol by the poplar tree, and promote plant growth that way. Thus, gene loss in *P. putida* W619 may have been selected to increase the production of acetoin available to the host.

Polyamines Polyamines (putrescine, spermidine and spermine) are a group of phytohormone-like natural compounds present in almost all living organisms, which have been shown to increase plant tolerance to a variety of abiotic stressors, including high and low temperatures, salinity, hyperosmosis, hypoxia and atmospheric pollutants (Gill and Tuteja 2010). Given that exogenous application of polyamines protects plants from damage under stress conditions, bacterially produced polyamines could potentially sustain plant growth in the face of abiotic

stress. *Azospirillum brasilense* can produce and secrete polyamines, and has the potential capacity to promote plant growth through these compounds (Perrig et al. 2007). Though not discussed in detail, a few authors note the existence of genes involved in polyamine synthesis (Bertalan et al. 2009; Han et al. 2011; Taghavi et al. 2010). For example Bertalan et al. (2009) suggest that the genes for synthesis and secretion of spermidine in the *G. diazotrophicus* Pal5 may contribute to the plant growth-promoting effects of this organism.

Although bacteria are known to produce other compounds with the potential to alter plant hormone balance such as gibberellins (Bottini et al. 1989), cytokinins (Guo et al. 2001), and salicylic acid (De Meyer et al. 1999), no search for genes involved in the biosynthesis those compounds were reported in the endophyte genome analyses. Gamma-aminobutyrate (GABA) might be involved in plant defense against invertebrate pests (Bown et al. 2006). Taghavi et al. (2009) reported the presence of genes required for GABA synthesis in *Enterobacter* sp. 638. However, because the genome lacks the gene for GABA permease, the bacterium probably does not export GABA as a protecting agent (Taghavi et al. 2010).

A better understanding of the patterns of gene retention, gain and loss that lead to the current constellation of phytohormone-related genes in the endophyte genome collection, and the significance of the presence/absence of such genes, requires a phylogenomic analysis that includes bacteria with various life styles. In addition, the relatively low frequency of known phytostimulation-related genes in the genomes of growth-promoting endophytes suggests that more endophytic phytostimulation pathways remain to be discovered.

4.5.2 Disease Protection and Induced Systemic Resistance

Endophytes can also stimulate plant growth indirectly by inhibiting phytopathogens (that would otherwise inhibit plant growth). They can do this through direct antagonism of microbial pathogens (fungal or bacterial), through competition for space and nutrients, or by activating the plant's natural defense mechanisms, called induced systemic resistance (ISR) (Bordiec et al. 2011).

ISR Bacterial determinants that are thought to produce ISRs include the cell wall outer membrane lipopolysaccharide (LPS), siderophores, flagella, antibiotics, quorum sensing N-acylhomoserine lactones, salicylic acid, and volatile compounds, such as acetoin and 2,3-butanediolsiderophores (Bordiec et al. 2011; van Loon et al. 2008). Many of these are present in the endophyte genomes as discussed above (N-acylhomoserine lactones, volatiles, flagella) or below (siderophores). Moreover, LPS genes were detected in *Enterobacter* sp. 638 (on the plasmid), in *Azoarcus* sp. BH72 (genes are clustered and most similar to LPS genes of gamma-proteobacteria), and *G. diazotrophicus* Pal5 (Taghavi et al. 2010; Krause et al. 2006; Bertalan et al. 2009).

Siderophores and iron uptake Siderophores may serve a dual purpose in plant disease protection, as elicitors of ISR, and through their role in competition for iron. Because the availability of iron often limits microbial growth, an efficient iron uptake system can outcompete pathogens, which has been demonstrated in the rhizosphere (Schippers et al. 1987; O'Sullivan and O'Gara 1992). Endophytic siderophore production has been associated with plant health (Sessitsch et al. 2004), and may contribute to protect the host plant against pathogenic infections, however the role of siderophores *in planta* is unknown, and it is also possible that endophytic siderophore production has a negative effect on the host plant. Bacteria have developed several distinct mechanisms to compete for iron, including (i) specific iron uptake transporters, (ii) the secretion of large numbers of diverse siderophores, and (iii) the synthesis of siderophore receptors to utilize siderophores released from other microorganisms.

Genome analysis suggest that *Enterobacter* sp. 638 is well-equipped to compete for iron; the genome contains genes for synthesis and secretion of the siderophore enterobactin, recovery of the iron-enterobactin complex, and extraction of the iron using an enterobactin esterase (Taghavi et al. 2010). The genes are located in a large cluster adjacent to genes encoding two ABC transporters involved in iron uptake (out of a total of nine). Furthermore, *Enterobacter* sp. 638 possesses 12 outer membrane ferric and ferric-related siderophore receptors, almost twice as many as *Esherichia coli* K12. Likewise, The *V. varivorax* S110 genome has 24 genes encoding siderophore receptors, 16 genes for siderophore biosynthesis and additional 12 genes potentially involved in iron transport and homeostasis (Han et al. 2011).

In contrast, *Azoarcus* sp. BH72 lacks biosynthetic pathways for known siderophores, and production of siderophores was not detected experimentally (Krause et al. 2006). However, *Azoarcus* sp. BH72 possesses 22 genes encoding proteins related to iron transport (more than other nitrogen fixing endosymbionts), including two genes not even present in the genome of *P. fluorescens* Pf5, which is known for its capacity to produce and take up a wide range of siderophores. The authors suggest that the strain is highly adapted to obtaining chelated iron from other microbes, possibly with an antagonizing effect on fungi and bacteria. Kaneko et al. (2010) noted that the number of iron transport proteins in the *Azospirillum* sp. B510 genome was lower than in other N₂-fixing endophytes, including *Azoarcus* sp. BH72 (22 genes). Finally, Wu et al. (2011) noted that *P. putida* W619 has a smaller number of siderophore receptors than its close relatives *P. putida* F1 and *P. putida* KT2440. Iron transport and siderophores were not discussed in the other endophyte genome publications.

Antimicrobial compounds *Enterobacter* sp. 638 has been shown to produce 2-phenylethanol, and carries two genes putatively involved in its synthesis, located in regions not syntenic with other closely related strains (Taghavi et al. 2010). *Enterobacter* sp. 638 also possesses the gene *ubiC*, putatively encoding an enzyme that degrades chorismate into the antimicrobial 4-hydroxybenzoate. The gene *ubiC* is also annotated in the genomes of *Azoarcus* sp. BH72, *H. seropedicae* SmR1, *K. pneumoniae* 342 and *S. maltophilia* R551-3.

P. putida W619 carries a gene encoding mannitol dehydrogenase, which might protect against fungal pathogens (Wu et al. 2011; Jennings et al. 2002). *S. proteamaculans* 568 and *K. pneumoniae* 342 also contain genes annotated as mannitol dehydrogenases.

Bacterial chitinases might protect the plant against fungal pathogens by lysing fungal cell walls, but could also have a role in triggering plant defense mechanisms (Ryan et al. 2008). A chitinase from *S. maltophilia* strain C5 suppresses fungal disease in Kentucky bluegrass by activating plant resistance genes (Kobayashi et al. 2002). *S. maltophilia* R551-3 encodes a homolog of this gene, as well as two other chitinases, and *S. proteamaculans* 568 encodes two chitinases, with possible role in fungal antagonism.

4.5.3 Nitrogen Fixation

Half of the endophytes whose genomes have been sequenced are able to fix nitrogen. Genome analyses confirm the presence of the structural and regulatory genes for the nitrogenase complex in *Azoarcus* sp. BH72, *Azospirillum* sp. B510, *G. diazotrophicus* Pal5, *H. seropedicae* SmR1, *K. pneumoniae* 342, *N. azollae* 0708, and in several cases, the absence of those same genes in close relatives. Although none of the tree-endophytes sequenced are nitrogen fixers, diazotrophic endophytes have been detected in e.g. poplar (Doty et al. 2009).

5 Secretion Systems in Endophytes

Many plant- and animal associated bacteria, both pathogenic and mutualistic, encode specific secretion systems used to translocate so called ‘effector’ molecules (usually proteins but sometimes DNA) that modulate interactions with the host. Effectors are secreted across the cell membrane(s), from the bacterial cell cytoplasm into the extracellular milieu, or directly injected into the cells of the eukaryote host. There are currently seven known distinct, non-general secretion systems (designated type I through type VII). While some of these (e.g. the type II secretion system), depend on the so called *Sec* pathway to have proteins exported across the inner membrane to the periplasm before translocation across the outer membrane, others use complex multi-component protein assemblies that directly translocate proteins from the cytoplasm to the extracellular environment. Although specialized secretion systems such as types III and IV are often associated with pathogenicity, it is the effector proteins they deliver, not the systems themselves that are harmful (Dale and Moran 2006). In fact, secretion systems are commonly used in mutualistic interactions between bacteria and both plants and animals (Deakin and Broughton 2009). For example, a type III secretion system is involved in the establishment of symbiosis by the tsetse fly endosymbiont *Sodalis glossinidius* (Dale et al. 2002), and a type IV secretion system is used by the nodule-forming plant symbiont

Mesorhizobium loti (Hubber et al. 2007). Collectively, the endophyte genomes encode the entire range of secretion systems.

Type I secretion systems are simple, composed of an ABC transporter, a multimeric Membrane Fusion Protein (MFP), and a specific outer membrane protein (OMP) (Delepelaire 2004). In rhizobia, proteins secreted via Type I secretion systems play indirect roles in symbiosis (Finnie et al. 1998). This secretion system is present in the genomes of *Azoarcus* sp. BH72, *K. pneumoniae* 342 and *P. putida* W619. The latter is homologous to the type I secretion system in *Pseudomonas fluorescens* WCS365 that exports a protein called LapA involved in biofilm formation (Hinsa et al. 2003). *P. putida* W619 contains an additional gene coding for an OMP component, and next to it, a putative adhesin and a surface-adhesion outer membrane-like protein with weak similarity to LapA (Wu et al. 2011). These two proteins are candidates for secretion by the Type I system, with possible roles in host colonization.

Type II secretion systems are often used by pathogens to export toxins and hydrolytic enzymes proteins from the periplasm across the outer membrane (Sandkvist 2001). The sugar-cane endophyte *G. diazotrophicus* grows on sucrose but does not directly metabolize it. Instead, the bacterium use a Type II secretion system to secrete levansucrase, which converts sucrose to beta-1,2-oligofructans and levan (Hernandez et al. 1995). Type II secretion systems were also annotated in the genomes of *K. pneumoniae* 342 and *Azoarcus* sp. BH72.

Type III secretion systems are related to bacterial flagella (Blocker et al. 2003) and present in many pathogenic bacteria of animals and plants (McCann and Guttman 2008), but also in endosymbionts as discussed above. Type III secretion systems are used to directly inject effectors into the cytoplasm of host cells (Mota et al. 2005). Phytopathogens use them to inject effector proteins that suppress plant innate immunity and enable colonization of plant tissue (Abramovitch et al. 2006; de Torres et al. 2006). Legume symbionts use them to inject so called 'nodulation outer proteins' involved in host range determination and nodulation (Deakin and Broughton 2009). Only one of the published endophyte genomes (that of *V. paradoxus* S110) encodes a putative Type III secretion system (Han et al. 2011). The paucity of this specific host-interaction system in e.g. *Azoarcus* sp. BH52 was interpreted as adaptive; the authors suggested that the absence of type III secretion systems may prevent the export of toxic proteins to the host (Krause et al. 2006). Similarly, Wu et al. (2011) noted that *P. putida* W619 and other non-pathogenic *P. putida* strains lack the type III secretion system present in the plant pathogen *P. syringae* (Wu et al. 2011). However, Type III secretion systems may have a significant role in the biology of beneficial *Pseudomonas* sp. as DNA hybridization analysis indicates their presence in many plant-colonizing, growth-promoting strains of this species (Preston et al. 2001). Interestingly, two of the unpublished genomes (those of *B. phytofirmans* PsJN and *H. seropedicae* SmR1) encode putative Type III secretion systems, suggesting a possible role of type III secretion in the interaction between these endophytes and their host plants.

Type IV secretion systems are used to translocate both DNA and protein substrates across bacterial membranes, and generally require direct contact with the target cell (Alvarez-Martinez and Christie 2009). There are three types of Type IV secretion systems with different functions; conjugation systems, effector translocator systems, and DNA release/uptake systems. In addition, Type IV secretion systems have been shown to be essential for adhesion to erythrocytes and to determine host range in the mammal-associated bacterium *Bartonella* (Vayssier-Taussat et al. 2010; Nystedt et al. 2008). The effector translocator system is important in bacterium-host interactions as they deliver proteins or DNA to eukaryote cells. However, because Type IV effector translocators have evolved from conjugation systems multiple times during evolution, and can do so in a relatively short time (Frank et al. 2005), the role of a newly discovered, uncharacterized Type IV secretion (conjugation vs. effector translocation) is difficult to predict. Type IV secretion systems are found in *Azospirillum* sp. B510, *G. diazotrophicus* Pal5 (four, all in accessory regions), *K. pneumoniae* 342 (present on integrated element and potentially part of a conjugal transfer system), *V. paradoxus* S110, and in the unpublished genome of *M. populi* BJ001.

The type V secretion pathway is often involved in pathogenesis, and encompasses the autotransporter proteins, the two-partner secretion system, and the Vc or AT-2 family of proteins (Henderson et al. 2004). The autotransporters are large proteins equipped with two translocator domains that enable their export across bacterial membranes. Autotransporters are present in *Azoarcus* sp. BH72, *Enterobacter* sp. 638 (on the plasmid), in *K. pneumoniae* 342, and in *P. putida* W619 (with a pectin/lyase/pertactin domain).

Type VI secretion systems are Sec-independent and related to bacteriophage DNA injection machines (Leiman et al. 2009). Although they can be used to deliver effectors into eukaryote cells (Schwarz et al. 2010b), for example by the pathogen *Vibrio cholerae* (Bingle et al. 2008), it appears as if most bacteria use them to defend against simple eukaryotic cells and other bacteria in the environment (Schwarz et al. 2010a). It has also been demonstrated that the Type VI secretion system is a determining factor for host-specificity in the symbiont *Rhizobium leguminosarum* (Van Brussel et al. 1986). Therefore, endophytic type VI secretion systems could have two possible ‘beneficial’ roles; in host interaction, or in plant defense by antagonizing pathogenic microbes in the endosphere. The type VI secretion system is annotated in *K. pneumoniae* 342, *P. putida* W619, *B. phytofirmans* PsJN, and *V. paradoxus* S110.

Thus, every type of non-general secretion system, except the Type VII secretion system recently discovered in *Mycobacteria* (Abdallah et al. 2007), is annotated in several of the endophyte genomes. Notably, *K. pneumoniae* 342, which can colonize the interior of a wide range of host plants with a very small inoculum dose, encodes all of them except a type III secretion system (Fouts et al. 2008). Endophytic secretion systems are prime candidates for involvement in various host-interaction processes including attachment, colonization, immune evasion and plant defense.

6 Horizontal Gene Transfer and Genomic Islands

In bacteria, horizontal gene transfer plays a key role in the adaptation to specific lifestyles and environmental niches. Horizontally transferred genes are often clustered together on chromosomes on clearly defined ‘genomic islands’ (GI) of genes that either are transferred as a group through transformation, conjugation or transduction, or that are hot spots of horizontal gene transfer (Juhas et al. 2009). Genomic islands are easily detected e.g. through comparative genomics due to the tendency of these regions to differ between otherwise closely related strains, or by the presence of typical features, such as duplicated portions of tRNA genes at their boundaries (created by integration into the tRNA gene), the presence of an integrase gene, or a GC content that is different from the rest of the chromosome (Juhas et al. 2009). Most sequenced bacterial genomes, except those of obligate symbionts, for which gene loss dominates over gain (Moran 2003) have genomic islands, and the expectation is therefore that all endophytes, unless in an obligate association with the host plant, have them. Several of the published endophyte genome papers include an analysis of genomic islands, however, the results are not directly comparable since different methods were used to detect them. Detected genomic island typically encode proteins involved in adaptation to the endophytic lifestyle.

For example, 31 putative genomic islands were identified for *P. putida* W619 (Wu et al. 2011). This species was isolated from a tree growing on a site with contaminated groundwater, and was therefore expected to possess the capacity to deal with heavy metals. Putative heavy metal-responsive genes, many of which are absent in other *P. putida* strains were found on genomic islands on the W619 chromosome. Interestingly, the mannitol dehydrogenase, putatively involved in defense against fungi, is located on a putative genomic island and is lacking from other *P. putida* strains. This gene is perhaps a recent addition to the genome, giving *P. putida* W619 a competitive edge in the endosphere.

In contrast, *Azoarcus* sp. BH72 contains few GIs compared to its soil-borne relative (Krause et al. 2006), and only eight islands of size 6–70 kb were located in *Azospirillum* sp. B510 genome (Kaneko et al. 2010). Twenty-eight genomic islands, altogether encoding >800 proteins, were predicted in the *G. diazotrophicus* Pal5 genome (Bertalan et al. 2009). Two of them appear to be important for adaptation to the endophytic niche, carrying genes involved in oxidative stress, proteases, biosynthesis of antimicrobial agents, amino acid metabolism and secondary metabolites, transport systems and transcriptional regulators.

Eighteen genomic islands were predicted in the *Enterobacter* sp. 638 genome, harboring genes related to sugar transport, adhesion, pectate utilization, iron uptake through siderophore receptors, nitrate reduction, pilus biosynthesis, transporters and regulators (Taghavi et al. 2010). One of the genomic islands is extraordinarily interesting, encoding genes involved sucrose transport and utilization next to genes for synthesis of the volatiles acetoin and 2,3-butanediol, suggesting a coupling in the expression of these two gene clusters. Indeed, quantitative RT-PCR demonstrated

that the production of acetoin and 2,3-butanediol is induced by the presence of sucrose in the growth medium (Taghavi et al. 2010). This result illustrates that genomic location can give clues to function. It also suggests a significant role of endophytic volatiles *in planta*, which is interesting given that several of the other endophyte genomes also encode them.

Eleven ‘site-specific integrated elements’ were identified in the genome of Kp342, including two putatively integrated plasmids (Fouts et al. 2008). One integrated element encodes a beta-(1,2)-glucan, similar to *ndvB*, a gene involved in nodule invasion in *Rhizobium meliloti*.

7 Mobile Genetic Elements

Mobile genetic elements (MGEs) such as insertion sequence (IS) elements, prophages, and plasmids tend to reflect the degree of plasticity of genomes (Frost et al. 2005). In other words, a genome with many mobile genetic elements is usually plastic and more amenable to genetic change. The *Azoarcus* sp. BH72 genome contains only eight MGE loci, in stark contrast to the closely related soil strain, *Azoarcus* sp. EbN1, which contains over 200 genes encoding transposases. Krause et al. (2006) suggest that the lack of MGEs in *Azoarcus* reflects adaptation to the stable, low-stress endophytic niche, whereas the higher number of MGEs in the soil-dwelling relative, and in rhizosphere bacteria in general, reflect a need for ongoing adaptation to a variable niche. Interestingly, the sugarcane endophyte *G. diazotrophicus* Pal5, which, like *Azoarcus* BH72, seems poorly adapted to the soil environment (similar to *Azoarcus*, it is rarely isolated from soil), contains 190 transposases, more than any other endophyte (Bertalan et al. 2009). Thus, the number of transposases in a genome might not reflect the need for continuous adaptation, and a low number of transposases may not necessarily be a result of adaptation to a stable niche. As pointed out by Bertalan et al. (2009), expansion of IS elements can be a non-adaptive consequence of host restriction (Moran and Plague 2004): when bacteria lose the ability to live in the environment, their effective population size is reduced, which in turn decreases the efficiency of purifying selection (the selective removal of alleles that are deleterious). However, the presence of other types of MGEs such as plasmids and prophages may still correlate with life style factors.

7.1 Prophages

Phage transduction is a major mechanism of horizontal gene transfer in bacteria, and genes transferred this way are often located adjacent to prophages (integrated bacteriophages) in bacterial genomes. The *Enterobacter* sp. 638 genome carries eight prophages, an unusually high number, and over 300 phage proteins (Taghavi et al.

2010). Six of the prophages are flanked by regions not present in closely related bacteria. Probably acquired through phage transduction, these regions contain genes important in bacterium-plant interactions such as amino-acid and iron transporters, and a hemagglutinin. Two prophages were detected in the *Azospirillum* sp. B510 chromosome, two in *K. pneumoniae* 342, and three in *P. putida* W619. The other genomes lacked prophages or were not investigated for their presence. Kaneko et al. (2010) were able to demonstrate release of phage particles from *Azospirillum* sp. B510 cells is induced by mitomycin C, and estimated the DNA in the released phage was to be 10 kb in size. Because none of the *Azospirillum* sp. B510 prophages are 10 kb long, the authors suggest that they may have overlooked some other prophages. Alternatively, the *Azospirillum* genome might encode a gene transfer agent or defective prophage that randomly packages 10 kb pieces of the host genome, as has been demonstrated in other Alphaproteobacteria, such as for example *Bartonella grahamii* (Berglund et al. 2009).

7.2 Plasmids

Plasmids extend the accessory gene content of bacteria, and can sometimes encode proteins that give a competitive edge in a particular niche. The *Enterobacter* sp. 638 genome includes one plasmid, which based on the analysis of the ‘backbone’ genes (those involved in plasmid replication and maintenance), belongs to a family of plasmids generally involved in host interaction and virulence (Taghavi et al. 2010). The *Enterobacter* sp. 638 plasmid carries many genes related to plant adhesion and colonization, suggesting a role in host-endophyte interaction. Other plasmid-encoded genes with relevance for the endophytic life style include genes for flagella and nitrogen fixation on *Azospirillum* sp. B510 plasmids, and Type IV secretion on a *G. diazotrophius* Pal5 plasmid (Kaneko et al. 2010; Bertalan et al. 2009).

8 Comparative Genomics

8.1 Comparisons with Close Relatives

Contrasting genomes of close relatives with different life styles is an efficient way to associate presence and absence of genes with life style factors such as niche specificity or host range. However, the result of such comparisons must be interpreted with caution; first it is impossible to know if differences have adaptive meaning or if they simply reflect neutral gains and losses of genes. Second, with only a few data points (e.g. a few genomes) we rarely have sample sizes big enough to draw statistically significant conclusions about the presence and absence of genes. Still this is a common and valid approach to detecting genes putatively responsible

for life in a particular niche. For example, the endophytic *K. pneumoniae* 343 was compared to a clinical strain and potential pathogen isolated from the human respiratory tract, *K. pneumoniae* MGH78578 (Fouts et al. 2008). This strain can colonize plants, but at a much lower efficiency than endophytic *K. pneumoniae* strains. Apart from the nitrogen fixation genes (present in the endophyte only), the two strains show differences in content of genes encoding transcription factors (more in the endophyte), signal transduction (more in the endophyte), surface-associated structures (different, more in the clinical isolate), and secretion systems (two type IV secretion systems present in the endophyte but absent from the clinical isolate). It is possible that the larger number of transcription factors in the endophyte is advantageous, if their purpose is to downregulate genes whose expression would otherwise interfere with plant colonization by inducing plant defense. Furthermore, the smaller number of surface-structures in the endophyte relative to the clinical isolate may reflect an adaptation to avoid the plant immune system, and the two Type IV systems present only in the endophyte may play a role in interaction with the host.

A similar approach was taken for analysis of the *G. diazotrophicus* Pal5 genome, but using the three phylogenetically closest genomes (Bertalan et al. 2009). Regions exclusive to the endophyte included genes for type IV secretion and flagella. Genome comparison between *Azoarcus* sp. BH72 and the related soil bacterium strain EbN1 revealed an unusually low degree of conserved gene order and content (Krause et al. 2006). Genes exclusive to or more common in the endophyte included those for pilus assembly and for other cell surface components potentially important for plant-microbe interactions. The latter were more closely related to those of plant-associated bacteria than to the *Azoarcus* relative, suggesting gene transfer. Moreover, the endophyte encoded more proteins related to iron-transport, and of course, to nitrogen fixation.

8.2 Comparative Endophyte Genomics

Attempts to identify endophyte-specific genes were not particularly successful. Bertalan et al. (2009) searched for common and exclusive coding sequences among nine endophytic genomes but found very few. Similarly, no uniquely shared genes were found between *K. pneumoniae* 342 and *Azoarcus* sp. (Fouts et al. 2008). The presence of unique ‘endophyte genes’ shared only among current endophyte genomes seems unlikely, given both the diversity of organisms occupying this broadly defined niche, and the diversity of hosts from which they were isolated.

8.3 Comparisons with Other Plant-Associated Bacteria

When the *K. pneumoniae* 342 genome was sequenced, only one other endophyte genome (that of *Azoarcus* sp. BH72) had been described. Therefore, to search for

proteins unique to phytobacteria and identify putative genes important for a plant-associated lifestyle, protein sequences of *K. pneumoniae* 342 were compared to those of 28 genomes of other types of bacteria that interact with plants, including plant pathogens and rhizobia (Fouts et al. 2008). Twenty-three proteins were detected, the majority of which were classified as “hypothetical proteins” with unknown function. The fact that most of these proteins have unknown functions likely reflect a gap in our knowledge of plant-bacteria relationships, and make them interesting targets for further functional studies.

Fouts et al. (2008) did another creative comparison, between the *K. pneumoniae* 342 gene complement and ‘plant-induced’ genes from other organisms. These were genes that in earlier studies had been shown to be turned on specifically during colonization or growth associated with plants. Homologs of over 200 such plant-induced proteins were searched for in the *K. pneumoniae* 342 genome. Genes appearing to be involved in adaptation to life in the endosphere were detected, including amino acid nucleotide biosynthesis (amino acids are limited within the plant), stress response genes (to evade plant defense), and plant attachment (e.g. hemagglutinins).

9 Conclusion

Genome analysis of currently sequenced endophyte genomes demonstrates that diverse genome characteristics and diverse sets of host interaction genes are used for colonization of the endophytic niche. Host interaction systems are shared with other phytobacteria as well as with bacteria infecting animals. This diversity likely reflects the fact that bacterial endophytes are broadly defined (as microbes living inside healthy plants), that sequenced endophyte genomes come from a diverse set of bacteria (alpha-, gamma- and beta-proteobacteria, and cyanobacteria) and that the isolates were obtained from a diverse set of hosts (from grasses to trees) and from a relatively diverse set of tissues (roots, stems and cuttings).

This diversity in strategies used to explore the niche inside plants demonstrates the existence of various evolutionary routes to becoming an endophyte. Moreover, the evolutionary routes taken by the bacteria sequenced so far are independent; they are all more closely related to non-endophytes than to each other, and there are significant differences in the type and number of genes used to interact with the plant host. Consequently, attempts to identify shared and unique genes among endophyte genomes have been unsuccessful, and may continue to be so as the number of sequenced endophyte genomes increase. Instead, identification and analysis of known host-interaction systems in endophytes suggest that among the many genes encoding hypothetical proteins, candidates for relevance to the endophytic niche (i.e. with roles in colonization, plant defense evasion and growth promotion) will sometimes be present in all endophyte genomes (e.g. motility genes), but more often be present only in a subset of all endophyte genomes (e.g. bacterial volatiles or type III secretion systems). Moreover, such genes are likely to be shared with other

phytobacteria (e.g. ACC deaminase and IAA biosynthesis genes), and sometimes other host-associated bacteria (e.g. secretion systems and hemagglutinins).

As more endophyte genomes are sequenced, hopefully from understudied hosts such as forest trees, and representing different transmission routes and degrees of interaction with the host (including vertically transmitted, obligate and intracellular endophytes), the known repertoire of genes underlying the adaptation to life within plants is likely to expand.

Acknowledgements This work was supported by UC Merced startup funds.

References

- Abdallah AM, Gey van Pittius NC, Champion PA et al (2007) Type VII secretion – mycobacteria show the way. *Nat Rev Microbiol* 5:883–891
- Abramovitch RB, Anderson JC, Martin GB (2006) Bacterial elicitation and evasion of plant innate immunity. *Nat Rev Mol Cell Biol* 7:601–611
- Ahmad F, Ahmad I, Khan MS (2008) Screening of free-living rhizospheric bacteria for their multiple plant growth promoting activities. *Microbiol Res* 163:173–181
- Alvarez-Martinez CE, Christie PJ (2009) Biological diversity of prokaryotic type IV secretion systems. *Microbiol Mol Biol Rev* 73:775–808
- Antunes LCM, Ferreira RBR, Buckner MMC et al (2010) Quorum sensing in bacterial virulence. *Microbiology* 156:2271–2282
- Balder R, Hassel J, Lipski S et al (2007) *Moraxella catarrhalis* strain O35E expresses two filamentous hemagglutinin-like proteins that mediate adherence to human epithelial cells. *Infect Immun* 75:2765–2775
- Barnhart MM, Chapman MR (2006) Curli biogenesis and function. *Annu Rev Microbiol* 60:131–147
- Belimov A, Hontzeas N, Safronova V et al (2005) Cadmium-tolerant plant growth-promoting bacteria associated with the roots of Indian mustard (*Brassica juncea* L. Czern.). *Soil Biol Biochem* 37:241–250
- Belimov AA, Dodd IC, Hontzeas N et al (2009) Rhizosphere bacteria containing 1-aminocyclopropane-1-carboxylate deaminase increase yield of plants grown in drying soil via both local and systemic hormone signalling. *New Phytol* 181:413–423
- Berglund EC, Frank AC, Calteau A et al (2009) Run-off replication of host-adaptability genes is associated with gene transfer agents in the genome of mouse-infecting *Bartonella grahamii*. *PLoS Genet* 5:e1000546
- Bertalan M, Albano R, de Padua V et al (2009) Complete genome sequence of the sugarcane nitrogen-fixing endophyte *Gluconacetobacter diazotrophicus* Pal5. *BMC Genomics* 10:450
- Bingle LE, Bailey CM, Pallen MJ (2008) Type VI secretion: a beginner's guide. *Curr Opin Microbiol* 11:3–8
- Blanco Y, Blanch M, Pinon D et al (2005) Antagonism of *Gluconacetobacter diazotrophicus* (a sugarcane endosymbiont) against *Xanthomonas albilineans* (pathogen) studied in alginate-immobilized sugarcane stalk tissues. *J Biosci Bioeng* 99:366–371
- Blocker A, Komoriya K, Aizawa S (2003) Type III secretion systems and bacterial flagella: insights into their function from structural similarities. *Proc Natl Acad Sci USA* 100:3027–3030
- Blomqvist K, Nikkola M, Lehtovaara P et al (1993) Characterization of the genes of the 2,3-butanediol operons from *Klebsiella terrigena* and *Enterobacter aerogenes*. *J Bacteriol* 175:1394–1404

- Bohm M, Hurek T, Reinhold-Hurek B (2007) Twitching motility is essential for endophytic rice colonization by the N₂-fixing endophyte *Azoarcus* sp. strain BH72. *Mol Plant Microbe Interact* 20:526–533
- Bordiec S, Paquis S, Lacroix H et al (2011) Comparative analysis of defence responses induced by the endophytic plant growth-promoting rhizobacterium *Burkholderia phytofirmans* strain PsJN and the non-host bacterium *Pseudomonas syringae* pv. pisi in grapevine cell suspensions. *J Exp Bot* 62:595–603
- Bottini R, Fulchieri M, Pearce D et al (1989) Identification of gibberellins A(1), A(3), and Iso-A(3) in cultures of *Azospirillum lipoferum*. *Plant Physiol* 90:45–47
- Bown AW, Macgregor KB, Shelp BJ (2006) Gamma-aminobutyrate: defense against invertebrate pests? *Trends Plant Sci* 11:424–427
- Brandl MT, Quinones B, Lindow SE (2001) Heterogeneous transcription of an indoleacetic acid biosynthetic gene in *Erwinia herbicola* on plant surfaces. *Proc Natl Acad Sci USA* 98:3454–3459
- Camilli A, Bassler BL (2006) Bacterial small-molecule signaling pathways. *Science* 311:1113–1116
- Cho SM, Kang BR, Han SH et al (2008) 2R,3R-butanediol, a bacterial volatile produced by *Pseudomonas chlororaphis* O6, is involved in induction of systemic tolerance to drought in *Arabidopsis thaliana*. *Mol Plant Microbe Interact* 21:1067–1075
- Dale C, Moran NA (2006) Molecular interactions between bacterial symbionts and their hosts. *Cell* 126:453–465
- Dale C, Plague GR, Wang B et al (2002) Type III secretion systems and the evolution of mutualistic endosymbiosis. *Proc Natl Acad Sci USA* 99:12397–12402
- De Meyer G, Capieau K, Audenaert K et al (1999) Nanogram amounts of salicylic acid produced by the rhizobacterium *Pseudomonas aeruginosa* 7NSK2 activate the systemic acquired resistance pathway in bean. *Mol Plant Microbe Interact* 12:450–458
- de Torres M, Mansfield JW, Grabov N et al (2006) *Pseudomonas syringae* effector AvrPtoB suppresses basal defence in *Arabidopsis*. *Plant J* 47:368–382
- Deakin WJ, Broughton WJ (2009) Symbiotic use of pathogenic strategies: rhizobial protein secretion systems. *Nat Rev Microbiol* 7:312–320
- Delepelaire P (2004) Type I secretion in gram-negative bacteria. *Biochim Biophys Acta* 1694:149–161
- Dong Y, Iniguez AL, Triplett EW (2003) Quantitative assessments of the host range and strain specificity of endophytic colonization by *Klebsiella pneumoniae* 342. *Plant Soil* 257:49–59
- Doty SL, Oakley B, Xin G et al (2009) Diazotrophic endophytes of native cottonwood and willow. *Symbiosis* 47:23–33
- Ettema TJ, Andersson SG (2009) The alpha-proteobacteria: the Darwin finches of the bacterial world. *Biol Lett* 5:429–432
- 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
- Finnie C, Zorreguieta A, Hartley NM et al (1998) Characterization of *Rhizobium leguminosarum* exopolysaccharide glycanases that are secreted via a type I exporter and have a novel heptapeptide repeat motif. *J Bacteriol* 180:1691–1699
- Fouts DE, Tyler HL, DeBoy RT et al (2008) Complete genome sequence of the N₂-fixing broad host range endophyte *Klebsiella pneumoniae* 342 and virulence predictions verified in mice. *PLoS Genet* 4:e1000141
- Frank AC, Alsmark CM, Thollesson M et al (2005) Functional divergence and horizontal transfer of type IV secretion systems. *Mol Biol Evol* 22:1325–1336
- Frost LS, Leplae R, Summers AO et al (2005) Mobile genetic elements: the agents of open source evolution. *Nat Rev Microbiol* 3:722–732
- Gill S, Tuteja N (2010) Polyamines and abiotic stress tolerance in plants. *Plant Signal Behav* 5:26–33
- Glick B (1995) The enhancement of plant growth by free-living bacteria. *Can J Microbiol* 42:109–117

- Glick B, Penrose D, Li J (1998) A model for the lowering of plant ethylene concentrations by plant growth-promoting bacteria. *J Theor Biol* 190:63–68
- Gottig N, Garavaglia BS, Garofalo CG et al (2009) A filamentous hemagglutinin-like protein of *Xanthomonas axonopodis* pv. citri, the phytopathogen responsible for citrus canker, is involved in bacterial virulence. *PLoS ONE* 4:e4358
- Guo M, Manulis S, Barash I et al (2001) The operon for cytokinin biosynthesis of *Erwinia herbicola* pv. gypsophilae contains two promoters and is plant induced. *Can J Microbiol* 47:1126–1131
- Han JI, Choi HK, Lee SW et al (2011) Complete genome sequence of the metabolically versatile plant growth-promoting endophyte *Variovorax paradoxus* S110. *J Bacteriol* 193:1183–1190
- Hardoim PR, van Overbeek LS, Elsas JD (2008) Properties of bacterial endophytes and their proposed role in plant growth. *Trends Microbiol* 16:463–471
- Harrison PW, Lower RPI, Kim NKD et al (2010) Introducing the bacterial ‘chromid’: not a chromosome, not a plasmid. *Trends Microbiol* 18:141–148
- Henderson IR, Navarro-Garcia F, Desvaux M et al (2004) Type V protein secretion pathway: the autotransporter story. *Microbiol Mol Biol Rev* 68:692–744
- Hernandez L, Arrieta J, Menendez C et al (1995) Isolation and enzymic properties of levansucrase secreted by *Acetobacter diazotrophicus* SRT4, a bacterium associated with sugar cane. *Biochem J* 309:113–118
- Herron SR, Benen JA, Scavetta RD et al (2000) Structure and function of pectic enzymes: virulence factors of plant pathogens. *Proc Natl Acad Sci USA* 97:8762–8769
- 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
- Hubber AM, Sullivan JT, Ronson CW (2007) Symbiosis-induced cascade regulation of the *Mesorhizobium loti* R7A VirB/D4 type IV secretion system. *Mol Plant Microbe Interact* 20:255–261
- Hurek T, Reinhold-Hurek B, Van Montagu M et al (1994) Root colonization and systemic spreading of *Azoarcus* sp. strain BH72 in grasses. *J Bacteriol* 176:1913–1923
- Iniguez AL, Dong Y, Carter HD et al (2005) Regulation of enteric endophytic bacterial colonization by plant defenses. *Mol Plant Microbe Interact* 18:169–178
- Jennings D, Daub M, Pharr D et al (2002) Constitutive expression of a celery mannitol dehydrogenase in tobacco enhances resistance to the mannitol-secreting fungal pathogen *Alternaria alternata*. *Plant J* 32:41–49
- Juhas M, van der Meer JR, Gaillard M et al (2009) Genomic islands: tools of bacterial horizontal gene transfer and evolution. *FEMS Microbiol Rev* 33:376–393
- Kaneko T, Nakamura Y, Sato S et al (2000) Complete genome structure of the nitrogen-fixing symbiotic bacterium *Mesorhizobium loti*. *DNA Res* 7:331–338
- Kaneko T, Nakamura Y, Sato S et al (2002) Complete genomic sequence of nitrogen-fixing symbiotic bacterium *Bradyrhizobium japonicum* USDA110. *DNA Res* 9:189–197
- Kaneko T, Minamisawa K, Isawa T et al (2010) Complete genomic structure of the cultivated rice endophyte *Azospirillum* sp. B510. *DNA Res* 17:37–50
- Kanzler BE, Pfannes KR, Vogl K et al (2005) Molecular characterization of the nonphotosynthetic partner bacterium in the consortium “*Chlorochromatium aggregatum*”. *Appl Environ Microbiol* 71:7434–7441
- Kitano H (2007) Towards a theory of biological robustness. *Mol Syst Biol* 3:137
- Kobayashi DY, Reedy RM, Bick J et al (2002) Characterization of a chitinase gene from *Stenotrophomonas maltophilia* strain 34 S1 and its involvement in biological control. *Appl Environ Microbiol* 68:1047–1054
- Krause A, Ramakumar A, Bartels D et al (2006) Complete genome of the mutualistic, N₂-fixing grass endophyte *Azoarcus* sp. strain BH72. *Nat Biotechnol* 24:1385–1391
- Lapierre P, Gogarten JP (2009) Estimating the size of the bacterial pan-genome. *Trends Genet* 25:107–110

- Lee S, Flores-Encarnacion M, Contreras-Zentella M et al (2004) Indole-3-acetic acid biosynthesis is deficient in *Gluconacetobacter diazotrophicus* strains with mutations in cytochrome c biogenesis genes. *J Bacteriol* 186:5384–5391
- Leiman PG, Basler M, Ramagopal UA et al (2009) Type VI secretion apparatus and phage tail-associated protein complexes share a common evolutionary origin. *Proc Natl Acad Sci USA* 106:4154–4159
- Looney W, Narita M, Mühlemann K (2009) *Stenotrophomonas maltophilia*: an emerging opportunist human pathogen. *Lancet Infect Dis* 9:312–323
- Lukjancenko O, Wassenaar TM, Ussery DW (2010) Comparison of 61 sequenced *Escherichia coli* genomes. *Microb Ecol* 60:708–720
- Madmony A, Chernin L, Pleban S et al (2005) *Enterobacter cloacae*, an obligatory endophyte of pollen grains of Mediterranean pines. *Folia Microbiol Praha* 50:209–216
- Manulis S, Haviv-Chesner A, Brandl MT et al (1998) Differential involvement of indole-3-acetic acid biosynthetic pathways in pathogenicity and epiphytic fitness of *Erwinia herbicola* pv. *gypsophylae*. *Mol Plant Microbe Interact* 11:634–642
- McCann HC, Guttman DS (2008) Evolution of the type III secretion system and its effectors in plant-microbe interactions. *New Phytol* 177:33–47
- Mehnaz S, Lazarovits G (2006) Inoculation effects of *Pseudomonas putida*, *Gluconacetobacter azotocaptans*, and *Azospirillum lipoferum* on corn plant growth under greenhouse conditions. *Microb Ecol* 51:326–335
- Miller MB, Bassler BL (2001) Quorum sensing in bacteria. *Annu Rev Microbiol* 55:165–199
- Moran NA (2003) Tracing the evolution of gene loss in obligate bacterial symbionts. *Curr Opin Microbiol* 6:512–518
- Moran NA, Plague GR (2004) Genomic changes following host restriction in bacteria. *Curr Opin Genet Dev* 14:627–633
- Morgan PW, Drew MC (1997) Ethylene and plant responses to stress. *Physiol Plant* 100:620–630
- Mota LJ, Sorg I, Cornelis GR (2005) Type III secretion: the bacteria-eukaryotic cell express. *FEMS Microbiol Lett* 252:1–10
- Navarro L, Dunoyer P, Jay F et al (2006) A plant miRNA contributes to antibacterial resistance by repressing auxin signaling. *Science* 312:436–439
- Nystedt B, Frank AC, Thollesson M et al (2008) Diversifying selection and concerted evolution of a type IV secretion system in *Bartonella*. *Mol Biol Evol* 25(2):287–300
- O'Sullivan DJ, O'Gara F (1992) Traits of fluorescent *Pseudomonas* spp. involved in suppression of plant root pathogens. *Microbiol Rev* 56:662–676
- Ochman H, Moran NA (2001) Genes lost and genes found: evolution of bacterial pathogenesis and symbiosis. *Science* 292:1096–1099
- Perrig D, Boiero ML, Masciarelli OA et al (2007) Plant-growth-promoting compounds produced by two agronomically important strains of *Azospirillum brasilense*, and implications for inoculant formulation. *Appl Microbiol Biotechnol* 75:1143–1150
- Perrine FM, Hocart CH, Hynes MF et al (2005) Plasmid-associated genes in the model micro-symbiont *Sinorhizobium meliloti* 1021 affect the growth and development of young rice seedlings. *Env Microbiol* 7:1826–1838
- Philippot L, Bru D, Saby NP et al (2009) Spatial patterns of bacterial taxa in nature reflect ecological traits of deep branches of the 16 S rRNA bacterial tree. *Env Microbiol* 11:3096–3104
- Philippot L, Andersson SG, Battin TJ et al (2010) The ecological coherence of high bacterial taxonomic ranks. *Nat Rev* 8:523–529
- Ping L, Boland W (2004) Signals from the underground: bacterial volatiles promote growth in *Arabidopsis*. *Trends Plant Sci* 9:263–266
- Preston GM, Bertrand N, Rainey PB (2001) Type III secretion in plant growth-promoting *Pseudomonas fluorescens* SBW25. *Mol Microbiol* 41:999–1014
- Ramsay JP, Sullivan JT, Stuart GS et al (2006) Excision and transfer of the *Mesorhizobium loti* R7A symbiosis island requires an integrase IntS, a novel recombination directionality factor RdfS, and a putative relaxase RlxS. *Mol Microbiol* 62:723–734

- Ran L, Larsson J, Vigil-Stenman T et al (2010) Genome erosion in a nitrogen-fixing vertically transmitted endosymbiotic multicellular cyanobacterium. *PLoS ONE* 5:e11486
- Reinhold-Hurek B, Hurek T (1998a) Interactions of gramineous plants with *Azoarcus* spp. and other diazotrophs: identification, localization and perspectives to study their function. *Crit Rev Plant Sci* 17:29–54
- Reinhold-Hurek B, Hurek T (1998b) Life in grasses: diazotrophic endophytes. *Trends Microbiol* 6:139–144
- Reinhold-Hurek B, Maes T, Gemmer S et al (2006) An endoglucanase is involved in infection of rice roots by the not-cellulose-metabolizing endophyte *Azoarcus* sp. strain BH72. *Mol Plant Microbe Interact* 19:181–188
- Rodriguez-Navarro DN, Dardanelli MS, Ruiz-Sainz JE (2007) Attachment of bacteria to the roots of higher plants. *FEMS Microbiol Lett* 272:127–136
- Rosenblueth M, Martinez-Romero E (2006) Bacterial endophytes and their interactions with hosts. *Mol Plant Microbe Interact* 19:827–837
- Rothballer M, Eckert B, Schmid M et al (2008) Endophytic root colonization of gramineous plants by *Herbaspirillum frisingense*. *FEMS Microbiol Ecol* 66:85–95
- Ryan RP, Germaine K, Franks A et al (2008) Bacterial endophytes: recent developments and applications. *FEMS Microbiol Lett* 278(1):1–9
- Ryu C-M, Farag MA, Hu C-H et al (2003) Bacterial volatiles promote growth in *Arabidopsis*. *Proc Natl Acad Sci USA* 100:4927–4932
- Ryu C-M, Farag MA, Hu C-H et al (2004) Bacterial volatiles induce systemic resistance in *Arabidopsis*. *Plant Physiol* 134:1017–1026
- Sanchez-Contreras M, Bauer WD, Gao M et al (2007) Quorum-sensing regulation in rhizobia and its role in symbiotic interactions with legumes. *Philos Trans R Soc Lond B Biol Sci* 362:1149–1163
- Sandkvist M (2001) Biology of type II secretion. *Mol Microbiol* 40:271–283
- Saravanan VS, Kalaiarasan P, Madhaiyan M et al (2007) Solubilization of insoluble zinc compounds by *Gluconacetobacter diazotrophicus* and the detrimental action of zinc ion (Zn²⁺) and zinc chelates on root knot nematode *Meloidogyne incognita*. *Lett Appl Microbiol* 44:235–241
- Schippers B, Bakker AW, Bakker PA (1987) Interactions of deleterious and beneficial rhizosphere microorganisms and the effect of cropping practises. *Ann Rev Phytopathol* 25:339–358
- Schwarz S, Hood RD, Mougous JD (2010a) What is type VI secretion doing in all those bugs? *Trends Microbiol* 18:531–537
- Schwarz S, West TE, Boyer F et al (2010b) *Burkholderia* type VI secretion systems have distinct roles in eukaryotic and bacterial cell interactions. *PLoS Pathog* 6:e1001068
- Sessitsch A, Reiter B, Berg G (2004) Endophytic bacterial communities of field-grown potato plants and their plant-growth-promoting and antagonistic abilities. *Can J Microbiol* 50:239–249
- Siddikee MA, Glick BR, Chauhan PS et al. (2011) Enhancement of growth and salt tolerance of red pepper seedlings (*Capsicum annuum* L.) by regulating stress ethylene synthesis with halotolerant bacteria containing 1-aminocyclopropane-1-carboxylic acid deaminase activity. *Plant Physiol Biochem. In press*
- Simon C, Daniel R (2011) Metagenomic analyses: past and future trends. *Appl Environ Microbiol* 77:1153–1161
- Spaepen S, Vanderleyden J (2010) Auxin and Plant-Microbe Interactions. *Cold Spring Harb Perspect Biol. In press*
- Spaepen S, Vanderleyden J, Remans R (2007) Indole-3-acetic acid in microbial and microorganism-plant signaling. *FEMS Microbiol Rev* 31:425–428
- Sun Y, Cheng Z, Glick BR (2009) The presence of a 1-aminocyclopropane-1-carboxylate (ACC) deaminase deletion mutation alters the physiology of the endophytic plant growth-promoting bacterium *Burkholderia phytofirmans* PsJN. *FEMS Microbiol Lett* 296:131–136

- Suzuki S, He Y, Oyaizu H (2003) Indole-3-acetic acid production in *Pseudomonas fluorescens* HP72 and its association with suppression of creeping bentgrass brown patch. *Curr Microbiol* 47:138–143
- Taghavi S, Garafola C, Monchy S et al (2009) Genome survey and characterization of endophytic bacteria exhibiting a beneficial effect on growth and development of poplar trees. *Appl Environ Microbiol* 75:748–757
- Taghavi S, van der Lelie D, Hoffman A et al (2010) Genome sequence of the plant growth promoting endophytic bacterium *Enterobacter* sp. 638. *PLoS Genet* 6:e1000943
- Tettelin H, Riley D, Cattuto C et al (2008) Comparative genomics: the bacterial pan-genome. *Curr Opin Microbiol* 11:472–477
- Toft C, Andersson SG (2010) Evolutionary microbial genomics: insights into bacterial host adaptation. *Nat Rev Genet* 11:465–475
- Van Brussel AA, Zaat SA, Cremers HC et al (1986) Role of plant root exudate and Sym plasmid-localized nodulation genes in the synthesis by *Rhizobium leguminosarum* of Tsr factor, which causes thick and short roots on common vetch. *J Bacteriol* 165:517–522
- van der Lelie D, Taghavi S, Monchy S et al (2009) Poplar and its bacterial endophytes: coexistence and harmony. *Crit Rev Plant Sci* 28:346–358
- van Loon LC, Bakker PA, van der Heijden WH et al (2008) Early responses of tobacco suspension cells to rhizobacterial elicitors of induced systemic resistance. *Mol Plant Microbe Interact* 21:1609–1621
- Vayssier-Taussat M, Le Rhun D, Deng HK et al (2010) The Trw type IV secretion system of *Bartonella* mediates host-specific adhesion to erythrocytes. *PLoS Pathog* 6:e1000946
- Vial L, Cuny C, Gluchoff-Fiasson K et al (2006) N-acyl-homoserine lactone-mediated quorum-sensing in *Azospirillum*: an exception rather than a rule. *FEMS Microbiol Ecol* 58:155–168
- Wooley JC, Godzik A, Friedberg I (2010) A primer on metagenomics. *PLoS Comput Biol* 6:e1000667
- Woyke T, Tighe D, Mavromatis K et al (2010) One bacterial cell, one complete genome. *PLoS ONE* 5:e10314
- Wu X, Monchy S, Taghavi S et al (2011) Comparative genomics and functional analysis of niche-specific adaptation in *Pseudomonas putida*. *FEMS Microbiol Rev* 35:299–323
- Zeidler D, Zähringer U, Gerber I et al (2004) Innate immunity in *Arabidopsis thaliana*: lipopolysaccharides activate nitric oxide synthase (NOS) and induce defense genes. *Proc Natl Acad Sci USA* 101:15811–15816