Dennis C. Gross · Ann Lichens-Park
Chittaranjan Kole *Editors*

Genomics of Plant-Associated Bacteria

Springer

Genomics of Plant-Associated Bacteria

Dennis C. Gross · Ann Lichens-Park
Chittaranjan Kole
Editors

Genomics of Plant-Associated Bacteria

Springer

Editors
Dennis C. Gross
Department of Plant Pathology
and Microbiology
Texas A&M University
College Station, TX
USA

Ann Lichens-Park
United States Department of Agriculture
National Institute of Food
and Agriculture
Washington, DC
USA

Chittaranjan Kole
Bidhan Chandra Krishi Viswavidyalaya
Mohanpur, West Bengal
India

ISBN 978-3-642-55377-6 ISBN 978-3-642-55378-3 (eBook)
DOI 10.1007/978-3-642-55378-3
Springer Heidelberg New York Dordrecht London

Library of Congress Control Number: 2014940933

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

Preface

The first genome of a plant pathogen (*Xylella fastidiosa* that causes citrus variegated chlorosis) was published in 2000. In that same year, the United States Department of Agriculture (USDA) first offered a competitive grants program to support sequencing of agriculturally relevant microorganisms. At that time, only a few microbes of agricultural significance were being sequenced. In 2001, the United States National Science Foundation (NSF) joined USDA, and the jointly-offered program was expanded to support the genomic sequencing of an even broader range of microorganisms. The partnership lasted until 2009, by which time sequencing costs had decreased dramatically and sequencing speed had increased enormously. USDA's support for microbial genomics shifted more towards functional analysis of genome sequences. During the 10 years that the genome sequencing program was offered, the genomes of a large number of agriculturally significant microorganisms were sequenced with support from the USDA, including bacteria, viruses, fungi, oomycetes, and even a nematode. The genomes of microbes relevant to basic science were sequenced with funding from the NSF. Three books are being published to describe the impact of some agriculturally relevant genomes and their analysis. In addition to this volume, two other volumes (edited by Dr. Ralph Dean, Dr. Ann Lichens-Park and Dr. Chittaranjan Kole) describe the genomic analysis of plant-associated fungi and oomycetes. These volumes are entitled "Genomics of Plant-Associated Fungi: Monocot Pathogens" and "Genomics of Plant-Associated Fungi and Oomycetes: Dicot Pathogens." This book describes how the availability of some agriculturally important plant-associated bacterial genomes, many of which were first sequenced with support from the USDA, have revolutionized our understanding of these bacteria and how they can be managed to improve the sustainability of agriculture worldwide.

Bacterial genome data is a launching pad (or scientific springboard) that, with appropriate functional analysis, can lead to greater knowledge of bacterial evolution and diversity, ecological adaptation, and environmental influences on gene expression and phenotype. Because

bacterial genomes are generally much smaller than the genomes of eukaryotic organisms, important insights about how the bacteria function can be obtained from appropriate functional analysis, often in less time than that needed for such analyses employed in higher eukaryotes. At the same time, studies aimed at understanding how bacteria interact with plants can lead to new insights about important plant processes and metabolic pathways, some of which may be exploited by bacterial pathogens to cause diseases. As is explained in the chapter on the genomics of *Xanthamonas oryzae*, the availability of xanthomonad genomic information has led to discoveries of new technologies such as transcription activator-like effector nucleases (TALENs) that facilitate targeted genome editing and will revolutionize approaches to genetic engineering in eukaryotes.

Each chapter in this book describes the genomic analysis of a particular bacterial genus, species or group of related bacteria about which the genome sequence and genomic analysis have led to significant new insights. Some chapters address bacterial pathogens that are readily tractable to genetic analysis and are, therefore, considered to be good model systems. *Pseudomonas syringae*, the subject of two chapters, is an important pathogen and a model system. Other chapters focus on bacterial pathogens that are less tractable genetically, but which are responsible for diseases that can result in devastating economic losses for growers. *Xanthamonas citri* causes citrus canker, a disease that can result in severe losses to citrus crops. One chapter focuses on *Pseudomonas flourescens*, which is a beneficial bacterial species that can help prevent diseases in plants. The very first plant pathogen, *Erwinia amylovora*, was first described in the 1800's and is the subject of one of the chapters. The fastidious phloem-limited bacteria, exemplified by the phytoplasmas and the liberibacters, were first described in recent years, and are each the subject of one chapter.

Several "themes" run through the chapters in this volume. These include the growing evidence for the importance of horizontal gene transfer or "nature's transgenes" in originating new bacterial strains and species. Advances in transcriptomic analysis are facilitating studies describing complex regulatory networks critical to expression of processes important in plant–microbe interactions. The technology is leading to identification of new bacterial factors or products that mediate communication with and establishment in the plant host. Genomic studies of plant-associated bacteria promise to lead to a better understanding of the natural microbial communities associated with plants (the phytobiome) and to innovative means of controlling diseases caused by plant pathogens.

We wish to express our thanks to the lead authors and co-authors of the chapters in this volume. They have done a marvelous job of

explaining the advances and significance of the new knowledge described in their chapters. We also wish to express our thanks to some special people who are current or former employees of USDA and NSF whose support has been critical to the microbial genomics program and to the existence of this volume. These people are Dr. Sonny Ramaswamy, Dr. Colien Hefferan, Ms. Betty Lou Gilliland, Ms. Erin Daly, Mr. Edward Nwaba, Dr. Deborah Sheely, Ms. Cynthia Montgomery, Dr. Michael Fitzner, Dr. Daniel Jones, Ms. Pushpa Kathir, Dr. Anna Palmisano, Dr. Mark Poth, Dr. Maryanna Henkart, and all of the USDA and NSF Program Officers and staff who worked with Dr. Lichens-Park while the Microbial Genome Sequencing Program was offered. Space limitations prevent us from describing each of the roles played by these individuals but they all have been significant and we are extremely grateful to all of them.

March 2014

Dr. Dennis C. Gross
Texas A&M University
College Station, TX, USA

Dr. Ann Lichens-Park
National Institute of Food and Agriculture
United States Department of Agriculture
Washington, DC, USA

Prof. Chittaranjan Kole
Bidhan Chandra Krishi Viswavidyalaya
West Bengal, India

Contents

Genomics of *Erwinia amylovora* and Related *Erwinia* Species Associated with Pome Fruit Trees

1

Youfu Zhao

1.1 Introduction

Erwinia amylovora, the causal agent of fire blight of apple, pear, quince, blackberry, raspberry, and other rosaceous plants, is of particular interest to plant bacteriologists, not only because it is the first bacterium demonstrated to cause disease in plants, but also because of its significant economic and political impact (Griffith et al. 2003; van der Zwet et al. 2012; Vanneste 2000). Recently, the pathogen was voted as one of the top 10 plant pathogenic bacteria in molecular plant pathology (Mansfield et al. 2012). In the last decade, several new pathogenic *Erwinia* species associated with pome fruit trees have been described, including *Erwinia pyrifoliae*, *Erwinia piriflorinigrans*, and *Erwinia uzenensis* (Kim et al. 1999; Matsuura et al. 2012; Lopez et al. 2011). Additionally, *Erwinia billingiae* and *Erwinia tasmaniensis* are two non-pathogenic *Erwinia* species associated with pome fruit trees (Geider et al. 2006; Mergaert et al. 1999). *E. amylovora* is the type species of the genus (Lelliott and Dickey 1984) and remains the most studied species of this genus. Less information is available on other *Erwinia* species associated with pome fruit trees and their genetic relationship to *E. amylovora* (Palacio-Bielsa et al. 2012). Resolution of the genetic compositions of these microorganisms has, therefore, dramatically increased our knowledge base of *E. amylovora* and its relatives (Zhao and Qi 2011). Complete and draft genome sequences for more than a dozen strains, belonging to five *Erwinia* species, including *E. amylovora*, *E. pyrifoliae*, *E. piriflorinigrans*, *E. tasmaniensis*, and *E. billingiae*, have been published (Kube et al. 2008, 2010; Mann et al. 2013; Park et al. 2011; Powney et al. 2011; Sebaihia et al. 2010; Smits et al. 2010a, b, 2013). These genome sequences provide almost complete genetic information about *E. amylovora* and other closely related species. In this chapter, we present (1) an updated review of *E. amylovora* and related species from genome sequencing efforts; (2) summarize the general characteristics of the pathogen, the disease it causes, and its genome; and (3) highlight current genome-enabled understanding of *E. amylovora* pathogenesis, including comparative genomic analyses and evolution, as well as genetic and functional genomic studies. Future perspectives and research directions for this important pathogen are also discussed.

1.1.1 Fire Blight and Related Diseases Associated with Pome Fruit Trees

Fire blight has been known as one of the most important plant bacterial diseases worldwide and is a devastating necrotic disease affecting apples, pears, and other rosaceous plants (Norelli et al.

Y. Zhao (✉)
University of Illinois at Urbana-Champaign,
Urbana, IL 61801, USA
e-mail: zhao888@illinois.edu

D. C. Gross et al. (eds.), *Genomics of Plant-Associated Bacteria*,
DOI: 10.1007/978-3-642-55378-3_1,

2003). Fire blight was first observed on rosaceous plants in 1780 and was considered the first recognized plant bacterial disease back in 1884 (van der Zwet and Keil 1979). Several bacterial diseases with symptoms resembling those of fire blight have been described for pome fruit trees in recent years. Bacterial shoot blight of pear, first described as a fire blight-like disorder, was observed on Asian pear (*Pyrus ussuriensis* cv. Mishirazu) and European pear (*Pyrus communis* L.) trees on Hokkaido island in Japan in the 1970s (Tanii et al. 1981), and the disease was considered eradicated in 1999 (Mizuno et al. 2000, 2010). A similar disease was described on Asian pear (*Pyrus pyrifolia* cvs. Shingo and Mansamgil) in Korea in the late 1990s (Kim et al. 1999). In 2007, bacterial black shoot disease of European pear cv. La France was found in Kaminoyama City, Yamagata Prefecture, Japan (Mizuno et al. 2010). At the same time, a necrotic pear blossom disease of European pear cvs. Ercolini (Coscia) and Tendral was observed in Valencia, Spain (Rosello et al. 2006). Due to their restricted distribution and limited access to disease samples for the fire blight research community, there is not much information about these new diseases beyond their original location. It is expected that these diseases may have a wider distribution than has been reported.

1.1.1.1 Fire Blight Disease and Its Current Distribution

Historically, a severe outbreak of fire blight in the summer of 1880 in Illinois led to the discovery of the first bacterial disease of plants by Thomas J. Burrill, a professor and former president of the University of Illinois at Champaign-Urbana (Griffith et al. 2003). The disease is indigenous to North America and was first observed on pear and quince (*Cydonia oblonga* Mill) in the lower Hudson Valley of New York in 1780 (Griffith et al. 2003). Since 1862, the disease has been widespread and destructive in commercial apple (*Malus sylvestris* Mill) and pear orchards in the northeastern United States. In 1882, the disease spread from New York to Illinois and to California in 1887 (Kado 2000; Pierce 1902). Another milestone in fire blight research occurred in 1891, when insects were demonstrated to be attracted to exudates from infected shoots in the spring and subsequently transmitted the disease to other trees (Kado 2000). This is also the first (bacterial) plant pathogen for which an insect vector was demonstrated.

Since its discovery, fire blight has spread to more than 50 countries (van der Zwet et al. 2012). Fire blight was first observed in Ontario, Canada, in 1904; in British Columbia in 1911; and in Mexico in 1921. Since 1924, the disease has spread to all pear- and apple-growing areas of Canada and Mexico. The reported long-distance spread of fire blight from America was to New Zealand in 1919; to the United Kingdom and Northern Europe in the late 1950s (Crosse et al. 1958); and to the Mediterranean Region and Northern Africa in the 1960s. Fire blight is now widespread across Europe, Northern Africa, and the Middle East including Iran, threatening the native origin of apple germplasm resources in central Asia (van der Zwet et al. 2012). Interestingly, fire blight has not been reported in Australia, China, South Africa, and any country in South America. However, fire blight-related diseases associated with pome fruit trees were reported in Japan and South Korea (Kim et al. 1999; Mizuno et al. 2010).

1.1.1.2 Economic Losses and Costs of Fire Blight Disease Management

Since its discovery, fire blight has been considered the most destructive disease and a limiting factor for apple and pear production (Palacio-Bielsa et al. 2012). Fire blight not only can greatly reduce crop yield and marketability in the current season by infecting blossoms and killing of fruit spurs, but also cause the loss of entire trees and orchards. The severity of fire blight outbreaks in California, and the Midwest prevents the commercial production of pears in these regions (Eastgate 2000; Pierce 1902). The increased replanting cost of high-density blocks (ca. $20,000 per ha) and losses due to international trade regulations on fire blight has resulted

in significant financial losses for many growers. Therefore, fire blight poses multifarious threats to the global pome fruit industry.

It is extremely difficult to accurately estimate the economic losses due to fire blight to an individual grower, a region, and on a global scale. However, some of the most traumatic fire blight epidemics recorded in recent years provide a glimpse into how fire blight can cause great economic losses. In 1998, losses were estimated to be in excess of $68 million in Washington and northern Oregon (van der Zwet et al. 2012). In 2000, the most widespread epidemic of fire blight occurred in southern Michigan and losses were estimated to be $42 million, including removal of 300,000 trees and replacement of 1,550 acres of young orchards (Longstroth 2000).

Fire blight has also become a great threat to the world's pome fruit industry. Losses were estimated at 10 million dollars (NZ) in the Hawke's Bay region of New Zealand (Vanneste 2000). In the 1990s, more than a half million trees were destroyed in Italy alone (Vanneste 2000). In 2007, a severe outbreak of fire blight occurred in Switzerland. The Swiss government reported a loss of $27.5 million dollars and 10 % of their apple acreage (Ashton 2008). Fire blight also affects the ornamental nursery business. The most popular ornamental plant in Europe in the 1970s was cotoneaster, a major host of fire blight. In 1975, more than 2 million cotoneasters were destroyed in nurseries and garden centers in the Netherlands (van der Zwet et al. 2012).

Since the 1970s, spray applications of streptomycin have been the most effective means of controlling blossom blight of apples and pears. However, the occurrence of streptomycin resistance in the USA, Canada, and elsewhere has rendered this antibiotic ineffective (Chiou and Jones 1995; Coyier and Covey 1975; McManus et al. 2002). Moreover, the use of streptomycin for control and the progressive accumulation of resistant strains have been estimated to bring additional losses of more than 100 million dollars (US) per year in the USA (Norelli et al. 2003). Furthermore, fire blight also results in great economic losses due to stringent quarantine and international trade regulations. For example, Australia, free of fire blight, reported spending $40 million Australian dollars for the eradication, diagnostics, loss of sales, and exports due to an unconfirmed presence of fire blight (Rodoni et al. 2006). Moreover, strict regulatory measures against *E. amylovora* are still imposed by many European countries, such as Switzerland, Belgium, and Germany (Deckers 1996; Duffy et al. 2005; EPPO 1992). In Germany, expenses for eradication, inspection, fruit losses, and chemical treatment of fire blight totaled $110 million from 1972 to 1990 (van der Zwet et al. 2012).

1.1.1.3 Disease Symptoms and Life Cycle

The disease name "fire blight" was first coined by William Coxe in 1817, when he described a disorder that "in a few hours, turn the leaves suddenly brown, as if they had passed through a hot flame" (Griffith et al. 2003). Now, we can easily recognize that this is the typical symptoms of fire blight, i.e., the appearance of blackening tissue as though they had been scorched by fire (Schroth et al. 1974). In nature, *E. amylovora* is capable of infecting blossoms, fruits, vegetative shoots, woody tissues, and rootstock crowns, leading to blossom blight, fruit blight, shoot blight, twig and trunk blight, and rootstock blight symptoms with frequent creamy ooze production (Eastgate 2000). Usually, blossom blight is the first symptom of the disease in early spring followed by infection of young fruits (Thomson 1986). Succulent shoots, twigs, and water sprouts are the next most susceptible parts of the tree. Infected young succulent shoots and twigs wither and turn brown and in most cases, the tip of the shoot bends in a characteristic fashion to form a "shepherd's crook." From the infected blossoms, shoots, or fruits, the disease spreads systemically through the spurs to larger twigs and branches to cause cankers. Disease then may spread into the scaffold limbs and the trunk, resulting in the death of the entire tree. The fire blight disease cycle has been extensively described by a number of authors

(Thomson 1986; van der Zwet and Keil 1979; van der Zwet et al. 2012).

Bacterial shoot blight disease, bacterial black shoot disease, and necrotic pear blossom disease induce very similar symptoms to those of fire blight, but major differences exist (Palacio-Bielsa et al. 2012). The symptoms of bacterial shoot blight disease of Asian pear, caused by *E. pyrifoliae*, include black to brown stripes in the leaf midribs, dark brown leaf spots, and necrotic petioles on large parts of the trees (Kim et al. 1999; Rhim et al. 1999). However, the symptoms of black shoot disease of European pear trees, caused by *E. uzenensis* (Matsuura et al. 2012), were different from those of bacterial shoot blight disease (Rhim et al. 1999) and fire blight. Necrotic symptoms were only observed in young shoots, and the development of lesions stopped within 20 cm from the base of the shoots and did not affect the branches (Mizuno et al. 2000). Some typical symptoms of fire blight and bacterial shoot blight disease (i.e., blossom blight, fruitlet blight, and formation of a shepherd's crook), were not observed for black shoot disease (Matsuura et al. 2012). Furthermore, the symptoms of necrotic pear blossom disease, caused by *E. piriflorinigrans*, were also different from typical fire blight disease (Rosello et al. 2006; Lopez et al. 2011). Only pear blossoms showed necrotic symptoms, but not pear shoots or fruitlets, apple trees, or other inoculated *Rosaceae* species (Rosello et al. 2006).

1.1.2 The Pathogen *E. amylovora* and Related *Erwinia* Species

1.1.2.1 General Characteristics

E. amylovora and related *Erwinia* species all belong to the family *Enterobacteriaceae* and the genus *Erwinia*. *E. amylovora* was initially described as *Micrococcus amylovorus* by Professor Burrill in 1883 and then *Bacillus amylovorus* (Burrill) by Trevisan in 1889, under the erroneous assumption that the pathogen hydrolyzes starch. In the early 1900s, it was renamed as *E. amylovora* (Burrill) by Winslow et al. (Lelliott and Dickey 1984). Though *E. amylovora* and related *Erwinia* species share many basic characteristics, such as Gram negative; rod shaped; motile with peritrichous flagella, facultative anaerobic growth, oxidase negative, catalase positive; and acid production from glucose, fructose, and galactose (Palacio-Bielsa et al. 2012), they differ in some phenotypic traits, host range, and virulence factors (Kim et al. 1999; Lopez et al. 2011; Mizuno et al. 2010; Rhim et al. 1999; Rosello et al. 2006; Shrestha et al. 2003).

Host Range of *E. amylovora* and Related *Erwinia* Species

One major difference between *E. amylovora* and related *Erwinia* species is their abilities to cause disease on different host plants (host range). *E. amylovora* has a wide host range within the family of *Rosaceae*. However, other pathogenic *Erwinia* species associated with pome fruit trees are host-specific, only infecting certain varieties of Asian or European pear trees, and some with tissue specificity. For example, *E. piriflorinigrans* only infects blossoms of European pear trees (Rosello et al. 2006), whereas *E. uzenensis* causes disease on young shoots of European pear trees with limited disease progress after infection (Mizuno et al. 2010). The host range of *E. pyrifoliae* may be broader, as disease symptoms are observed on several varieties of Asian pears and after inoculation, on several commercial European pear cultivars and apple (*Malus domestica*) (Kim et al. 2001; Mizuno et al. 2010). Furthermore, some *E. amylovora* strains isolated from *Rubus* plants within the subfamily *Rosoideae* are also host-specific, which can only infect *Rubus* plants. Interestingly, natural *E. amylovora* strains isolated from *Maloideae* (*Spiraeoideae*) with a wide host range show differential virulence on different apple cultivars (Lee et al. 2010; Wang et al. 2010a). These observations and early genetic studies suggest that *E. amylovora* strains may further be divided into different species or subspecies with distinct host ranges, i.e., strains isolated from *Maloideae* and *Rosoideae*.

The host range of *E. amylovora* includes more than 180 species from 39 genera in the family of *Rosaceae* (van der Zwet et al. 2012) and all four subfamilies (old classification system): *Maloideae* (syn. *Pomoideae*), *Rosoideae*, *Amygdaloideae* (syn. *Prunoideae*), and *Spiraeoideae*. Of the 39 genera, eight are fruit crops: *Malus*, *Pyrus*, *Cydonia, Eriobotrya, Fragaria, Mespilus, Prunus,* and *Rubus* (van der Zwet 1995). The remaining genera are mostly ornamental plants and trees, including those most susceptible and economically important hosts: *Cotoneaster*, *Crataegus*, *Pyracantha*, and *Sorbus*. Fire blight has been described in raspberry (*Rubus idaeus*) (Starr et al. 1951), in *Rosa rugosa* in Germany (Vanneste et al. 2002), and in chokeberry and strawberry in Bulgaria (Bobev et al. 2007). The complete list of host plants for fire blight could be found in van der Zwet et al. (2012) and van der Zwet and Keil (1979).

Taxonomic Position and Phylogenetic Relationships

Taxonomically, the genus *Erwinia* belongs to the *γ-Proteobacteria* (Order *Enterobacteriales*, Family *Enterobacteriaceae*) (Starr and Chatterjee 1972). Based on sequence analysis of the 16S rRNA genes, *E. amylovora* and related species associated with pome fruit trees are closely related to other enterobacteria such as *Escherichia coli*, *Salmonella enterica*, and *Yersinia pestis*. A phylogenetic tree based on 16S rRNA gene sequences showed that all strains of *E. amylovora* formed a separate clade within the genus *Erwinia* (Matsuuza et al. 2012). Housekeeping genes, such as *atpD*, *gyrB*, *infB,* and *rpoB* are commonly used for the phylogenetic analysis of *Enterobacteriaceae* (Sarkar and Guttman 2004). A phylogenetic tree reflecting their evolutionary relationship from concatenated sequences of four housekeeping proteins (AcnB, GltA, GyrB, and RpoD) is presented in Fig. 1.1 (Zhao and Qi 2011). The topologies of the phylogenetic trees based on 16S rRNA gene and housekeeping genes are very similar. Furthermore, the topology is identical to a phylogenetic tree generated using core genomes of the

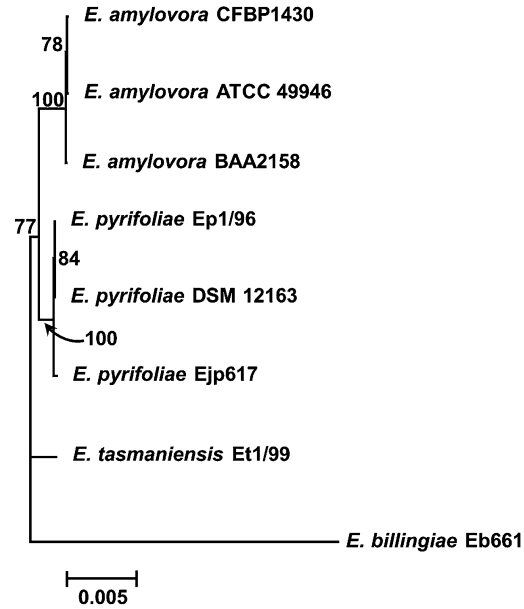

Fig. 1.1 Evolutionary relationship of *Erwinia amylovora* and related *Erwinia* species from Zhao and Qi (2011). The phylogenetic tree was constructed from concatenated sequences (2222 aa) of four housekeeping proteins (AcnB, GltA, GyrB and RpoD) using the neighbor-joining (NJ) method. Bootstrap scores greater than 60 are given at each node. The scale bar represents 0.005 amino acids substitutions per site. *E. billingiae* strain Eb661 was used as an outgroup

sequenced *E. amylovora* strains (Mann et al. 2013). In addition, *E. uzenensis* and *E. piriflorinigrans* cluster more closely to *E. pyrifoliae* strains than to *E. amylovora* strains (Matsuuza et al. 2012). Within the species of *E. amylovora*, phylogenetic analyses of the core genomes of *E. amylovora* strains indicate that the *Spiraeoideae*-infecting strains exhibit much less diversity than the *Rubus*-infecting strains (Mann et al. 2013; Rezzonico et al. 2012). This grouping is consistent with previous studies using rep-PCR, carbon utilization, and phylogeny based on *rpoB* (Mann et al. 2013; Rezzonico et al. 2012).

1.1.2.2 Virulence Factors and Regulation

E. amylovora is a highly virulent necrogenic vascular pathogen. Extensive genetic studies over the past several decades have explored the

molecular mechanism of *E. amylovora* pathogenesis (Khan et al. 2012; Zhao and Qi 2011). One of the highlights is the identification of an essential virulence system, namely the hypersensitive response and pathogenicity (*hrp*)-type III secretion system (T3SS), which is clustered on an ostensible pathogenicity island 1(PAI1) (Oh and Beer 2005; Oh et al. 2005). The current understanding is that two virulence factors are strictly required for *E. amylovora* to cause disease, i.e., the T3SS and its effectors and the acidic exopolysaccharide (EPS) amylovoran (Zhao and Qi 2011; Zhao et al. 2009a, b). Other virulence determinants found in *E. amylovora* that contribute to virulence and plant colonization include levan, protease, the iron-scavenging siderophore desferrioxamine, and genes involved in sorbitol and sucrose metabolism (Aldridge et al. 1997b; Bogs and Geider 2000; Dellagi et al. 1999; Du and Geider 2002; Smits and Duffy 2011).

Type III Secretion and Effectors

The discovery of the *hrp* gene cluster, which encodes a T3SS common to Gram-negative bacterial pathogens, is a breakthrough in understanding the molecular mechanism of bacterial pathogenesis and is a landmark in modern molecular plant pathology (Lindgren et al. 1986). The most noticeable phenotype of *hrp* mutants is their lost pathogenicity on susceptible host plants, and inability to elicit the hypersensitive response (HR) on resistant cultivars of host plants and non-host plants (Barny et al. 1990; Lindgren et al. 1986; Steinberger and Beer 1988). In *E. amylovora*, the PAI1 *hrp* gene cluster is located on an approximately 60-kb genomic region (Bogdanove et al. 1996; Kim and Beer 2001). The *hrp* gene cluster can be divided into three subregions: the *hrp/hrc* region, the Hrc effector and elicitors (HEE) region, and the Hrp-associated enzymes (HAE) region (Oh and Beer 2005; Oh et al. 2005). The *hrp/hrc* region contains 25 genes, including four regulatory genes (*hrpL*, *hrpS* and *hrpXY*) and genes encoding structural components of T3SS. Among them, nine *hrc* genes constitute the core structural components of the T3SS, including HrpA, a pilin protein (Jin et al. 2001; Kim et al. 1997). The seven gene HEE regions encode two harpin proteins (HrpN and HrpW) and several effectors and chaperones, including disease-specific protein DspE/A and its chaperone DspF/B (Bogdanove et al. 1998a, b; Gaudriault et al. 1997). The HAE region contains five genes, including three *hrp*-associated systemic virulence genes (*hsvABC*) and *hrpK*, which encodes a putative translocator (Oh et al. 2005). Therefore, T3SS proteins can be categorized into three groups, regulatory proteins (e.g., HrpL sigma factor), secretion apparatus proteins (e.g., HrpA pilin), and extracellularly secreted effector proteins (e.g., HrpN harpin and DspE).

E. amylovora has been developed as a model pathogen for studying plant–microbe interactions because the first cell-free elicitor (HrpN, harpin) was identified in 1992 (Wei et al. 1992). The T3SS of *E. amylovora* secretes at least 15 virulence-associated proteins, including HrpA, HrpN, HrpW, HrpJ, HrpK, HopAK1 (Eop2), DspE, HopC1, HopX1 (Eop3); AvrRpt2 (Eop4), and Eop1(EopB, OrfB) (Bogdanove et al. 1998a, b; Gaudriault et al. 1997, 1998; Nissinen et al. 2007; Zhao et al. 2005, 2006). Among them, harpins (HrpN and HrpW) are glycine-rich hydrophilic proteins and elicit an HR when infiltrated into intercellular spaces of some plants (Wei et al. 1992). HrpJ, HrpK, and HopAK1 (Eop2, Eam_2780) are putative translocators for delivery of effector proteins or Hop proteins (*H*rp *o*uter *p*roteins). Together with HrpN, they may form the translocon apparatus.

Many studies including genome sequencing have reached the conclusion that only five effector or *hop* genes [*eop1*, *eop3* (*hopX1*), *eop4* (*avrRpt2*), *dspA/E*, and *hopPtoC* (*hopC1*)] and several chaperones (*dspF/B, esc1* (*orfC*), and *esc3*) exist in the genome of *E. amylovora* (Nissinen et al. 2007; Zhao et al. 2005, 2006). Among them, *eop3* (*Eam_2190*), *eop4* (*Eam_0423*), *eop2*, and *hopC1* (*Eam_2679*) are located outside of the T3SS PAI1 in the genome. *DspA/E*, *avrRpt2*, and *hopC1* have been demonstrated to be induced in immature pear fruit,

indicating that they may play a major role in virulence (Zhao et al. 2005, 2006). Eop1 and Eop3 are AvrRxv/YopJ and HopX family proteins, respectively. Eop3 (HopX1) may function as an avirulence gene in apple (Bocsanczy et al. 2012), whereas Eop1 may function as a factor in limiting host range (Asselin et al. 2011).

Exopolysaccharide Amylovoran

E. amylovora produces two types of EPSs, the heteroexopolysaccharide amylovoran and the homoexopolysaccharide levan (Bellemann and Geider 1992; Geier and Geider 1993). Levan is a homopolymer of fructose residues and is synthesized by the enzyme levansucrase. Mutants deficient in levan production are reduced in their virulence (Geier and Geider 1993; Gross et al. 1992). However, strains deficient in levan production have also been found in nature and are virulent (Bereswill et al. 1997). Amylovoran may have multiple functions because mutants deficient in amylovoran biosynthesis are nonpathogenic (Bellemann et al. 1994; Bernhard et al. 1993). Furthermore, the ability of individual *E. amylovora* strains to produce amylovoran is positively correlated with the degree of virulence (Wang et al. 2010a). More recent studies also revealed the indispensable functions of amylovoran in bacterial biofilm formation in plant xylem and for survival under stress conditions (Koczan et al. 2009; Ordax et al. 2010).

Amylovoran was first isolated from bacterial ooze (Bennet and Billing 1980). It is a complex, high molecular weight (50–150 mDa), acidic capsular EPS, consisting of galactose, glucose, and pyruvate residues (Nimtz et al. 1996). Amylovoran biosynthetic genes are located within a 12-gene amylovoran biosynthetic (*ams*) operon, from *amsA* to *amsL*, with *amsG* as the first gene in the operon (Aldridge et al. 1997a; Bernhard et al. 1993; Bugert and Geider 1995). Another two genes, *galF* and *galE*, which are located on the right adjacent to the *ams* cluster, are involved in amylovoran precursor formation. Studies have proposed that products of the *amsGBCDEJK* genes play roles in glycosyl transfer for the repeating unit (Langlotz et al. 2011). Until recently, new evidence has shown that ams-II (*amsG2*) and ams-III (*amsO-amsL2*) may also play a role in side-chain modification (Wang et al. 2012c).

Regulatory Systems

In prokaryotes, gene expression is regulated primarily at the level of transcription initiation. In *E. amylovora*, transcription of the *hrp*-T3SS genes is activated by the master regulator HrpL, a member of the ECF subfamily of sigma factors (Wei and Beer 1995). HrpL binds to a consensus sequence known as the *hrp* box (GGAACC-N_{16}-CCACNNA) in *hrp* gene promoters. Most T3SS and effector genes are subject to direct HrpL regulation (McNally et al. 2012, Nissinen et al. 2007). A hidden Markov model has identified about 30 *hrp* promoters in the genome of *E. amylovora* strain Ea273, which contain the *hrp* box recognized by HrpL (McNally et al. 2012; Bocsanczy et al. 2012).

The expression of *hrpL* is believed to be activated by both HrpS and a two-component regulatory system HrpX (sensor) and HrpY (response regulator) (Wei et al. 2000). Further domain structure analysis indicated that HrpX contains two PAS domains (initially found in *P*ER, *A*RNT, and *S*IM proteins) within the N-terminal sensor region, suggesting that HrpX is a soluble and cytoplasmic protein that may sense intracellular signals. Recent studies also found that *hrpXY* mutants remain virulent (Zhao et al. 2009b), which is different from a previous report that analyzed Tn*5*-insertional *hrpXY* mutants (Wei et al. 2000). An early report suggested that HrpS, a member of the NtrC family of σ^{54} enhancer-binding proteins, only partially controls *hrpL* expression (Wei and Beer 1995). However, recent findings indicated that HrpS, YhbH, and alternative sigma factor RpoN (σ^{54}) are absolutely required for *hrpL* expression (Ancona et al. 2014; Zhao et al. 2009b) and the *hrpL* gene also contains a σ^{54} consensus sequence in its promoter region. Furthermore, global regulators such as GrrSA and EnvZ/OmpR two-component systems may also be involved in regulating *hrp* gene expression (Li et al. 2014).

Several key regulators of amylovoran biosynthesis have been characterized earlier, including RcsA (Bernhard et al. 1990; Coleman et al. 1990) and RcsB (Bereswill and Geider 1997) as well as interactions between RcsA and RcsB (Kelm et al. 1997; Wehland and Bernhard 2000; Wehland et al. 1999). Recently, the RcsCDB system has been demonstrated to be essential for virulence (Wang et al. 2009, 2011b; Zhao et al. 2009b). The Rcs phosphorelay system is a unique enterobacterial-specific two-component system, and phosphorylated RcsB could form RcsB-RcsB homodimers or interact with RcsA to form RcsAB heterodimers, which then bind to an "RcsAB box" to regulate gene expression, including the promoter of the *ams* operon involved in amylovoran biosynthesis (Pristovsek et al. 2003; Wehland et al. 1999, Wehland and Bernhard 2000). A Hidden Markov model identified about 60 genes in the genome of *E. amylovora* strain Ea273 (ATCC49946), which contains the RcsAB box, and half of these genes were directly regulated by RcsBC (Wang et al. 2012a). Genome-wide screening of two-component-system mutants identified four groups of mutants that exhibited varying levels of amylovoran production in vitro, indicating that two-component systems in *E. amylovora* play a major role in regulating amylovoran production, and may form a regulatory network to govern the production of amylovoran (Wang et al. 2011c; Zhao et al. 2009b).

Additional novel and global regulatory genes for amylovoran biosynthesis have also been identified through genetic screening. These included Lon protease, global regulator H-NS, RcsF, DjlA, and AmyR (YbjN) (Eastgate et al. 1995; Hildebrand et al. 2006; Wang et al. 2011c, 2012b). Both RcsF and DjlA are activators of RcsC, whereas H-NS binds to the promoter of *rcsA* and suppresses *rcsA* gene expression. Furthermore, the RcsA protein is subject to Lon-dependent degradation, which is a heat-shock protein. In addition, AmyR, an *amy*lovoran *r*epressor, is an enterobacterial-specific orphan protein and was recently characterized as a novel negative regulator of EPS production in both *E. coli* and *E. amylovora* (Wang et al. 2011a, 2012b).

Based on current knowledge, a simple model for *E. amylovora* virulence gene expression centered on T3SS and amylovoran production is presented in Fig. 1.2. Upon initiating plant infection, *E. amylovora* senses the unknown host/environmental signals. This process activates a sigma factor cascade that regulates T3SS gene expression and at the same time, several two-component signal transduction systems that regulate both T3SS and amylovoran production (Li et al. 2014; Zhao et al. 2009b). In the sigma factor cascade, σ^{54} enhancer-binding protein HrpS forms a hexamer and binds to the upstream DNA activator sequences (UAS) of the *hrpL* promoter. Meanwhile, RpoN and a core RNA polymerase (RNAP) forms a σ^{54}-RNAP complex that binds to the σ^{54} promoter of the *hrpL* gene, but remains transcriptionally silent. With the assistance of integration host factor (IHFα/β, not shown), HrpS contacts the σ^{54}-RNAP-promoter complex via the consensus GAFTGA motif and by DNA looping, with the energy provided by ATP hydrolysis of HrpS AAA+ -domain. This triggers the opening of the σ^{54}-RNAP-promoter complex and DNA melting. This process also requires a ribosome-associated protein YhbH with an unknown mechanism (Ancona et al. 2014). HrpL/RNAP complex then recognizes the "*hrp* box" at the promoter regions of HrpL-dependent operons or genes and regulates *hrp* gene expression (Wei et al. 2010). On the other hand, phosphorylated RcsB dimer or RcsAB heterodimer binds directly to the *amsG* promoter in the *ams* operon and regulates amylovoran production (Wang et al. 2009). The GrrSA two-component system specifically regulates small regulatory RNA *rsmB*, which stabilizes RNA-binding protein RsmA (Ancona and Zhao 2013; Li et al. 2014). However, the molecular mechanism as how GrrSA along with EnvZ/OmpR system negatively regulate T3SS and amylovoran production is unknown. Furthermore, the identity of the signals and how the bacterium senses the signals remain unsolved mysteries.

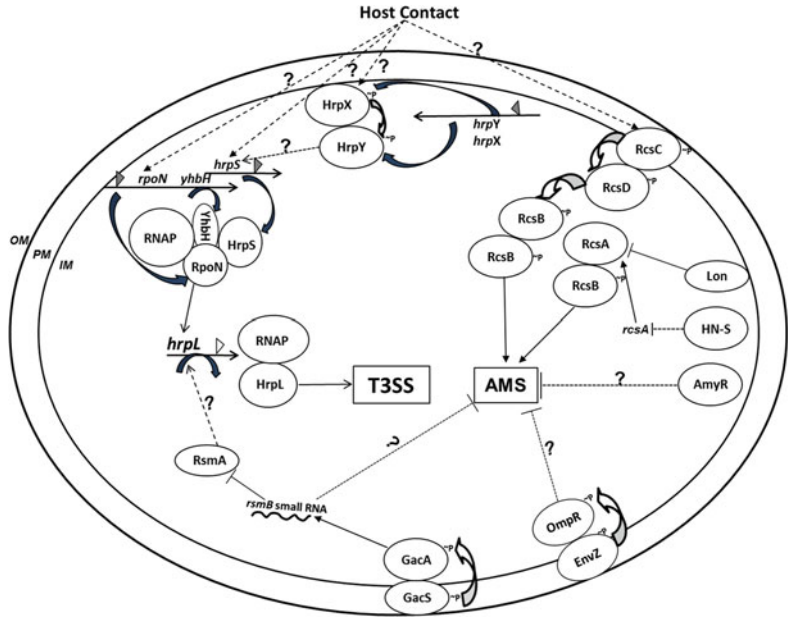

Fig. 1.2 A model for virulence gene regulation in *Erwinia amylovora*. AmyR orphan protein, ortholog of *E. coli* YbjN; AMS amylovoran biosynthesis, HN-S nucleoid-associated protein; HrpL; an ECF sigma factor, HrpS a σ^{54}-dependent enhancer-binding protein, HrpX/HrpY, OmpR/EnvZ, GacS/GacA, RcsABCD two-component regulatory systems, Lon protease, RpoN a sigma 54 factor, RNAP RNA polymerase, RsmA RNA-binding protein, *rsmB* a small regulatory RNA, T3SS type III secretion system, YhbH σ^{54} modulation protein (ribosome-associated protein). *OM* outer membrane, *PM* plasma membrane, *IM* inner membrane, *P* phosphorylation, *filled triangle*, σ^{70} promoter, *open triangle* σ^{54} promoter. Positive regulation is indicated by an *arrow* and down-regulation by T sign; question mark and *dash line* unknown mechanism; *Thick arrow* lines show gene or operons, *oval* and *circles* indicate proteins

1.1.2.3 Tools for Molecular Genetics Studies

Traditionally, genetic studies for plant pathogenic bacteria utilize mobile genetic elements such as Tn*5*, Mu, and Tn*10* to generate mutant libraries (Vanneste et al. 1990). However, because of the randomness, we may not obtain a mutant for a specific gene of interest. Alternatively, standard homologous DNA recombination systems using suicide vectors could also be used to generate mutants (Zhao et al. 2005); however, standard recombination techniques require extensive and time-consuming in vitro cloning steps (Zhao et al. 2005). Recently, we have successfully adopted a PCR-based one-step inactivation of chromosomal genes and also referred to as the Red-cloning technique, to generate mutants in *E. amylovora* (Datsenko and Wanner 2000; Zhao et al. 2006, 2009a, b). We found that this technique is not only easy and fast (capable of mutant generation within a week without any cloning steps), but can also generate deletion mutants for a single gene, an operon or a genomic island (Zhao et al. 2009a, b). These mutants provide basic tools to study gene function and signal transduction in *E. amylovora*.

1.2 Genomics

1.2.1 Genome Sequencing of *E. amylovora* and Related *Erwinia* Species Associated with Pome Fruit Trees

In recent years, complete genomes for five species from the genus *Erwinia*, including two *E. amylovora* strains, three *E. pyrifoliae* strains, one *E. tasmaniensis* strain, one *E. billingiae* strain, and one strain of *E. piriflorinigrans* from

Table 1.1 Overview of complete genome sequencing of *Erwinia amylovora* and related *Erwinia* species associated with pome fruit trees

Strains	Origin	Size[a] (Mb)	G + C content	Total proteins	Plasmid #s	Host	Accession #s
E. amylovora CFBP1430	France 1972	3.81	53.6	3,706	1	*Crataegus*	FN434113-114
E. amylovora ATCC 49946	New York 1973	3.81	53.6	3,565 (3,712)[b]	2	Apple	FN666575-577
E. pyrifoliae Ep1/96	South Korea 1996	4.03	53.4	3,697	4	Asian pear	FP236842 FP928999 FP236827-29
E. pyrifoliae DSM 12163 (Ep16/96)	South Korea 1996	4.03	53.4	4,038	4	Asian pear	FN392235-39
E. pyrifoliae Ejp617	Japan	3.91	53.6	3,672	5	Asian pear	CP002124-29
E. tasmaniensis Et1/99	Australia 1999	3.88	53.7	3,622	5	Apple flower	CU468128, 30-33, 35
E. billingiae Eb661	UK	5.10	55.2	4,917	2	Tree	FP236826, 30, 43
E. piriflorinigrans CFBP 5888	Spain 2000	3.97	49.8	3,857	1	Pear	CAHS0100001-25; HE792893

Data adapted from Kube et al. (2008, 2010), Mann et al. (2012, 2013), Park et al. (2011), Powney et al. (2011), Sebaihia et al. (2010), Smits et al. (2010a, b, 2013)

[a] Chromosomal size; plasmids are not included

[b] See Mann et al. 2013

Spain, have been sequenced (Table 1.1). In addition, draft genome sequences have been obtained for an additional 10 *E. amylovora* strains, including three strains isolated from *Rubus* spp. (Table 1.2). These genome sequences provide abundant scientific information about the genetic composition of these species and facilitate functional and comparative genomic studies to determine how genomes of closely related *Erwinia* species have evolved (Mann et al. 2013; Zhao and Qi 2011).

1.2.1.1 General Features and Distinguishing Characteristics

The first genome sequence of *E. amylovora* strain Ea273 (ATCC49946), funded by the United States Department of Agriculture, was started in the early 2000s and published in 2010 (Sebaihia et al. 2010). At the same time, the complete genome sequence of *E. amylovora* strain CFBP1430 was also published (Smits et al. 2010b). Draft genome sequences were later obtained and published for 10 additional *E. amylovora* strains, including three isolated from host-specific *Rubus* spp. (Mann et al. 2013; Powney et al. 2011). These sequenced strains represent various host plants, geographical origins, and temporal distributions to exemplify distinct bacterial populations of *E. amylovora* (Tables 1.1 and 1.2). Other general information, including genome size, GC content, total proteins, and plasmid content, is listed in Tables 1.1 and 1.2.

The genomes of *E. amylovora* and related *Erwinia* species associated with pome fruit trees range from 3.8 to 5.1 Mbp, with *E. amylovora* containing the smallest genome compared to other pathogenic enterobacteria sequenced so far (up to 5.5 Mbp) (Toth et al. 2006). The genome of *E. billingiae* is larger than that of pathogenic *Erwinia* species and the non-pathogenic *E. tasmaniensis*. A comparison of genomes of *E. amylovora* strains CFBP1430 and ATCC49946 shows that the two genomes share more than 99.9 % identity at the nucleotide level, indicating that *E. amylovora* is a relatively homogeneous

Table 1.2 Overview of draft genome sequencing of *Erwinia amylovora*

Strains	Origin	Size[a] (Mb)	G + C content	Total proteins	Plasmid #s	Host	Accession #s
E. amylovora CFBP1232	UK 1959	3.77	53.6	3,780	1	Pear	CAPB01000001-42; HF560650
E. amylovora 01SFR-BO	Italy 1991	3.77	53.6	3,744	1	*Sorbus*	CAPA01000001-11,HF560647
E. amylovora ACW 56400	Switzerland 2007	3.77	53.6	3,758	2	Pear	AFHN01000001-22, CP002951, AFHN01000023
E. amylovora UPN527	Spain 1997	3.77	53.6	3,746	0	Apple	CAPC01000001-18
E. amylovora CFBP2585 (Ea495)	Ireland 1986	3.77	53.6	3,734	2	*Sorbus*	CAOZ01000001-12, HF560645-46
E. amylovora EA266 (E4001A)	Canada	3.76	53.6	3,804	1	Apple	CAOY01000001-38, HF560644
E. amylovora Ea356 (Ea1/79)	Germany 1979	3.76	53.6	3,744	1	Cotoneaster	CAOX01000001-14,HF560643
E. amylovora BAA2158 (IL5)	Illinois 1972	3.81	53.6	3,827	3	*Rubus*	FR719181 to FR719212
E. amylovora Ea644	MA, USA 2003	3.80	53.3	3,937	1	*Rubus*	CAPD01000001-40, HF560648
E. amylovora MR-1(Ea574)	Michigan	3.79	53.4	4,042	1	*Rubus*	CAPE01000001-29, HF560649

Data adapted from Kube et al. (2008, 2010), Mann et al. (2012, 2013), Park et al. (2011), Powney et al. (2011), Sebaihia et al. (2010), Smits et al. (2010a, b)

[a] Chromosome size; plasmids are not included

species (Smits et al. 2010b; Zhao and Qi 2011). The chromosomes of another seven *Spiraeoideae*-infecting strains are also highly homogeneous. Greater genetic diversity is observed between *Spiraeoideae*- and *Rubus*-infecting strains than among individual *Rubus*-infecting strains; this suggests that further delineation of *E. amylovora* species with different host specificity may be warranted (Mann et al. 2013; Zhao and Qi 2011). Similarly, the genomes of the two *E. pyrifoliae* strains from Korea (Ep1/96 and DSM12163) are almost identical (Kube et al. 2010; Park et al. 2011; Smits et al. 2010a; Thapa et al. 2013).

One interesting characteristic of the *E. amylovora* genome is that five of seven copies of the rRNA operon have a 99-bp insertion within helix 45 of the 23S rRNA gene, which represents an intervening sequence (IVS) as described for some species in the *Enterobacteriaceae* family. Although IVSs are absent in *E. coli*, *Pectobacterium*, and *Dickeya* spp. (McGhee et al. 2002; Pronk and Sanderson 2001), IVS sequences are present in all seven copies of the 23S rRNA gene of *E. pyrifoliae* DSM 12163T, but are absent in copies of the 23S rRNA gene of *E. tasmaniensis* Et1/99 (Kube et al. 2008; Smits et al. 2010a). The rRNA fragmentation pattern from the excision of IVSs is shown in Fig. 1.3 for *E. amylovora*.

1.2.1.2 Genome Structure and Rearrangements

Whole-genome structure comparisons of *E. amylovora*, *E. pyrifoliae*, *E. piriflorinigrans*, *E. billingiae* and *E. tasmaniensis* with complete genome sequences reveal many large-scale chromosomal re-organizations and inversions,

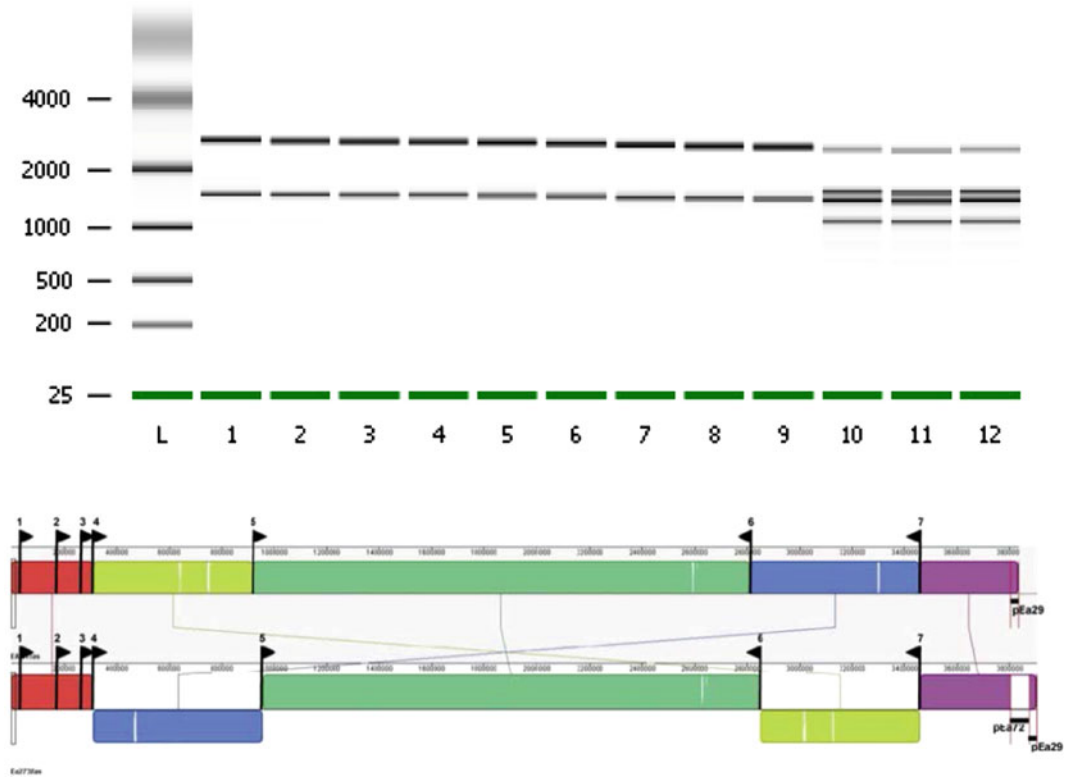

Fig. 1.3 *Top* Comparison of *E. coli* and *E. amylovora* RNAs showing the 5:2 ratio of 23S rRNA fragmentation pattern using the Agilent 2100 Bioanalyzer (Agilent Technologies, Palo Alto, CA, USA). *Lanes 1 to 9 E. coli*, *Lanes 10 to 12 E. amylovora*. *23S rRNA* 2.9 kb, *16S rRNA* 1.5 kb, *two extra bands* 1.7 and 1.2 kb, respectively. *Below* Structure comparison of the genomes of *E. amylovora* CFBP 1430 and ATCC 49946, indicating genome re-arrangements (reprinted from Smits et al. 2010b). *Vertical lines* indicate ribosomal (r)RNA operons. Numbers indicate the rRNA operon numerator, as used in the alignments. Arrows indicate the transcriptional direction of the rRNA operons

probably resulting from homologous recombination events (Kube et al. 2010; Smits et al. 2010a, b, 2013). However, there is only one single large-scale re-arrangement of the genomes for *E. amylovora* strains CFBP 1430 and ATCC49946, which represents two independent recombination events, and may have occurred within copies of the ribosomal (r)RNA operon (Fig. 1.3). Analyses of the genome structures of the remaining seven *Spiraeoidea*-infecting strains reveal that genome structures for five strains (all from Europe) are identical to that of CFBP1430, and the other two strains (CFBP2585 from Ireland and Ea266 from Canada) are identical to that of ATCC49946 (Mann et al. 2013). In silico analyses revealed the exact fragment sizes, which classify *E. amylovora* CFBP 1430 as a Pt-3 genotype pattern and ATCC49946 as a Pt-4 genotype using pulsed-field gel electrophoresis (PFGE) analysis of *Xba*I- digested genomic DNA (Smits et al. 2010b; Zhang and Geider 1997). In addition, the ITS regions between the 16S and 23S rRNA genes of three operons of *E. amylovora* CFBP 1430 contain the transfer (t)RNA-Ala and tRNA-Ile genes, and four operons contain the tRNA-Glu genes. In contrast, the ITS regions between the 16S and 23S rRNA genes of four operons of the *E. amylovora* ATCC49946 genome are composed of tRNA-Ala/tRNA-Ile-genes and three operons of tRNA-Glu-genes, further suggesting that the recombination event

Table 1.3 Overview of plasmids in *Erwinia amylovora* and related *Erwinia* species associated with pome fruit trees

Plasmid	Species	Strain	Size (kb)	Host	Origin	Accession
pEA29	*E. amylovora*	All[a]	28.2	Various	Various	FN666576
pEA72	*E. amylovora*	ATCC49946	71.5	Apple	NY	FN666577
pEI70	*E. amylovora*	ACW56400	65.8	Pear	Switzerland	CP002951
pEL60	*E. amylovora*	Leb66	60	Apple	Lebanon	NC-005246
pEA34	*E. amylovora*	CA11	34	Apple	MI	NA
pEA30/ pEU30	*E. amylovora*	CFBP2585 UTRJ2	29.6	Sorbus Apple	Ireland Utah	HF560646 NC_005247
pEA8.7	*E. amylovora*	CA3R	8.7	Apple	CA	NA
pEAR5.2	*E. amylovora*	BAA-2158	5.25	*Rubus*	IL	FR719211
pEAR4.3	*E. amylovora*	BAA-2158	4.37	*Rubus*	IL	FR719210
pEA2.8	*E. amylovora*	IL-5	2.8	*Rubus*	IL	AY123047
pEA1.7	*E. amylovora*	IH3-1	1.7	*Crataegus*	LA	AY123046
pEP36	*E. pyrifoliae*	Ep1/96, 16/ 96	35.9	pear	Korea	FN392238
pEP5 pEP3 pEP2.6	*E. pyrifoliae*	Ep1/96, 16/ 96, Ejp617	4.96, 3.0, 2.6	pear	Korea Japan	FN392239 FN392237 FN392236
pJE01	*E. pyrifoliae*	Ejp617	30.9	pear	Japan	CP002125
pJE03	*E. pyrifoliae*	Ejp617	6.4	pear	Japan	CP002127
pET9, 35, 45, 46, 49	*E. tasmaniensis*	Et1/99	9.3, 35.4, 44.7, 46.1, 48.8	Apple flower	Australia	CU468128 CU468130 CU468131 CU468132 CU468133
pEb102, pEB170	*E. billingiae*	Eb661	102, 170	tree	UK	FP236826 FP236830
pEPIR37	*E. piriflorinigrans*	CFBP5887	37	pear	Spain	HE792893

Data adapted from Foster et al. (2004), Kube et al. (2008, 2010), Llop et al. (2011), Mann et al. (2012, 2013), Park et al. (2011), Powney et al. (2011), Sebaihia et al. (2010), Smits et al. (2010a, b, 2013)

[a] Except strain UPN527

may occur within the rRNA operon (Smits et al. 2010b). It is remarkable to discover how the *E. amylovora* genome structure has changed since the spread of the bacterium from North America to Europe about 60 years ago.

1.2.1.3 Plasmids

Comparative genomic analyses of several species of plant pathogenic bacteria revealed that the majority of strain-specific genes are plasmid-borne, indicating that acquisition and maintenance of plasmids may represent a major mechanism for bacteria to change their genetic composition and acquire new virulence factors (Zhao and Qi 2011). Plasmids also may contribute to genetic diversity of the *Spiraeoideae*-infecting strains of *E. amylovora* (Mann et al. 2013; Zhao and Qi 2011). A total of 11 plasmids have been reported in *E. amylovora* (Table 1.3) of which six are present among 12 sequenced *E. amylovora* genomes (Foster et al. 2004; Llop et al. 2011; Mann et al. 2013). The nearly ubiquitous plasmid pEA29 is present in all sequenced strains except UPN527 (Llop et al. 2006; Mann et al. 2013; McGhee and Jones 2000). Plasmid pEA29 contains genes encoding for thiamine biosynthesis, and loss of the *thiOSGF* genes results in thiamine auxotrophy (Llop et al. 2012; McGhee and Sundin 2008).

The major genetic distinction of *E. amylovora* strain ATCC49946 is the presence of plasmid pEA72, which is not present in the 11 *E. amylovora* genomes sequenced. Plasmid pEA72 contains a type IV secretion system that may be involved in conjugative transfer of the plasmid (Llop et al. 2012). In addition, plasmids pEA30, pEI70, and two small plasmids (pEAR5.2 and pEAR4.3) are present in strain CFBP 2585, ACW 56400, and ATCC BAA-2158, respectively (Table 1.3). In three sequenced *E. pyrifoliae* strains, three plasmids are common among them; and the fourth, pEp36, is not present in the Japanese strain, Ejp617 (Table 1.3). Instead, strain Ejp617 contains two extra plasmids (pJE01 and pJE03) (Kube et al. 2010; Park et al. 2011). Furthermore, *E. piriflorinigrans, E. billingiae,* and *E. tasmaniensis* strains contain one, two, and five plasmids, respectively (Table 1.3) (Kube et al. 2008, 2010; Smits et al. 2013).

1.2.2 Pan-Genome of *E. amylovora*

A pan-genome includes the full complement of genes in a species, which consists of the "core genome" containing genes present in all strains, a "dispensable or accessory genome" containing genes present in two or more strains, and finally "unique genes" specific to a single strain (Medini et al. 2005). It is predicted that the pan-genome of *E. amylovora* is still "open" based on two complete and 10 draft genome sequences (Smits et al. 2011; Mann et al. 2013).

1.2.2.1 Protein-Coding Genes, Core Genome, and Pan-Genome

The numbers of protein-coding genes in the genome of *E. amylovora* and related *Erwinia* species are listed in Tables 1.1 and 1.2. A comparison of genomes of *E. amylovora* strains CFBP1430 and ATCC 49946 shows that the two genomes share more than 99.9 % identity at the nucleotide level. However, based on initial annotation, the total predicted proteins in strain ATCC 49946 and CFBP1430 are 3565 and 3706, respectively (Sebaihia et al. 2010; Smits et al. 2010b). In a recent report, the number of predicted proteins for strain ATCC49946 is revised to 3712 (Mann et al. 2013). The numbers of predicted proteins in *Rubus*-infecting strains Ea644 and MR-1 are slightly higher as compared to those observed for *Spiraeoideae*-infecting strains of *E. amylovora* (Mann et al. 2013; Powney et al. 2011). Similarly, the genomes of the two *E. pyrifoliae* strains from Korea (Ep1/96 and DSM 12163 (Ep16/99) are almost identical; however, due to similar discrepancies in annotation, the total predicted proteins are 3697 and 4038 in Ep1/96 and DSM 12163, respectively (Kube et al. 2010; Smits et al. 2010a). On the other hand, *E. tasmaniensis* and *E. billingiae*, the two genetically most distant *Erwinia* species associated with pome fruit trees, contain 3,622 and 4,917 predicted proteins, respectively (Table 1.1) (Kube et al. 2008, 2010).

Using the subtractive hybridization-based mGenomeSubtractor program, which compares the reference genome against multiple bacterial genomes for in silico comparative genomic analyses, Zhao and Qi (2011) found that the number of conserved proteins with homology values greater than 0.81 is about 2,100. This indicates that the corresponding genes probably constitute the "core genome" among sequenced *E. amylovora* and related species (Zhao and Qi 2011). When *E. amylovora* strains CFBP1430 or ATCC 49946 are compared to ATCC BAA2158, a *Rubus*-infecting strain more closely related to the *Spiraeoideae*-infecting strains, more than 3,400 of the 3,500 conserved proteins (98 %) have homology values of 1. This indicates that the genomes of these *E. amylovora* strains are identical (Zhao and Qi 2011). When compared to all 12 sequenced *E. amylovora* genomes, including the genetically diverse *Rubus*-infecting strains (MR-1 and Ea644), about 3414 coding sequences are identified as core genes (Mann et al. 2013).

When *E. amylovora* strains are compared to sequenced *E. pyrifoliae, E. tasmaniensis,* and *E.*

billingiae strains, the numbers of conserved proteins are about 2,800, 2,600, and 2,200, respectively; and the number of proteins with homology values of 1 drops dramatically to 1,200 and below (Zhao and Qi 2011). This indicates that more diversification occurs for these pathogenic/saprophytic microorganisms and that *E. amylovora* and *E. pyrifoliae* may be evolutionally derived from two separate sources, one in North America and the other in Asia. Similar conclusions could also be drawn for strains of *E. pyrifoliae* from Japan and Korea, whereby about 85 % of conserved proteins (2,800 out of 3,300) are identical (Zhao and Qi 2011). In contrast, the number of strain-specific proteins varies among genomes (Zhao and Qi 2011). The majority of specific proteins among *Erwinia* species that have homology values of 0 are plasmid-borne, indicating that acquisition and maintenance of plasmids may represent a major mechanism for erwinias to change their genetic composition. The ever expanding pan-genome of *E. amylovora* is currently calculated to contain 5751 coding sequences based on 12 genome sequences (Mann et al. 2013; Smits et al. 2011).

1.2.2.2 Genomic Islands and Pathogenicity Islands

Genomic islands (GIs) are defined as clusters of genes in prokaryotic genomes, which may be acquired by horizontal gene transfer, and include prophages, integrated plasmids, integrative conjugative elements, integrons, and conjugative transposons (Langille et al. 2010). Typically, GIs contain mobility-related genes and may also carry "cargo" genes that can be involved in virulence, resistance, and ecological fitness (Seth-Smith and Croucher 2009). Pathogenicity islands (PAIs) are generally regarded as large regions of chromosomal or plasmid DNA containing multiple virulence genes, which are flanked by repeated sequences and are characteristically distinct in GC content from the rest of the genome (Hacker et al. 1997). Among the 12 sequenced *E. amylovora* strains, 12 GIs and three PAIs have been identified, and the former represent the majority of the genetic variation observed within the chromosomal component of the pan-genome (Mann et al. 2013; Zhao and Qi 2011). The majority of coding sequences within the GIs of the *E. amylovora* pan-genome consists of hypothetical and mobility-related genes, including genes involved in replication, transfer, and integration of mobile elements (Mann et al. 2013). In contrast, three T3SS genes and some effectors are located within the three PAIs of *E. amylovora* (Zhao et al. 2009a).

Among the GIs in the genome of *E. amylovora* strains, one 34.5-kb GI is present in the *Rubus*-infecting strains Ea644 and MR1, but a different GI of 23.4-kb occupies the same locus in the *Spiraeoideae*-infecting strains and ATCC BAA-2158 (Mann et al. 2013). Strains Ea644 and MR1 contain a type 1 restriction modification system, which protects the host DNA by adding methyl groups to recognition sites of expressed restriction enzymes. The *Spiraeoideae*-infecting strains encode a DNA degradation (Dnd) host-specific modification system, which incorporates sulfur into the DNA backbone to prevent restriction recognition (Mann et al. 2013). In addition, all three sequenced *Rubus*-infecting strains contain one 20-kb GI, which is absent in all nine sequenced *Spiraeoideae*-infecting strains. This locus encodes three polyketide synthases (PKS), a non-ribosomal peptide synthase (NRPS), and a putative transporter, which may represent a novel NRPS/PKS system for metabolite production (Mann et al. 2013). Remnants of this locus are found in CRISPR region 1 (CRR1) of the *Spiraeoideae*-infecting strains, suggesting that this GI in *Rubus*-infecting strains may be ancestral to CRR1 of the *Spiraeoideae*-infecting strains (see below) (Rezzonico et al. 2011).

There are large differences in the island transfer (IT) region between each of the *Rubus* strains and the *Spiraeoideae*-infecting strains (Mann et al. 2012, 2013). The IT region, an integrative conjugative element (ICE), is next to the *hrp* PAI1 HEE region (Oh and Beer 2005). The IT regions in *Spiraeoideae*-infecting strains are highly conserved, but the IT regions of the *Rubus*-infecting strains, *E. piriflorinigrans,* and

Table 1.4 Virulence-associated traits and their distribution in *E. amylovora* and related *Erwinia* species associated with pome fruit trees

Strains	*E. amylovora*			*E. pyrifoliae*			*E. tasmaniensis*	*E. billingiae*
Traits	CFBP 1430	ATCC 49946	BAA 2158	DSM 12163	EP 1/96	Ejp 617	Et1/99	Eb661
T3SS PAI1	+	+	+	+	+	+	+ (P)	–
T3SS PAI2	+	+	+	+	+	+	+	–
T3SS PAI3	+	+	+	–	–	–	+ (P)	–
Flagella 1 (S)	+	+	+	+	+	+	+	+
Flagella 2 (C)	+	+	+	+	+	+	–	–
Amylovoran biosynthesis[a]	+	+	+	+	+	+	+ (E)	+ (E)
Levansucrase (*lsc*)	+	+	+	–	–	–	+	+
Regulators of levansucrase (*rlsABC*)	+	+	+	+	+	+	+	–
Sorbitol metabolism (*srlAEBDMR*)	+	+	+	+	+	+	–	+
Protease A (*prtADEF*)	+	+	+	–	–	–	–	–
Siderophore biosynthesis (*dfoA*)	+	+	+	+	+	+	+	–
hopC1(hopPtoC)	+	+	+	–	–	–	–	–
hopAK1 (eop2)	+	+	+	–	–	–	–	–
hopX1 (eop3)[b]	+	+	+[b]	+	+	+	–	–
avrRpt2(eop4)[b]	+	+	+[b]	–	–	–	–	–
eop1/esc1[b]	+	+	+[b]	+	+	+	+	–
HrpK, HsvABC	+	+	+[b]	+	+	+	–	–

P partial, *S* separated, *C* clustered, *E* In Et1/99 and Eb661, the *amsE* gene is missing, but additional genes are present

[a] Some genes such as *amsCDE* are very diverse among different species of *Erwinia*

[b] Sequence diversification found in different species and even between *Spiraeoideae*- and *Rubus*-infecting strains. HrpK is truncated in strain ATCC BAA-2158. A single-base deletion at position 165 of AvrRpt2 found in strains Ea644 and MR-1; an amino acid substitution (C156S) of AvrRpt2 found in some *E. amylovora* strains isolated from Canada and US (Ea110, Ea400). Data adapted from: Kube et al. (2010), Mann et al. (2013), Vogt et al. (2013), Smits et al. (2010a, b, 2011), Zhao and Qi (2011)

E. pyrifoliae strains vary in gene content and length, showing a mosaic structure (Mann et al. 2012, 2013). It appears that the IT regions have undergone significant genome reduction in the *Spiraeoideae*-infecting strains, being more than 30 kb shorter in length than all sequenced *Rubus*-infecting strains (Mann et al. 2013). The IT region in *E. piriflorinigrans* is the largest (72 kb) (Smits et al. 2013).

Besides the well-characterized *hrp* T3SS PAI1, analysis of the complete genome sequence of *E. amylovora* strains revealed two extra non-flagellar T3SS PAIs (PAI2 and PAI3) and two flagellar T3SS systems (Flg-1 and Flg-2) (Table 1.4). PAI2 and PAI3 are 38 and 43 kb in length and contain 24 and 26 genes, respectively. Both PAI2 and PAI3 have a significantly lower %G+C content (38.4 and 43.4 % mol G+C, respectively) and similar gene organization with the known Inv/Spa-like T3SS PAIs of the insect endosymbiont *Sodalis glossinidius* and *ysa* of *Yersinia enterocolitica* (Zhao et al. 2009a). The non-flagellar T3SSs can be divided into at least five groups based on phylogenetic

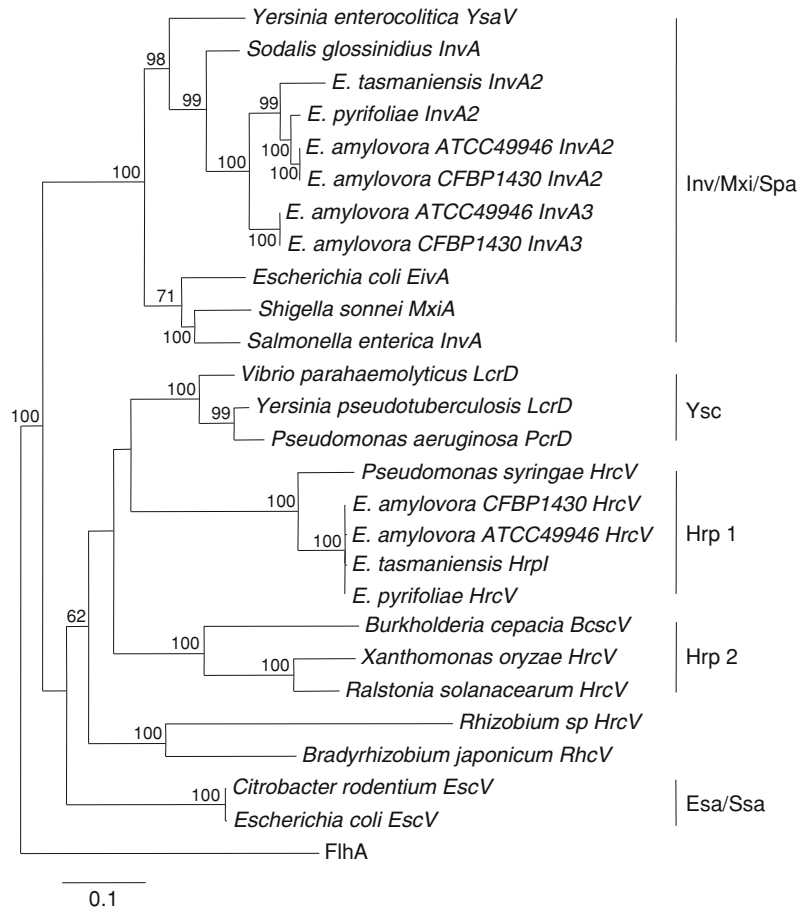

Fig. 1.4 Phylogenetic tree based on aligned amino acid sequences of HrcV/FlhA homologs from Zhao et al. (2011). The number at each node is bootstrap confidence values from 1,000 replicate maximum-likelihood trees. Similar copies of *invA* from *Erwinia* spp. (PAI2 and PAI3) are indicated by the number. The five T3SS groups are indicated on the *right*. FlhA was used as an outgroup

analysis (He et al. 2004). Phylogenetic trees based on the HrcV or InvA protein sequences revealed the PAI1 belongs to the Hrp1 group, whereas PAI2 and PAI3 belong to Inv/Mxi/Spa group (Fig. 1.4). The function of PAI2 and PAI3 is still unknown, but they are not directly involved in virulence to plants (Zhao et al. 2009a).

Analyses of genome sequences of closely related *Erwinia* species indicated that most PAIs are present in *E. piriflorinigrans, E. pyrifoliae,* and *E. tasmaniensis* (Table 1.4). However, PAI2 is not present in *E. billingiae* (Kube et al. 2008, 2010; Smits et al. 2010a, b, 2013), PAI3 is absent in the genome of *E. pyrifoliae* and *E. billingiae*, and only parts of PAI3 are present in *E. tasmaniensis* (Kube et al. 2008, 2010; Smits et al. 2010a, b). In addition, *E. pyrifoliae, E. piriflorinigrans,* and *E. tasmaniensis* strain Et1/99 contain *hrp* T3SS PAI1, but *E. piriflorinigrans* and *E. tasmaniensis* lack the HAE region that includes the effector *hrpK* and *hsvABC* genes (Kube et al. 2008; Smits et al. 2013). Furthermore, Flg-1 (separated into four gene clusters in the genome) is present in all *Erwinia* genomes; however, Flg-2 (clustered) is only present in *E. amylovora* and *E. pyrifoliae* strains (Smits et al. 2011; Zhao and Qi 2011).

1.2.2.3 Repetitive DNA and CRISPR Elements

Clustered regularly interspaced short palindromic repeats (CRISPR) represent a family of short DNA repeat sequences found in most archaeal and bacterial genomes (Horvath and Barrangou 2010). CRISPR typically consists of several non-contiguous direct repeats of 21–47 bp in size and

are separated by stretches of variable sequences (spacers). CRISPR is often adjacent to *cas* (CRISPR-associated) and *cse* (CRISPR Cascade complex) genes. Many of the spacer sequences associated with CRISPRs share sequence identity with bacteriophage, plasmid, and other laterally transferred DNA sequences. Thus, CRISPR/Cas systems function as CRISPR RNA (crRNA)-mediated adaptive immunity systems against bacteriophages and conjugative plasmids for sequence-specific detection and silencing of foreign DNAs, similar to RNA interference (RNAi) pathways in eukaryotes (Marraffini and Sontheimer 2010; Wiedenheft et al. 2012). In particular, Cas9, a DNA nuclease, has been shown to use dual-RNAs for site-specific DNA cleavage, which highlights the potential to exploit the system for RNA-programmable genome editing (Cho et al. 2013; Jinek et al. 2012).

Three *C*RISPR *r*epeat *r*egions (CRR1, CRR2, and CRR4) have been identified in genomes of *E. amylovora* and related species regardless of host range (McGhee and Sundin 2012; Rezzonico et al. 2011; Smits et al. 2010b). *E. pyrifoliae* strains contain four CRRs (CRR1 to CRR4), and CRR1, CRR2, and CRR4 repeats share 100 % sequence identity to those of CRRs in *E. amylovora*. *E. piriflorinigrans* and *E. tasmaniensis* contain two CRRs (CRR3 and CRR4). The repeats of both CRR1 and CRR2 are 29 bp in length, and only two nucleotide substitutions (GA to AT) at positions 14 and 15 differentiate CRR1 and CRR2 repeats (McGhee and Sundin 2012). The repeats of CRR3 and CRR4 are significantly different from those of CRR1 and CRR2 and are 28 bp in length. The CRR3 repeats of *E. pyrifoliae* and *E. tasmaniensis* are identical to each other, but differ by one bp from the CRR4 repeats of *E. amylovora* and *E. pyrifoliae* (Rezzonico et al. 2011).

In *E. amylovora*, the majority of spacers are characteristically 32 bp in length, ranging from 30 to 34 bp with some exception (Rezzonico et al. 2011). The number of spacers within CRR1 and CRR2 is variable among strains and ranges from 12 to 98 and 23 to 49 within CRR1 and CRR2, respectively (McGhee and Sundin 2012; Rezzonico et al. 2011). All strains contain five 34–35-bp spacers in CRR3 and CRR4. A total of 588 unique spacers have been identified in 85 *E. amylovora* strains (McGhee and Sundin 2012). Among them, approximately 23 % of the spacers match known sequences, including 16 % plasmids and 5 % bacteriophage. The plasmid pEU30, isolated from *E. amylovora* strains from the Western USA, is targeted by 55 spacers (Foster et al. 2004; McGhee and Sundin 2012). Interestingly, spacers from *E. pyrifoliae* do not share homology with those of *E. amylovora* (McGhee and Sundin 2012).

Both *E. amylovora* and *E. pyrifoliae* contain eight *E. coli*-type *cse* and *cas* genes between CRR1 and CRR2, and housekeeping genes between CRR2 or CRR3 and CRR4, respectively (Rezzonico et al. 2011). However, four housekeeping genes within a 2.56-kb region between the 3′ end of the CRR1 and the *cas3* gene is missing in *E. pyrifoliae, E. amylovora Rubus* strains, and some *E. amylovora* strains from the Western USA (CRISPR group III) (McGhee and Sundin 2012). In contrast, *E. pyrifoliae, E. piriflorinigrans,* and *E. tasmaniensis* have *Y. pestis* subtype *csy* genes between CRR3 and CRR4. The *csy* genes and CRR3 are apparently lost in *E. amylovora*, leaving only CRR4 as a relic (Rezzonico et al. 2011).

1.2.3 Genomic Resources

The EMBL/GenBank accession numbers for *E. amylovora* genome and plasmid sequences are listed in Tables 1.1, 1.2 and 1.3. The *E. amylovora* microarray design and HrpL regulon microarray data are available at ArrayExpress Web site (http://www.ebi.ac.uk/arrayexpress/; accessions: Microarray #A-MEXP-2000, and Dataset #E-TABM-1137) (McNally et al. 2012). Other microarray data are available at the National Center for Biotechnology Information (NCBI) Gene Expression Omnibus database (http://www.ncbi.nlm.nih.gov/geo/) (accession numbers GSE30751 (RcsBC regulons), GSE28630, GSE37064 (YbjN/AmyR regulons), and GSE45602 (T3SS inhibitors) (Wang et al. 2011a, 2012a, b; Yang et al. 2014).

1.3 Genome-Enabled Understanding of *E. amylovora* Virulence

During the past two decades, research on fire blight has made great strides in elucidating the genetic, molecular, and physiological basis for pathogenesis (Oh and Beer 2005). However, many key questions remain to be answered due to insufficient genetic information regarding *E. amylovora* and related *Erwinia* species, despite extensive studies on *E. amylovora* and fire blight disease for more than a century. Two puzzling questions for the fire blight research community are why natural isolates of *E. amylovora* display differential virulence and what are the molecular mechanisms underlying the host specificity of *Erwinia* strains (Zhao and Qi 2011). The recent revelation of the genetic composition of these microorganisms provides opportunities to employ both comparative and functional genomic approaches to understand the pathogen, its ability to cause disease, and its interaction between host plants and insect vectors.

1.3.1 Evolutionary Insights into Genome Sequencing of *E. amylovora* and Related Species

In prokaryotes, small subunit ribosomal RNA (rRNA) genes, especially 16S rRNA, have been universal phylogenetic markers for reconstructing the evolutionary relationships between microorganisms. Relatedness among closely related organisms can also be distinguished by analyzing the phylogenetic relationship of housekeeping genes (Fig. 1.1) (Sarkar and Guttman 2004; Zhao and Qi 2011). A global phylogeny of 191 fully sequenced organisms, representing the three domains of life, has been reconstructed based on a concatenated alignment of 31 single universal protein families related to ribosomal function (Ciccarelli et al. 2006). Phylogenomic trees have also been generated from more specialized data sets. A phylogenetic tree based on 14 flagellar proteins is congruent with a bacterial species tree using 25 single-copy proteins present in 249 genomes, indicating core components of the bacterial flagella may originate from a single ancestral gene through duplication and diversification (Liu and Ochman 2007). Furthermore, phylogenetic trees reconstructed from a concatenation of the seven core set of two-component systems from enterobacteria agreed well with that of the 16S rRNA gene (Qi et al. 2010).

1.3.1.1 Comparative Genomics and Evolution of *E. amylovora* and Related Species

In order to identify genes that are responsible for basic biology, virulence, and evolution of *E. amylovora* and its related species, and to determine the molecular mechanisms of the differential virulence and host specificity, comparisons of the complete genomes of *E. amylovora* strains and related *Erwinia* species have been performed (Kube et al. 2010; Mann et al. 2012, 2013; De Maayer et al. 2011; Rezzonico et al. 2012; Smits et al. 2010a, b, 2011, 2013; Thapa et al. 2013; Zhao and Qi 2011). These comparative genomic studies provided a preliminary scientific basis for determining the relatedness and evolution of genes/proteins within the genomes of *E. amylovora* and closely related *Erwinia* species (Smits et al. 2011; Zhao and Qi 2011).

Based on comparative genomic studies, a hypothesis of an evolutionary history of genome-sequenced *E. amylovora* strains and related *Erwinia* species has been proposed (Mann et al. 2012; Rezzonico et al. 2012; Smits et al. 2011). According to this hypothesis, the genealogy within the genus *Erwinia* agrees largely with the phylogeny (Fig. 1.1). From the ancestral *Erwinia* in the enterobacterial ancestor, it takes several evolutionary steps to reach the common ancestor for the genome-sequenced *Erwinia* species, which separate from other *Erwinia* spp. From there, the non-pathogenic *E. tasmaniensis* separates from the pathogenic *Erwinia* ancestor, which further differentiates into *E. piriflorinigrans, E. amylovora,* and *E. pyrifoliae* (Smits et al. 2010a). This hypothesis is supported by

specific features or traits discovered in the comparative genomic analyses (Malhony et al. 2012; Smits et al. 2010a, b). Most of the specific features or traits are summarized in Table 1.4 and some discussed below in detail.

Comparative genomic studies have identified the following ancestral origins of virulence-associated traits, including three T3SS PAIs (PAI1 to PAI3) (Mann et al. 2012; Zhao et al. 2009a), three type VI secretion systems (T6SS) (De Maayer et al. 2011), two flagellar systems (Zhao et al. 2011), one type I secretion system (Palacio-Bielsa et al. 2012), CRISPR repeat sequences and associated *cas*/*cse* genes (McGhee and Sundin 2012; Rezzonico et al. 2011), lipopolysaccharide, amylovoran and levan biosynthesis genes (Rezzonico et al. 2012; Smits et al. 2011), and genes involved in sorbitol metabolism (Mann et al. 2013; Smits et al. 2011). The major difference within the *hrp*-T3SS PAI1 between pathogenic *Erwinia* species and *E. tasmaniensis* is the IT and HAE regions; however, gene sequence diversification is also found within the HEE region as exemplified by the *eop1* gene among *E. amylovora* strains with different host ranges (Table 1.4).

Type VI secretion system (T6SS) has been identified in many Gram-negative bacteria (Records 2011); however, its role in virulence has not been documented in *E. amylovora*. In *E. amylovora*, three T6SS gene clusters (1–3) have been identified (De Maayer et al. 2011; Smits et al. 2010b). Variation between strains of *E. amylovora* is primarily found within the non-conserved *hcp* and *vgrG* islands of T6SS-1 regions II and IV, and T6SS-3 region IV (Mann et al. 2013). T6SS-1 and T6SS-2 are highly similar to the T6SS clusters of *E. pyrifoliae*, *E. piriflorinigrans, E. tasmaniensis*, and *E. billingiae* and some differences in the genes encoding VgrG proteins in T6SS-1. T6SS-2 also shows variations in gene content (Kube et al. 2010; Smits et al. 2010a), where a frame shift is found in one of the genes in *E. pyrifoliae* and a gene is lost in the corresponding cluster in *E. tasmaniensis* (Smits et al. 2010a, b). T6SS-3 is only identified in *E. amylovora*, but absent in other related *Erwinia* species (Smits et al. 2010b, 2013).

Lipopolysaccharides (LPS) are major components of the cell surface of Gram-negative bacteria. An LPS biosynthesis gene (*waaL*) in *Spiraeoideae*-infecting strains of *E. amylovora* has been shown to be involved in virulence (Berry et al. 2009). Based on genome sequences, *Spiraeoideae*-infecting strains contain three glycosyltransferases and an LPS ligase (*Spiraeoideae*-type *waaL*), whereas only two glycosyltransferases and a different LPS ligase (*Rubus*-type *waaL*) for *Rubus*-infecting strains exist in the core region of the LPS biosynthetic gene cluster (Rezzonico et al. 2012). These coding sequences for LPS biosynthesis genes share little to no homology at the amino acid level between *Rubus*- and *Spiraeoideae*-infecting strains (Mann et al. 2013). In addition, *amsCDE* within the *ams* biosynthetic operon are very diverse among different species of *Erwinia* (Langlotz et al. 2011).

Other genes seem to have been acquired after divergence of pathogenic species, including a second flagellar T3SS (Flg-2). Phylogenetic trees based on concatenation of 14 conserved flagellar proteins showed that both Flg-1 and Flg-2 are clustered with those of enterobacteria, indicating that these flagellar systems may be originated from other enterobacteria (Fig. 1.5) (Zhao et al. 2011). The Flg-1 system is much closer to the phylogeny of species trees than that of Flg-2, which is closely related to those of *Sodalis glossinidius*, suggesting that Flg-2 along with PAI2 and PAI3 may be acquired from a similar enterobacterial source by horizontal gene transfer (Figs. 1.4 and 1.5). Interestingly, PAI2 and PAI3 are also clustered together and closely related to those of *S. glossinidius* (Fig. 1.4) (Zhao et al. 2011).

Virulence factors such as type III effectors (Eop2, HopC1 and AvrRpt2) are present in *E. amylovora* strains, but not in *E. pyrifoliae* strains, and these effectors may contribute to host specificity (Khan et al. 2012; Zhao et al. 2005, 2006). The EPS levan is a virulence factor in *E. amylovora*; however, the levansucrase gene (*lsc*) is absent in the genome of *E. pyrifoliae* strains, which can be used to differentiate *E. amylovora* from *E. pyrifoliae* (Zhao and Qi

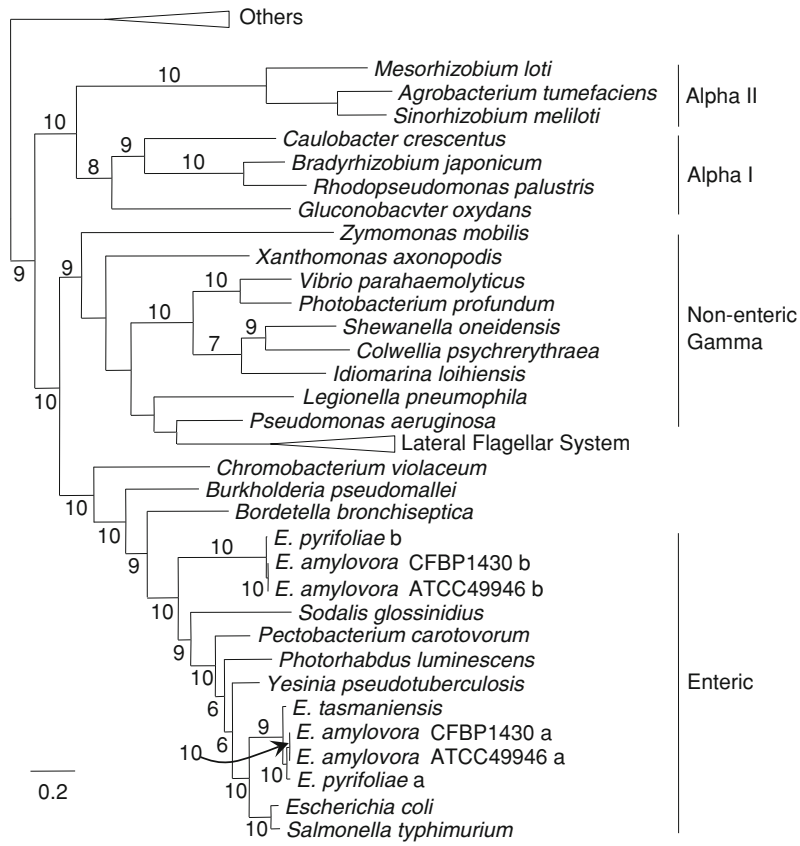

Fig. 1.5 Phylogenetic tree based on a concatenated alignment of 14 flagellar proteins that are present in all primary and lateral flagellar systems. "a" and "b" refer to separated (Flg-1) and clustered (Flg-2) flagellar systems, respectively from Zhao et al. (2011)

2011). A type I secretion system (T1SS, PrtA-DEF) is only present in the genome of *E. amylovora*, but not in the other related species (Table 1.4). The T1SS encodes a secreted protease (PrtA), which is a virulence factor, and its export system (PrtDEF) (Smits et al. 2011; Zhang and Geider 1999; Zhang et al. 1999). In summary, these virulence-associated traits may also be determinants of host specificity.

1.3.1.2 Understanding the Biology and Virulence of *E. amylovora* in Host–Pathogen Interactions

The *hrp*-T3SS and its effectors are one of the primary determinants in *E. amylovora* to cause fire blight disease (Oh and Beer 2005). In the past two decades, effector biology has been the center stage for studying the molecular function of secreted proteins, which are widely recognized as an essential component for understanding the pathogen infection process. Knowledge of the effector repertoire of *E. amylovora* and related *Erwinia* species from genome sequence projects promotes an understanding of *Erwinia*–host interactions. In *E. amylovora*, recent studies have begun to elucidate how type III effectors modulate plant susceptibility and promote the growth and dissemination of the pathogen (Khan et al. 2012; Zhao and Qi 2011).

The first T3SS effectors identified in *E. amylovora* were harpins (HrpN and HrpW) and a unique disease-specific protein DspA/E, all located within the *hrp*-T3SS PAI1 (Wei et al. 1992; Bogdanove et al. 1998a, b; Gaudriault et al. 1997, 1998). The large multi-domain protein DspE is essential for *E. amylovora* pathogenesis, as mutation of *dspA/E* resulted in a non-pathogenic phenotype and no growth of the bacterium on host plants (Bogdanove et al. 1998a, b; Gaudriault et al. 1997, 1998). Secretion of DspA/E requires a specific chaperone

DspB/F, which prevents DspA/E from degradation (Gaudriault et al. 2002; Oh et al. 2010). DspB/F functions as a classical T3SS chaperone by binding to amino acids 51 through 100 of DspA/E and also interacting with the C-terminal half of DspA/E (Oh et al. 2010; Triplett et al. 2009, 2010).

DspE/A is a member of conserved bacterial effector family and members of this family have been found to contribute to disease development by inhibiting salicylic acid-mediated innate immunity, such as blocking of callose deposition and suppression of PR1 gene expression (Boureau et al. 2006; DebRoy et al. 2004). Transient expression of DspE/A in apple, tobacco, and *Nicotiana benthamiana* led to necrosis, which is dependent on NbSGT-1, a protein required for programmed cell death (Boureau et al. 2006; Oh et al. 2007). Host targets for DspE in apple (*Malus x domestica*) have recently been identified using the yeast two-hybrid system (Meng et al. 2006; Oh and Beer 2007). The N-terminus 4.7-kb fragment of DspE/A interacts with four similar putative leucine-rich repeat (LRR) receptor-like serine/threonine kinases (RLK) from apple, named as DspE-interacting proteins from *Malus* (DIPM) (Meng et al. 2006). The DspE 3.2-kb fragment from the C-terminus interacts with a cytoplasmic precursor, ferredoxin, which serves as an electron carrier in photosystem I (PS-I) (Bonasera et al. 2006). These studies suggested that DspE/A may inhibit signal transduction and photosynthesis by targeting chloroplasts to facilitate pathogenesis.

HrpN and HrpW belong to the harpin family, and HrpN was the first cell-free elicitor identified in 1992 (Wei and Beer 1993; Wei et al. 1992). Wei et al. (1992) first purified the harpin protein from *E. amylovora* and found that it was required to trigger HR when infiltrated into tobacco, tomato, and *Arabidopsis thaliana*. HrpN plays an essential role in virulence and avirulence activities of *E. amylovora* as the *hrpN* mutant was incapable of causing disease in pear and induced a reduced oxidative burst in apple, suggesting that HrpN contributes to the generation of oxidative stress in host plants (Degrave et al. 2008; Venisse et al. 2001, 2003). However, HrpN does not trigger cell death on apple leaves or cells (Reboutier et al. 2007). HrpW contains a pectate lyase domain homologous with type 3 pectate lyase and is partially required for virulence (Kim and Beer 1998). Both HrpN and HrpW contain functional N-terminal type III translocation signals, indicative of injection of harpins into the plant cell (Bocsanczy et al. 2008; Boureau et al. 2011; Perino et al. 1999).

HrpN plays a critical role in facilitating its own translocation process into the host plant cell (Sinn et al. 2008) and other effectors, including DspE/A and HrpW. This is based on a drastic reduction in translocation of DspA/E:CyaA, HrpN:CyaA, and HrpW:CyaA fusion proteins in a *hrp*N mutant background (Bocsanczy et al. 2008; Boureau et al. 2011). In addition, purified HrpN triggered opposite ion flux responses in non-host *Arabidopsis thaliana* cells and host apple cells, which indicates that HrpN may be perceived differently by host and non-host plants (Reboutier et al. 2007). However, purified HrpN fails to induce oxidative stress in apple leaves or cells, suggesting that oxidative stress induction may depend on HrpN-dependent translocation of other effectors, such as DspE (Boureau et al. 2011; Degrave et al. 2008). A host target protein for HrpN (named HIPN) has been identified in apple (Oh and Beer 2007), and interestingly, AtHIPN, the homolog of HIPN, inhibits plant growth in *Arabidopsis*. This evidence indicates that translocation of HrpN promotes both plant growth and infection by the pathogen (Khan et al. 2012).

The first T3SS effectors identified that are located outside the *hrp*-T3SS PAI1 in *E. amylovora* genome were AvrRpt2 (Eop4) and HopPtoC (HopC1) (Zhao et al. 2005, 2006). Mutation in *hopPtoC* does not result in a noticeable virulence defect (Zhao et al. 2005), whereas AvrRpt2 acts both as a virulence factor in susceptible hosts (Zhao et al. 2006) and an avirulence factor on resistant *Malus X robusta 5* apple in a gene-for-gene manner (Vogt et al. 2013). AvrRpt2 from *E. amylovora* is a homolog of AvrRpt2 from *P. syringae* pv. *tomato* and genetically interacts with the RPS2 disease-

resistance protein in *A. thaliana* when expressed in *P. syringae* pv. *tomato* DC3000 (Zhao et al. 2006). The effector protein AvrRpt2 and disease-resistance protein RPS2 in *Arabidopsis* is a well-studied example of gene-for-gene interactions (Axtell and Staskawicz 2003; Day et al. 2005; Mackey et al. 2003). AvrRpt2 is a cysteine protease and encodes a 28-kD effector protein delivered into plant cells during infection (Axtell et al. 2003). The RPS2 protein is a plasma membrane-associated intracellular nucleotide-binding (NB)-leucine-rich repeat (LRR) protein (Bent et al. 1994; Kunkel et al. 1993). Several elegant studies provide experimental support for the so-called guard hypothesis to describe the interaction between AvrRpt2 and RPS2 proteins (Coaker et al. 2005, 2006).

One of the virulence targets of AvrRpt2 in *Arabidopsis* is the RIN4 (RPM1-interacting protein 4), which is guarded by the RPS2 NB-LRR protein (Belkhadir et al. 2004; Chisholm et al. 2005; Mackey et al. 2002, 2003). In healthy *Arabidopsis* plants, RIN4, as a negative regulator, binds and suppresses the activity of RPS2 (Belkhadir et al. 2004). After secretion into plant cells by the T3SS, the AvrRpt2 interacts with host protein ROC1 (cyclophilins) which binds AvrRpt2 at four sites, thus activating the self-cleavage function of AvrRpt2 (Coaker et al. 2005, 2006; Jin et al. 2003). The activated AvrRpt2 protease not only cleaves its own secretion signal peptide, but also eliminates RIN4, thus activating the RPS2 protein (Coaker et al. 2006; Day et al. 2005). In *E. amylovora*, deletion of *avrRpt2* overcomes the resistance conferred by *Malus* X *robusta* 5 (Mr5) (Vogt et al. 2013). Interestingly, an amino acid substitution (C156S) of AvrRpt2 found in some *E. amylovora* strains isolated from Canada and the USA also overcomes the resistance of Mr5 (Vogt et al. 2013). However, co-expression of AvrRpt2 with RIN4 from *A. thaliana* indicated that only the S-allele of AvrRpt2 led to partial disappearance of RIN4 and caused stronger HR response in *N. benthamiana*. This suggested that the mechanism of resistance may be different from *Arabidopsis*, although a homolog of RIN4 has been detected in Mr5 (Vogt et al. 2013). Furthermore, a single-base deletion at nucleotide 165 (amino acid 55) of AvrRpt2, which occurs in the *Rubus*-infecting strains Ea644 and MR-1, results in a frame shift and truncated AvrRpt2 without the N-terminal secretion and translocation signal (Mann et al. 2013).

Comparison of effector homologs in the pan-genome revealed variation of the effector proteins HopX1 and Eop1 among different strains of *E. amylovora* (Asselin et al. 2011; Bocsanczy et al. 2012; Mann et al. 2013). The consistent variation observed among *Spiraeoideae*-infecting and *Rubus*-infecting strains of *E. amylovora* suggested that both proteins are strong candidates as determinants of host specificity, although deletion of both genes in their corresponding host strains did not result in reduced virulence (Asselin et al. 2011; Mann et al. 2013). Nevertheless, these effectors are indeed secreted by *E. amylovora* (Nissinen et al. 2007).

1.3.1.3 Population Genetics and Diversity, and Antibiotic Resistance Tracking

E. amylovora is a relatively homogeneous bacterial species (Triplet et al. 2006; Zhao and Qi 2011). Early studies, however, showed that *E. amylovora* strains differ in their ability to synthesize dihydrophenylalanine, and in their profiles for carbon utilization and fatty acid composition (Gehring and Geider 2012; van der Zwet and Wells 1993). *E. amylovora* can also be differentiated based on genetic DNA polymorphisms, including pEA29-based PCR (Llop et al. 2006; McGhee and Jones 2000), Rep-PCR (McManus and Jones 1995), *hrp* gene restriction fragment length polymorphism (RFLP) (Giorgi and Scortichini 2005), random amplified DNA (RAPD) analysis (Momol et al. 1997), amplified fragment length polymorphism (AFLP) analysis (Rico et al. 2004), amplified rRNA gene restriction analysis (ARDRA) (Rezzonico et al. 2011), PFGE (Jock and Geider 2004; Jock et al. 2002), and ribotyping (Donat et al. 2007). These studies consistently differentiate *E. amylovora* strains isolated from *Spiraeoideae* (*Maloideae*) and those from *Rosoideae* (*Rubus* spp.),

suggesting a greater diverse population of *Erwinia* strains within its North American origin. Genome sequence analyses of *E. amylovora* strains isolated from *Spiraeoideae* (*Maloideae*) and those from *Rosoideae* (*Rubus* spp.) further confirm these observations (Mann et al. 2013). In addition, some *E. amylovora* strains isolated from the Netherlands are deficient in levan synthesis (Bereswill et al. 1997). In Spain, *E. amylovora* strain UPN527 is deficient in the ubiquitous pEA29 plasmid, but the virulence of this pEA29-deficient strain is not reduced (Llop et al. 2006). The nearly ubiquitous plasmid pEA29 has been used for diagnostics purpose and has plasmid variants implicated in streptomycin resistance (McGhee et al. 2011; Russo et al. 2008).

Among strains isolated from *Spiraeoideae*, differential virulence has been observed on various apple cultivars (Norelli et al. 1984, 1986). *E. amylovora* strain Ea273 causes higher disease severity on most apple/crabapple genotypes including "Holly," and little or no disease on genotypes "Quinte," "Novole," and *Malus* x *robusta* 5. In contrast, *E. amylovora* strain Ea266 appears to be less virulent on "Holly," but causes severe disease on "Quinte," whereas strain EA4001A is virulent on all three genotypes. Based on severity of virulence on different apple/crabapple genotypes, *E. amylovora* strains have been classified as highly virulent (Ea4001a), moderately virulent (Ea273), and low virulent (Ea321) (Norelli et al. 1984). Similar results have been reported for *E. amylovora* strains isolated from Bulgaria, and from Idaho, Michigan, and Pennsylvania (Lee et al. 2010; Thoelen et al. 2008; Wang et al. 2010a). Lee et al. (2010) reported a highly virulent strain (HKN06P1) in Pennsylvania, which secreted more DspE than less virulent strains. Wang et al. (2010a) reported the molecular signature of differential virulence in four natural isolates of *E. amylovora*, including two strains from Michigan and New York (Ea110 and Ea273) and two strains from Europe (CFBP1430 and Ea1189). A positive correlation between bacterial virulence on relatively resistant genotypes and the expression/production of major virulence factors such as HrpL, DspE, and amylovoran in *E. amylovora* WT strains has been demonstrated (Wang et al. 2010a). However, from genome sequence, strains Ea273 and CFBP1430 share 99.9 % sequence identity; therefore, unknown factors could influence the expression of virulence genes (Wang et al. 2010a).

CRISPR spacer diversity among *E. amylovora* strains varying in geographical origin, host range, and streptomycin sensitivity/resistance have recently been characterized (McGhee and Sundin 2012; Rezzonico et al. 2011). Based on CRISPR spacer diversity, *Spiraeoideae*-infecting strains of *E. amylovora* can be differentiated into three main groups of different geographical origin (McGhee and Sundin 2012; Rezzonico et al. 2011). CRISPR group I contains strains from Europe, the Middle East, New Zealand, and the East Coast of North America, suggesting that it is plausible that an *E. amylovora* strain from the eastern USA is the likely source of fire blight disease spread into New Zealand and Europe as initially reported (McGhee and Sundin 2012; van der Zwet 2004; van der Zwet et al. 2012). Both CRISPR group II and III strains are from the Western USA to suggest that relatedness of spacer content could differentiate strains isolated in the Eastern USA from those recovered in the west. Interestingly, CRISPR genotyping correlated with ribotype, PFGE, and *groEL* sequence groupings (Jock et al. 2002; McManus and Jones 1995; McGhee et al. 2011). In addition, *E. amylovora* strains from *Rubus* and Indian hawthorn contain mostly unique spacers compared to apple and pear strains, while strains from loquat share 79 % of spacers with apple and pear strains (McGhee and Sundin 2012). Currently, all sequenced *Spiraeoideae*-infecting strains are of CRISPR group I (Mann et al. 2013). It is possible that genome sequences of *E. amylovora* strains from CRISPR groups II and III might uncover more genetic information about this pathogen.

Analysis of CRISPR spacer pattern has been found to be very useful in tracking streptomycin resistance in Michigan (McGhee and Sundin 2012). CRISPR spacer patterns for most streptomycin-resistant (Sm^R) strains from Michigan

exhibit an identical CRISPR genotype, whereas diverse CRISPR spacer patterns are found among CRR1 and CRR2 for streptomycin-sensitive (Sm^S) strains, with 10 genotypes observed among 17 strains (McGhee and Sundin 2012). Similar CRISPR genotypes were detected in Michigan in Sm^S *E. amylovora* populations and in corresponding Sm^R strains, which contain either Tn*5393* or a spontaneous Sm^R mutation (McGhee et al. 2011). Thus, the Sm^R populations in Michigan may have evolved from indigenous populations of *E. amylovora* with antibitoic resistance arising from locally adapted genotypes. It will be interesting to determine whether genotyping of CRISPR spacer sequences can be used for Sm^R strain tracking on a regional level.

1.3.2 Functional Genomic Studies

1.3.2.1 Transcriptomics Using Bacterial and Host Microarrays

Based on the complete genome sequences of *E. amylovora*, a 60-mer oligonucleotide microarray has been recently developed; thus, genome-wide transcriptomic studies become possible (McNally et al. 2012; Wang et al. 2012a; Yang et al. 2014). The microarray consists of 3,483 chromosomal genes from *E. amylovora* strain ATCC 49946 and 483 sequences from known *E. amylovora* plasmids (McNally et al. 2012). McNally et al. (2012) identified 19 genes positively regulated by HrpL, including known T3SS and effector genes, and five negatively regulated genes in cells of *E. amylovora* grown in a *hrp*-inducing medium. A novel virulence gene, *EAM_2938*, was identified that encodes a putative membrane protein and contributes to virulence (McNally et al. 2012).

Using the same microarray, the regulons of RcsBC of *E. amylovora* were also determined both in vitro and in vivo (Wang et al. 2012a). A total of 648 genes differentially regulated by RcsCB in vitro and in vivo were identified, including amylovoran biosynthesis and regulatory genes, cell wall and cell envelope (membrane) genes, and several novel genes (Wang et al. 2012a). Consistent with our previous findings, RcsB acts as a positive regulator in both conditions, whereas RcsC positively controls expression of amylovoran biosynthetic genes in vivo, but negatively in vitro (Wang et al. 2009, 2011b, 2012a). Interestingly, the expression of *EAM_2938* was also regulated by the RcsBCD phosphorelay system, suggesting possible cross-talk between amylovoran biosynthesis and T3SS. In addition, an orphan protein AmyR, a functionally conserved ortholog of YbjN in *E. coli*, is a novel negative regulator of amylovoran production in *E. amylovora* (Wang et al. 2011a, 2012b). Microarray studies also revealed that amylovoran biosynthesis and related membrane protein-encoding genes, such as *EAM_0255*, were highly expressed in the *amyR* mutant, but down-regulated in the *amyR* over-expression strains in vitro (Wang et al. 2012b). Further characterization of novel genes identified by microarray will provide a more comprehensive understanding of *E. amylovora* pathogenesis and biology (Malnoy et al. 2012).

Identification of genes differentially expressed in resistant and susceptible apple cultivars, as well as apple flowers and rootstocks challenged with *E. amylovora,* were reported (Baldo et al. 2010; Jensen et al. 2012; Sarowar et al. 2011b). Using an apple microarray of 40,000 genes representing different tissues and conditions, including fire blight challenged tissues, Sarowar et al. (2011b) identified about 3,500 genes that were significantly modulated in response to *E. amylovora* after blossoms were spray inoculated. The majority of up-regulated genes are involved in metabolism, signal transduction, transport, and stress response. More than 150 genes were commonly modulated by comparing wild-type versus *T3SS* and *dspE* mutant-challenged blossoms, suggesting these genes may be specific to T3SS, and indicating that T3SS may play a significant role in the early stage of infection (Sarowar et al. 2011a). Jensen et al. (2012) identified 690 transcripts whose steady-state expression levels correlated with the degree of fire blight susceptibility of the scion/rootstock combinations using a NimbleGen™ array

consisting of 55,230 probes. Results determined that 39 transcripts had expression levels that correlated with fire blight resistance in the breeding population. Since both apple and pear genome sequences have become available (Velasco et al. 2010; Wu et al. 2012), the next generation of microarray (AryAne2 developed on the basis of the annotation of the apple genome) or next-generation sequencing methods, such as RNA-seq, will provide new opportunities for genome-wide studies on host–pathogen interactions for fire blight disease (Malnoy et al. 2012).

1.3.2.2 Proteomics

Compared to other organisms, proteomics approaches have not been widely used to study *E. amylovora* and related species. So far, only two comprehensive proteomics studies have been reported for *E. amylovora* (Nissinen et al. 2007; Wu et al. 2013). Nissinen et al. (2007) reported the first type III secretome of *E. amylovora* strain Ea273 using mass spectrometry and identified a total of twelve T3SS-secreted proteins, including known effectors and chaperones such as DspE, HrpW, HrpN, and HrpA. Six new proteins whose secretion in vitro is dependent on a functional T3SS were also identified (Nissinen et al. 2007). Among them, HrpJ, a homolog of YopN of *Y. pestis* (Bogdanove et al. 1996), may participate together with HrpK and HrpN to form the translocon apparatus. Homologs to known T3SS effectors in other phytopathogenic bacteria identified in the secretome included Eop1 to Eop4. Eop1 (OrfB or EopB) has significant homology to the AvrRXv/YopJ effector family and is believed to act as a cysteine protease (Asselin et al. 2011; Oh and Beer 2005). Eop2 shares homology with the HopPmaH/HopAK family proteins, and Eop3 (HopX1) is homologous to AvrPphE (HopX) (Bocsanczy et al. 2012). Eop4 (AvrRpt2) is homologous to AvrRpt2, a cysteine protease and an avirulence protein of *P. syringae* (Zhao et al. 2006).

Wu et al. (2013) recently reported the first lysine acetylome in *E. amylovora* strains Ea1189 and Ea273 using liquid chromatography and tandem mass spectrometry (LC-MS/MS) and matrix-assisted laser desorption/ionization-time-of-flight mass spectrometer (MALDI-TOF). Protein lysine acetylation (LysAc), a dynamic and reversible protein posttranslational modification, has recently been demonstrated to be widespread in *E. coli* and *Salmonella* (Wang et al. 2010b; Zhang et al. 2009) and was shown to broadly regulate bacterial physiology and metabolism. A total of 141 LysAc sites in 96 proteins that function in a wide range of biological pathways were identified in two *E. amylovora* strains (Wu et al. 2013). Consistent with previous reports, 44 % of the proteins are involved in metabolic processes, including central metabolism, lipopolysaccharide, nucleotide, and amino acid metabolism. Interestingly, several proteins involved in *E. amylovora* virulence, including amylovoran biosynthesis (i.e., AmsJ, GalE) and type III secretion-associated proteins (i.e., HsvB, YopH), were found, for the first time, to be lysine acetylated. These results suggest that LysAc may play a major role in bacterial virulence. The study by Wu et al. (2013) also demonstrated that LysAc profiles in *E. amylovora* were strongly altered under different growth conditions and that differential LysAc profiles were observed for two *E. amylovora* strains that differ in their ability to cause disease on different apple genotypes (Wu et al. 2013). Collectively, these results reinforce the notion that lysine acetylation of proteins is widespread in bacterial metabolism and virulence.

1.3.2.3 From Genome to Gene Function

Before the genome sequence of *E. amylovora* became available, numerous genome-level studies had been conducted to elucidate the molecular mechanisms of pathogenicity of *E. amylovora* with emphasis on genes involved in symptom development, colonization, and plant–microbe interaction (Triplett et al. 2006; Wang and Beer 2006; Zhao et al. 2005). Zhao et al. (2005) used a modified in vivo gene expression technique (IVET) system and identified 394 unique *p*ear *f*ruit-*i*nduced (*pfi*) genes

of *E. amylovora* that are activated during infection of immature pear tissue, a process that requires expression of the major pathogenicity factors of this organism. These genes included host–microbe interactions (3.8 %), stress responses (5.3 %), regulation (11.9 %), cell surface properties (8.9 %), transport (13.5 %), mobile elements (1.0 %), metabolism (20.3 %), nutrient acquisition and synthesis (15.5 %), and unknown or hypothetical proteins (19.8 %). Known virulence genes, including *hrp/hrc* components of the type III secretion system, the major effector gene *dspE*, levansucrase *lsc*, and regulators of levansucrase and amylovoran biosynthesis were up-regulated during pear tissue infection (Zhao et al. 2005). New effectors, including HopPtoC and AvrRpt2, were also identified in this screening (Zhao et al. 2005, 2006).

Based on the genome sequence of *E. amylovora* strain ATCC49946, Zhao et al. (2009b) utilized a systems approach and identified about 50 two-component systems (TCS) and related signal transduction genes in *E. amylovora*. Comparative genomic analysis of 53 genomes of 16 enterobacterial species revealed that enterobacteria contain eight pairs of core TCSs (Qi et al. 2010). Phylogenetic trees reconstructed from the core set of TCSs showed that TCS protein trees agreed well with that of the corresponding 16S rRNA gene. Several core TCSs were quite ancient and universal based on phylogenomic analysis of protein structures (Qi et al. 2010). These results suggest that enterobacteria may have maintained their ancient core TCSs and might acquire specific new TCSs for their survival in different environments or hosts, or may have evolved new regulatory functions for adaptation to different ecological niches (Qi et al. 2010).

Genome-wide TCS gene-knockout and mutant screening revealed that the TCS plays a major role in *E. amylovora* virulence and regulation of amylovoran biosynthesis, including the Rcs phosphorelay system (Nakka et al. 2010a, b; Wang et al. 2010a, 2011b, 2012a; Zhao et al. 2009b). Both negative and positive regulators of amylovoran biosynthesis were identified in support ofthe presence of regulatory networks governing expression of critical virulence genes in *E. amylovora* (Zhao et al. 2009b). In addition, new regulators for amylovoran biosynthesis and T3SS were discovered based on genome sequences, including AmyR (Wang et al. 2011b), a ribosome-associated protein YhbH (Ancona et al. 2014), RNA-binding protein RsmA (Ancona and Zhao 2013), and small RNA chaperone Hfq and regulatory RNAs *rprA* and *ryhA* (Zeng et al. 2013). Combined with other functional genomic approaches such as RNA-seq, it is expected that broader identification and increased discovery of new virulence and regulatory genes in *E. amylovora* are possible in the near future.

1.4 Future Perspectives and Research Directions

In summary, genome sequences of *E. amylovora* and related *Erwinia* species associated with pome fruit trees have provided a near-complete picture of the genetic composition and/or the core/pan-genome of these erwinias. It can be concluded that *E. amylovora* is closely related to both pathogenic and non-pathogenic erwinias associated with pome fruit trees. Both *E. amylovora* and *E. pyrifoliae* exhibit very low levels of genetic diversity and have both been considered as homogeneous species. Genome sequence and comparative genomic analyses reinforce these assumptions, support the current classification scheme of *Erwinia* species associated with pome fruit trees, and provide new clues about the evolution of *E. amylovora* and related *Erwinia* species. Genetic and functional genomic studies have demonstrated that the key determinants of *E. amylovora* pathogenesis are the EPS amylovoran and the *hrp*-T3SS and its associated effectors. In recent years, many novel virulence factors or regulators have been discovered after the genome sequence became available. Studies on host–pathogen interactions, especially effector functions, have started to reveal the molecular mechanisms of *Erwinia* pathogenesis and the host defense response. The

recent genome sequence of the apple and pear host plants will create new opportunities for studying the host–pathogen interactions at a new level that will help identify host-resistance genes. There will be tremendous progress in the next decade or so in studying fire blight and related plant diseases, which will ultimately lead to the development of environmentally sound disease management strategies.

Unfortunately, the function of about half of the genes is still unknown for any given bacterial genome, including *E. amylovora* and related species. These genes are annotated as genes of unknown function, and their products are typically referred to as "conserved" or "hypothetical" proteins. These unknown proteins provide opportunities to better understand the biology of a particular organism, and open up potentially new biomedical and commercial opportunities. Novel approaches such as protein structure-based computational prediction could be used to understand the function of unknown proteins, and the "in silico" (computer-aided) predictions could be validated by additional means such as enzyme assays and X-ray crystallography. On the other hand, complete annotation of a given bacterial genome is still a very daunting task, and many small proteins and small regulatory RNAs are poorly annotated in any given bacterial genome. These small proteins and small regulatory RNAs play significant roles in virulence and survival during host–pathogen interactions. Therefore, the fire blight research community should take on the challenge of improving the annotation of *Erwinia* genomes in the near future.

Whole-genome sequencing and functional genomic studies are powerful hypothesis generators, which will help to better understand the pathogen, its virulence mechanisms, and host specificity. A holistic systems biology approach should be employed to investigate the complicated interactions of *Erwinia* with both plant hosts and insect vectors, including the identification of the complex and dynamic exchanges of signals between host and pathogen and how these signals are sensed. Reconstructing regulatory or signaling networks and metabolic networks (biochemical pathways) will lead to a better understanding of disease development and resistance mechanisms. These regulatory/signaling networks could be defined at different levels, including DNA/RNA–protein and protein–protein interactions, posttranscriptional and protein posttranslational modifications. Furthermore, comprehensive understanding of host–pathogen interactions requires not only the genomics of the pathogen, but also the genomics of the host. Recent technological advances such as next-generation sequencing-based transcription analysis (RNA-seq) will revolutionize investigations of prokaryotic as well as eukaryotic transcriptomes. A conceptually feasible "dual" RNA-seq of both pathogen and host could monitor gene expression simultaneously in two organisms, which could provide simultaneous information to a high level of accuracy and depth during host–pathogen interactions.

Finally, some obvious and imminent questions facing the fire blight research community include: What are the host specificity determinants? What causes the differential virulence phenomenon of *Erwinia* strains? What are the signals sensed by the pathogen to trigger the expression of T3SS and amylovoran biosynthesis genes? What is the molecular mechanism of effector protein function, such as DspE/A and AvrRpt2, when they are translocated inside plant cells and what are their host targets or potential resistance genes? What is the function of the type VI secretion systems in erwinias? What are the functions of the T3SS PAI2 and PAI3 during interactions with insect vectors and their relationship within the microbiome of insect vectors? Ultimately, answering these questions will lead to a common goal, namely to find alternative control measures to decrease losses due to fire blight disease, and to develop novel and improved disease management strategies.

Acknowledgements This project was supported by the Agriculture and Food Research Initiative Competitive Grants Program Grant no. 2010-65110-20497 from the USDA National Institute of Food and Agriculture (Microbial Functional Genomics Program).

References

Aldridge P, Bernhard F, Bugert P, Coplin DL, Geider K (1997a) Characterization of a gene locus from *Erwinia amylovora* with regulatory functions in exopolysaccharide synthesis of *Erwinia* spp. Can J Microbiol 44:657–666

Aldridge P, Metzger M, Geider K (1997b) Genetics of sorbitol metabolism in *Erwinia amylovora* and its influence on bacterial virulence. Mol Genom Genet 256:611–619

Ancona V, Zhao YF (2013) CsrA is a positive regulator of virulence factors in *Erwinia amylovora*. Phytopahtology 103(S2):6

Ancona V, Li WT, Zhao YF (2014) Alternative sigma factor RpoN and its modulator protein YhbH are indispensable for *Erwinia amylovora* virulence. Mol Plant Pathol 15:58–66

Ashton D (2008) Fire blight in Switzerland. Aust Fruitgrower 2:5

Asselin JE, Bonasera JM, Kim JF, Oh CS, Beer SV (2011) Eop1 from a *Rubus* strain of *Erwinia amylovora* functions as a host-range limiting factor. Phytopathology 101:935–944

Axtell MJ, Staskawicz BJ (2003) Initiation of RPS2-specified disease resistance in *Arabidopsis* is coupled to the AvrRpt2-directed elimination of RIN4. Cell 112:369–377

Axtell MJ, Chisholm ST, Dahlbeck D, Staskawicz BJ (2003) Genetic and molecular evidence that the *Pseudomonas syringae* type III effector protein AvrRpt2 is a cysteine protease. Mol Microbiol 49:1537–1546

Baldo A, Norelli JL Jr, Farrell RE, Bassett CL, Aldwinckle HS, Malnoy M (2010) Identification of genes differentially expressed during interaction of resistant and susceptible apple cultivars (*Malus* × *domestica*) with *Erwinia amylovora*. BMC Plant Biol 10:1

Barny MA, Guinebretiere MH, Marcais B, Coissac E, Paulin JP, Laurent J (1990) Cloning of a large gene cluster involved in *Erwinia amylovora* CFBP1430 virulence. Mol Microbiol 4:777–786

Belkhadir Y, Nimchuk Z, Hubert DA, Mackey D, Dangl JL (2004) *Arabidopsis* RIN4 negatively regulates disease resistance mediated by RPS2 and RPM1 downstream or independent of the NDR1 signal modulator and is not required for the virulence functions of bacterial type III effectors AvrRpt2 or AvrRpm1. Plant Cell 16:2822–2835

Bellemann P, Geider K (1992) Localization of transposon insertions in pathogenicity mutants of *Erwinia amylovora* and their biochemical characterization. J Gen Microbiol 138:931–940

Bellemann P, Bereswill S, Berger S, Geider K (1994) Visualization of capsule formation by *Erwinia amylovora* and assays to determine amylovoran synthesis. Int J Biol Macromol 16:290–296

Bennet RA, Billing E (1980) Origin of the polysaccharide component of ooze from plants infected with *Erwinia amylovora*. J Gen Microbiol 116:341–349

Bent AF, Kunkel BN, Dahlbeck D, Brown KL, Schmidt R, Giraudat J, Leung J, Staskawicz BJ (1994) RPS2 of *Arabidopsis thaliana*: a leucine-rich repeat class of plant disease resistance genes. Science 265:1856–1860

Bereswill S, Geider K (1997) Characterization of the *rcsB* gene from *Erwinia amylovora* and its influence on exopolysaccharide synthesis and virulence of the fire blight pathogen. J Bacteriol 179:1354–1361

Bereswill S, Jock S, Aldridge P, Jansen JD, Geider K (1997) Molecular characterization of natural *Erwinia amylovora* strains deficient in levan synthesis. Physiol Mol Plant Pathol 51:215–225

Bernhard F, Poetter K, Geider K, Coplin DL (1990) The *rcsA* gene from *Erwinia amylovora*: identification, nucleotide sequence, and regulation of exopolysaccharide biosynthesis. Mol Plant Microbe Interact 3:29–437

Bernhard F, Coplin DL, Geider K (1993) A gene cluster for amylovoran synthesis in *Erwinia amylovora*: characterization and relationship to *cps* genes in *Erwinia stewartii*. Mol Gen Genet 239:158–168

Berry MC, McGhee GC, Zhao YF, Sundin GW (2009) Effect of a *waaL* mutation on lipopolysaccharide composition, oxidative stress survival, and virulence in *Erwinia amylovora*. FEMS Microbiol Lett 291:80–87

Bobev SG, van Vaernbergh J, Maes M (2007) First report of fire blight on *Pyrus elaeagrifolia* and *Amelanchier sp.* in Bulgaria. Plant Dis 91:110

Bocsanczy AM, Nissinen RM, Oh C-S, Beer SV (2008) HrpN of *Erwinia amylovora* functions in the translocation of DspA/E into plant cells. Mol Plant Pathol 9:425–434

Bocsanczy AM, Schneider DJ, DeClerck GA, Cartinhour S, Beer SV (2012) HopX1 in *Erwinia amylovora* functions as an avirulence protein in apple and is regulated by HrpL. J Bacteriol 194:553–560

Bogdanove AJ, Wei ZM, Zhao L, Beer SV (1996) *Erwinia amylovora* secretes harpin via a type III pathway and contains a homolog of *yopN* of *Yersinia* spp. J Bacteriol 178:1720–1730

Bogdanove AJ, Bauer DW, Beer SV (1998a) *Erwinia amylovora* secretes DspE, a pathogenicity factor and functional AvrE homolog, through the Hrp (type III secretion) pathway. J Bacteriol 180:2244–2247

Bogdanove AJ, Kim JF, We ZM, Kolchinsky I, Charkowski AO, Conlin AK, Collmer A, Beer SV (1998b) Homology and functional similarity of an *hrp*-linked pathogenicity locus, *dspEF*, of *Erwinia amylovora* and the avirulence locus *avrE* of *Pseudomonas syringae* pathovar tomato. Proc Natl Acad Sci USA 95:1325–1330

Bogs J, Geider K (2000) Molecular analysis of sucrose metabolism of *Erwinia amylovora* and influence on bacterial virulence. J Bacteriol 182:5351–5358

Bonasera JM, Meng X, Owens T, Beer SV (2006) Interaction of DspE/A, a pathogenicity/avriulence protein of *Erwinia amylovora*, with pre-ferredoxin from apple and its relationship to photosynthetic efficiency. Acta Hortic 704:473–477

Boureau T, El Maarouf-Bouteau H, Garnier A, Brisset MN, Perino C, Pucheu I, Barny MA (2006) DspA/E, a type III effector essential for *Erwinia amylovora* pathogenicity and growth in planta, induces cell death in host apple and nonhost tobacco plants. Mol Plant Microbe Interact 19:16–24

Boureau T, Siamer S, Perino C, Gaubert S, Patrit O, Degrave A, Fagard M, Chevreau E, Barny MA (2011) The HrpN effector of *Erwinia amylovora*, which is involved in type III translocation, contributes directly or indirectly to Callose elicitation on apple leaves. Mol Plant Microbe Interact 24:577–584

Bugert P, Geider K (1995) Molecular analysis of the ams operon required for exopolysaccharide synthesis of *Erwinia amylovora*. Mol Microbiol 15:917–933

Chiou C-S, Jones AL (1995) Molecular analysis of high-level streptomycin resistance in *Erwinia amylovora*. Phytopathology 85:324–328

Chisholm ST, Dahlbeck D, Krishnamurthy N, Day B, Sjolander K, Staskawicz BJ (2005) Molecular characterization of proteolytic cleavage sites in the *Pseudomonas syringae* effector AvrRpt2. Proc Natl Acad Sci USA 102:2087–2092

Cho SW, Kim S, Kim JM, Kim JS (2013) Targeted genome engineering in human cells with the Cas9 RNA-guided endonuclease. Nat Biotechnol 31: 230–232

Ciccarelli FD, Doerks T, von Mering C, Creevey CJ, Snel B, Bork P (2006) Toward automatic reconstruction of a highly resolved tree of life. Science 311:1283–1287

Coaker G, Falick A, Staskawicz B (2005) Activation of a phytopathogenic bacterial effector protein by a eukaryotic cyclophilin. Science 308:548–550

Coaker G, Zhu G, Ding Z, van Doren SR, Staskawicz B (2006) Eukaryotic cyclophilin as a molecular switch for effector activation. Mol Microbiol 61:1485–1496

Coleman M, Pearce R, Hitchin E, Busfield F, Mansfield JW, Roberts IS (1990) Molecular cloning, expression and nucleotide sequence of the *rcsA* gene of *Erwinia amylovora*, encoding a positive regulator of capsule expression: evidence for a family of related capsule activator proteins. J Gen Microbiol 136:1799–1806

Coyier DL, Covey RP (1975) Tolerance of *Erwinia amylovora* to streptomycin sulfate in Oregon and Washington. Plant Dis Rep 59:849–852

Crosse JE, Bennett M, Garrett CME (1958) Fire blight of pear in England. Nature 182:1530

Datsenko KA, Wanner BL (2000) One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. Proc Natl Acad Sci USA 97:6640–6645

Day B, Dahlbeck D, Huang J, Chisholm ST, Li D, Staskawicz BJ (2005) Molecular basis for the RIN4 negative regulation of RPS2 disease resistance. Plant Cell 17:1292–1305

De Maayer P, Venter SN, Kamber T, Duffy B, Coutinho TA, Smits THM (2011) Comparative genomics of the type VI secretion systems of *Pantoea* and *Erwinia* species reveals the presence of putative effector islands that may be translocated by the VgrG and Hcp proteins. BMC Genom 12:576

DebRoy S, Thilmony R, Kwack YB, Nomura K, He SY (2004) A family of conserved bacterial effectors inhibits salicylic acid-mediated basal immunity and promotes disease necrosis in plants. Proc Natl Acad Sci USA 101:9927–9932

Deckers T (1996) Fire blight: the present state of its occurrence in Belgium and phytosanitory measures to control the problem. Parasitica 52:127–131

Degrave A, Fagard M, Perino S, Brisset MN, Gaubert S, Laroche S, Patrit O, Barny MA (2008) *Erwinia amylovora* type three-secreted proteins trigger cell death and defense responses in *Arabidopsis thaliana*. Mol Plant Microbe Interact 21:1076–1086

Dellagi A, Reis D, Vian B, Expert D (1999) Expression of the ferrioxamine receptor gene of *Erwinia amylovora* CFBP1430 during pathogenesis. Mol Plant Microbe Interact 12:463–466

Donat V, Biosca EG, Penalver J, Lopez MM (2007) Exploring diversity among Spanish strains of *Erwinia amylovora* and possible infection sources. J Appl Microbiol 103:1639–1649

Du Z, Geider K (2002) Characterization of an activator gene upstream of lsc, involved in levan synthesis of *Erwinia amylovora*. Physiol Mol Plant Pathol 60:9–17

Duffy B, Schärer HJ, Vogelsanger J, Schoch B, Holliger E (2005) Regulatory measures against *Erwinia amylovora* in Switzerland. EPPO Bull 35:239–244

Eastgate JA (2000) *Erwinia amylovora*: the molecular basis of fire blight disease. Mol Plant Pathol 1:325–329

Eastgate JA, Taylor N, Coleman MJ, Healy B, Thompson L, Robert IS (1995) Cloning, expression and characterization of the *lon* gene of *Erwinia amylovora*: Evidence for a heat shock response. J Bacteriol 177:932–937

EPPO (1992) Quarantine procedure: *Erwinia amylovora*—sampling and test methods. EPPO Bull 22:225–231

Foster GC, McGhee GC, Jones AL, Sundin GW (2004) Nucleotide sequences, genetic organization, and distribution of pEU30 and pEL60 from *Erwinia amylovora*. Appl Environ Microbiol 70:7539–7544

Gaudriault S, Malandrin L, Paulin JP, Barny MA (1997) DspA, an essential pathogenicity factor of *Erwinia amylovora* showing homology with AvrE of *Pseudomonas syringae*, is secreted via Hrp secretion pathway in a DspB-dependent way. Mol Microbiol 26:1057–1069

Gaudriault S, Brisset MN, Barny MA (1998) HrpW of *Erwinia amylovora*, a new Hrp-secreted protein. FEBS Lett 428:224–228

Gaudriault S, Paulin JP, Barny MA (2002) The DspB/F protein of *Erwinia amylovora* is a type III secretion

chaperone ensuring efficient intrabacterial production of the Hrp-secreted DspA/E pathogenicity factor. Mol Plant Pathol 3:313–320

Gehring I, Geider K (2012) Differentiation of *Erwinia amylovora* and *Erwinia pyrifoliae* strains with single nucleotide polymorphisms and by synthesis of dihydrophenylalanine. Curr Microbiol 65:73–84

Geider K, Auling G, Du Z, Jakovljevic V, Jock S, Voilksch B (2006) *Erwinia tasmaniensis* sp nov., a non-phytopathogenic bacterium from apple and pear trees. Int J Syst Evol Microbiol 56:2937–2943

Geier G, Geider K (1993) Characterization and influence on virulence of the levansucrase gene from the fire blight pathogen *Erwinia amylovora*. Physiol Mol Plant Pathol 42:387–404

Giorgi S, Scortichini M (2005) Molecular characterization of *Erwinia amylovora* strains from different host plants through RFLP analysis and sequencing of *hrpN* and *dspA/E* genes. Plant Pathol 54:789–798

Griffith CS, Sutton TB, Peterson PD (2003) Fire blight: the foundation of phytobacteriology. APS Press, St. Paul MN, p 144

Gross M, Geier G, Rudolph K, Geider K (1992) Levan and levansucrase synthesized by the fire blight pathogen *Erwinia amylovora*. Physiol Mol Plant Pathol 40:371–381

Hacker J, Blum-Oehler G, Muhldorfer I, Tschape H (1997) Pathogenicity islands of virulent bacteria: structure, function and impact on microbial evolution. Mol Microbiol 23:1089–1097

He SY, Nomura K, Whittam TS (2004) Type III protein secretion mechanism in mammalian and plant pathogens. Biochem Biophys Acta 1694:181–206

Hildebrand M, Aldridge P, Geider K (2006) Characterization of *hns* genes from *Erwinia amylovora*. Mol Genet Genomics 275:310–319

Horvath P, Barrangou R (2010) CRISPR/Cas, the immune system of bacteria and archaea. Science 327:167–170

Jensen PJ, Halbrendt N, Fazio G, Makalowska I, Altman N, Praul C, Maximova SN, Ngugi HK, Crassweller RM, Travis JW, McNellis TW (2012) Rootstock-regulated gene expression patterns associated with fire blight resistance in apple. BMC Genom 9:13

Jin Q, Hu W, Brown I, McGhee G, Hart P, Jones AL, He SY (2001) Visualization of secreted Hrp and Avr proteins along the Hrp pilus during type III secretion in *Erwinia amylovora* and *Pseudomonas syringae*. Mol Microbiol 40:1129–1139

Jin P, Wood MD, Wu Y, Xie Z, Katagiri F (2003) Cleavage of the *Pseudomonas syringae* type III effector AvrRpt2 requires a host factor(s) common among eukaryotes and is important for AvrRpt2 localization in the host cell. Plant Physiol 133:1072–1082

Jinek M, Chylinski K, Fonfara I, Hauer M, Doudna JA, Charpentier EA (2012) Programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. Science 337:816–821

Jock S, Geider K (2004) Molecular differentiation of *Erwinia amylovora* strains from North America and of two Asian pear pathogens by analyses of PFGE patterns and *hrpN* genes. Environ Microbiol 6:480–490

Jock S, Donat V, Lopez MM, Bazzi C, Geider K (2002) Following spread of fire blight in Western, Central and Southern Europe by molecular differentiation of *Erwinia amylovora* strains with PFGE analysis. Environ Microbiol 4:106–114

Kado CI (2000) Plant bacteriology. APS Press, St Paul

Kelm O, Kiecker C, Geider K, Bernhard F (1997) Interaction of the regulator proteins RcsA and RcsB with the promoter of the operon for amylovoran biosynthesis in *Erwinia amylovora*. Mol Gen Genet 256:72–83

Khan MA, Zhao YF, Korban SS (2012) Molecular mechanisms of pathogenesis and resistance to the bacterial pathogen *Erwinia amylovora*, causal agent of fire blight disease in Rosaceae. Plant Mol Biol Rep 30:247–260

Kim JF, Beer SV (1998) HrpW of *Erwinia amylovora*, a new harpin that contains a domain homologous to pectate lyases of a distinct class. J Bacterial 180:5203–5210

Kim JF, Beer SV (2001) Molecular basis of the Hrp pathogenicity of the fire blight pathogen *Erwinia amylovora*: a type III protein secretion system encoded in a pathogenicity island. Plant Pathol J 17:77–82

Kim JF, Wei ZM, Beer SV (1997) The *hrpA* and *hrpC* operons of *Erwinia amylovora* encode components of a type III pathway that secrets harpin. J Bacteriol 179:1690–1697

Kim WS, Gardan L, Rhim SL, Geider K (1999) *Erwinia pyrifoliae* sp. nov., a novel pathogen that affects Asian pear trees (*Pyrus pyrifolia* Nakai). Int J Syst Bacteriol 49:899–906

Kim WS, Hildebrand M, Jock S, Geider K (2001) Molecular comparison of pathogenic bacteria from pear trees in Japan and the fire blight pathogen *Erwinia amylovora*. Microbiology 147:2951–2959

Koczan JM, McGrath MJ, Zhao YF, Sundin GW (2009) Contribution of *Erwinia amylovora* exopolysaccharides amylovoran and levan to biofilm formation: implications in pathogenicity. Phytopathology 99:1237–1244

Kube M, Migdoll AM, Müller I, Kuhl H, Beck A, Reinhardt R, Geider K (2008) The genome of *Erwinia tasmaniensis* strain Et1/99, a non-pathogenic bacterium in the genus *Erwinia*. Environ Microbiol 10:2211–2222

Kube M, Migdoli AM, Gehring I, Heltmann K, Mayer Y, Kuhl H, Knaust F, Geider K, Reinhardt R (2010) Genome comparison of the epiphytic bacteria *Erwinia billingiae* and *E. tasmaniensis* with the pear pathogen *E. pyrifoliae*. BMC Genom 11:393

Kunkel BN, Bent AF, Dahlbeck D, Innes RW, Staskawicz BJ (1993) *RPS2*, an *Arabidopsis* disease resistance locus specifying recognition of *Pseudomonas syringae*

strains expressing the avirulence gene *avrRpt2*. Plant Cell 5:865875

Langille MG, Hsiao WW, Brinkman FS (2010) Detecting genomic islands using bioinformatics approaches. Nat Rev Microbiol 8:373–382

Langlotz C, Schollmeyer M, Coplin DL, Nimtz M, Geider K (2011) Biosynthesis of the repeating units of the exopolysaccharides amylovoran from *Erwinia amylovora* and stewartan from *Pantoea stewartii*. Physiol Mol Plant Pathol 75:163–169

Lee SA, Ngugi HK, Halbrendt NO, O'Keefe G, Lehman B, Travis JW, Sinn JP, McNellis TW (2010) Virulence characteristics accounting for fire blight disease severity in apple trees and seedlings. Phytopathology 100:539–550

Lelliott RA, Dickey RS (1984) Genus VII, *Erwinia*. In: Krieg NR (ed) Bergey's manual of systematic bacteriology, vol 1. Williams and Wilkins, Baltimore, pp 469–476

Li WT, Ancona V, Zhao YF (2014) Co-regulation of polysaccharide production, motility, and expression of type III secretion genes by EnvZ/OmpR and GrrS/GrrA systems in *Erwinia amylovora*. Mol Genet Genom 1:63–75

Lindgren PB, Peet RC, Panopoulos NJ (1986) Gene cluster of *Pseudomonas syringae pv. phaseolicola* controls pathogenicity of bean plants and hypersensitivity on nonhost plants. J Bacteriol 168:512–522

Liu R, Ochman H (2007) Stepwise formation of the bacterial flagellar system. Proc Natl Acad Sci USA 104:7116–7121

Llop P, Donat V, Rodríguez M, Cabrefiga J, Ruz L, Jalomo JL, Montesinos E, López MM (2006) An indigenous virulent strain of *Erwinia amylovora* lacking the ubiquitous plasmid pEA29. Phytopathology 96:900–907

Llop P, Cabrefiga J, Smits THM, Dreo T, Barbe S, Pulawska J, Bultreys A, Blom J, Duffy B, Montesinos E, López MM (2011) *Erwinia amylovora* novel plasmid pEI70: complete sequence, biogeography, and role in aggressiveness in the fire blight phytopathogen. PLoS One 6:e28651

Llop P, Barbé S, López MM (2012) Functions and origin of plasmids in *Erwinia* species that are pathogenic to or epiphytically associated with pome fruit trees. Trees Struct Funct 26:31–46

Longstroth M (2000) Fire blight epidemic strikes SW Michigan. Fruit Grow News July, pp 17–18

López MM, Roselló M, Llop P, Ferrer S, Christen R, Gardan L (2011) *Erwinia piriflorinigrans* sp. nov., a novel pathogen that causes necrosis of pear blossoms. Int J Syst Evol Microbiol 61:561–567

Mackey D, Holt BF, Wiig A, Dangl JL (2002) RIN4 interacts with *Pseudomonas syringae* type III effector molecules and is required for RPM1-mediated resistance in *Arabidopsis*. Cell 108:743–754

Mackey D, Belkhadir Y, Alonso JM, Ecker JR, Dangl JL (2003) *Arabidopsis* RIN4 is a target of the type III virulence effector AvrRpt2 and modulates RPS2-mediated resistance. Cell 112:379389

Malnoy M, Martens S, Norelli JL, Barny MA, Sundin GW, Smits TH, Duffy B (2012) Fire blight: applied genomic insights of the pathogen and host. Annu Rev Phytopathol 50:475–494

Mann RA, Blom J, Buhlmann A, Plummer KM, Beer SV, Luck JE, Goesmann A, Frey JE, Rodoni BC, Duffy B, Smits THM (2012) Comparative analysis of the Hrp pathogenicity island of *Rubus*- and Spiraeoideae infecting *Erwinia amylovora* strains identifies the IT region as a remnant of an integrative conjugative element. Gene 504:6–12

Mann RA, Smits THM, Buhlmann A, Blom J, Goesmann A, Frey JE, Plummer KM, Beer SV, Luck JE, Duffy B, Rodoni BC (2013) Comparative genomics of 12 strains of *Erwinia amylovora* identifies a pan-genome with a large conserved core. PLoS One 8:e55644

Mansfield J, Genin S, Magori S, Citovsky V, Sriariyanum M, Ronald P, Dow M, Verdier V, Machado MA, Toth I, Salmond G, Foster GD (2012) Top 10 plant pathogenic bacteria in molecular plant pathology. Mol Plant Pathol 13:614–629

Marraffini LA, Sontheimer EJ (2010) Self versus non-self discrimination during CRISPR RNA-directed immunity. Nature 463:568–571

Matsuura T, Mizuno A, Tsukamoto T, Shimizu Y, Saito N, Sato S, Kikuchi S, Uzuki T, Azegami K, Sawada H (2012) *Erwinia uzenensis* sp. nov., a novel pathogen that affects European pear trees (*Pyrus communis* L.). Int J Syst Evol Microbiol 62:1799–1803

McGhee GC, Jones AL (2000) Complete nucleotide sequence of ubiquitous plasmid pEA29 from *Erwinia amylovora* strain Ea88: gene organization and intraspecies variation. Appl Environ Microbiol 66: 4897–4907

McGhee GC, Sundin GW (2008) Thiamin biosynthesis and its influence on exopolysaccharide production: A new component of the virulence identified on *Erwinia amylovora* plasmid pEA29. Acta Hortic 793:271–277

McGhee GC, Sundin GW (2012) *Erwinia amylovora* CRISPR elements provide new tools for evaluating strain diversity and for microbial source tracking. PLoS One 7:e41706

McGhee GC, Schnabel EL, Maxson-Stein K, Jones B, Stromberg VK, Lacy GH, Jones AL (2002) Relatedness of chromosomal and plasmid DNAs of *Erwinia pyrifoliae* and *Erwinia amylovora*. Appl Environ Microbiol 68:6182–6192

McGhee GC, Guasco J, Bellomo LM, Blumer-Schuette SE, Shane WW, Irish-Brown A, Sundin GW (2011) Genetic analysis of streptomycin-resistant (Sm(R)) strains of *Erwinia amylovora* suggests that dissemination of two genotypes is responsible for the current distribution of Sm^R *E. amylovora* in Michigan. Phytopathology 101:182–191

McManus PS, Jones AL (1995) Genetic fingerprinting of *Erwinia amylovora* strains isolated from tree-fruit crops and *Rubus* spp. Phytopathology 85:1547–1553

McManus PS, Stockwell VO, Sundin GW, Jones AL (2002) Antibiotic use in plant agriculture. Annu Rev Phytopathol 40:443–465

McNally RR, Toth IK, Cock PJA, Pritchard L, Hedley PE, Zhao YF, Sundin GW (2012) Genetic characterization of the HrpL regulon of the fire blight pathogen *Erwinia amylovora* reveals novel virulence factors. Mol Plant Pathol 13:160–173

Medini D, Donati C, Tettelin H, Masignani V, Rappuoli R (2005) The microbial pan-genome. Curr Opin Genet Dev 15:589–594

Meng XD, Bonasera JM, Kim JF, Nissinen RM, Beer SV (2006) Apple proteins that interact with DspA/E, a pathogenicity effector of *Erwinia amylovora*, the fire blight pathogen. Mol Plant Microbe Interact 19:53–61

Mergaert J, Hauben L, Cnockaert MC, Swings J (1999) Reclassification of non-pigmented *Erwinia herbicola* strains from trees as *Erwinia billingiae* sp. nov. Int J Syst Bacteriol 49:377–383

Mizuno A, Sato S, Kawai A, Nishiyama K (2000) Taxonomic position of the causal pathogen of bacterial shoot blight of pear. J Gen Plant Pathol 66:48–58

Mizuno A, Tsukamoto T, Shimizu Y, Ooya H, Matsuura T, Saito N, Sato S, Kikuchi S, Uzuki T, Azegami K (2010) Occurrence of bacterial black shoot disease of European pear in Yamagata Prefecture. J Gen Plant Pathol 76:43–51

Momol MT, Momol EA, Lamboy WF, Norelli JL, Beer SV, Aldwinckle HS (1997) Characterization of *Erwinia amylovora* strains using random amplified polymorphic DNA fragments (RAPDs). J Appl Microbiol 82:389–398

Nakka S, Qi M, Zhao YF (2010a) The PmrAB system in *Erwinia amylovora* renders the pathogen more susceptible to polymyxin B and more resistance to excess iron. Res Microbiol 161:153–157

Nakka S, Qi M, Zhao YF (2010b) The *Erwinia amylovora* PhoPQ system is involved in resistance to antimicrobial peptide and suppresses gene expression of two novel type III secretion systems. Microbiol Res 165:665–673

Nimtz M, Mort A, Domke T, Wray V, Zhang Y, Qiu F, Coplin D, Geider K (1996) Structure of amylovoran, the capsular exopolysaccharide from the fire blight pathogen *Erwinia amylovora*. Carbohydr Res 287:59–76

Nissinen RM, Ytterberg AJ, Bogdanove AJ, van Wijk K, Beer SV (2007) Analyses of the secretomes of *Erwinia amylovora* and selected *hrp* mutants reveal novel type III secreted proteins and an effect of HrpJ on extracellular harpin levels. Mol Plant Pathol 8:55–67

Norelli JL, Aldwinckle HS, Beer SV (1984) Differential host × pathogen interactions among cultivars of apple and strains of *Erwinia amylovora*. Phytopathology 74:136–139

Norelli JL, Aldwinckle HS, Beer SV (1986) Differential susceptibility of *Malus* spp. cultivars robusta 5, Novole, and Ottawa 523 to *Erwinia amylovora*. Plant Dis 70:1017–1019

Norelli JL, Jones AL, Aldwinckle HS (2003) Fire blight management in the twenty-first century: using new technologies that enhance host resistance in apple. Plant Dis 87:756–765

Oh CS, Beer SV (2005) Molecular genetics of *Erwinia amylovora* involved in the development of fire blight. FEMS Microbiol Lett 253:185–192

Oh CS, Beer SV (2007) AtHIPM, an ortholog of the apple HrpN-interacting protein, is a negative regulator of plant growth and mediates the growth-enhancing effect of HrpN in *Arabidopsis*. Plant Physiol 145:426–436

Oh CS, Kim JF, Beer SV (2005) The Hrp pathogenicity island of *Erwinia amylovora* and identification of three novel genes required for systemic infection. Mol Plant Pathol 6:125–138

Oh CS, Martin GB, Beer SV (2007) DspA/E, a type III effector of *Erwinia amylovora*, is required for early rapid growth in *Nicotiana benthamiana* and causes NbSGT1-dependent cell death. Mol Plant Pathol 8:255–265

Oh CS, Carpenter CD, Hayes ML, Beer SV (2010) Secretion and translocation signals and DspB/F-binding domains in the type III effector DspA/E of *Erwinia amylovora*. Microbiology 156:1211–1220

Ordax M, Marco-Noales E, López MM, Biosca EG (2010) Exopolysaccharides favor the survival of *Erwinia amylovora* under copper stress through different strategies. Res Microbiol 161:549–555

Palacio-Bielsa A, Rosello M, Llop P, Lopez MM (2012) *Erwinia* spp. from pome fruit trees: similarities and differences among pathogenic and non-pathogenic species. Trees 26:13–29

Park DH, Thapa SP, Choi BS, Kim WS, Hur JH, Cho JM, Lim J, Choi I, Lim CK (2011) Complete genome sequence of Japanese *Erwinia* strain Ejp617, a bacterial shoot blight pathogen of pear. J Bacteriol 193:586–587

Perino C, Gaudriault S, Vian B, Barny MA (1999) Visualization of harpin secretion in planta during infection of apple seedlings by *Erwinia amylovora*. Cell Microbiol 1:131–141

Pierce NB (1902) Pear blight in California. Science 16:193–194

Powney R, Smits THM, Sawbridge T, Frey B, Blom J, Frey JE, Plummer KM, Beer SV, Luck J, Duffy B, Rodoni B (2011) Genome sequence of an *Erwinia amylovora* strain with pathogenicity restricted to *Rubus* plants. J Bacteriol 193:785–786

Pristovsek P, Sengupta K, Lohr F, Schafer B, von Trebra MW, Ruterjans H, Bernhard F (2003) Structural analysis of the DNA-binding domain of the *Erwinia amylovora* RcsB protein and its interaction with the RcsAB box. J Biol Chem 278:17752–17759

Pronk LM, Sanderson KE (2001) Intervening sequences in *rrl* genes and fragmentation of 23S rRNA in genera of the family *Enterobacteriaceae*. J Bacteriol 183:5782–5787

Qi M, Sun F, Caetano-Anolles G, Zhao YF (2010) Comparative genomic and phylogenetic analyses reveal the evolution of core two-component signal

transduction systems in enterobacteria. J Mol Evol 70:167–180

Reboutier D, Frankart C, Briand J, Biligui B, Laroche S, Rona JP, Barny MA, Bouteau F (2007) The $HrpN_{ea}$ harpin from *Erwinia amylovora* triggers differential responses on the nonhost *Arabidopsis thaliana* cells and on the host apple cells. Mol Plant Microbe Interact 20:94–100

Records AR (2011) The type VI secretion system: a multipurpose delivery system with a phage-like machinery. Mol Plant Microbe Interact 24:751–757

Rezzonico F, Smits THM, Duffy B (2011) Diversity, evolution, and functionality of clustered regularly interspaced short palindromic repeat (CRISPR) regions in the fire blight pathogen *Erwinia amylovora*. Appl Environ Microbiol 77:3819–3829

Rezzonico F, Braun-Kiewnick A, Mann RA, Rodoni B, Goesmann A, Duffy B, Smits THM (2012) Lipopolysaccharide biosynthesis genes discriminate between *Rubus*- and Spiraeoideae-infective genotypes of *Erwinia amylovora*. Mol Plant Pathol 13:975–984

Rhim S-L, Volksch B, Gardan L, Paulin JP, Langlotz C, Kim WS, Geider K (1999) *Erwinia pyrifoliae*, an *Erwinia* species different from *Erwinia amylovora*, causes a necrotic disease of Asian pear trees. Plant Pathol 48:514–520

Rico A, Ortiz-Barredo A, Ritter E, Murillo J (2004) Genetic characterization of *Erwinia amylovora* strains by amplified fragment length polymorphism. J Appl Microbiol 96:302–310

Rodoni BC, Merriman PR, McKirdy SJ, Wittwer G (2006) Costs associated with fire blight incursion management and predicted costs of future incursions. Acta Hortic 704:55–61

Rosello M, Penalver J, Llop P, Gorris MT, Chartier R, Garcıa F, Monton C, Cambra M, Lopez MM (2006) Identification of an *Erwinia* sp. different from *Erwinia amylovora* and responsible for necrosis on pear blossoms. Can J Plant Pathol 28:30–41

Russo NL, Burr TJ, Breth DI, Aldwinckle HS (2008) Isolation of streptomycin-resistant isolates of *Erwinia amylovora* in New York. Plant Dis 92:714–718

Sarkar SF, Guttman DS (2004) Evolution of the core genome of *Pseudomonas syringae*, a highly clonal, endemic plant pathogen. Appl Environ Microbiol 70:999–2012

Sarowar S, Guerra R, Wang D, Zheng DM, Korban SS, Zhao YF (2011a) Transcriptome analysis of apple blossom after challenging with the fire blight pathogen *Erwinia amylovoran* wild type and mutant strains. Acta Hortic 896:245–251

Sarowar S, Zhao YF, Soria-Guerra RE, Ali S, Zheng D, Wang DP, Korban SS (2011b) Expression profiles of differentially regulated genes during the early stages of apple flower infection with *Erwinia amylovora*. J Exp Bot 62:4851–4861

Schroth MN, Thomson SV, Hildebrand DC (1974) Epidemiology and control of fire blight. Annu Rev Phytopathol 12:389–412

Sebaihia M, Bocsanczy AM, Biehl BS, Quail MA, Perna NT, Glasner JD, DeClerck GA, Cartinhour S, Schneider DJ, Bentley SD, Parkhill J, Beer SV (2010) Complete genome sequence of the plant pathogen *Erwinia amylovora* strain ATCC 49946. J Bacteriol 192:2020–2021

Seth-Smith H, Croucher NJ (2009) Genome watch: breaking the ICE. Nat Rev Microbiol 7:328–329

Shrestha R, Koo JH, Park DH, Hwang I, Hur JH, Lim CK (2003) *Erwinia pyrifoliae*, a causal endemic pathogen of shoot blight of Asian pear tree in Korea. Plant Pathol J 19:294–300

Sinn JP, Oh CS, Jensen PJ, Carpenter SC, Beer SV, McNellis TW (2008) The C-terminal half of the HrpN virulence protein of the fire blight pathogen *Erwinia amylovora* is essential for its secretion and for its virulence and avirulence activities. Mol Plant Microbe Interact 21:1387–1397

Smits THM, Duffy B (2011) Genomics of iron acquisition in the plant pathogen *Erwinia amylovora*: insights in the biosynthetic pathway of the siderophore desferrioxamine E. Arch Microbiol 193: 693–699

Smits THM, Jaenicke S, Rezzonico F, Kamber T, Goesmann A, Frey JE, Duffy B (2010a) Complete genome sequence of the fire blight pathogen *Erwinia pyrifoliae* DSM 12163(T) and comparative genomic insights into plant pathogenicity. BMC Genom 11:2

Smits THM, Rezzonico F, Kamber T, Blom J, Goesmann A, Frey JE, Duffy B (2010b) Complete genome sequence of the fire blight pathogen *Erwinia amylovora* CFBP 1430 and comparison to other *Erwinia* spp. Mol Plant Microbe Interact 23:384–393

Smits THM, Rezzonico F, Duffy B (2011) Evolutionary insights from *Erwinia amylovora* genomics. J Biotechnol 155:34–39

Smits THM, Rezzonico F, López MM, Blom J, Goesmann A, Frey JE, Duffy B (2013) Phylogenetic position and virulence apparatus of the pear flower necrosis pathogen *Erwinia piriflorinigrans* CFBP 5888(T) as assessed by comparative genomics. Syst Appl Microbiol 36:449–456

Starr MP, Chatterjee AK (1972) The genus *Erwinia*: Enterobacteria pathogenic to plants and animals. Annul Rev Microbiol 26:389–426

Starr MP, Cardona C, Folsom D (1951) Bacterial fire blight of raspberry. Phytopathology 41:915–919

Steinberger EM, Beer SV (1988) Creation and complementation of pathogenicity mutants of *Erwinia amylovora*. Mol Plant Microbe Interact 1:135–144

Tanii A, Tamura O, Ozaki M (1981) Causal pathogen of fire blightlike symptoms of pear. Ann Phytopath Soc Jpn 47:102

Thapa SP, Park DH, Kim WS, Choi BS, Lim JS, Choi IY, Hur JH, Lim CK (2013) Comparative genomics of Japanese *Erwinia pyrifoliae* strain Ejp617 with closely related erwinias. Genome 56:83–90

Thoelen M, Noben JP, Robben J, Valcke R, Deckers T (2008) Comparative proteome analysis of four

Erwinia amylovora strains with different pathogenicity. Acta Hort 793:183–185

Thomson SV (1986) The role of the stigma in fire blight infections. Phytopathology 76:476–482

Toth IK, Pritchard L, Birch PRJ (2006) Comparative genomics reveals what makes an enterobacterial plant pathogen. Ann Rev Phytopathol 44:305–306

Triplett L, Zhao YF, Sundin GW (2006) Genetic differences among blight-causing *Erwinia* species with differing host specificities identified by suppression subtractive hybridization. Appl Environ Microbiol 72:7359–7364

Triplett LR, Melotto M, Sundin GW (2009) Functional analysis of the N terminus of the *Erwinia amylovora* secreted effector DspA/E reveals features required for secretion, translocation, and binding to the chaperone DspB/F. Mol Plant Microbe Interact 22:1282–1292

Triplett LR, Wedemeyer WJ, Sundin GW (2010) Homology-based modeling of the *Erwinia amylovora* type III secretion chaperone DspF used to identify amino acids required for virulence and interaction with the effector DspE. Res Microbiol 161:613–618

van der Zwet T (1995) First report of *Erwinia amylovora* on new host species in the genus *Sorbus*. Plant Dis 79:424

van der Zwet T (2004) Present worldwide distribution of fire blight and closely related diseases. Acta Hort 704:35

van der Zwet T, Keil HL (1979) Fire blight: a bacterial disease of rosaceous species. United States Department of Agriculture Handbook 510

van der Zwet T, Wells JM (1993) Application of fatty acid class analysis for the detection and identification of *Erwinia amylovora*. Acta Hortc 388:233

van der Zwet T, Orolaza-Halbrendt N, Zeller W (2012) Fire blight: history, biology, and management. APS Press, St Paul, MN 420 pp

Vanneste JL (2000) Fire blight: the disease and its causative agent, *Erwinia amylovora*. CABI Publishing, New York

Vanneste JL, Paulin JP, Expert D (1990) Bacteriophage MU as a genetic tool to study *Erwinia amylovora* pathogenicity and hypersensitive reaction on tobacco. J Bacteriol 172:932–941

Vanneste JL, Lex S, Vermeulen M, Berger F (2002) Isolation of *Erwinia amylovora* from blighted plums (*Prunus domestica*) and potato roses (*Rosa rugosa*). Acta Hortc 590:89–94

Velasco R, Zharkikh A, Affourtit J, Dhingra A, Cestaro A, Kalyanaraman A, Fontana P, Bhatnagar SK, Troggio M, Pruss D, Salvi S, Pindo M, Baldi P, Castelletti S, Cavaiuolo M, Coppola G, Costa F, Cova V, Dal Ri A, Goremykin V, Komjanc M, Longhi S, Magnago P, Malacarne G, Malnoy M, Micheletti D, Moretto M, Perazzolli M, Si-Ammour A, Vezzulli S, Zini E, Eldredge G, Fitzgerald LM, Gutin N, Lanchbury J, Macalma T, Mitchell JT, Reid J, Wardell B, Kodira C, Chen Z, Desany B, Niazi F, Palmer M, Koepke T, Jiwan D, Schaeffer S, Krishnan V, Wu C, Chu VT, King ST, Vick J, Tao Q, Mraz A, Stormo A, Stormo K, Bogden R, Ederle D, Stella A, Vecchietti A, Kater MM, Masiero S, Lasserre P, Lespinasse Y, Allan AC, Bus V, Chagné D, Crowhurst RN, Gleave AP, Lavezzo E, Fawcett JA, Proost S, Rouzé P, Sterck L, Toppo S, Lazzari B, Hellens RP, Durel CE, Gutin A, Bumgarner RE, Gardiner SE, Skolnick M, Egholm M, van de Peer Y, Salamini F, Viola R (2010) The genome of the domesticated apple (*Malus* × *domestica* Borkh.). Nat Genet 42:833–839

Venisse JS, Gullner G, Brisset MN (2001) Evidence for the involvement of an oxidative stress in the initiation of infection of pear by *Erwinia amylovora*. Plant Physiol 125:2164–2172

Venisse JS, Barny MA, Paulin JP, Brisset MN (2003) Involvement of three pathogenicity factors of *Erwinia amylovora* in the oxidative stress associated with compatible interaction in pear. FEBS Lett 537:198–202

Vogt I, Wöhner T, Richter K, Flachowsky H, Sundin GW, Wensing A, Savory EA, Geider K, Day B, Hanke MV, Peil A (2013) Gene-for-gene relationship in the host-pathogen system *Malus* × *robusta* 5-*Erwinia amylovora*. New Phytol 197:1262–1275

Wang L, Beer SV (2006) Application of signature-tagged mutagenesis to the study of virulence of *Erwinia amylovora*. FEMS Microbiol Lett 265:164–171

Wang DP, Korban SS, Zhao YF (2009) The Rcs phosphorelay system is essential for pathogenicity in *Erwinia amylovora*. Mol Plant Pathol 10:277–290

Wang DP, Korban SS, Zhao YF (2010a) Molecular signature of differential virulence in natural isolates of *Erwinia amylovora*. Phytopathology 100:192–198

Wang Q, Zhang Y, Yang C, Xiong H, Lin Y, Yao J, Li H, Xie L, Zhao W, Yao Y, Ning ZB, Zeng R, Xiong Y, Guan KL, Zhao S, Zhao GP (2010b) Acetylation of metabolic enzymes coordinates carbon source utilization and metabolic flux. Science 327:1004–1007

Wang DP, Calla B, Vimolmangkang S, Wu X, Korban SS, Huber SC, Clough SJ, Zhao YF (2011a) The orphan gene *ybjN* conveys pleiotropic effects on multicellular behavior and survival of *Escherichia coli*. PLoS One 6:e25293

Wang DP, Korban SS, Pusey L, Zhao YF (2011b) Characterization of the RcsC sensor kinase from *Erwinia amylovora* and other enterobacteria. Phytopathology 101:710–717

Wang DP, Korban SS, Sundin GW, Clough S, Toth I, Zhao YF (2011c) Regulatory genes and environmental regulation of amylovoran biosynthesis in *Erwinia amylovora*. Acta Hort 896:195–202

Wang DP, Qi MS, Calla B, Korban SS, Clough SJ, Cock P, Sundin GW, Toth I, Zhao YF (2012a) Genome-wide identification of genes regulated by the Rcs phosphorelay system in *Erwinia amylovora*. Mol Plant Microbe Interact 25:6–17

Wang DP, Korban SS, Pusey L, Zhao YF (2012b) AmyR is a novel negative regulator of amylovoran production in *Erwinia amylovora*. PLoS One 7:e45038

Wang X, Yang F, von Bodman SB (2012c) 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

Wehland M, Bernhard F (2000) The RcsAB box. Characterization of a new operator essential for the regulation of exopolysaccharide biosynthesis in enteric bacteria. J Bio Chem 275:7013–7020

Wehland M, Kiecker C, Coplin DL, Kelm O, Saenger W, Bernhard F (1999) Identification of an RcsA/RcsB recognition motif in the promoters of exopolysaccharide biosynthetic operons from *Erwinia amylovora* and *Pantoea stewartii* subspecies *stewartii*. J Biol Chem 274:3300–3307

Wei ZM, Beer SV (1993) Harpin of *Erwinia amylovora* functions in secretions of harpin and is a member of a new protein family. J Bacteriol 175:7958–7967

Wei Z, Beer SV (1995) HrpL activates *Erwinia amylovora hrp* gene transcription and is a member of the ECF subfamily of sigma factors. J Bacteriol 177:6201–6210

Wei ZM, Laby RJ, Zumoff CH, Bauer DW, He SY, Collmer A, Beer SV (1992) Harpin, elicitor of the hypersensitive response produced by the plant pathogen *Erwinia amylovora*. Science 257:85–88

Wei Z, Kim JF, Beer SV (2000) Regulation of *hrp* genes and type III protein secretion in *Erwinia amylovora* by HrpX/HrpY, a novel two-component system, and HrpS. Mol Plant Microbe Interact 11:1251–1262

Wiedenheft B, Sternberg SH, Doudna JA (2012) RNA-guided genetic silencing systems in bacteria and archaea. Nature 482:331–338

Wu J, Wang Z, Shi Z, Zhang S, Ming R, Zhu S, Khan MA, Tao S, Korban SS, Wang H, Chen NJ, Nishio T, Xu X, Cong L, Qi K, Huang X, Wang Y, Zhao X, Wu J, Deng C, Gou C, Zhou W, Yin H, Qin G, Sha Y, Tao Y, Chen H, Yang Y, Song Y, Zhan D, Wang J, Li L, Dai M, Gu C, Wang Y, Shi D, Wang X, Zhang H, Zeng L, Zheng D, Wang C, Chen M, Wang G, Xie L, Sovero V, Sha S, Huang W, Zhang S, Zhang M, Sun J, Xu L, Li Y, Liu X, Li Q, Shen J, Wang J, Paull RE, Bennetzen JL, Wang J, Zhang S (2012) The genome of the pear (*Pyrus bretschneideri* Rehd.). Genome Res 23:396–408

Wu X, Vellaichamy A, Wang D, Zamdborg L, Kelleher NL, Huber SC, Zhao YF (2013) Differential lysine acetylation profiles of *Erwinia amylovora* strains revealed by proteomics. J Proteomics 79:60–71

Yang F, Korban SS, Pusey L, Elofsson M, Sundin GW, Zhao YF (2014) Small molecule inhibitors suppress expression of both type III secretion and amylovoran biosynthesis genes in *Erwinia amylovora*. Mol Plant Pathol 15:44–57

Zeng Q, McNally RR, Sundin GW (2013) Global small RNA chaperone Hfq and regulated small RNAs control virulence in the fire blight pathogen *Erwinia amylovora*. J Bacteriol 195:1706–1717

Zhang Y, Geider K (1997) Differentiation of *Erwinia amylovora* strains by pulsed-field gel electrophoresis. Appl Environ Microbiol 63:4421–4426

Zhang Y, Geider K (1999) Molecular analysis of the *rlsA* gene regulating levan production by the fire blight pathogen *Erwinia amylovora*. Phys Mol Plant Pathol 54:187–201

Zhang Y, Bak DD, Heidi H, Geider K (1999) Molecular characterization of a protease secreted by *Erwinia amylovora*. J Mol Biol 289:1239–1251

Zhang JM, Sprung R, Pei JM, Tan XH, Kim S, Zhu H, Liu CF, Grishin NV, Zhao YM (2009) Lysine acetylation is a highly abundant and evolutionarily conserved modification in *Escherichia coli*. Mol Cell Prot 8:215–225

Zhao YF, Qi M (2011) Comparative genomics of *Erwinia amylovora* and related *Erwinia* Species-What do we learn? Genes 2:627–639

Zhao YF, Blumer SE, Sundin GW (2005) Identification of *Erwinia amylovora* genes induced during infection of immature pear tissue. J Bacteriol 187:8088–8103

Zhao YF, He SY, Sundin GW (2006) The *Erwinia amylovora avrRpt2*$_{EA}$ gene contributes to virulence on pear and AvrRpt2$_{EA}$ is recognized by *Arabidopsis* RPS2 when expressed in *Pseudomonas syringae*. Mol Plant Microbe Interact 19:644–654

Zhao YF, Sundin GW, Wang D (2009a) Construction and analysis of pathogenicity island deletion mutants of *Erwinia amylovora*. Can J Microbiol 55:457–464

Zhao YF, Wang D, Nakka S, Sundin GW, Korban SS (2009b) Systems level analysis of two-component signal transduction systems in *Erwinia amylovora*: role in virulence, regulation of amylovoran biosynthesis and swarming motility. BMC Genom 10:245

Zhao YF, Qi M, Wang D (2011) Evolution and function of flagellar and non-flagellar type III secretion systems in *Erwinia amylovora*. Acta Hort 896: 177–184

Genomics of Plant-Associated Bacteria: The Soft Rot *Enterobacteriaceae*

2

Amy O. Charkowski, Jenna Lind, and Isael Rubio-Salazar

2.1 Introduction

2.1.1 Soft Rot *Enterobacteriaceae* Taxonomy and Phylogeny

The two genera in the soft rot *Enterobacteriaceae* (SRE), *Pectobacterium* and *Dickeya*, have had many names over the past 110 years, with each change reflecting growth and precision gained in knowledge about phytopathogens. Like many bacterial genera, the SRE were originally described as *Bacillus* (Jones 1900, 1901), but they were soon moved to the new genus, *Erwinia*, the name of which was derived from one of the founders of phytobacteriology, Erwin Frink Smith (Winslow et al. 1917). For many decades, the genus *Erwinia* was used for all plant pathogenic *Enterobacteriaceae* despite clear evidence that these species were not monophyletic. As early as 1945, taxonomists proposed to move *Erwinia carotovora* and other SRE into a new genus, *Pectobacterium*, to reflect differences in its physiology (Waldee 1945), but this new name was not generally accepted until more extensive phylogenetic work was completed, almost 100 years after this group of plant pathogens was first described (Gardan et al. 2003; Hauben et al. 1998; Kwon et al. 1997) (Table 2.1). Soon afterward, another SRE species, *Erwinia chrysanthemi*, was placed into the new genus *Dickeya* and at the same time divided into multiple species (Brady et al. 2012; Samson et al. 2005).

The two genera that comprise the SRE may be phylogenetically separated by the genus *Brennaria,* which contains non-pectolytic strains that infect mainly trees (Brady et al. 2012). Although *Brennaria* genome sequences are available, the genomic differences among the *Pectobacterium*, *Dickeya*, and *Brennaria* remain little explored. Little is known about the biology of *Brennaria,* and since this genus mainly infects landscape trees, it has not been developed as a model system for the study of phytopathogens.

Although *Pectobacterium* and *Dickeya* are in different genera, they are often discussed together because they produce high levels of plant-cell-wall-degrading enzymes (PCWDE) and cause similar wilt and decay diseases on a wide range of monocot and dicot plant species (Ma et al. 2007). It is the activity of the PCWDE pectate lyase that is used to isolate these genera from plant, water, and soil samples; researchers typically isolate SRE on pectate-containing media, on which the SRE form pitting colonies (Hyman et al. 2001). Since this method can only isolate strains that encode pectate-inducible pectinases, if non-pectolytic members of these

A. O. Charkowski (✉) · J. Lind · I. Rubio-Salazar
Department of Plant Pathology,
University of Wisconsin-Madison,
1630 Linden Dr., Madison, WI 53706, USA
e-mail: acharkowski@wisc.edu

J. Lind
e-mail: lind@wisc.edu

I. Rubio-Salazar
e-mail: rubiosalazar@wisc.edu

D. C. Gross et al. (eds.), *Genomics of Plant-Associated Bacteria*,
DOI: 10.1007/978-3-642-55378-3_2,

Table 2.1 *Pectobacterium* and *Dickeya* species

Species	Known host range	References
Dickeya chrysanthemi	Broad host range[a]	Samson et al. (2005)
Dickeya paradisiaca	Banana, maize	Samson et al. (2005)
Dickeya dadantii	Broad host range	Samson et al. (2005), Brady et al. (2012)
Ssp. dadantii		
Ssp. dieffenbachiae		
Dickeya dianthocola	Broad host range	Samson et al. (2005)
Dickeya zeae	Broad host range	Samson et al. (2005)
Pectobacterium aroiderium	Ornamental monocots[a], potato	Nabhan et al. (2013), Yishay et al. (2008)
Pectobacterium atrosepticum	Potato, pepper, sunflower	Gardan et al. (2003, Hauben et al. (1998)
Pectobacterium betavasculorum	Sugar beet	Gardan et al. (2003), Hauben et al. (1998)
Pectobacterium cacticida	Cacti[a]	Alcorn et al. (1991), Hauben et al. (1998)
Pectobacterium carotovorum		Duarte et al. (2004), Hauben et al. (1998)
ssp.carotovorum	Broad host range among dicots	
ssp. brasiliense	Potato	
ssp. odoriferum	Potato, chicory	
Pectobacterium wasabiae	Wasabi, potato	Gardan et al. (2003), Hauben et al. (1998)

[a] Four or more known hosts

genera exist, they will remain uncaptured by standard isolation methods.

Like other *Enterobacteriaceae*, the SRE are gram-negative rod-shaped non-spore forming facultative anaerobes with peritrichous flagella. They are easily grown in culture, and many mutagenesis methods and plasmids developed for the model bacterial species *Escherichia coli* also work with the SRE. Numerous SRE genome sequences are available, and a genial worldwide research community has collaborated on SRE studies for many years (Bell et al. 2004; Glasner et al. 2008, 2011; Nykyri et al. 2012; Pritchard et al. 2013b). Together, these attributes have allowed the SRE to become an important model for phytopathogenesis and a common teaching tool in plant disease laboratory courses.

2.1.2 Agricultural Relevance of *Pectobacterium* and *Dickeya* Genomics

The SRE are widespread and are found in all agricultural regions on numerous fruit, vegetable, and ornamental crops (Ma et al. 2007). SRE species differ in economic significance in different agricultural regions and on different crops, but the genomic contributions to host range and geographical distribution remain unknown.

Most of the SRE research to date has focused on its impact on potato, which is the most important dicot food crop worldwide. The SRE can cause disease on potato at planting, where they decay seed pieces, during the growing season, where they may kill stems and tubers, and in storage, where they can rot large volumes of potatoes in warehouses or during shipping. On tubers, the disease often starts either from the stem end, where the pathogen often appears to have entered the tuber through the stolon, or from the bud end, where it may have entered due to incomplete periderm formation (Fig. 2.1). The bacteria thrive in these low-oxygen environments, and often, there is a vase-like portion of undecayed tissue surrounding the diseased portion, as though the pathogen was creating an anaerobic chamber inside the tuber. Similar core rots are also seen in carrots infected with SRE (Fig. 2.1). The SRE can infect foliage from infected tubers or from wounds on the stem, and once inside stems, it causes a brown or black decay that can kill the plant (Fig. 2.1).

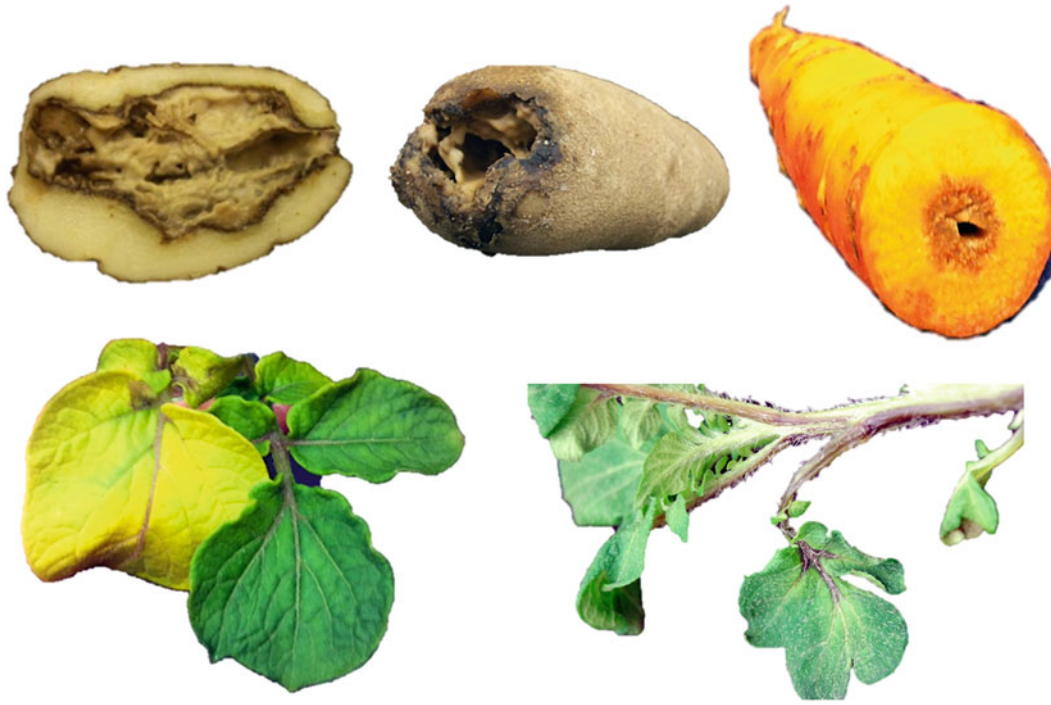

Fig. 2.1 Typical symptoms of SRE on *potato* and *carrot*. *Top Row* Soft rot in *potato* and *carrot* is often only found in the center of the *tuber* or *root*, highlighting the pathogenesis of these bacteria under anaerobic conditions. *Bottom Row* Chlorotic symptoms are characteristic of *Pectobacterium*, but not *Dickeya*. The *left leaf* was inoculated via the petiole with *P. carotovorum,* and the *leaf* on the *right* is from an uninoculated control. The SRE cause wilt and decay symptoms on *potato stems,* and they can invade the *leaves* from the *stems* via the petiole. The *plant* on the *left* grew from a *potato tuber* naturally infected with *P. carotovorum*

Pectobacterium also typically causes plants to turn yellow, a symptom that is less apparent in plants infected with *Dickeya* (Fig. 2.1).

There are no curative methods for plants infected with SRE. Copper sprays provide limited control when conditions are conducive for infection, but copper does not affect the bacteria once they have colonized the inside of the plant. Recent studies with sanitizers suggest that treating tubers with bleach or benzoic acid provides useful control (Czajkowski et al. 2013). Thus, sanitation, exclusion, crop rotation, and, to a limited extent, plant resistance remain the most important controls for the SRE (Czajkowski et al. 2011).

The SRE are usually described as brute force broad host range pathogens. However, a phylogenetic analysis of all reported hosts of the SRE initiated several years ago by Arthur Kelman suggests some host specificity among strains and genera (Ma et al. 2007). For example, *Dickeya* species infect many grain crops (Poales), but there are no reports of *Pectobacterium* attacking major crops such as rice or maize. Similarly, *Pectobacterium* appears to be a significant pathogen on brassica crops, whereas *Dickeya* has not yet been reported on these widely grown crops. Recent phylogenetic work supports this host species analysis. For example, *Pectobacterium* strains that infect monocot ornamentals are genetically distinct from those infecting dicots (Nabhan et al. 2013; Yishay et al. 2008).

Some SRE species have been reported to have narrow host ranges, but the genomic basis for host range in SRE remains obscure. For example, *Pectobacterium wasabiae* had been reported as a wasabi pathogen (Goto and Matsumoto 1987), but sequence analysis of strains from diseased plants later showed that this species is widespread on potato (Baghaee-Ravari et al. 2011;

Moleleki et al. 2013; Pitman et al. 2010). Similarly, *Pectobacterium atrosepticum*, which causes blackleg on potato, is considered a narrow host range pathogen, but it has been reported on crops such as sunflower (Bastas et al. 2009), and it can cause disease on some non-solanaceous plants in controlled environments (Marquez-Villavicencio et al. 2011).

Future work will likely show that like many bacterial plant pathogens, the genera *Pectobacterium* and *Dickeya* have broad host ranges, but individual SRE strains have restricted host ranges. Genomic data have been used recently to develop simple and relatively inexpensive methods to detect and differentiate all SRE species (Pritchard et al. 2013a). However, because home owners or farmers generally do not make control decisions based on SRE species or genus, there is rarely incentive for them to pay for these sophisticated diagnostic tests that could aid in understanding SRE epidemiology.

Because we know very little about how SRE genomes affect regional distribution or host range of these pathogens, we also know little about factors that affect emergence of novel SRE species or strains in agriculture. However, a recent outbreak of a novel *Dickeya* species, *Dickeya solani*, on potato has provided much insight into these questions. *D. solani* was unknown on potato until the past decade, when in emerged swiftly in the Netherlands on potato, and was then spread via seed potatoes to the Middle East, North Africa, and Asia (Degefu et al. 2013; Slawiak et al. 2009; Toth et al. 2011). *D. solani* has now replaced *P. atrosepticum* as the most important SRE on potato in Europe. Unlike *P. atrosepticum* and other well-studied SRE, *D. solani* strains are clonal, suggesting a recent emergence. The available data suggest that *D. solani*, which is also an efficient pathogen of ornamental bulb crops, was spread into potato in the Netherlands due to rotation of potatoes with ornamental bulbs, supporting the hypothesis that crop rotation promotes SRE emergence on new host species. The Netherlands exports large volumes of both ornamental bulbs and potatoes, which resulted in swift dispersal of this pathogen around the world within a few years.

D. solani has not yet been found in the Americas or in South Africa. In these locations, *Pectobacterium carotovorum* subsp. *brasiliense* appears to be more common (Duarte et al. 2004; Ma et al. 2007; van der Merwe et al. 2010). *D. solani* also has not yet been found in some seed-producing regions of Europe. Its emergence in Europe resulted in quarantines in *D. solani*-free regions, such as potato-producing regions in Scotland in an effort to exclude this pathogen (Kerr et al. 2010). In other countries, such as the USA, this pathogen is essentially ignored. Indeed, despite the lessons learned in Europe, potato producers in Idaho have recently begun rotating potato with ornamental bulbs, which is likely to result in a similar spread of novel SRE in North America.

The *D. solani* scenario describes the largest practical impact of SRE genomics on agriculture today. Due to the availability of genomic sequences, researchers were able to quickly identify a swiftly emerging potato pathogen and to develop specific and sensitive diagnostic tools to detect this pathogen (Degefu et al. 2013; Pritchard et al. 2013a). These tools are now being used to maintain *D. solani*-free regions in some seed potato-producing regions of Europe and to understand *D. solani* epidemiology.

2.1.3 Educational Relevance of *Pectobacterium* and *Dickeya* Genomics

A growing human population and climate change increase our need for expertise in food and fiber production, and the need for expertise in plant protection will certainly grow, even as the pool of people available to teach in this area continues to shrink. In a time of shrinking resources, the SRE are useful teaching tools for students interested in learning about bacterial pathogenesis. The freely available SRE genomic resources allow development of laboratory exercises with the bacteria or in silico. Many features of bacterial physiology and pathogenesis, including Koch's postulates, bacterial isolation on selective media, plant cell wall degradation, biofilm

formation, quorum sensing, insect vectoring, bacterial mutagenesis, pigment production, motility, the plant hypersensitive response, autophagy, antibiosis, bacterial iron acquisition, phage responses, comparative genomic analyses, and contact-dependent inhibition (CDI), can be demonstrated with the SRE. They grow quickly and cause dramatic plant symptoms, and because they do not infect vertebrates, these genera are among the safest bacteria for students to work with. They are also ubiquitous, so teachers can easily obtain local isolates, eliminating the need for maintenance or transfer permits. Finally, the development of ASAP (Glasner et al. 2003), which is free intuitive peer-reviewed database that allows students to enter new findings about SRE genomes, provides a way for undergraduate and graduate students to learn about genomics and, at the same time, to contribute to phytobacteriological research (Glasner et al. 2006).

2.1.4 The SRE Life Cycle

The SRE can cause disease at all stages of production, from planting through storage. Plants may become infected or contaminated with SRE at any point in the production cycle. The SRE are found in irrigation water, soil, and on insects and can latently colonize plants, so it is difficult to keep plants free of SRE in greenhouses or in the field. Fortunately, the SRE are not as aggressive as some viral, fungal, and oomycete pathogens, so complete losses due to SRE during the growing season are rare. However, even relatively, small losses can complicate crop management. For example, poor plant emergence in a potato field due to soft rot of seed tubers makes the potato crop difficult to manage since the plants adjacent to the rotted tubers will grow larger and the harvested tubers will have an uneven size profile. The SRE appear to be significant only in agricultural environments since epidemics in natural environments have not been reported, unlike, for example, as occurs with some fungal tree diseases, such as Dutch elm disease or chestnut blight.

The most significant SRE-caused losses are in storage, where the bacteria from a few decaying vegetables can quickly spread to entire piles of stored vegetables. Unlike many bacterial pathogens, the SRE are facultative anaerobes, and some of their virulence genes are upregulated under anaerobic conditions (Babujee et al. 2012; James and Hugouvieux-Cotte-Pattat 1996). This may explain why they can devastate stored vegetables, where the rotting piles quickly become anaerobic. The high virulence of these strains under anaerobic conditions is often evident in symptom development in vegetables, which are often only rotten in the center, with the ring of tissue outside of the vascular ring remaining undecayed (Fig. 2.1).

The SRE life cycle was once considered to be very simple, and SRE were described as brute force pathogens, with the bacteria colonizing wounded or stressed plants and decaying them through the action of PCWDE. A major impact of genomic studies on the SRE has been a greater appreciation for the subtleties of SRE interactions with plants, insects, and other microbes (Costechareyre et al. 2010, 2013; Llama-Palacios et al. 2002; Toth and Birch 2005). The brute force view of the SRE life cycle is likely the major reason that more effort has not been put into understanding plant resistance over the past century. A long-term impact of genomic research will likely be an increased appreciation for multiple mechanisms required for SRE pathogenicity and resurgence in studying plant resistance to the SRE.

2.1.5 Will Insights from SRE Genomes Lead to Disease Control?

The genetics of SRE pathogenesis were first explored in the early 1970s (Beraha and Garber 1971; Beraha et al. 1974; Chatterjee and Starr 1972), and hundreds of genetic and genomic articles about SRE pathogenesis have been published since then. Despite this, we still use the same control recommendations that farmers have been using for generations—sanitation,

exclusion, crop rotation, and resistance. Many possibilities for disease control that were fostered by genomic studies remain on the horizon, either because the methods proposed, such as transgenic plants, are expensive to implement and not acceptable to consumers or because we still lack detailed information needed about plant and insect interactions with the SRE that could lead to useful control methods. As described above, the most significant contribution today from SRE genomics is an improved ability to detect and differentiate SRE taxa. The most significant contribution to control on the horizon will likely be improvements in breeding for SRE resistance and in control of insect vectors of the SRE.

2.2 Recent Insights into Pathogenesis Made from SRE Genomics

More *Enterobacteriaceae* species genomes are available than from any other family of organisms. The numerous *Enterobacteriaceae* and SRE genomes available have allowed useful applications of genomic data in control and treatment of diseases caused by this important bacterial family. However, much remains to be done in development of new controls, understanding the evolution and pathogenesis in this bacterial family, and exploring the ecology of these species.

P. atrosepticum SCRI 1043 was the first SRE genome sequenced (Bell et al. 2004). The *Dickeya dadantii* 3937 genome was sequenced at approximately the same time as a large community project (Glasner et al. 2011). Although the progress toward publication of the 3937 genome was slow, the collaborative nature of the project and associated genome meetings were effective at building rapport among this research community. In recent years, numerous additional *Pectobacterium* and *Dickeya* genomes have been sequenced by individual lab groups and deposited in public databases (Glasner et al. 2008, 2011; Nykyri et al. 2012; Pritchard et al. 2013b). These genome sequences have aided in resolving phylogenetic relationships within genera, but, perhaps due to high rates of horizontal gene transfer, it has been difficult to determine how the SRE and related genera are placed within the *Enterobacteriaceae* tree.

The SRE genomes consist of a single circular chromosome of just under 5 Mb in size. None of the sequenced strains have sequenced plasmids, although small plasmids and other types of extrachromosomal elements have been reported in the SRE (Nomura et al. 1996). Like many *Enterobacteriaceae* species, the *Pectobacterium* genome is largely conserved, with indels accounting for much of the genome differences among species (Bell et al. 2004; Glasner et al. 2008). Similarly, the *Dickeya* species backbone is also largely conserved (Glasner et al. 2011; Nykyri et al. 2012; Pritchard et al. 2013b). For the most part, the contribution of the various indels to SRE pathogenicity remains unknown, although there are several examples where the genes encoded by these regions contribute to virulence (Evans et al. 2010; Glasner et al. 2008; Hommais et al. 2008; Nykyri et al. 2012, 2013; Pérez-Mendoza et al. 2011; Williamson et al. 2010). At least some of these indels are able to excise from the genome and replicate extrachromosomally (Vanga et al. 2012).

Much of the SRE genome-enabled research over the past few years has focused on identification and characterization of the roles of these indels to SRE virulence. Through this work, we have learned about SRE adhesion (Jahn et al. 2011; Pérez-Mendoza et al. 2011), pigments (Williamson et al. 2010), potential toxins (Bell et al. 2004), contact-dependent secretion (Aoki et al. 2010), and even novel plant-cell-wall-degrading enzymes (Rondelet and Condemine 2012). The SRE genome sequences have also led to insights into insect vectoring, and infection by SRE and the realization that genes required for insect interactions are among the least conserved portions of the SRE genomes (Acosta Muniz et al. 2007; Basset et al. 2003; Costechareyre et al. 2010, 2013; Quevillon-Cheruel et al. 2009).

SRE genomes have enabled a limited number of gene expression studies via promoter-identification studies, microarrays, or chromatin

immunoprecipitation microarrays. Most of these gene expression experiments were designed to catalog genes controlled by well-known central regulators, and the essential discovery was that most of these regulators control large numbers of virulence and central metabolism genes (Hommais et al. 2008; Koiv et al. 2013; Liu et al. 2008; Monson et al. 2013; Rodionov et al. 2004; Venkatesh et al. 2006; Yap et al. 2008). A contrast to this is an assessment of genes controlled by the type III secretion system sigma factor HrpL in *Pectobacterium*, where HrpL was found to control mainly the type III secretion system (T3SS) and genes encoding harpins and a single known effector, DspE (Hogan et al. 2013), rather than being integrated into control of a wide range of other virulence or central metabolism genes.

A few SRE gene expression studies have focused on conditions experienced by SRE and provided insights into bacterial mechanisms used to survive under diverse stresses, such as survival inside insects or during a shift from aerobic to anaerobic conditions. These experiments led to novel insights about the SRE and their host plants or insects. For example, *D. dadantii* upregulates genes required for detoxifying antimicrobial peptides, an antibiosis mechanism not previously thought to be important for the aphid immune system (Costechareyre et al. 2013). Antimicrobial peptides have long been known to be important for plant defenses, but only recently, through gene expression analyses, have we learned that SRE responses to these peptides are controlled via the PhoPQ two-component system (Rio-Alvarez et al. 2012). An investigation into SRE responses to oxygen availability clearly showed that although *Pectobacterium* and *Dickeya* share many regulatory genes, their response to a shift to low-oxygen conditions differed (Babujee et al. 2012), with *Pectobacterium* being better suited to growing in low-oxygen conditions. These data also supported a theme seen many times in comparisons between these two genera; although they cause similar symptoms, the regulatory pathways used to control similar virulence genes differ among *Pectobacterium* and *Dickeya* [for details see (Charkowski et al. 2012)]. Thus, conclusions about regulatory networks cannot necessarily be generalized across the SRE.

2.2.1 Genomics and Secretion

Genome sequences revealed that the SRE, as a group, encode six of the seven protein secretion systems known in gram-negative bacteria (Fig. 2.2), and their role in virulence was recently reviewed (Charkowski et al. 2012). Because several of these secretion systems play key roles in virulence, they have been explored with multiple post-genomic methods, such as gene expression arrays and proteomic surveys. The SRE also secrete or produce extracellular polysaccharides that play important roles in biofilm formation and diseases.

T1SS. The SRE strains encode type I secretion systems (T1SS), which they likely use to secrete proteases and adhesins (Pérez-Mendoza et al. 2011). With the exception of the single report by Pérez-Mendoza et al. (2011), these systems have been almost completely unexamined in the past decade.

T2SS. All SRE encode a type II secretion system, which secretes PCWDE, such as pectinases and cellulases, as well as other virulence-related proteins. Secretome analyses with SRE have focused in large part on proteins that travel the T2SS (Kazemi-Pour et al. 2004). The action of the pectinases is required for isolation of SRE on pectate-containing media, so due to the isolation method used, the presence of these genes in SRE is expected. Some SRE encode a second T2SS, which has only one known target, a cell-bound pectin lyase (Ferrandez and Condemine 2008).

Hrp T3SS. Many, but not all, SRE encode a Hrp T3SS, which appears to secrete a single effector protein, DspE (Hogan et al. 2013). Genome sequences were instrumental in demonstrating that one or, at most, only a few effectors travel the T3SS. How this single effector promotes virulence remains unknown; it plays no apparent role in suppressing plant

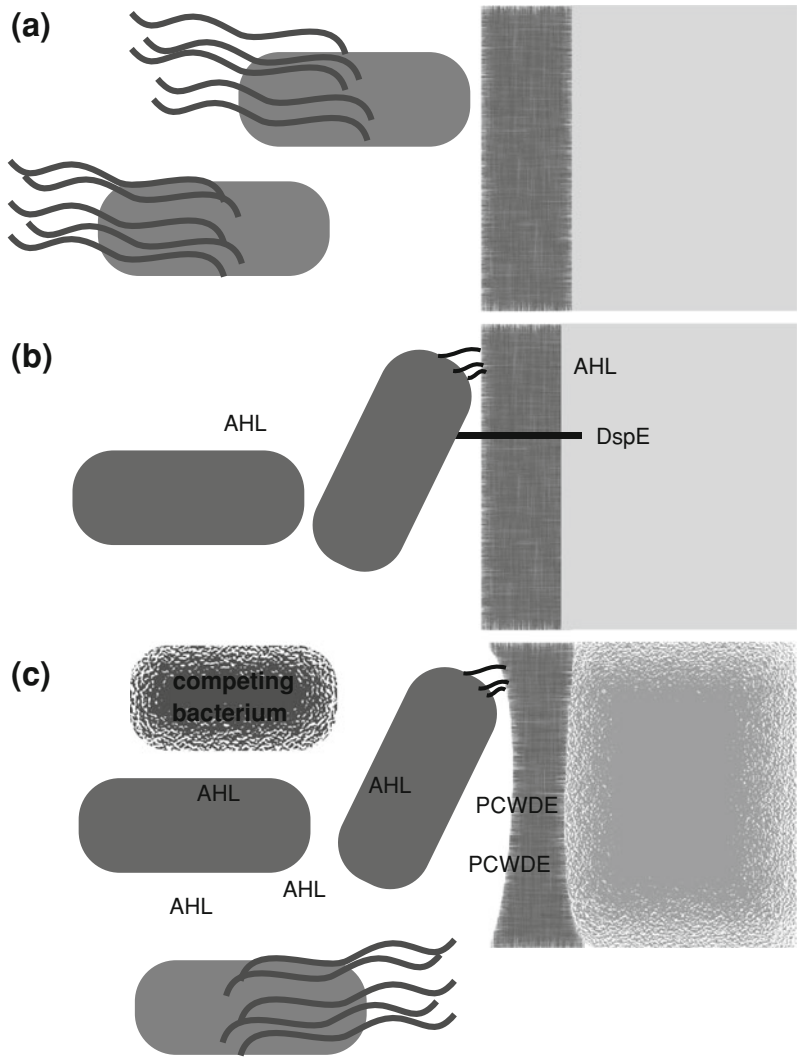

Fig. 2.2 SRE infection of plant tissue. **a** The SRE have flagella and use chemotaxis to move toward wounded plant tissue. **b** Once adjacent to plant cells, they adhere to the cells with secreted adhesin proteins. Some SRE elicit plant cell death by translocating their single known T3SS effector, DspE, into plant cells. **c** They sense that they are in plant cells by integrating signals, such as bacterial AHL and plant cell wall fragments, and then produce massive amounts of plant-cell-wall-degrading enzymes (PCWDE) to macerate plant cell walls. The SRE encode multiple antimicrobial systems and can kill other bacterial strains and species with antibiotics, the contact-dependent inhibition system, or bacteriocins. The SRE are motile in macerated lesions. They multiply and spread from the initial infection zone both locally from cell to cell and systemically via the plant vascular system. Multiple cell types per lesion (e.g., motile and sessile) and cell motility complicate gene expression analysis and growth analysis of SRE in infected plants

defenses. Among the SRE, *P. wasabiae* is notable for its lack of a Hrp T3SS, although strains in other clades also lack this system (Kim et al. 2009; Nykyri et al. 2012; Pitman et al. 2010). Some T3SS contribute to bacterial aggregation in culture; this phenotype for the T3SS was first noted with *Dickeya* (Yap et al. 2005).

Flagellar T3SS. The SRE are motile via a flagellar T3SS. Multiple flagellins are commonly

encoded by SRE genomes, but whether the different flagellins affect virulence remains unknown. Unlike many plant pathogens, the SRE are often motile in diseased plants. The flagellar regulator FliA is required for *Dickeya* virulence (Jahn et al. 2008), but the contribution of motility to virulence has been almost unexplored in our post-genomic era. Swimming may play a role in addition to motility since SRE swimming increases the efficacy of PCWDE or signaling by mixing enzymes or signaling compounds in decaying plant material.

T4SS. The type IV secretion system is only encoded by some strains, and its targets and role in SRE virulence remain unclear (Bell et al. 2004; Glasner et al. 2008; Nykyri et al. 2012).

T5SS. Compared to the other secretion systems in the SRE, the type V secretion system is little explored. It appears to contribute to *Dickeya* adherence and aggregation during pathogenicity (Rojas et al. 2002).

T6SS. The SRE also encode type VI secretion systems, but their role in SRE biology remains unclear (Nykyri et al. 2012).

Pili. The SRE encode pili genes, but they have been little examined, and their role in disease remains unknown (Nykyri et al. 2013).

Cellulose. Bacterial cellulose is used for biofilm formation by many genera. The SRE encode two different types of cellulose synthesis operons, with *Pectobacterium* encoding synthesis genes commonly found in other *Enterobacteriaceae* and *Dickeya* encoding a cellulose synthesis gene cluster that appears to have been recently horizontally acquired. The cellulose synthesis cluster in *Dickeya* contributes to biofilm formation (Jahn et al. 2011), but makes no known contribution to virulence.

2.2.2 Bacterial Signal Molecules and Disease Control

Disruption of virulence gene expression is an attractive mechanism for control of diseases caused by the SRE. *Pectobacterium* and *Dickeya* regulate virulence genes by integrating responses to inter- and intracellular bacterial signals and to plant signals [recently reviewed in Charkowski et al. (2012)]. Although the symptoms caused by these two genera are similar and their main pathogenicity factor is production of a massive amount of PCWDE, regulation of these enzymes, including the signals that the bacteria produce and respond to, has diverged within the SRE.

In both genera, acyl-homoserine lactone (AHL)-mediated quorum sensing plays a role in virulence gene activation, although its importance varies greatly among strains (Ham et al. 2004; Hussain et al. 2008; Mhedbi-Hajri et al. 2011). *Pectobacterium* was the first pathogen in which AHL-mediated quorum sensing was shown to control virulence genes (Pirhonen et al. 1993) and AHL appears to be universally important in this genus. In *Pectobacterium*, AHL-mediated quorum sensing regulates approximately one-quarter of the genes through its regulation of the RsmA-*rsmB* RNA degradation pathway (Chatterjee et al. 2002; Liu et al. 2008). When AHL binds to the DNA-binding proteins ExpRI or ExpR2, it reduces the affinity of ExpR for DNA, resulting in decreased RmsA production (Cui et al. 2005, 2006). Since RsmA targets virulence protein mRNAs for degradation, less RsmA results in higher production of some proteins, including many virulence genes (Chatterjee et al. 1995; Mukherjee et al. 1996).

Genome sequence-enabled studies have resulted in identification of additional bacterial signal molecules in the SRE. These intracellular signaling molecules are widespread among bacteria, such as the autoinducer 2 (AI2) system and cyclic-di-GMP (c-di-GMP)-mediated signaling. Both *Pectobacterium* and *Dickeya* encode *luxS*, which is required for production of the AI2 signal (Crepin et al. 2012), but the role this signal plays in the SRE life cycle remains unknown. The c-di-GMP system is an example of an intracellular signaling system that was revealed in large part through genomic data (Ryan et al. 2012). In bacteria, enzymatic modification of c-di-GMP is involved in cell differentiation in response to environmental or physiological cues. Regulation may occur through sensing of c-di-GMP levels, which are

altered by diguanylate cyclases or phosphodiesterases, or via riboswitches upstream of key genes required for adhesion or motility (Sudarsan et al. 2008). Recently, seven of the 18 proteins in these classes were mutated in *D. dadantii*. Two of the mutant strains (*ecpB* and *ecpC*) had enhanced biofilm formation and reduced virulence, motility, pectate lyase production, and T3SS gene expression (Yi et al. 2010). These same two genes also contribute to virulence in *Pectobacterium*, where they regulate motility and expression of a T1SS, which secretes an adhesion (Pérez-Mendoza et al. 2011). Since these intracellular signaling systems are widespread in both plant and animal pathogens, they are unlikely to be useful targets for disease control in agriculture.

SRE genomics enabled discovery of less conserved signaling systems in *Dickeya*, including the Vfm system and a putative auxin-mediated system. At least, some *Dickeya* strains appear to produce auxin, and it may be a crucial signaling molecule since mutants unable to produce auxin are non-pathogenic (Yang et al. 2007). The auxin signaling pathway acts via the RsmA-*rsmB* pathway, but how auxin is sensed by *Dickeya* remains unknown. Like auxin, the Vfm pathway is only present in *Dickeya* and not *Pectobacterium* (Nasser et al. 2013). The 25 kb Vfm locus, which is located adjacent to the AHL locus encoding ExpR and ExpI in *D. dadantii*, is required for PCWDE production by this pathogen. This locus produces a signal that regulates both virulence gene expression and expression of Vfm locus genes, reminiscent of the feedback regulation of AHL-mediated quorum sensing. These regulatory systems, which appear to be specific to *Dickeya*, may be useful targets for control of this genus.

The key role of bacterial signaling in virulence makes these systems obvious targets for disruption of pathogenicity, but implementation of these methods in agriculture has not yet occurred. To date, the only exploration of signal disruption has been with AHL-mediated signaling. Researchers have examined the use of AHL antagonists and of plants that degrade or produce AHL. Regardless of their efficacy in culture, AHL antagonists are unlikely to be effective under most field conditions since there is a short window of time during infection that these compounds are effective at disrupting pathogenicity (Palmer et al. 2011). However, it is possible that such compounds could be useful under very specific conditions, such as treating plant cuttings just prior to planting.

Transgenic plants that produce an AHL lactonase derived from *Bacillus*, AiiA, are able to constantly degrade AHL and thereby resist soft rot caused by *Pectobacterium* (Dong et al. 2000, 2001). As expected, transgenic potato plants that produce AHL through targeting an AHL synthase to the plant chloroplast are more susceptible to tuber soft rot and stem decay (Toth et al. 2004), further supporting the importance of AHL to SRE pathogenicity. Although plants that disrupt AHL signaling resist bacterial soft rot and AHL signaling is common among rhizosphere bacteria, these plants do not have detectable effects on bacterial colonization of roots (D'Angelo-Picard et al. 2011; Dong et al. 2000, 2001).

Similar future signal disruption experiments with antagonists or transgenic plants will likely provide important information about the various bacterial signaling systems important for virulence. However, transgenic plants that target any particular signaling system may not provide long-term control of SRE diseases since these plants would likely select for strains that use alternate signaling pathways.

2.2.3 Genomics, Bacterial Metabolism, and Disease Control

Multiple intracellular metabolites aid in integration of information about bacterial environment and physiology and thereby play key roles in SRE gene expression during pathogenicity. Unfortunately, for the most part, large holes remain in our knowledge of SRE metabolism. Of those signals that are known, most have only been studied individually and not as part of a comprehensive metabolic model. Although novel control methods have not yet been developed based upon this

metabolic information, these results help us explain how some of our current control methods work.

To date, glucose metabolism has been most closely examined. Like typical *Enterobacteriaceae*, the SRE grow efficiently on glucose and repress other metabolic pathways via the cyclic AMP receptor protein, CRP, including the production of enzymes required to digest plant cell walls, when glucose is present (Reverchon et al. 1997). In turn, KdgR, a repressor that regulates pectinases and other virulence genes also regulates gluconeogenic enzymes (Rodionov et al. 2004), providing a mechanism to coordinate metabolism and plant cell wall degradation. There are several other hints that central metabolism is closely tied to virulence. For example, the starvation signal (p)ppGpp is required for PCWDE production (Wang et al. 2007). Similarly, mutation of gluconate metabolic genes results in hyper-maceration, a lack of motility, and mis-regulation of KdgR and the flagellar regulator FlhD (Mole et al. 2010). Genomics will undoubtedly provide new insights into ties between metabolism and virulence now that we have a more comprehensive inventory of the metabolic pathways present in the SRE and tools to examine expression of metabolic pathway genes in response to different stimuli. As an example, Babujee et al. (2012) recently reported the numerous and divergent effects that a shift to low-oxygen conditions has on *Pectobacterium* and *Dickeya* gene expression, demonstrating that conserved metabolic pathways may respond differently to the same stimulus even in closely related genera.

A comprehensive metabolic model for the SRE will also likely provide insights into host range and the rationale for the effects of various fertilizers on SRE-caused diseases. For example, recent observations that phosphorus increases susceptibility of calla lilies to soft rot may be tied to the effects of the PhoPQ regulon on virulence gene expression, providing an example of how genomic studies combined with greenhouse experiments could lead to useful recommendations for farmers (Gracia-Garza et al. 2004; Llama-Palacios et al. 2003, 2005).

2.2.4 Genomics, Plant Signal Molecules, and Disease Control

The SRE have long been known to activate virulence genes in response to plant cell wall fragments, but until recently, little else was known about how they perceived plant cells and respond to plant cell signals. An exciting recent finding aided by genomics is the discovery that *D. dadantii* is attracted to and swims toward the wound hormone, jasmonic acid (Antunez-Lamas et al. 2009). The SRE encode numerous methyl-accepting chemotaxis (Mcp) receptors, which are membrane proteins that allow bacterial cells to perceive and move along a chemical gradient. The identity of the Mcp that senses jasmonic acid remains unknown as do the targets of nearly all of the over 40 different Mcp proteins encoded by the SRE.

The SRE also sense other organic acids commonly produced by plants. For example, *o*-coumaric acid and *t*-cinnamic acid, both of which are intermediates in the salicylic acid biosynthesis pathway of plants, activate the *D. dadantii* T3SS (Yang et al. 2008), suggesting that regulators controlling the T3SS sense plant-derived chemicals. There are likely to be many additional plant-produced signals waiting to be discovered, and some of these may explain differences in host susceptibility to the SRE.

2.2.5 Genomics, Microbial Antibiosis, and Disease Control

Antagonism between two microorganisms may be based on one or more mechanisms such as nutrient competitiveness, secretion of growth inhibitory compounds, and antibiotic production or through inducing systemic resistance in plants. Several SRE genes dedicated to compete with other microbes were known before any SRE genomes were sequenced, but SRE genome sequence analysis highlighted that much of the variable portion of SRE genomes is dedicated to producing or protecting against antimicrobial metabolites, proteins, or phage (Bell et al. 2004; Glasner et al. 2008).

The systems studied to date in the SRE include antibiotic production (McGowan et al. 2005), bacteriocins (Grinter et al. 2012; Roh et al. 2010; Tovkach 1998), iron competition (Expert 1999), export of bacterial antimicrobial compounds (Llama-Palacios et al. 2002), and the contact-dependent inhibition systems (Aoki et al. 2010). Few of these systems have been evaluated in greenhouse or field experiments, and global studies of the contribution of an entire SRE genome to microbial competition or of the evolution of genes involved in microbial competition have not yet been done.

Among the most exciting new fundamental discoveries in antibiosis studies derived from genomics is the discovery of the CDI system. Recently, the SRE pathogen *D. dadantii* was used to demonstrate that Rhs proteins and related YD-peptide repeat proteins, which are present in a wide range of bacterial species, including other SRE, inhibit growth of neighboring bacteria. Rhs carry polymorphic C-terminal toxin domains which are predicted to be deployed into target cells using the type VI secretion system. *D. dadantii* 3937 is able to protect itself from the cognate toxins and autoinhibition by encoding sequence-diverse immunity proteins (RhsIB) (Aoki et al. 2010; Koskiniemi et al. 2013; Poole et al. 2011). Since the CDI system is usually only functional when the bacteria are in association with host cells and SRE virulence assays are simpler and less expensive that model pathogens that require animals, the SRE will likely continue to be an important model system for studying CDI system function.

The most likely avenue for disease control via antibiosis is through phage therapy. Phage therapy trials with *Pectobacterium* have been conducted on calla lily tubers in greenhouse settings and demonstrated to reduce disease incidence by up to fifty percent (Ravensdale et al. 2007). Recently, two phages, vB_DsoM_LIMEstone1 and vB_DsoM_LIMEstone2, were isolated and characterized for the control of *D. solani*, an aggressive biovar 3 variant of *Dickeya dianthicola* (Adriaenssens et al. 2013). In this case, phage therapy reduced disease incidence and severity in greenhouse trials with inoculated potatoes and resulted in higher yields in potato field trials. The species within the SRE are diverse, and multiple taxa are commonly found in fields (Kim et al. 2009; Yap et al. 2004), and thus, treatments with individual phage are unlikely to be successful in most cases. However, increased efficiency of phage therapy may be achieved through the use of phage cocktails containing multiple phage types (Adriaenssens et al. 2013; Jones et al. 2007). In the future, genomic-enabled detection methods and phage analysis may allow us to assess an agricultural environment or seed lot and determine the appropriate phage cocktail for SRE disease control.

2.2.6 Plant Resistance and Disease Control

Although some examples of single dominant genes conferring resistance to SRE have been known for decades (Lewellen et al. 1978; Whitney and Lewellen 1978), the mechanisms of resistance to soft rot disease remain mostly unknown and were little studied until recently. Data from both cultivated and wild potato suggest that pre-formed resistances occur such as differences in plant cell walls, protease inhibitors, antimicrobial peptides, or other antimicrobial compounds. These genes may appear as quantitative trait loci in genetic experiments designed to identify SRE resistance genes.

Much of what we know about plant resistance to bacterial pathogens is derived from *Pseudomonas syringae—Arabidopsis* interactions or *Xanthomonas*—rice interactions. Unfortunately, these hemibiotrophic pathogens use very different virulence strategies. Consequently, models based on these pathosystems provide little insight into resistance to SRE. Similarly, many pathosystems examined in *Arabidopsis* use a few key hormones, namely salicylic acid, jasmonic acid, and ethylene for defense pathway signaling. But again these systems may not play an important role in resistances observed in some plant species or varieties in the field and

greenhouse. Also, plant disease resistance has mainly been studied with leaf pathogens. The SRE are stem, root, and tuber colonizers, so models based on foliar pathogens that colonize leaf surfaces may only be tangentially useful for understanding SRE pathogenesis.

Only a few researchers have attempted to map SRE resistance in plants, and this work has mainly been done with *P. atrosepticum*. In potato, there are genes that contribute to leaf or tuber resistance to *P. atrosepticum* on all 12 chromosomes (Zimnoch-Guzowska et al. 2000). The available data suggest that leaf and tuber resistance is independent, highlighting the complexity of studying resistance to this pathogen (Zimnoch-Guzowska et al. 2000). Whether the same genes provide resistance to other *Pectobacterium* species or to *Dickeya* is unknown.

Wild species likely encode novel SRE resistance genes. In an assessment of tubers from 123 accessions of wild potato species, *P. carotovorum*-resistant plants were found, but there was no tie to potato taxonomy or strong ties to biogeography of the accessions (Chung et al. 2011). Despite the difficulty in predicting where to find resistant wild species, these results show that novel resistance genes are available in wild relatives of an important crop species.

2.2.7 Genomics Contributions to Pathogen Detection and Epidemiology

The field of study that most separates plant pathology from other research focused on microbial symbiosis is pathogen epidemiology. Until recently, nearly all epidemiological work on the SRE focused on *Pectobacterium* in potatoes grown in temperate climates. Almost nothing is known about the epidemiology of the SRE in tropical and subtropical climates. In the past few years, considerable work has focused on the epidemiology of *D. solani* due to its emergence as an important potato pathogen in Europe. We know that the SRE can be found in irrigation water, agricultural soils, insects, snails, and in latent associations with the roots of numerous plant species. They are not typically found on leaf surfaces nor do they tend to cause diseases of trees or emergence problems with plants grown from true seed. Even on inoculated potato in field trials, SRE disease is sporadic, suggesting that much remains to be learned about environmental components of SRE–plant interactions.

One of the most significant contributions now used to control SRE can be tied to disruption of SRE epidemiology, and specifically in disruption of insect vectoring of SRE. The reduction of *Dickeya* stalk rot in maize as a result of transgenic insect-resistant plants is a significant example of control through disruption of the vector life cycle (Dalmacio et al. 2007). Maize stalk rot used to be a significant concern in North America (Thind and Payak 1976; Thind and Singh 1976), but no work has been published from North America on this disease in decades. The reason for the quiet disappearance of this disease from the literature is likely due to a combination of the prevalence of insect-resistant transgenic maize, the lack of phytobacteriologists working on SRE diseases in North America, the almost complete lack of routine survey data for most plant diseases, including all SRE diseases, and the nearly complete control of research and maize seed production by a handful of large corporations.

Although reports of SRE–insect associations go back over 80 years (Leach 1926, 1931, 1933), very little is known about SRE–insect interactions. However, this has begun to change, and it is one of the largest contributions of SRE genomics toward understanding of SRE ecology and epidemiology. The relationship between SRE and insects is complex, with evidence for insects acting as SRE vectors (Molina et al. 1974) and for SRE being a deadly insect pathogen (Costechareyre et al. 2010). Prior to availability of SRE genome sequences, the gene, Evf, which contributes to *Pectobacterium* fitness on fruit flies (*Drosophila melanogaster*), was identified (Basset et al. 2003). Evf is a lipid-binding protein that requires palmitoylation for function, but how it contributes to bacterial colonization of fly guts remains unknown (Quevillon-Cheruel

et al. 2009). Although Evf is required for *Pectobacterium* colonization of fruit flies, it is not conserved among *Pectobacterium* strains. Whether only a subset of *Pectobacterium* colonizes fruit flies or whether another gene has an analogous role in colonization remains unknown.

The effect of *Dickeya* on insects is much more dramatic since some strains infect and kill aphids. A locus encoding four tandem repeats of a cytotoxin and possibly the ability to detoxify antimicrobial peptides contribute to *D. dadantii* insect pathogenicity (Costechareyre et al. 2010, 2013). *D. dadantii* has a limited host range on insects, and its virulence is affected by aphid life stage, suggesting an intricate interaction between these two organisms. Like Evf, the *Dickeya* cytotoxin genes are not widespread in the SRE, making it clear that we have much to learn about the diversity of SRE–insect interactions and their impact on SRE epidemiology.

A third set of genes likely to be involved in SRE–insect interactions is widespread in the SRE and in many other bacteria. The action of these genes, detected through the Voges–Proskauer assay, has been used to differentiate environmental *Enterobacteriaceae* from coliform *Enterobacteriaceae* for decades (Levine 1916). The potential role of this gene cluster, denoted as *budAB*, in insect interactions was brought to the forefront by genomic studies. The *budAB* operon encodes two of the three enzymes required for the 3-hydroxy-2-butanone (3H2B) pathway, which is an alternate fermentation pathway (Huang et al. 1999; Lopez et al. 1975). This operon is expressed in plants and required for virulence, particularly under anaerobic conditions (Effantin et al. 2011; Marquez-Villavincencio et al. 2011). There are two possible volatile compounds produced by this pathway, 3H2B, and the related compound 2,3-butanediol (23B). Both can function as kairomones for a wide variety of insects, including sap beetles (*Carpophilus hueralis*) (Nout and Bartelt 1998), lygus bugs (*Lygus* sp.) (Buttery et al. 1984), cockroaches (*Nauphoeta cinerea*) (Moore and Moore 1999; Moore et al. 2002), Melanesian rhinoceros beetles (*Scapanes australis*) (Rochat et al. 2002), sorghum chafers (*Pachnoda interrupta*) (Bengtsson et al. 2009), and Mexican fruit flies (*Anastrepha ludens*) (Robacker and Lauzon 2002). The mechanism for attraction has only been described for fruit flies, to date. These insects normally avoid CO_2, which is emitted by ripening fruit, but 23B inhibits antenna neurons sensitive to CO_2 (Turner and Ray 2009). Fruit flies and other insects are often found in and around decaying fruit and vegetables, with massive numbers of fruit flies commonly found in warehouses when soft rot symptoms develop on stored vegetables. Whether the volatile molecules produced by the *budAB* operon are part of what attracts insects to decaying plants in warehouses, compost piles, and kitchens remains unknown.

2.3 Future Prospects for Disease Control Simplified by Genomics

Scientists studying SRE spent the past several decades focused on the genetics, and later the genomics, of SRE–plant interactions and have learned a tremendous amount about how this pathogen deconstructs plant cells. Concurrently, epidemiological and phylogenetic studies have clarified where SRE are in agricultural environments and how these species are related to each other and to other *Enterobacteriaceae*. And yet, we have made little progress in control of diseases caused by SRE. The largest recent contributions toward disease control have been in testing methods that allow faster disease diagnosis and in the serendipitous effects of the BT transgene in maize in essentially eliminating *Dickeya* stalk rot.

The SRE are closely related to *E. coli* and *Salmonella*, two animal pathogens that are very well characterized by a large research community. Despite our large body of genomic knowledge for these human pathogens, we still use antibiotics to treat diseases caused by these microbes, just as we have for the past 70 years. This suggests that significant breakthroughs in chemistry, metabolism, or signaling must be made before bacterial genomic studies lead to control of SRE diseases through manipulation of the bacteria. However, a combination of

information from SRE and host plant genomics is likely to lead to clearer understanding of how some plant species resist SRE; this work will likely lead to control through plant resistance in the near future. Importantly, SRE-resistant plants represent a sustainable, environmentally friendly disease control approach, and breeding and deployment of resistant plants does not require significant technological or theoretical breakthroughs. But it does require sustained funding for plant breeding and plant–microbe interaction studies.

2.3.1 Development of SRE-Resistant Plants

Those attempting to control plant pathogens have one large advantage that researchers studying human *Enterobacteriaceae* pathogens, such as *E. coli*, lack. We are able to breed the host for disease resistance. Although plant resistance is always among the best options for disease control, scientists have focused very little effort on understanding resistance to the SRE. One reason for this lack of research is that the SRE attack mainly vegetable and ornamental crops, and SRE resistance is often not the most important characteristic to breed for. Also, compared to major food crops, little is known about the genomics of these plants. However, this is quickly changing, and genome sequences of numerous vegetable crop plants are now available. In addition, researchers have defined the SRE as brute force secondary pathogens for which genetic resistance might not be available, and therefore, many researchers simply have not searched for it. However, SRE-resistant wild relatives of major vegetable crops are fairly easy to identify. Recently available genomic tools should allow identification of the types of genes that confer resistance to SRE.

Potato, one of the major hosts of the SRE, is a good example of underutilized SRE-resistant germplasm and availability of genomic tools. Unlike most other SRE hosts, potato is a major food crop with an active worldwide research community. But, unlike other major crops, such as maize, soybean, and canola, there is relatively little corporate interest in potato; this crop remains mostly in the hands of family farms and public researchers. Thus, there has been little investment in potato genomics compared to maize and soybean, and until very recently, no genome sequence was available (Xu et al. 2011). Commercial lines of potato are tetraploid, which makes them difficult to sequence and to assess through standard genetic methods. Of necessity, sequencing efforts initially focused on an inbred potato that bears little resemblance to commercial potato. Essentially, no disease resistance work has been completed by the larger research community with this sequenced line.

Fortunately, the potato genus, *Solanum*, is highly diverse, and many species within this genus are easily hybridized. In addition, discoveries in potato genetics now allow fairly simple breeding of diploid potato capable of self-fertilization, which allows development of inbred lines more similar to commercial potato and suitable for disease resistance studies (Lindhout et al. 2011). Since there are multiple wild *Solanum* species that are resistant to SRE, this line of work holds promise both for understanding how plants resist the SRE and for development of SRE-resistant varieties.

2.3.2 Insect Control

SRE genomics has provided new insights into the epidemiology and ecology of these pathogens, and as a result, it has drawn new researchers to this field. For example, SRE–insect interactions have been sporadically studied for the past 80 years, but the genetics and epidemiology of these relationships remain little characterized. Recently, decoded genomes of SRE and other plant pathogenic *Enterobacteriaceae* have clarified that insects play a large role in dissemination of these pathogens and suggest that several of these genera have intimate interactions with insects. Control of insects through transgenic plants has essentially eliminated SRE disease in maize, suggesting that insect resistance would benefit other crops as

well. It is possible that modern synthetic pesticides have also reduced SRE diseases, but we have almost no historical or current survey data for SRE disease incidence on any crop.

Despite the apparent relationship with insect vectors, SRE genes known to be involved in insect interactions are not well-conserved among SRE species or strains, suggesting that conserved genes required for insect interactions remain uncharacterized and/or that only some SRE strains are vectored by insects. Identification of major potential insect SRE vectors and development of systems to study these vectors would aid in determining if vector control would reduce SRE incidence in vegetable and ornamental crops.

2.3.3 Antibiosis

These methods are likely to be based on species- or strain-specific antibiosis techniques that remain almost entirely unexplored at the farm level, such as phage-mediated control or contact-dependent inhibition. If simple detection methods could be followed up by targeted control through phage or other similar methods in seed lots, soils, or water supplies known to carry aggressive SRE strains, a useful level of SRE control might be achieved. The accurate and sensitive diagnostic tools that allow researchers to differentiate among SRE species and strains in order to study their epidemiology in detail have recently been developed, so one factor required for effective control through antibiosis is in place. Expertise in phage biology and biocontrol is sporadic among researchers studying SRE–plant interactions, so both funding and training will likely be required to further develop these control methods.

2.3.4 Inhibitors of SRE Pathogenicity or Metabolism Genes

Interference with bacterial quorum sensing provided promising results for control of soft rot in potato, but this solution is not commercially viable for a number of reasons, one of which is the expense of development and lack of acceptance for transgenic vegetable crops. Oddly enough, consumers find it acceptable for vegetable crops to be sprayed with a range of synthetic chemicals, including endocrine inhibitors and obesogens in order to control pests and manage diseases. In general, though, crops are not sprayed with antibiotics, which can effectively control bacterial diseases, at least in the short term, because these chemicals are reserved for treatment of disease of humans and other animals.

Researchers are now searching for chemicals that are not antibiotics, but rather that inhibit specific virulence systems (Li et al. 2009). For the most part, plant and animal pathogens use similar virulence and regulatory systems for pathogenicity, so even if useful chemicals are discovered, they will likely be reserved for treatment of animal pathogens. One of the few types of virulence genes unique to the SRE and other similar bacterial decay pathogens, like the pectolytic pseudomonads, are the PCWDE. As yet, no chemical inhibitor of PCWDE has been discovered.

2.3.5 From Genomics to Sustainable Plant Production

By obtaining and analyzing SRE genome sequences, we now know far more about the genetics, evolution, and molecular virulence strategies of SRE pathogenicity than we did a decade ago, but we still have much work to do to put our knowledge into practice in order to improve sustainability of food, fiber, fuel, and ornamental plant production. If we are willing to invest in the necessary research and training, a combination of plant and microbe genomics has the potential to transform plant production in the near future. Control of SRE soft rot and stem rot in potato is an excellent example of this. Potato farmers of the future could feasibly be planting their fields only with hybrid diploid tubers grown from true seed that was planted into hydroponic systems rather than with tetraploid potato that has been through several generations

of field production. The simplification of the seed system would reduce land, water, and fuel use, would reduce the risk of losses due to inclement weather, would simplify planning, and would eliminate the many SRE disease problems that occur in the field and in storage. In addition, these diploid potatoes, for which the entire genome sequence could be known, could carry resistance genes to SRE and many other pests and pathogens that were derived from wild relatives. Genome sequences of these wild relatives would aid in marker-assisted selection of breeding lines encoding these resistance genes. If a new SRE strain emerges, as recently happened with *D. solani* in Europe, its genome could be sequenced and detection tools could be developed and tested within a few weeks. Farmers could use these detection tools to detect SRE strains within their seed supply, soil, and water to enable them to improve their planting and storage decisions. Genome sequences could also be used to make decisions about control options, such as whether a commercially available phage could be used to help control the SRE pathogen in water, soil, or on plants.

Together, these genomics-enabled decisions on variety development, pathogen detection, and pathogen control have the potential to increase efficiency and reduce year-to-year variability, which would simplify farm management and food systems planning. But, little of this is likely to happen if our society does not invest in research and training in both fundamental and transformational plant and microbial sciences.

References

Acosta Muniz C, Jaillard D, Lemaitre B, Boccard F (2007) *Erwinia carotovora* Evf antagonizes the elimination of bacteria in the gut of *Drosophila* larvae. Cell Microbiol 9:106–119

Adriaenssens EM, Vaerenbergh JV, Vandenheuvel D, Dunon V, Ceyssens PJ, De Proft M, Kropinski AM, Noben JP, Maes M, Lavigne R (2013) T4-related bacteriophage LIMEstone isolates for the control of soft rot on potato caused by *Dickeya solani*. PLoS One 7:e33227

Alcorn SM, Orum TV, Steigerwalt AG, Foster JL, Fogleman JC, Brenner DJ (1991) Taxonomy and pathogenicity of *Erwinia cacticida* sp. nov. Int J Syst Bacteriol 41:197–212

Antunez-Lamas M, Cabrera E, Lopez-Solanilla E, Solano R, González-Melendi P, Chico JM, Toth I, Birch P, Prichard L, Liu H, Rodriguez-Palenzuela P (2009) Bacterial chemoattraction towards jasmonate plays a role in the entry of *Dickeya dadantii* through wounded tissues. Mol Microbiol 74:662–671

Aoki SK, Diner EJ, de Roodenbeke C, Burgess BR, Poole SJ, Braaten BA, Jones AM, Webb JS, Hayes CS, Cotter PA, Low DA (2010) A widespread family of polymorphic contact-dependent toxin delivery systems in bacteria. Nature 468(7322):439–442

Babujee L, Apodaca J, Balakrishnan V, Liss P, Kiley PJ, Charkowski AO, Glasner JD, Perna NT (2012) Evolution of the metabolic and regulatory networks associated with oxygen availability in two phytopathogenic enterobacteria. BMC Genomics 13(1):110. doi:10.1186/1471-2164-13-110

Baghaee-Ravari S, Rahimian H, Shams-Bakhsh M, Lopez-Solanilla E, Antunez-Lamas M, Rodriguez-Palenzuela P (2011) Characterization of *Pectobacterium* species from Iran using biochemical and molecular methods. Eur J Plant Pathol 129:413–425

Basset A, Tzou P, Lemaitre B, Boccard F (2003) A single gene that promotes interaction of a phytopathogenic bacterium with its insect vector, *Drosophila melanogaster*. EMBO Rep 4:205–209

Bastas KK, Hekimhan H, Maden S, Tor M (2009) First report of bacterial stalk and head rot disease caused by *Pectobacterium atrosepticum* on sunflower in Turkey. Plant Dis 93:1352

Bell KS, Sebaihia M, Pritchard L, Holden MTG, Hyman LJ, Holeva MC, Thomson NR, Bentley SD, Churcher LJC, Mungall K, Atkin R, Bason N, Brooks K, Chillingworth T, Clark K, Doggett J, Fraser A, Hance Z, Hauser H, Jagels K, Moule S, Norbertczak H, Ormond D, Price C, Quail MA, Sanders M, Walker D, Whitehead S, Salmond GPC, Birch PRJ, Parkhill J, Toth IK (2004) Genome sequence of the enterobacterial phytopathogen *Erwinia carotovora* subsp *atroseptica* and characterization of virulence factors. Proc Natl Acad Sci USA 101:11105–11110

Bengtsson JM, Wolde-Hawariat Y, Khbaish H, Negash M, Jembere B, Seyoum E, Hansson BS, Larsson MC, Hillbur Y (2009) Field attractants for *Pachnoda interrupta* selected by means of GC-EAD and single sensillum screening. J Chem Ecol 35:1063–1076

Beraha L, Garber ED (1971) Avirulence and extracellular enzymes of *Erwinia carotovora*. Phytopath Z 70:335–344

Beraha L, Garber ED, Billeter BA (1974) Enzyme profiles and virulence in mutants of *Erwinia carotovora*. Phytopath Z 81:15–22

Brady CL, Cleenwerck I, Denman S, Venter SN, Rodrıguez-Palenzuela P, Coutinho TA,De Vos P (2012) Proposal to reclassify *Brenneria quercina* (Hildebrand and Schroth 1967) Hauben et al. 1999 into a new genus, *Lonsdalea* gen. nov., as *Lonsdalea quercina* comb. nov., descriptions of *Lonsdalea*

quercina subsp. *quercina* comb. nov., *Lonsdalea quercina* subsp. *iberica* subsp. nov. and *Lonsdalea quercina* subsp. *britannica* subsp. nov., emendation of the description of the genus *Brenneria*, reclassification of *Dickeya dieffenbachiae* as *Dickeya dadantii* subsp. *dieffenbachiae* comb. nov., and emendation of the description of *Dickeya dadantii Intern.* J Syst Evol Microbiol 62:1592–1602

Buttery RG, Kamm JA, Ling LC (1984) Volatile components of red clover leaves, flowers, and seed pods: possible insect attractants. J Agric Food Chem 32:254–256

Charkowski A, Blanco C, Condemine G, Expert D, Franza T, Hayes C, Hugouvieux-Cotte-Pattat N, López Solanilla E, Low D, Moleleki L, Pirhonen M, Pitman A, Perna N, Reverchon S, Rodriguez-Palenzuela P, San Francisco MJ, Toth I, Tsuymu S, van der Waals JE, van der Wolf J, Van Gijsegem F, Yang C-H, Yedidia I (2012) The role of secretion systems and small molecules in soft rot *Enterobacteriaceae* pathogenicity. Ann Rev Phytopathol 50:425–449

Chatterjee AK, Starr MP (1972) Genetic transfer of episomic elements among *Erwinia* species and other Enterobacteria: F'*lac*+. J Bacteriol 111:169–176

Chatterjee A, Cui Y, Liu Y, Dumenyo CK, Chatterjee AK (1995) Inactivation of *rsmA* leads to overproduction of extracellular pectinases, cellulase, and proteases in *Erwinia carotovora* subsp. *carotovora* in the absence of the starvation/cell density-sensing signal, N-(3-oxohexanoyl)-L-homoserine lactone. Appl Environ Microbiol 61:1959–1967

Chatterjee A, Cui Y, Chatterjee AK (2002) RsmA and the quorum-sensing signal, N-[3-Oxohexanoyl]- L-homoserine lactone, control the levels of rsmB RNA in *erwinia carotovora* subsp. *carotovora* by affecting its stability. J Bacteriol 184:4089–4095

Chung YS, Holmquist K, Spooner DM, Jansky SH (2011) A test of taxonomic and biogeographic predictivity: resistance to soft rot in wild relatives of cultivated potato. Phytopathology 101:205–212

Costechareyre D, Dridi B, Rahbé Y, Condemine G (2010) Cyt toxin expression reveals an inverse regulation of insect and plant virulence factors of *Dickeya dadantii*. Environ Microbiol 12:3290–3301

Costechareyre D, Chich J-F, Strub J-M, Rhabe Y, Condemine G (2013) Transcriptome of *Dickeya dadantii* infecting *Acrythosiphon pisum* reveals a strong defense against antimicrobial peptides. PLoS One 8:e54118. doi:10.1371/journal.pone/oo54118

Crepin A, Barbey C, Beury-Cirou A, Helias V, Taupin L, Reverchon S, Nasser W, Faure D, Dufour A, Orange N, Feuiloley M, Heurlier K, Burini J-F, Latour X (2012) Quorum sensing signaling molecules produced by reference and emerging soft-rot bacteria (*Dickeya* and *Pectobacterium* spp.). PLoS One 7(4):e35176. doi:10.1371/journal.pone.0035176:

Cui Y, Chatterjee A, Hasegawa S, Dixit V, Leigh N, Chatterjee AK (2005) ExpR, a LuxR homolog of *Erwinia carotovora* subsp. *carotovora*, activates transcription of *rsmA*, which specifies a global regulatory RNA-binding protein. J Bacteriol 187:4792–4803

Cui Y, Chatterjee A, Hasegawa S, Chatterjee AK (2006) *Erwinia carotovora* subspecies produce duplicate variants of ExpR, LuxR homologs that activate *rsmA* transcription but differ in their interactions with N-acylhomoserine lactone signals. J Bacteriol 188:4715–4726

Czajkowski R, Perombelon MCM, van Veen JA, van der Wolf JM (2011) Control of blackleg and tuber soft rot of potato caused by *Pectobacterium* and *Dickeya* species: a review. Plant Pathol 60:999–1013

Czajkowski R, de Boer WJ, van der Wolf JM (2013) Chemical disinfectants can reduce potato blackleg caused by *Dickeya solani*. Eur J Plant Pathol 136:419–432

D'Angelo-Picard C, Chapelle E, Ratet P, Faure D, Dessaux Y (2011) Transgenic plants expressing the quorum quenching lactonase AttM do not significantly alter root-associated bacterial populations. Res Microbiol 162:951–958

Dalmacio SC, Lugod TR, Serrano EM, Munkvold GP (2007) Reduced incidence of bacterial rot on transgenic insect-resistant maize in the Philippines. Plant Dis 91:346–351

Degefu Y, Potrykus M, Golanowska M, Virtanen E, Lojkowska E (2013) A new clade of *Dickeya* spp. plays a major role in potato blackleg outbreaks in North Finland. Ann Appl Biol 162:231–241

Dong YH, Xu JL, Li XZ, Zhang LH (2000) AiiA, an enzyme that inactivates the acylhomoserine lactone quorum-sensing signal and attenuates the virulence of *Erwinia carotovora*. Proc Natl Acad Sci USA 97:3526–3531

Dong Y-H, Wang L-H, Xu J-L, Zhang H-B, Zhang X-F, Zhang L-H (2001) Quenching quorum-sensing-dependent bacterial infection by an *N*-acyl homoserine lactonase. Nature 411:813–817

Duarte V, De Boer SH, Ward LJ, de Oliveira MC (2004) Characterization of atypical *Erwinia carotovora* strains causing blackleg of potato in Brazil. J Appl Microbiol 96:535–545

Effantin G, Rivasseau C, Gromova M, Bligny R, Hugouvieux-Cotte-Pattat N (2011) Massive production of butanediol during plant infection by phytopathogenic bacteria of the genera *Dickeya* and *Pectobacterium*. Mol Microbiol 82:988–997

Evans TJ, Coulthurst SJ, Komitopoulou E, Salmond GPC (2010) Two mobile *Pectobacterium atrosepticum* prophages modulate virulence. FEMS Microbiol Lett 304:195–202

Expert D (1999) Withholding and exchanging iron: interactions between *Erwinia* spp. and their plant hosts. Ann Rev Phytopathol 37:307–334

Ferrandez Y, Condemine G (2008) Novel mechanism of outer membrane targeting of proteins in gram-negative bacteria. Mol Microbiol 69:1349–1357

Gardan L, Gouy C, Christen R, Samson R (2003) Elevation of three subspecies of *Pectobacterium carotovorum* to species level: *Pectobacterium*

atrosepticum sp. nov., *Pectobacterium betavasculorum* sp. nov. and *Pectobacterium wasabiae* sp. nov. Int J Syst Evol Microbiol 53:381–391

Glasner JD, Liss P, Plunkett G, Darling A, Prasad T, Rusch M, Byrnes A, Gilson M, Biehl B, Blattner FR, Perna NT (2003) ASAP, a systematic annotation package for community analysis of genomes. Nucl Acids Res 31:147–151

Glasner JD, Rusch M, Liss P, Plunkett G 3rd, Cabot EL, Darling A, Anderson BD, Infield-Harm P, Gilson MC, Perna NT (2006) ASAP: a resource for annotating, curating, comparing, and disseminating genomic data. Nucl Acids Res 34:D41–D45

Glasner JD, Marquez-Villavicencio M, Kim H-S, Jahn CE, Ma B, Biehl BS, Rissman AI, Mole B, Yi X, Yang C-H, Dangl JL, Grant SR, Perna NT, Charkowski AO (2008) Niche-specificity and the variable fraction of the *Pectobacterium* pan-genome. Mol Plant-Microbe Interact 21:1549–1560

Glasner JD, Yang C-H, Reverchon S, Hugouvieux-Cotte-Pattat N, Condemine G, Bohin J-P, Van Gijsegem F, Yang S, Franza T, Expert D, Plunkett G, San Francisco MJ, Charkowski AO, Py B, Bell K, Rauscher L, Rodriguez-Palenzuela P, Toussaint A, Holeva MC, He SH, Douet V, Boccara M, Blanco C, Toth I, Anderson BD, Biehl BS, Mau B, Flynn SM, Barras F, Lindeberg M, Birch PRJ, Tsuyumu S, Shi X, Hibbing M, Yap M-N, Carpentier M, Dassa E, Umehara M, Kim JF, Rusch M, Soni P, Mayhew GF, Fouts DE, Steven R, Gill SR, Blattner FR, Keen NT, Perna NT (2011) Genome sequence of the plant pathogenic bacterium *Dickeya dadantii* 3937. J Bacteriol 193:2076–2077

Goto M, Matsumoto K (1987) *Erwinia carotovora* subsp. *wasabiae* subsp. no. isolated from diseased rhizomes and fibrous roots of Japanese horseradish (*Eutrema wasabi* Maxim.). Int J Syst Bacteriol 37:130–135

Gracia-Garza JA, Blom TJ, Brown W, Roberts DP, Schneider K, Freisen M, Gombert D (2004) Increased incidence of *Erwinia* soft-rot on calla lilies in the presence of phosphorous. Eur J Plant Pathol 110:293–298

Grinter R, Milner J, Walker D (2012) Ferredoxin containing bacteriocins suggest a novel mechanism of iron uptake in *Pectobacterium* spp. PLoS One 7:e33033

Ham J-H, Yaya C, Alfano JR, Rodríguez-Palenzuela P, Rojas CM, Chatterjee AK, Collmer A (2004) Analysis of *Erwinia chrysanthemi* EC16 *pelE:uidA*, *pelL:uidA*, and *hrpN:uidA* mutants reveals strain-specific atypical regulation of the Hrp type III secretion system. Mol Plant Microbe Interact 17:184–194

Hauben L, Moore ERB, Vauterin L, Steenackers M, Mergaert J, Verdonck L, Swings J (1998) Phylogenetic position of phytopathogens within the *Enterobacteriaceae*. Syst Appl Microbiol 21:384–397

Hogan CS, Mole BM, Grant SR, Willis DK, Charkowski AO (2013) The type III secreted effector DspE Is required early in *Solanum tuberosum* leaf infection by *Pectobacterium carotovorum* to cause cell death, and requires Wx((3-6))D/E motifs. PLoS One 8:e65534

Hommais F, Oger-Desfeux C, Van Gijsegem F, Castang S, Ligori S, Expert D, Nasser W, Reverchon S (2008) PecS is a global regulator of the symtomatic phase in the phytopathogenic bacterium *Erwinia chrysanthemi* 3937. J Bacteriol 190:7508–7522

Huang M, Oppermann-Sanio FB, Steinbüchel A (1999) Biochemical and molecular characterization of the *Bacillus subtilis* acetoin catabolic pathway. J Bacteriol 181:3837–3841

Hussain MBBM, Zhang H-B, Xu J-L, Liu Q, Jiang Z, Zhang L-H (2008) The acyl-homoserine lactone-type quorum-sensing system modulates cell motility and virulence of *Erwinia chrysanthemi* pv. zeae. J Bacteriol 190:1045–1053

Hyman LJ, Sullivan L, Toth IK, Perombelon MCM (2001) Modified crystal violet pectate medium (CVP) based on a new polypectate source (Slendid) for the detection and isolation of soft rot erwinias. Potato Res 44:265–270

Jahn CE, Willis DK, Charkowski AO (2008) The flagellar sigma factor FliA is required for *Dickeya dadantii* virulence. Mol Plant Microbe Interact 21:1431–1442

Jahn CE, Selimi D, Barak JD, Charkowski AO (2011) The *Dickeya dadantii* biofilm matrix consists of cellulose nanofibres, and is an emergent property dependent upon the type III secretion system and the cellulose synthesis operon. Microbiology 157: 2733–2744

James V, Hugouvieux-Cotte-Pattat N (1996) Regulatory systems modulating the transcription of the pectinase genes of *Erwinia chrysanthemi* are conserved in *Escherichia coli*. Microbiology 142:2613–2619

Jones LR (1900) A soft rot of carrot and other vegetables caused by *Bacillus carotovorus* Jones. Vt Agric Exp Sta Rep 13:299–332

Jones LR (1901) *Bacillus carotovorus* n. sp., die ursache einer weichen faulnis der mohre. Zbl Bakt (Abt. II) 7:12–21, 61–68

Jones JB, Jackson LE, Balogh B, Obradovic A, Iriarte FB, Momol MT (2007) Bacteriophages for plant disease control. Ann Rev Phytopathol 45:245–262

Kazemi-Pour N, Condemine G, Hugouvieux-Cotte-Pattat N (2004) The secretome of the plant pathogenic bacterium *Erwinia chrysanthemi*. Proteomics 4:3177–3186

Kerr J, Spiers J, Saddler GS (2010) *Dickeya*: swift policy response by the Scottish Government to tackle this new bacterial threat. Asp Appl Biol 7–12

Kim H-S, Ma B, Perna NT, Charkowski AO (2009) Prevalence and virulence of natural type III secretion system deficient *Pectobacterium* strains. Appl Environ Microbiol 75:4539–4549

Koiv V, Andersen L, Broberg M, Frolova J, Somervuo P, Auvinen P, Pirhonen M, Tenson T, Mae A (2013) Lack of RsmA-mediated control results in constant hypervirulence, cell elongation, and hyperflagellation in *Pectobacterium wasabiae*. PLoS One 8:e54248

Koskiniemi S, Lamoureux JG, Nikolakakis KC, t'Kint de Roodenbeke C, Kaplan MD, Low DA, Hayes CS (2013) Rhs proteins from diverse bacteria mediate intercellular competition. Proc Natl Acad Sci USA 110:7032–7037

Kwon SW, Go SJ, Kang HW, Ryu JC, Jo JK (1997) Phylogenetic analysis of *Erwinia* species based on 16S rRNA gene sequences. Int J Syst Bacteriol 47:1061–1067

Leach JG (1926) The relation of the seed-corn maggot (*Phorbia fuciceps* Zett.) to the spread and development of potato blackleg in Minnesota. Phytopathology 16:149–176

Leach JG (1931) Further studies on the seed-corn maggot and bacteria with special reference to potato blackleg. Phytopathology 21:387–406

Leach JG (1933) The method of survival of bacteria in the puparia of the seed-corn maggot (*Hylemyia cilicrura* Rond.). Zeitschr. fur angewondte Entomologie 20:150–161

Levine M (1916) On the significance of the Voges-Proskauer reaction. J Bacteriol 1:153–164

Lewellen RT, Whitney ED, Goulas CK (1978) Inheritance of resistance to *Erwinia* root rot in sugarbeet. Phytopathology 68:947–950

Li Y, Peng Q, Selimi D, Wang Q, Charkowski AO, Chen X, Yang C-H (2009) The plant phenolic compound p-coumaric acid represses gene expression in the *Dickeya dadantii* type III secretion system. Appl Environ Microbiol 75:1223–1228

Lindhout P, Meijer D, Schotte T, Hutten RCB, Visser RGF, van Eck HJ (2011) Towards F-1 hybrid seed potato breeding. Potato Res 54:301–312

Liu H, Coulthurst SJ, Pritchard L, Hedley PE, Ravensdale M, Humphris S, Burr T, Takle G, Brurberg M-B, Birch PRJ, Salmond GPC, Toth IK (2008) Quorum sensing coordinates brute force and stealth modes of infection in the plant pathogen *Pectobacterium atrosepticum*. PLoS Pathog 4:e1000093. doi:10.1371/journal.ppat.1000093

Llama-Palacios A, Lopez-Solanilla E, Rodriguez-Palenzuela P (2002) The *ybiT* gene of *Erwinia chrysanthemi* codes for a putative ABC transporter and is involved in competitiveness against endophytic bacteria during infection. Appl Environ Microbiol 68:1624–1630

Llama-Palacios A, Lopez-Solanilla E, Poza-Carrion C, Garcia-Olmedo F, Rodriguez-Palenzuela P (2003) The *Erwinia chrysanthemi* PhoP-PhoQ operon plays an important role in growth at low pH, virulence and bacterial survival in plant tissue. Mol Microbiol 49:347–357

Llama-Palacios A, Lopez-Solanilla E, Rodriguez-Palenzuela P (2005) Role of the PhoP-PhoQ system in the virulence of *Erwinia chrysanthemi* strain 3937: involvement in sensitivity to plant antimicrobial peptides, survival at acid pH, and regulation of pectolytic enzymes. J Bacteriol 187:2157–2162

Lopez JM, Thoms B, Rehbein H (1975) Acetoin degradation in *Bacillus subtilis* by direct oxidative cleavage. Eur J Biochem 57:425–430

Ma B, Hibbing ME, Kim H-S, Reedy RM, Yedidia I, Breuer J, Breuer J, Glasner JD, Perna NT, Kelman A, Charkowski AO (2007) The host range and molecular phylogenies of the soft rot enterobacterial genera *Pectobacterium* and *Dickeya*. Phytopathology 97:1150–1163

Marquez-Villavicencio M, Groves RL, Charkowski AO (2011) Soft rot disease severity is affect by potato physiology and *Pectobacterium* taxa. Plant Dis 95:232–241

Marquez-Villavincencio M, Weber B, Witherell RA, Willis DK, Charkowski AO (2011) The 3-hydroxy-2-butanone pathway is required for *Pectobacterium carotovorum* pathogenesis. PLoS One 6:e22974

McGowan SJ, Barnard AML, Bosgelmez G, Sebaihia M, Simpson NJL, Thomson NR, Todd DE, Welch M, Whitehead NA, Salmond GPC (2005) Carbapenem antibiotic biosynthesis in *Erwinia carotovora* is regulated by physiological and genetic factors modulating the quorum sensing-dependent control pathway. Mol Microbiol 55:526–545

Mhedbi-Hajri N, Malfatti P, Pédron J, Gaubert S, Reverchon S, Van Gijsegem F (2011) PecS is an important player in the regulatory network governing the coordinated expression of virulence genes during the interaction between *Dickeya dadantii* 3937 and plants. Environ Microbiol 13:2901–2914

Mole B, Habibi S, Dangl JL, Grant SR (2010) Gluconate metabolism is required for virulence of the soft-rot pathogen *Pectobacterium carotovorum*. Mol Plant Microbe Interact 23:1335–1344

Moleleki LN, Onkendi EM, Mongae A, Kubheka GC (2013) Characterisation of *Pectobacterium wasabiae* causing blackleg and soft rot diseases in South Africa. Eur J Plant Pathol 135:279–288

Molina JM, Harrison D, Brewer JW (1974) Transmission of *Erwinia carotovora* var. *atroseptica* by *Drosophila melanogaster* Meig. I. Acquisition and transmission of the bacterium. Am Potato J 51:245–250

Monson R, Burr T, Carlton T, Liu H, Hedley P, Toth I, Salmond GPC (2013) Identification of genes in the VirR regulon of *Pectobacterium atrosepticum* and characterization of their roles in quorum sensing-dependent virulence. Environ Microbiol 15:687–791

Moore AJ, Moore PJ (1999) Balancing sexual selection through opposing mate choice and male competition. Proc R Soc Lond B Biol Sci 266:711–716

Moore AJ, Haynes KF, Preziosi RF, Moore PJ (2002) The evolution of interacting phenotypes: genetics and evolution of social dominance. Am Nat 160:5186–5197

Mukherjee A, Cui Y, Liu Y, Dumenyo CK, Chatterjee AK (1996) Global regulation in *Erwinia* species by *Erwinia carotovora rsmA*, a homologue of *Escherichia coli csrA*: repression of secondary metabolites

pathogenicity and hypersensitive reaction. Microbiology 142:427–434

Nabhan S, De Boer SH, Maiss E, Wydra K (2013) *Pectobacterium aroidearum* sp. nov., a soft rot pathogen with preference for monocotyledonous plants. Int J Syst Evol Microbiol 63:2520–2525

Nasser W, Dorel C, Wawrzyniak J, Van Gijsegem F, Groleau M-C, Deziel E, Reverchon S (2013) Vfm a new quorum sensing system controls the virulence of *Dickeya dadantii*. Environ Microbiol 15:865–880

Nomura N, Yamashita M, Murooka Y (1996) Genetic organization of a DNA-processing region required for mobilization of a non-self-transmissible plasmid, pEC3, isolated from *Erwinia carotovora* subsp *carotovora*. Gene 170:57–62

Nout MJR, Bartelt RJ (1998) Attraction of a flying nitidulid (*Carpophilus humeralis*) to volatiles produced by yeasts grown on sweet corn and a corn-based medium. J Chem Ecol 24:1217–1239

Nykyri J, Niemi O, Koskinen P, Nokso-Koivisto J, Pasanen M, Broberg M, Plyusnin I, Toronen P, Holm L, Pirhonen M, Palva ET (2012) Revised phylogeny and novel horizontally acquired virulence determinants of the model soft rot phytopathogen *Pectobacterium wasabiae* SCC3193. PLoS Pathog 8:e1003013

Nykyri J, Mattinen L, Niemi O, Adhikari S, Koiv V, Somervuo P, Fang X, Auvinen P, Mae A, Palva ET, Pirhonen M (2013) Role and regulation of the Flp/Tad pilus in the virulence of *Pectobacterium atrosepticum* SCRI1043 and *Pectobacterium wasabiae* SCC3193. PloS One 8(9):e73718

Palmer AG, Streng E, Blackwell HE (2011) Attenuation of virulence in pathogenic bacteria using synthetic quorum-sensing modulators under native conditions on plant hosts. ACS Chem Biol 6:1348–1356

Pérez-Mendoza D, Coulthurst SJ, Humphris S, Campbell E, Welch M, Toth IK, Salmond GPC (2011) A multi-repeat adhesin of the phytopathogen, *Pectobacterium atrosepticum*, is secreted by a type I pathway and is subject to complex regulation involving a non-canonical diguanylate cyclase. Mol Microbiol 82:719–733

Pirhonen M, Flego D, Heikinheimo R, Palva ET (1993) A small diffusible signal molecule is responsible for the global control of virulence and exoenzyme production in the plant pathogen *Erwinia carotovora*. EMBO J 12:2467–2476

Pitman AR, Harrow SA, Visnovsky SB (2010) Genetic characterisation of *Pectobacterium wasabiae* causing soft rot disease of potato in New Zealand. Eur J Plant Pathol 126:423–435

Poole SJ, Diner EJ, Aoki SK, Braaten BA, t'Kint de Roodenbeke C, Low DA, Hayes CS (2011) Identification of functional toxin/immunity genes linked to contact-dependent growth inhibition (CDI) and rearrangement hotspot (Rhs) systems. PLoS Genet 7:e1002217

Pritchard L, Humphris S, Saddler G, Parkinson NM, Bertrand V, Elphinstone JG, Toth IK (2013a) Detection of phytopathogens of the genus *Dickeya* using a PCR primer prediction pipeline for draft bacterial genome sequences. Plant Pathol 62:587–596

Pritchard L, Humpris S, Baeyen S, Maes M, Van Vaerenbergh J, Elphinstone J, Saddler G, Toth I (2013b) Draft genome sequences of four *Dickeya dianthicola* and four *Dickeya solani* strains. Genome Announcements 1. 10.1128/genomeA.00087-12

Quevillon-Cheruel S, Leulliot N, Muniz CA, Vincent M, Gallay J, Argentini M, Cornu D, Boccard F, Lemaitre B, van Tilbeurgh H (2009) Evf, a virulence factor produced by the *Drosophila* pathogen *Erwinia carotovora*, is an S-palmitoylated protein with a new fold that binds to lipid vesicles. J Biol Chem 284:3552–3562

Ravensdale M, Blom TJ, Svircev AM, Smith RJ (2007) Bacteriophages and the control of *Erwinia carotovora* subsp. *carotovora*. Can J Plant Pathol 29:121–130

Reverchon S, Expert D, Robert-Baudouy J, Nassar W (1997) The cyclic AMP receptor protein is the main activator of the pectinolysis genes in *Erwinia chrysanthemi*. J Bacteriol 179:3500–3508

Rio-Alvarez I, Rodriguez-Herva JJ, Cuartas-Lanza R, Toth I, Pritchard L, Rodriguez-Palenzuela P, Lopez-Solanilia E (2012) Genome-wide analysis of the response of *Dickeya dadantii* 3937 to plant antimicrobial peptides. Mol Plant Microbe Interact 25:523–533

Robacker DC, Lauzon CR (2002) Purine metabolizing capability of *Enterobacter agglomerans* affects volatiles production and attractiveness to Mexican fruit fly. J Chem Ecol 28:1549–1563

Rochat D, Morin JP, Kakul T, Beaudoin-Ollivier L, Prior R, Renou M, Malosse I, Stathers T, Embupa S, Laup S (2002) Activity of male pheromone of Melanesian rhinoceros beetle *Scapanes australis*. J Chem Ecol 28:479–500

Rodionov DA, Gelfand MS, Hugouvieux-Cote-Pattat N (2004) Comparative genomics of the KdgR regulon in *Erwinia chrysanthemi* 3937 and other gamma-proteobacteria. Microbiology 150:3571–3590

Roh E, Park T-H, Kim M-I, Lee S, Ryu S, Oh C-S, Rhee S, Kim D-H, Park B-S, Heu S (2010) Characterization of a new bacteriocin, carocin D, from *Pectobacterium carotovorum* subsp. *carotovorum* Pcc21. Appl Environ Microbiol 76:7541–7549

Rojas CM, Ham J-H, Deng W-L, Doyle JJ, Collmer A (2002) HecA is a member of a class of adhesins produced by diverse pathogenic bacteria and contributes to the attachment, aggregation, epidermal cell killing, and virulence phenotypes of *Erwinia chrysanthemi* EC16 on *Nicotiana clevelandii* seedlings. Proc Natl Acad Sci USA 99:13142–13147

Rondelet A, Condemine G (2012) SurA is involved in the targeting to the outer membrane of a Tat signal sequence-anchored protein. J Bacteriol 194:6131–6142

Ryan RP, Tolker-Nielsen T, Dow JM (2012) When the PilZ don't work: effectors for cyclic di-GMP action in bacteria. Trends Microbiol 20:235–242

Samson R, Legendre JB, Christen R, Fischer-Le Saux M, Achouak W, Gardan L (2005) Transfer of

Pectobacterium chrysanthemi (Burkholder et al. 1953) Brenner et al. 1973 and *Brenneria paradisiaca* to the genus *Dickeya* gen. nov as *Dickeya chrysanthemi* comb. nov and *Dickeya paradisiaca* comb. nov and delineation of four novel species, *Dickeya dadantii* sp nov., *Dickeya dianthicola* sp nov., *Dickeya dieffenbachiae* sp nov and *Dickeya zeae* sp nov. Int J Syst Evol Microbiol 55:1415–1427

Slawiak M, van Beckhoven JRCM, Speksnijder AGCL, Czajkowski RL, Grabe G, van der Wolf JM (2009) Biochemical and genetical analysis reveal a new clade of biovar 3 *Dickeya* spp. strains isolated from potato in Europe. Eur J Plant Pathol 125:245–261

Sudarsan N, Lee ER, Weinberg Z, Moy RH, Kim JN, Link KH, Breaker RR (2008) Riboswitches in eubacteria sense the second messenger cyclic di-GMP. Science 321:411–413

Thind BS, Payak MM (1976) A laboratory method for evaluating maize germplasm to *Erwinia carotovora* var. *zeae*. Plant Dis Rep 60:414–416

Thind BS, Singh N (1976) Maize borer [*Chilo partellus* (Swinhoe)] as carrier of *Erwinia carotovora* var. *zeae*, the causal agent of bacterial stalk rot of maize. Curr Sci 45:117–118

Toth IK, Birch PRJ (2005) Rotting softly and stealthily. Curr Opin Plant Biol 8:424–429

Toth IK, Newton JA, Hyman LJ, Lees AK, Daykin M, Ortori C, Williams P, Fray RG (2004) Potato plants genetically modified to produce *N*-acylhomoserine lactones increase susceptibility to soft rot erwiniae. Mol Plant Microbe Interact 17:880–887

Toth IK, van der Wolf JM, Saddler G, Lojkowska E, Helias V, Pirhonen M, Lahkim TL, Elphinstone JG (2011) *Dickeya* species: an emerging problem for potato production in Europe. Plant Pathol 60:385–399

Tovkach FI (1998) Biological properties and classification of *Erwinia carotovora* bacteriocins. Microbiology 67:636–642

Turner SL, Ray A (2009) Modification of CO2 avoidance behaviour in *Drosophila* by inhibitory odorants. Nature 461:277–281

van der Merwe JJ, Coutinho TA, Korsten L, van der Waals JE (2010) *Pectobacterium carotovorum* subsp *brasiliensis* causing blackleg on potatoes in South Africa. Eur J Plant Pathol 126:175–185

Vanga BR, Butler RC, Toth IK, Ronson CW, Pitman AR (2012) Inactivation of PbTopo IIIB causes hyper-excision of the pathogenicity island HAI2 resulting in reduced virulence of *Pectobacterium atrosepticum*. Mol Microbiol 84:648–663

Venkatesh B, Babujee L, Liu H, Hedley P, Fujikawa T, Birch P, Toth I, Tsuyumu S (2006) The *Erwinia chrysanthemi* 3937 PhoQ sensor kinase regulates several virulence determinants. J Bacteriol 188:3088–3098

Waldee EL (1945) Comparative studies of some peritrichous phytopathogenic bacteria. Iowa St Coll J Sci 19:435–484

Wang J, Gardiol N, Burr T, Salmond GPC, Welch M (2007) RelA-Dependent (p)ppGpp production controls exoenzyme synthesis in *Erwinia carotovora* subsp. *atroseptica*. J Bacteriol 189:7643–7653

Whitney ED, Lewellen RT (1978) Bacterial vascular necrosis and rot of sugarbeet: genetic vulnerability and selecting for resistance. Phytopathology 68:657–661

Williamson NR, Commander PM, Salmond GP (2010) Quorum sensing-controlled Evr regulates a conserved cryptic pigment biosynthetic cluster and a novel phenomycin-like locus in the plant pathogen, *Pectobacterium carotovorum*. Environ Microbiol 12:1811–1827

Winslow C-EA, Broadhurst J, Buchanan RE, Krumwiede CJ, Rogers LA, Smith GH (1917) The families and genera of the bacteria. Preliminary report of the Committee of the Society of American Bacteriologists on Characterization and Classification of Bacterial Types. J Bacteriol 2:505–566

Xu X, Pan S, Cheng S et al (2011) Genome sequence and analysis of the tuber crop potato. Nature 475:189–195

Yang S, Zhang Q, Guo J, Charkowski AO, Glick BR, Ibekwe M, Cooksey DA, Yang C-H (2007) Global effect of indole-3-acetic acid biosynthesis on multiple virulence factors of *Erwinia chrysanthemi* 3937. Appl Environ Microbiol 73:1079–1088

Yang S, Peng Q, San Francisco MJ, Wang Y, Zeng Q, Yang C-H (2008) Type III secretion system genes of *Dickeya dadantii* 3937 are induced by plant phenolic acids. PloS One 3:e2973. doi:10.1371/journal.pone.0002973

Yap M-N, Barak JD, Charkowski AO (2004) Genomic diversity of *Erwinia carotovora* subsp. *carotovora* and its correlation with virulence. Appl Environ Microbiol 70:3013–3023

Yap M-N, Yang CH, Barak JD, Jahn CE, Charkowski AO (2005) The *Erwinia chrysanthemi* type III secretion system is required for multicellular behavior. J Bacteriol 187:639–648

Yap MN, Yang C-H, Charkowski AO (2008) The response regulator HrpY of *Dickeya dadantii* 3937 regulates virulence genes not linked to the *hrp* cluster. Mol Plant Microbe Interact 21:304–314

Yi X, Yamazaki A, Biddle E, Zeng Q, Yang C-H (2010) Genetic analysis of two phosphodiesterases reveals cyclic diguanylate regulation of virulence factors in *Dickeya dadantii*. Mol Microbiol 77:787–800

Yishay M, Burdman S, Valverde A, Luzzatto T, Ophir R, Yedidia I (2008) Differential pathogenicity and genetic diversity among *Pectobacterium carotovorum* ssp *carotovorum* isolates from monocot and dicot hosts support early genomic divergence within this taxon. Environ Microbiol 10:2749–2759

Zimnoch-Guzowska E, Marczewski W, Lebecka R, Flis B, Schafer-Pregl R, Salamini F, Gebhardt C (2000) QTL analysis of new sources of resistance to *Erwinia carotovora* ssp. *atroseptica* in potato done by AFLP, RFLP, and resistance-gene-like markers. Crop Sci 40:1156–1167

Ecological Genomics of *Pseudomonas syringae*

3

David A. Baltrus, Tory A. Hendry, and Kevin L. Hockett

Pseudomonas syringae is a facultative bacterial phytopathogen of many (if not all) plant species, but strains can also survive and thrive outside of their plant hosts across many different environments (Hirano and Upper 2000). Given an abundance of recent reviews that focus largely on studies of pathogenicity (Block and Alfano 2011; Lindeberg et al. 2012; O'Brien et al. 2011; Studholme 2011), we take this opportunity to summarize and explore what is known about genomic features that influence ecological dynamics for *P. syringae* inside and outside of plant hosts. Recent sequencing efforts have yielded a wealth of data about genomic diversity throughout this species, but have also given rise to numerous overarching questions about evolutionary dynamics for *P. syringae* and related bacteria. We view this chapter as an incomplete roadmap that both summarizes current knowledge of the genomic basis for ecological diversity throughout the species and highlights unexplained patterns that arise from comparison across genomes.

D. A. Baltrus (✉) · T. A. Hendry · K. L. Hockett
School of Plant Sciences, The University of Arizona, Tucson, AZ 85721-0036, USA
e-mail: baltrus@email.arizona.edu

3.1 Taxonomy

Taxonomic relationships across phytopathogens and other soil-associated microbes often change, with nomenclatural disputes sparking heated disagreements across disciplines. The focus of this chapter will be what is currently referred to as *P. syringae* but which includes other potential species such as *P. savastanoi* and *P. cannabina* (Ramos et al. 2012; Sarris et al. 2013; Young 2010). We reference genome sequences from within these isolates as a point of context for comparison, but support those that would chose to split off species names given appropriate phenotypic contexts. *P. syringae* as a whole contains upward of 50 different pathovars spanning pathogens of important crops as well as wild plants (O'Brien et al. 2011; Young 2010). In recent years, due to extensive environmental sampling, greater appreciation has been given to environmental isolates that are not known to be virulent on any host (and thus lack pathovar designations) but which phylogenetically cluster within *P. syringae* (Diallo et al. 2012a; Kniskern et al. 2010; Morris et al. 2010). One clade in particular, referred to below as MLST group 2C, stands out as having shifted its ecological niche compared to closely related phytopathogenic strains. For the moment, we consider these all to be *P. syringae* isolates and suggest that elucidation of the ecological roles

D. C. Gross et al. (eds.), *Genomics of Plant-Associated Bacteria*,
DOI: 10.1007/978-3-642-55378-3_3, © Springer-Verlag Berlin Heidelberg 2014

for these environmental isolates is an intriguing future research direction.

Historically, *P. syringae* has been classified based on phenotypic responses to the LOPAT test, with positive interactions for levan synthesis and tobacco hypersensitive response and negative responses for arginine dihydrolase, oxidase production, and potato rot as distinguishing features (Young 2010). Strains may be further classified into pathovars based on phenotypic information concerning disease symptoms, hosts of isolation, and pathogenicity on a variety of alternative hosts, in addition to nutritional characteristics (Young 2010). While fruitful in providing a context for rapid and inexpensive classification, phenotype-based methods are prone to errors due to incorrect scoring of phenotypes and convergent evolution of phenotypes relevant for classification. For instance, the MLST group 2C clade tests negative for the tobacco hypersensitive response, but genome sequences indicate that they are a clear subclade within *P. syringae* (Diallo et al. 2012a). Moreover, there are multiple instances where pathovar designation is polyphyletic, with convergence of disease states on the same host (Baltrus et al. 2011; O'Brien et al. 2012). For these reasons, the focus of classification has shifted to genotypic methods that provide a much richer framework for interpreting evolutionary relationships and patterns between isolates.

Genotypic classification of strains began with random PCR amplification-based methods, but the first leap forward occurred when Gardan and colleagues used DNA–DNA hybridization to classify isolates into eight genomospecies by genomic similarity (Gardan et al. 1999; Young 2010). More recently, multi-locus sequence type (MLST)-based comparisons have become the default technique for characterizing strains and pathovars (Almeida et al. 2010; Hwang et al. 2005; Morris et al. 2010). MLST characterization relies on sequencing portions of conserved "housekeeping" loci such as *rpoD, gapA, gltA* (also known as *cit*), and *gyrB*. To further facilitate rapid MLST comparisons, a database and Web-based server has been established (Almeida et al. 2010). According to these MLST comparisons, *P. syringae* strains can be subdivided into at least five distinct clades with complete genome sequences available for strains within three of these groups (group I: *P. syringae* pv. *tomato* DC3000, *Pto*DC3000; group 2: *P. syringae* pv. *syringae* B728a, *Psy*B728a; and group 3: *P. syringae* pv. *phaseolicola* 1448a, *Pph*1448a), draft genome sequences for isolates from the other two main groups (group 4: *P. syringae* pv. *oryzae* 1_6, *Por*1_6 and group 5: *P. cannabina* pv. *alisalensis, Pcal*), and numerous other strains throughout the phylogeny (Baltrus et al. 2011; O'Brien et al. 2011; Sarris et al. 2013; Studholme Studholme 2011). For the remainder of this chapter, we will refer to the strains by their MLST group designations where possible and secondarily by their genomospecies (see Fig. 3.1) and will focus on important insights to be gained and questions to be asked by focusing on genome evolution between and within MLST groups.

3.2 Genome Features

Although genomic features within *P. syringae* have been covered extensively elsewhere, to facilitate further discussion, we highlight a handful of interesting nuances within this chapter that have arisen from comparison across diverse strains. The average genomic content for most *P. syringae* strains is roughly 5–6 Mb, coding for slightly more than 5,000 open reading frames (O'Brien et al. 2011). Strain *Pto*DC3000 contains the most thoroughly vetted genome annotation, the product of extensive hand curation and multiple RNAseq experiments (Filiatrault et al. 2010, 2011), and we caution that one should always be wary when analyzing numerous draft genomes due to their fractured nature.

Genomic comparisons across *P. syringae* have already raised numerous ecological questions, the number of which will only increase with additional genomic information. For instance, group I strains contain approximately 500 additional genes compared to any of the other group, with some functional classes

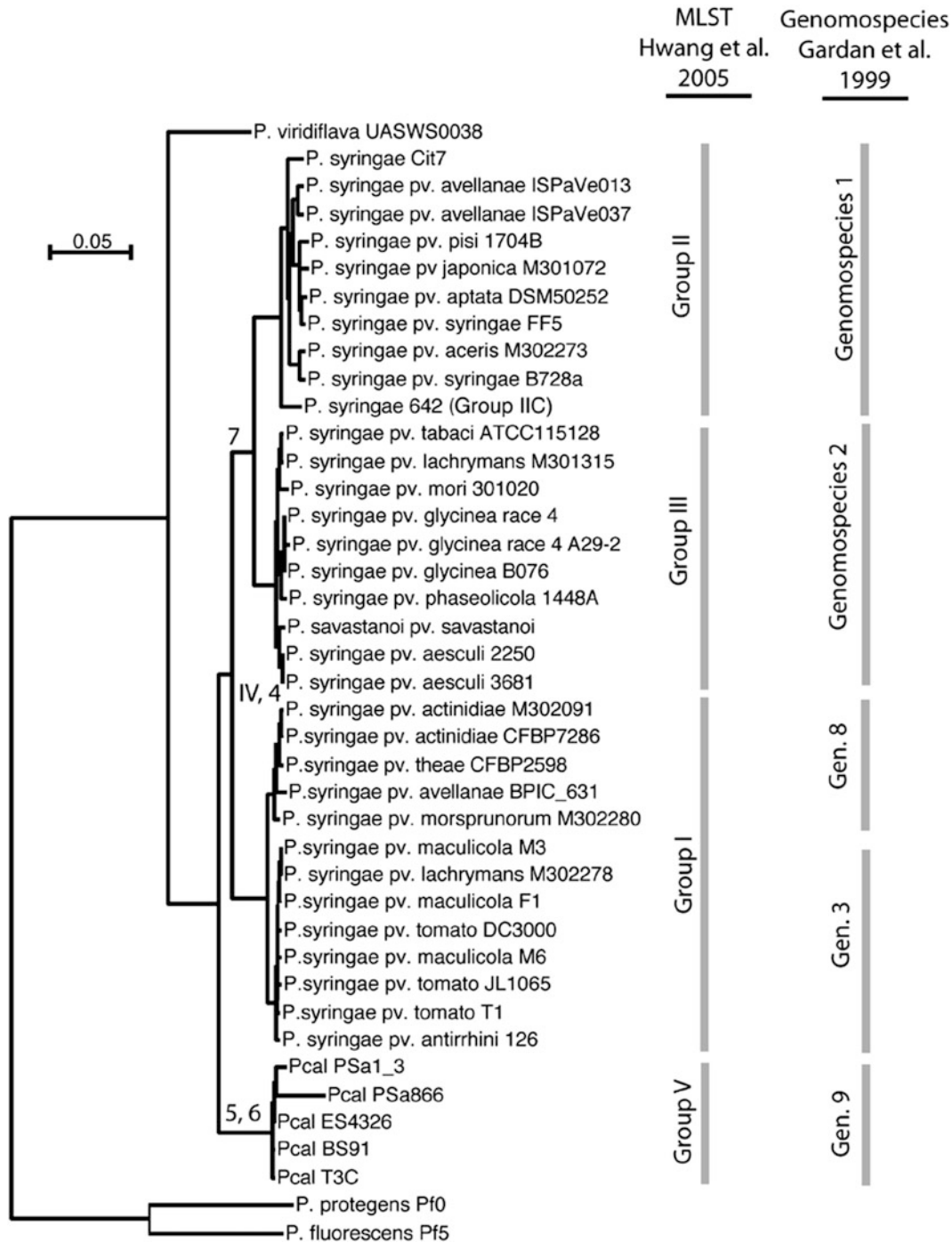

Fig. 3.1 Phylogeny of *Pseudomonas syringae* strains and related species. Phylogeny is based on draft and whole-genome sequences and has been modified from (Sarris et al. 2013). MLST groups and genomospecies for each strain are shown at the *right*. In cases where MLST groups (IV) or genomospecies (*4*, *5*, *6*, *7*) are not represented within the tree, numbers within the tree highlight estimated branching points for these groups

overrepresented within this suite compared to the remaining genomic distribution (Baltrus et al. 2011). While this pattern is unimpressive when compared to the considerable genomic diversity of plant-associated pseudomonads, it could signal important ecological differences when considered within the context of *P. syringae* as a species. Likewise, plasmids are found within most *P. syringae* isolates and likely play a primary role in structuring ecologically and virulence-related genomic diversity throughout the species, while also enabling rapid evolutionary shifts through horizontal transfer across strains. Of the complete genomes, *Pto*DC3000 and *Pph*1448a both contain two plasmids that house numerous virulence genes, but *Psy*B728a lacks such plasmids altogether (Buell et al. 2003; Feil et al. 2005; Joardar et al. 2005). Plasmids have also been observed within many other strains (Ma et al. 2007; Vivian et al. 2001). Since plasmids are notorious for containing repetitive elements, which confound draft genome assemblies, it is difficult to pinpoint which contigs within these incomplete sequences are associated with extrachromosomal elements. However, most strains with draft genomes do appear to contain genomic features that are correlated with the presence of mobile elements (Baltrus et al. 2011). Given the important role horizontal gene transfer plays in the evolution of virulence (Jackson et al. 2011; Sundin 2007), it remains to be seen what effects, if any, lack of plasmids has on strain evolutionary and ecological dynamics (Feil et al. 2005). One of the most interesting plasmids is found within a small number of group III strains from pathovar *lachrymans* (Baltrus et al. 2011). These strains have recently acquired an independently replicated, circular megaplasmid that increases total genomic content by $\sim$1 Mb. Although it contains hundreds of additional genes little else is known about the functions of this element; however, that megaplasmids often code for niche-specific pathways implies that undescribed ecological variability exists within this strain (Harrison et al. 2010).

3.3 Genomic Factors that Limit Host Range of *P. syringae*

Although *P. syringae* has a wide host range as a species, each individual isolate only causes disease on a limited number of plants. Moreover, disease symptoms can vary from leaf and fruit spots to trunk cankers to galls depending on strain, pathovar, and host context (Arnold et al. 2011; Preston 2000; Ramos et al. 2012; Scortichini et al. 2012). Our current understanding of interactions between plant immune responses and bacterial virulence mechanisms can be described, at best, as complicated. However, at a primary level, *P. syringae* growth is limited by the environment within a plant host regardless of dedicated immune responses. For instance, it has long been known that isolates of *P. syringae* differ in their abilities to use a variety of substrates for nutrition. Plant species can vary dramatically in the quantitative and qualitative distributions of these substrates, and the importance of manipulating host metabolism seems to be a growing theme across phytopathogen species [i.e., *Xanthomonas* and SWEET genes (Verdier et al. 2012)]. It is likely that nutrient concentrations limit growth of *P. syringae in planta*, so different metabolite compositions within leaves could contribute to host range limitation (Rico and Preston 2008). With this idea in mind, it seems particularly noteworthy that multiple clades of *P. syringae* that invade woody plant species have convergently acquired pathways involved in the breakdown of catechols (Green et al. 2010; Marcelletti et al. 2011; Rodríguez-Palenzuela et al. 2010). Furthermore, Thlaspi caerulescens can hyperaccumulate zinc, nickel, and cadmium at high enough concentrations to prevent bacterial pathogenesis (Fones et al. 2010). In the absence of high concentrations of these metals, *Pcal*ES4326 is virulent on *T. caerulescens*. Transposon mutagenesis can be used to isolate mutants with higher metal tolerance, and these mutations enabled *Pcal*ES4326 to sustain virulence in *T. caerulescens* plants as concentration of toxic metals increased. Although this result is tantalizing and suggestive,

it remains unclear how much natural variation exists for metal tolerance within *P. syringae* and how much of an impact such environmental context plays in structuring natural bacterial host ranges.

The first layer of induced plant immune responses against any phytopathogen involves dedicated receptor-like kinase proteins that directly recognize conserved peptides (termed pathogen-associated molecular patterns or PAMPs) triggering basal defenses (Thomma et al. 2011). Although we have learned extensive amounts about the molecular mechanisms underlying PAMP recognition, it is likely that this is only the tip of the iceberg and that there is much more to discover as to how these receptors vary in presence and specificity at the level of host phylogeny. Although the most highly reactive and widely studied PAMP within *P. syringae* is flagellin (specifically flg22), a library of other proteins has been implicated in triggering basal defenses (also known as pathogen-triggered immunity, PTI) (McCann et al. 2012; Thomma et al. 2011). Moreover, at least in a subset of plant species like tomato, other portions of flagellin (flg28) can also be recognized (Cai et al. 2011). Although natural variation in PAMP regions exists across *P. syringae* strains and pathovars, and these regions appear to have been the target of diversifying selection, the contribution of amino acid diversity of bacterial PAMPs to structuring host range remains unclear (McCann et al. 2012).

An emerging trend, however, is that receptor-like kinases that respond to specific peptides may not be conserved across all plant species, so that PAMP recognition may differ in a qualitative way across plant hosts (Segonzac and Zipfel 2011; Thomma et al. 2011). For example, bacterial elongation factor Tu (EfTu) can act as a PAMP (peptides are named elf18 and elf26) within *Brassicaceae* hosts, but does not in Solanaceous hosts like tobacco and tomato (Lacombe et al. 2010). Therefore, *P. syringae* isolates that are pathogenic on *Brassicaceae* must overcome recognition that isolates from other pathovars may not experience. The ability of such plant clade-specific receptors to structure host range is apparent, and indeed, genetic modification is already focused on moving specific PAMP receptors across plant families with the hopes of creating durable resistance (Lacombe et al. 2010). While the total number and phylogenetic distributions of PAMP receptors remain unknown, it is highly likely that there are additional receptors like EFR that can limit natural bacterial host range.

Immune responses to PTI include callose deposition, production of reactive oxygen species, and deployment of antimicrobial compounds (Jones and Dangl 2006; Nicaise et al. 2009). *P. syringae* strains must be able to cope with or avoid these stresses during infection, and it is possible that clade- or isolate-specific variation in proteins like catalase could mediate differential interactions inside of hosts. For instance, glucosinolate production within *Arabidopsis* strongly limits the growth of *P. syringae* (Fan et al. 2011). Interestingly, glucosinolate expression is not thought to be involved in the PTI response, as precursors are held in vacuoles until these containers are disrupted by mechanical forces like chewing insects. Bacterial isolates that are pathogenic on *Arabidopsis* have acquired multiple efflux systems through horizontal transfer that enable higher tolerance toward glucosinolates, but these systems are not found extensively throughout *P. syringae*. The overall contribution of such plant species with specific responses to host range limitation has not been thoroughly investigated across hosts and bacterial isolates.

Pseudomonas syringae uses a type III secretion system to translocate effector proteins into host cells, which is essential for virulence as described below, but recognition of these effector proteins (or their actions) by plant R genes can limit host range (referred to as effector-triggered immunity, ETI) (Jones and Dangl 2006; Thomma et al. 2011). The outcome of ETI is the plant hypersensitive response (HR), which involves localized cell death to limit pathogen growth and spread. At one level, ETI is responsible for establishing race structure within pathovars of *P. syringae* based on differential recognition of variable effector proteins across

cultivars of a given plant species (Taylor et al. 1996). At another level, ETI can contribute to limiting the host range of pathovars across plant species. For instance, recognition of the effector AvrPtoB has been demonstrated to limit the growth and decrease symptoms of multiple pathovars of *P. syringae* on tomato (Chien et al. 2013; Lin and Martin 2007). Likewise, recognition of HopQ1 renders *Pto*DC3000 avirulent and may partially explain lack of growth of *Pph*1448a on tobacco (Ferrante et al. 2009; Wei et al. 2007). When considering the contribution of ETI to host range limitation overall, however, one must consider the distribution and conservation of known avirulence factors across all isolates within a pathovar and comparisons of diversity for R genes across plant species and populations. Effectors have been shown to trigger HR reactions in every tested cultivar of a plant species for only a handful of cases, but if such trends hold up with further sampling, these effector proteins may truly limit host range at the level of plant species (Arnold et al. 2001; Wroblewski et al. 2009). Consideration of how such interactions structure host range may be especially important when two pathovars have recently diverged to infect different hosts, as with HopC1 in pathovars *Pgy* and *Pph* (Baltrus et al. 2012), as a recent report suggests that ETI is the first stage of host differentiation between pathovars (Schulze-Lefert and Panstruga 2011).

3.4 Genomic Factors that Promote Host Range Expansion

All known phytopathogenic *P. syringae* isolates require a type III secretion system for virulence (O'Brien et al. 2011; Tampakaki et al. 2010). This TTSS was acquired by an immediate progenitor of all *P. syringae* and *P. cannabina*, likely from a clade within *P. viridiflava* (Araki et al. 2006; Sarris et al. 2013). Although localized recombination within the TTSS has been observed, the system as a whole has been vertically inherited since this introduction (Sarris et al. 2013). While avirulence due to TTEs limits host range in an R gene-dependent way, in the absence of ETI, these translocated proteins are absolutely essential for bacterial growth and disease progression *in planta* (Block and Alfano 2011; Lindeberg et al. 2012). Moreover, the actions of some TTE can cover up ETI triggered by other TTEs (Jones and Dangl 2006).

Each strain of *P. syringae* possesses a handful of conserved effector genes that, although they may not be functional within a strain, have been vertically inherited from an ancestor of all virulent *P. syringae* including *hopI1, hopAH2, avrE, hopAA1,* and *hopM1* (Baltrus et al. 2011; O'Brien et al. 2011). Such conservation implies a foundational role for these proteins across a wide array of plant hosts. In addition to this small number of conserved effector proteins, each strain possesses from 3 to 36 (or more) additional effector proteins which are frequently lost through deletion or gained by horizontal gene transfer (Baltrus et al. 2011; Lindeberg et al. 2012). While an increasing amount of information about the precise functions of effector proteins has been summarized elsewhere, common themes emerge by comparing these isolates and exploring what these trends say about the evolution of virulence and host range.

Studies within one strain, *Pto*DC3000, across multiple host plants have demonstrated that effector proteins can be clustered into tiers based on function (Cunnac et al. 2011). Some TTEs can further be classified into redundant effector groups (REG) based on overlapping virulence function and ability to complement phenotypic virulence effects (Kvitko et al. 2009). Single knockouts of genes within a REG yield no (or very small) virulence defects, so that only by disrupting each member of a REG is virulence noticeably changed. Two main REGs have been identified at present within *P. syringae*: AvrE/HopM/HopR appears to disrupt secretion systems involved in plant defenses, while AvrPto/AvrPtoB disrupts perception and signaling after the recognition of bacterial PAMPs (Kvitko et al. 2009; Lin and Martin 2005; Badel et al. 2006). As a demonstration of effector tiers, Cunnac et al. (2011) showed that, if single members of either of these REGs were present

within *Pto*DC3000, a variety of other type III effector proteins could quantitatively and interchangeably contribute to virulence. It is unclear how conserved REG functions are across strains and host plants. However, as Baltrus and colleagues showed, the placement of AvrPto from the soybean pathogen *Pgy* into a virulent bean pathogen (*Pph*1448a) could increase virulence on the target host species even though *Pph*1448a already contains AvrPtoB (Baltrus et al. 2012). It will be interesting to see how the REG concept and tiers of effector proteins are refined with future studies.

Type III effectors are not the sole determinant of success and failure during infections by *P. syringae* as many strains also produce toxins and phytohormones that manipulate plant physiology (Bender et al. 1999; Melotto et al. 2006; Schellenberg et al. 2010), including polyketide synthase (PKS) and non-ribosomal peptide synthase (NRPS) pathways. One of the best studied of these PKS-derived phytohormone mimics is coronatine, which binds to jasmonic acid receptors, and is the product of two multi-gene pathways that can be horizontally transferred across strains (Alarcón-Chaidez et al. 1999; Melotto et al. 2006). While the best known function of coronatine involves stomatal manipulation, which can also be altered by the TTE AvrB or the toxin syringolin, additional evidence points to secondary virulence effects for coronatine after bacterial cells reach the apoplast (Geng et al. 2012; Melotto et al. 2006). One unexplained trend that has emerged from genomic comparisons across *P. syringae* is that group II strains contain a lower number of effector proteins on average than any of the other groups. Parsimonious reconstruction across TTE groups suggests that the ancestor of group II contained additional TTE, so that this trend is the outcome of gene loss [at a minimum loss of gene families *hopR*, *hopAS*, *hopQ*, and *hopD,* with more complicated histories for *avrpto* and *hopAB* (Baltrus et al. 2011)]. This pattern is not an artifact of sampling, but correlates with the presence of toxin pathways for syringolin, syringomycin, and syringopeptin. It is tantalizing to speculate that these strains have replaced virulence functions of TTE with those of toxins, especially because syringolin has recently been implicated both in manipulating stomatal function and in preventing ETI in *Psy*B728a (Schellenberg et al. 2010). Furthermore, an independent data point suggesting that toxins replace TTE function involves pv. *pisi* strains which appear to have secondarily lost all three of these toxins but have gained back numerous effector groups (Baltrus et al. 2011). It is difficult to know whether these toxins can each identically replace the "lost" effector groups as well as how toxins factor into strain specificity. Also of note when discussing phytohormone manipulation by *P. syringae* is that strains that form galls on woody plants contain additional copies of pathways to produce the phytohormone auxin, which directly contributes to plant tissue growth and is required for gall formation along with the TTSS (Rodríguez-Palenzuela et al. 2010). Other toxins include phaseolotoxin, mangotoxin, and tabtoxin, which are found in a limited number of strains, and all appear to contribute to disease symptoms like chlorosis but not necessarily growth *in planta* (Bender et al. 1999; Carrion et al. 2013; Groll et al. 2008).

A small subclade within group II strains (referred to as group IIC) has replaced the original TTSS with a phage sequence and has actually gained an alternative type III secretion system at a different location of the genome (Clarke et al. 2010; Mohr et al. 2008). While divergent in protein sequence from the canonical *P. syringae* TTSS, this second system most closely resembles structures used to translocate proteins across plant cells and these strains were shown to be able to grow *in planta* in a TTSS-dependent way (Clarke et al. 2010). Moreover, group IIC strains can be found epiphytically on plants across a wide range of environments and hosts (Diallo et al. 2012b; Kniskern et al. 2010). The precise function of this alternative type III secretion system is not known, but it could contribute to pathogenicity during this epiphytic stage, playing a more subtle ecological role than the canonical system, or may be used in an entirely different way during interactions with non-plant hosts. Of note is that this secondary

system within group IIC does not appear to be as finely regulated at a transcriptional level when compared to the canonical system (Clarke et al. 2010). Lastly, some strains within groups III and IV have been shown to harbor a second type III secretion system within their genomes (O'Brien et al. 2011). This system does not appear to be involved in phytopathogenicity, and its true function is not known.

3.5 Expansion of Host Range Outside of Plants

Although the most widely studied aspects of *P. syringae* evolution and ecology involve phytopathogenesis, a growing body of work suggests that certain members of this species can survive and thrive in non-plant hosts (Stavrinides et al. 2009). Indeed, such interactions may be a critical aspect of environmental persistence and could structure ecological relationships between strains and species in previously unrecognized ways. For instance, *P. syringae* is routinely exposed to insects throughout the environment and can be isolated from inside-surface-sterilized insects that have fed on plants (Stavrinides et al. 2009). Insects are a prevalent part of the life cycles for numerous plant pathogens, but the potential for insects to act as vectors or hosts of *P. syringae* has received comparatively little attention (Nadarasah and Stavrinides 2011). However, multiple strains possess a demonstrated capacity to grow to high densities inside of pea aphids (*Acyrthosiphon pisum*) and can be deposited on new plants through excreted honeydew. Intriguingly, strain-specific entomopathogenicity has also been demonstrated, as *Psy*B728a but not *Pto*DC3000 has the potential to kill up to 95 % of aphids within 48 h of infection (Stavrinides et al. 2009).

The ecological importance and relevance of these interactions are currently unclear, but these early results suggest fascinating directions for future study. Firstly, hemipteran insects, like aphids, are excellent candidates for bacterial hosts/vectors because they feed on and excrete carbohydrate-rich plant sap. They also continually probe plant surfaces during feeding, which increases the chance of encountering epiphytic bacteria (Stavrinides et al. 2009). However, it is currently unknown whether *P. syringae* can successfully infect hemipteran insects in addition to aphids. Also unknown is the genetic and genomic basis of persistence inside of insects and entomopathogenicity. While Stavrinides et al. (2009) were able to demonstrate that a functioning flagellum is important for establishing insect infection, they also ruled out contributions from all recognizable entomotoxins within the genome sequence compared to other insect pathogens. Given the wealth of genomic information available for *P. syringae,* a promising future research direction could involve evolutionary and phylogenetic comparisons of entomopathogenicity. Such experiments could shed light on both the genetic basis and evolutionary mechanisms enabling insect killing. Are similar genes involved in virulence on plants versus insects, or is insect virulence a newly evolved trait? Do strains that are able to make use of insect hosts have specific adaptations associated with the phenotype?

P. syringae genomes also harbor multiple independent pathways that mediate interactions with fungi. Notably, many group II strains are demonstrated producers of toxins like syringomycin, and the bioassay for production of these molecules involves suppression of growth of the fungus *Geotrichum candidum* (Quigley and Gross 1994; Scholz-Schroeder et al. 2001). This toxin is assembled through an NRPS pathway and also contributes to plant infection and to disease symptom development. Another gene present within many *P. syringae* isolates, *phcA,* has been shown to induce a cell death response in *Neurospora crassa,* but an ecological role for this protein remains unclear (Wichmann et al. 2008). Lastly, most (if not all) *P. syringae* genomes contain at least one type VI secretion system (HSI-I) and associated effector molecules, while strains *Pto*DC3000 and *P. cannabina* pv. *alisalensis* contain a second type VI secretion system (HSI-II) (Records and Gross 2010; Sarris et al. 2010, 2013). Both HSI-I and HSI-II enable *P. syringae* to outcompete fungi,

but very little is known about how such pathways structure natural communities (Haapalainen et al. 2012).

3.6 Competition with Other Bacteria

Syringopeptin and HSI-II can also be used to influence interactions between *P. syringae* and other bacterial species (Haapalainen et al. 2012; Scholz-Schroeder et al. 2001), and all isolates likely contain additional pathways to produce narrow spectrum antimicrobial compounds termed bacteriocins. Bacteriocins can take a variety of forms, but all are active only against strains closely related to the producer. Readers are directed to recent reviews (and references therein) on bacteriocins produced by phytopathogens, including *P. syringae,* for background on bacteriocin biology (Grinter et al. 2012; Holtsmark et al. 2008). In particular, bacteriocin activity has been demonstrated within multiple *P. syringae* isolates including pathovars *syringae*, *morsprunorum*, *glycinea*, *phaseolicola*, *ciccaronei*, and *tomato* (Barreteau et al. 2009; Garrett et al. 1966; Ghequire et al. 2012; Lavermicocca et al. 1999; Vidaver et al. 1972). Additionally, various bacteriocins have been predicted within the genomes of *Psy*B728a, *Pph*1448A, *Pto*DC3000, and *P. syringae* pv. aptata (Ghequire et al. 2012; Holtsmark et al. 2008; Parret and De Mot 2002).

Bacteriocins can influence the phyllosphere community as well as plant disease in several ways. Most directly, they could promote disease by providing potential pathogens with an epiphytic advantage by preferentially excluding closely related strains [which are likely to occupy a similar or identical niche (Wilson and Lindow 1994)]. Conversely, nonpathogenic strains could be used to reduce plant disease through preferential exclusion of pathogens. Indeed, Lavermicocca et al. (2002) found that application of a currently undetermined bacteriocin from *P. syringae* pv. *ciccaronei* reduced olive knot disease caused by *P. savastanoi* (Lavermicocca et al. 2002). Additionally, Garrett et al. (1966) found a general trend where *P. syringae* isolates recovered from citrus tended to produce bacteriocins active against isolates recovered from pear and vice versa. Whether such a pattern is common, is the result of long-term evolutionary forces, or is stochastic will require further research.

In addition to their direct killing effect, bacteriocins may influence the ecology of *P. syringae* by targeting receptors important for nutrient uptake, like TonB-dependent siderophore receptors (Davies and Reeves 1975; Denayer et al. 2007). Indeed, selection for bacteriocin resistance negatively affected the ability of *Erwinia chrysanthemi* (*Dickeya dadantii*) to cause soft rot in *Saintpaulia* plants, with a correlated loss of low-iron-induced outer membrane proteins (Expert and Toussaint 1985). In this way, bacteriocins could influence evolutionary patterns across *P. syringae* genomes, by directly altering selective pressures on targets like siderophore uptake receptors. While, to our knowledge, there are currently no receptors described as mediating bacteriocin sensitivity in *P. syringae*, identification of such receptors will significantly aid in understanding the influences of bacteriocins on *P. syringae* ecology.

3.7 Survival in the Environment

Pseudomonas syringae strains can survive epiphytically on the outside of plant leaves and can be isolated from environmental sites such as leaf litter, rivers, lakes, snowpack, and even from clouds (Morris et al. 2008, 2010). While some genetic factors that facilitate survival across such diverse habitats have been well studied [i.e., UV resistance (Kim and Sundin 2000; Zhang and Sundin 2004)], whole-genome analyses and intense sampling schemes have only raised additional ecological questions about persistence of *P. syringae* outside of disease-causing infections. Indeed, each genome contains many ECF sigma factors, regulatory targets of which are unknown, that appear to be dispensable for phytopathogenesis but could enable environmental survival (Oguiza et al. 2005; Thakur et al. 2013).

Although epiphytic and apoplastic populations of *P. syringae* experience very different environments, the type III secretion system and associated effectors appear to play critical roles in epiphytic survival. Recent work by Yu et al. (2013) has begun to illuminate the traits harbored by *P. syringae* that contribute to both epiphytic and apoplastic fitness, as well as those traits that appear to be utilized in only one of the two environments. Confirming previous research that showed that motility was an important epiphytic trait, genes encoding flagellar components, chemosensory and chemotaxis proteins, *rhlA*, which encodes, and enzyme that produces 3-(3-hydroxyalkanouloxy) alkanoic acid (HAA), a swarming motility-enabling surfactant (Burch et al. 2012), were upregulated in the epiphytic environment compared to the apoplastic environment. Taken together, the data suggest that motility is important for the epiphytic colonization, but does not play a prominent role in apoplastic colonization [with the caveat that motility aids in initial invasion (Hattermann and Ries 1989; Panopoulos and Schroth 1974)]. Supporting the idea that group II strains have transitioned from relying on diverse TTEs to phytotoxins as mediators of virulence, secondary metabolites, including many NRPSs and PKSs, were all more highly expressed apoplastically than epiphytically. Interestingly, genes associated with mobile elements (prophage, insertion sequences, and transposases) were also more highly expressed within the apoplast than on the leaf surface. This finding correlates with recent findings in *Pph*1448A, where a genomic island (PPHGI-1) readily transferred between strains co-infiltrated into bean leaves, but did not readily transfer between strains co-inoculated into culture media (Lovell et al. 2009). Taken together with the knowledge that horizontal gene transfer has significantly contributed to evolution across *P. syringae* isolates, these results might suggest that gene transfer readily occurs between bacteria co-colonizing a leaf apoplast. As might be expected, metabolism was also influenced by association with the plant host. Pathways involved in the uptake and catabolism of GABA, phenylalanine, as well as other amino acids and sugars were induced *in planta* compared to basal conditions (Yu et al. 2013). Some pathways were more highly induced in the apoplast (GABA), while others were more highly induced epiphytically (phenylalanine). Perhaps one of the most dominant environmental forces inside and outside of leaves is water limitation, a trend which stands out in both phenotypic assays and transcriptome studies, but very little is known across strains about survival under water stress. While there is now a wealth of data regarding expression patterns in two important plant environments (the leaf surface and leaf interior), it will be interesting to compare such patterns across strains as well as plant hosts to determine which expression profiles are largely conserved or are more specific to a particular strain or plant host. For instance, are traits that are generally conserved across *P. syringae* pathovars equally conserved in their regulation?

A phenotype often associated with virulence and success in plants, the ability to attach to and aggregate on surfaces, is also required for *P. syringae's* success in epiphytic and environmental habitats (Hirano and Upper 2000). Attachment and aggregation on surfaces not only prevent cells from being physically removed from a surface (i.e., being washed away), but also help *P. syringae* resist a variety of environmental stresses such as desiccation, exposure to UV light and reactive oxygen species, and competition from other microbes. However, the ability to aggregate on leaf surfaces is not required for *in planta* virulence and all strains are not equally proficient at forming environmental aggregations (Hirano and Upper 2000; Lindow et al. 1993).

Exactly how *P. syringae* attaches to surfaces is not well understood across strains, but can involve the secretion of extracellular polymeric substances (EPS) to form aggregations or biofilms (Hirano and Upper 2000; Laue 2006; Lindeberg et al. 2008). A variety of polymers have been proposed to be important in *P. syringae* biofilm formation. Epiphytic fitness has been best linked to production of the polysaccharide alginate (Laue 2006; Quiñones et al. 2005), the production of which is upregulated in

epiphytic populations compared to apoplastic populations within strain *Psy*B728a (Yu et al. 2013), although other polysaccharides such as levan or cellulose may also be important (Laue 2006; Lindeberg et al. 2008; Ude et al. 2006).

A major problem with uncovering the genomic basis for surface aggregation in *P. syringae* is that many of the genes likely to be involved are widespread and conserved across strains. However, although regulation of EPS pathways seems to be tied to quorum sensing and cell density (Quiñones et al. 2005), differences in expression levels of pathways involved in bacterial aggregation may ultimately underlie phenotypic differences. For instance, while *Pto*DC3000 and the notable biofilm-forming human pathogen *P. aeruginosa* both possess full operons for cellulose production, they differ in regulatory genes for the operon (Lindeberg et al. 2008). Possibly, pairing transcriptome analyses with genomic analyses will help to identify adaptations useful in surface aggregation. This approach would be particularly useful when compared across strains that vary in environmental niches. Another area that remains relatively unexplored is the formation of aggregations on other environmental surfaces besides plant leaves. Environmentally isolated *P. syringae* strains have been found to use cellulose as a biofilm matrix (Ude et al. 2006), but it is not known whether they also use polysaccharides produced on leaves and how the production of these compounds may be regulated in diverse environments.

3.8 Dispersal

Global dispersal of *P. syringae* is thought to be intricately linked to the water cycle, where long-distance transfer takes place by cells or aggregations of bacteria being swept up into clouds and transported by meteorological forces (Morris et al. 2008). One of the most important pathways contributing to this dispersal process for many plant-associated bacteria is ice nucleation proteins (Morris et al. 2010). However, there is substantial variability in the protein composition and nucleation properties of ice crystals across strains, with some isolates like *Pph*1448a lacking production capabilities due to disruption by insertion elements. Moreover, some strains like *Pto*DC3000 display ice nucleation capabilities even though they lack *inaZ* (Feil et al. 2005; Hwang et al. 2005; Joardar et al. 2005).

Some pathovars of *P. syringae* may rely more heavily on seed-borne rather than environmental dispersal, and it is currently unknown how such ecological changes manifest at the level of the genome. For example, pathovar *phaseolicola* is readily dispersed through seeds and has a substantially smaller nutritional range than other strains like *P. syringae* pv. *syringae* B728a or pv. *tomato* DC3000 (Arnold et al. 2011; Rico and Preston 2008). These changes could be the product of selection for mutational deterioration of unused molecular pathways, but may also reflect differences in population size and genetic drift. Although just a correlation at this point, such changes in addition to the disruption of *inaZ* could explain why subclades within group III strains in particular appear to be under-sampled when it comes to environmental reservoirs of *P. syringae* (Morris et al. 2008, 2010). Also unclear are specific virulence factors that enable strains to facilitate seed dispersal.

That iron limitation is a strong selective force for environmental pseudomonad populations is reflected in the diverse array of siderophores that have evolved to scavenge this scarce element. Most *P. syringae* isolates studied to date have the genetic capability to produce two, if not more, high-affinity iron uptake systems (see below). The most widely studied siderophore produced by *P. syringae* is pyoverdine (Fig. 3.2), which is encoded by a NRPS pathway and imparts the characteristic fluorescence to strains of this species as well as related "fluorescent" pseudomonads (Visca et al. 2007). This class of siderophores has been studied for several decades, largely in *Pseudomonas aeruginosa*, where much has been learned with regard to the genetics, genomics, and biochemistry of pyoverdine production and transport [reviewed in Budzikiewicz (2004), Gross and Loper

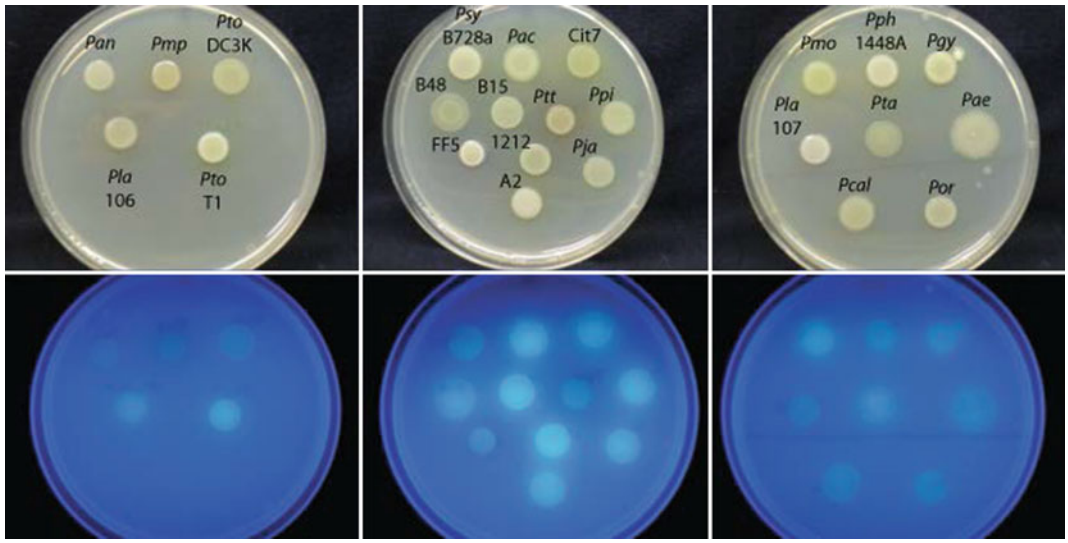

Fig. 3.2 Diverse colony morphologies and fluorescent pigment (pyoverdine) production across *Pseudomonas syringae* pathovars. Five microliters of overnight KB broth cultures was inoculated onto solid KB medium and incubated at room temperature for 48 h. Pictures of plates were taken without (*top panels*) or with UV illumination (*bottom panels*). MLST group I strains: *Pan* (pv. *actinidiae*), *Pmp* (pv. *morsprunorum*), *Pto* DC3 K (pv. *tomato* DC3000), *Pla* 106 (pv. *lachrymans* 106), and *Pto* T1 (pv. *tomato* T1); MLST group II strains: *Psy* B728a (pv. *syringae* B728a), *Pac* (pv. *aceris*), Cit7 (*P. syringae* Cit7), B48 (*P. syringae* B48), B15 (*P. syringae* B15), *Ptt* (pv. *aptata*), *Ppi* 1704 (pv. *pisi*), FF5 (pv. *syringae* FF5), 1212 (pv. *syringae* 1212), *Pja* (pv. *japonica*), and A2 (pv. *syringae* A2); MLST group III strains: *Pmo* (pv. mori), *Pph* 1448A (pv. *phaseolicola* 1448A), *Pgy* (pv. *glycinea*), *Pla* 107 (pv. *lachrymans* 107), *Pta* (pv. *tabaci*), and *Pae* (pv. *aesculi*); MLST group V *Pcal* (pv. *alisalensis,* but also known as *P. syringae pv. maculicola* ES4326), MLST group IV *Por* (pv. oryzae)

(2009), Meyer (2000), Visca et al. (2007)]. Interestingly, although three structurally different pyoverdines are produced among different strains of *P. aeruginosa* (Cornelis et al. 1989), which is mirrored by allelic diversity in the genes responsible for pyoverdine synthesis and uptake (Bodilis et al. 2009; Smith et al. 2005; Spencer et al. 2003), pyoverdines produced by different strains of *P. syringae* are largely identical (Bultreys and Gheysen 2000; Bultreys et al. 2001, 2003; Jülich et al. 2001). These observations suggest that while pyoverdine synthesis and uptake are under diversifying selection in *P. aeruginosa*, such selection does not occur in *P. syringae*. The difference in selection pressure between these two species is likely a result of differences between their ecologies.

While pyoverdine produced across *P. syringae* pathovars is identical (or nearly so), and presumably the underlying genetics governing synthesis, export, and import is highly similar, the regulation of pyoverdine production and perception may differ between pathovars (see Fig. 3.1 for example). For instance, evidence suggests that production of pyoverdine and achromobactin in *P. syringae* pv. *syringae* 22d/93 was much more sensitive to culture conditions (presumably oxygen availability) than *P. syringae* pv. *glycinea* 1a/96 (Wensing et al. 2010). Indeed, the one clear example of pyoverdine playing a role in virulence toward its host comes from pv. *tabaci*, where production of pyoverdine was shown to play a signaling role (Taguchi et al. 2010). Taguchi et al. (2010) demonstrated that disruption of pyoverdine production influenced the regulation of EPS production, surfactant production, swarming motility, AHL production, antibiotic resistance, and tabtoxin production. Presumably, this global regulatory role of pyoverdine production is not

generally conserved among all pathovars, or one would expect similar results to be observed by Jones and Wildermuth (2011).

P. syringae isolates may also contain one or several non-NRPS-derived siderophores including achromobactin (pv. *actinidiae*, pv. *syringae*, and pv. *phaseolicola*), yersiniabactin (pv. *tomato*, pv. *actinidiae*, and pv. *phaseolicola*), haemin (pv. *actinidiae*), enterobactin (pv. *actinidiae* and pv. *aesculi*), and citrate (Berti and Thomas 2009; Buell et al. 2003; Green et al. 2010; Jones and Wildermuth 2011; Owen and Ackerley 2011; Scortichini et al. 2012; Wensing et al. 2010). In addition to siderophore prediction from whole-genome analyses, yersiniabactin has either been detected directly using HPLC or is predicted to be encoded in several pathovars belonging to genomospecies 2, 3, 7, and 8 (Bultreys et al. 2001). The diversity of non-pyoverdine siderophores produced across *P. syringae* pathovars may be evolutionarily analogous to the diversity of pyoverdines produced across *P. aeruginosa* isolates.

In addition to receptors mediating uptake of endogenously produced siderophores, strains of *P. syringae* encode numerous receptors predicted to import non-native siderophores (Cornelis and Bodilis 2009; Cornelis and Matthijs 2002). Comparison of *Pto*DC3000, *Pph*1448A, and *Psy*B728a indicated that they share a common core set of 13 TonB-dependent uptake receptors, but collectively encode 29 such receptors. While cognate siderophores that interact with these receptors are largely unknown, it is tempting to speculate that this diversity is linked to each organism's distinct ecology.

Siderophores have been demonstrated to play a role in epiphytic fitness, where they likely function to aid *P. syringae* in acquiring otherwise unavailable iron (Karamanoli et al. 2011; Wensing et al. 2010). Indeed, whole-cell biosensors employed by Joyner and Lindow indicated that, like many nutrients, biologically available iron is heterogeneously distributed over a bean leaf surface (Joyner and Lindow 2000). Additionally, plants that harbor greater amounts of leaf surface polyphenolics (including tannin, an iron chelator) induced greater production of pyoverdine in *Psy*B728a, as well as inhibited growth of an isogenic mutant unable to produce pyoverdine, than plants with lower levels of leaf surface polyphenolics (Karamanoli et al. 2011). In contrast to *P. aeruginosa*, where siderophore production contributes to virulence (Meyer et al. 1996; Takase et al. 2000), siderophore production by *P. syringae* appears to be largely dispensable for infection of plant hosts (Cody and Gross 1987; Jones and Wildermuth 2011; Owen and Ackerley 2011). This is likely a result of sufficient available iron within the leaf interior [see references cited in (Jones and Wildermuth 2011)]. Indeed, recent transcriptomic work in *Psy*B728a found that iron acquisition systems were largely uninduced in the plant apoplast when compared with an iron-limiting medium (Yu et al. 2013). While no work has directly addressed the role of siderophores in the life history of *P. syringae* beyond the phyllosphere, presumably these high-affinity iron uptake systems are important for these stages of its life cycle. Moreover, it may be that the alternate siderophores produced across pathovars are uniquely adaptive to each strain's particular life cycle. Clearly, more experimental and comparative research is required to understand the adaptive basis for the diversity of alternate (non-pyoverdine) siderophores encoded by *P. syringae* pathovars.

While siderophores, and their associated uptake systems, are well understood for their role in acquiring scarce iron (and potentially other limiting metals) from the environment, one aspect of siderophore receptors that is completely unstudied is their role as receptors mediating phage and bacteriocin sensitivity. Similar to bacteriocins described above, phage–*P. syringae* interaction is likely to be an area where dedicated research is likely to improve our understanding of diversity-generating processes. Several prophage regions have been predicted in all of the fully sequenced *P. syringae* strains, though whether any of these regions is active (i.e., gives rise to an infective phage) is currently unreported in the literature. Previous research has indicated that some *P. syringae*

strains harbor active prophage (Garrett and Crosse 1963; Minor et al. 1996; Nordeen et al. 1983; Prior et al. 2007; Sato 1983), but the lack of genome sequences for these strains and prophage limits understanding of the underlying biology. It is tempting to speculate that *P. syringae*–phage interactions may be important in several ways given the recent results indicating that genes encoded in a prophage region of *Psy*B728a are induced under environmentally relevant conditions (Hockett et al. 2013; Yu et al. 2013). Additionally, recent work has demonstrated that there is an abundance of infective phage that can be isolated from the interior of horse chestnut leaves, which can replicate on *P. syringae* hosts (Koskella et al. 2011). Future research should aid in answering questions about the possible role(s) of phages in affecting the ecology and evolution of *P. syringae*. For instance, are phages important contributors to genetic exchange among *P. syringae* pathovars, similar to plasmids?

3.9 Concluding Remarks

The plummeting cost of DNA sequencing has enabled an explosion of genomic data for *P. syringae* strains and related species. While these sequences have provided numerous insights into the mechanisms of pathogenicity and environmental persistence for a handful of strains, ecological dynamics and mechanisms of survival outside of plant hosts remain a black box for the species as a whole. Moreover, it is far too easy to fall into treating *P. syringae* strains monolithically and assume what is true for one strain is true for all. In reality, vast phenotypic diversity exists across this species and ecological interactions could dramatically differ even for close relatives. While realization of these differences can enable more thorough ecological explorations, genomic-scale comparisons must be matched by greater sampling of strains outside of plant hosts and experimental tests of phenotypes to understand what traits are critical for environmental survival, persistence, and dispersal. Each lineage of *P. syringae* exhibits a highly complex evolutionary trajectory where selection acts on survival inside and outside of plant hosts and only incorporation of lineage-specific ecologies and appreciation of differences between strains can enable a complete understanding of genomic patterns and evolutionary dynamics across this species.

References

Alarcón-Chaidez FJ, Peñaloza-Vázquez A, Ullrich M, Bender CL (1999) Characterization of plasmids encoding the phytotoxin coronatine in *Pseudomonas syringae*. Plasmid 42:210–220

Almeida NF, Yan S, Cai R, Clarke CR, Morris CE, Schaad NW et al (2010) PAMDB, a multilocus sequence typing and analysis database and website for plant-associated microbes. Phytopathology 100:208–215

Araki H, Tian D, Goss EM, Jakob K, Halldorsdottir SS, Kreitman M, Bergelson J (2006) Presence/absence polymorphism for alternative pathogenicity islands in *Pseudomonas viridiflava*, a pathogen of *Arabidopsis*. Proc Natl Acad Sci USA 103:5887–5892

Arnold DL, Jackson RW, Fillingham AJ, Goss SC, Taylor JD, Mansfield JW, Vivian A (2001) Highly conserved sequences flank avirulence genes: isolation of novel avirulence genes from *Pseudomonas syringae* pv. *pisi*. Microbiology 147:1171–1182

Arnold DL, Lovell HC, Jackson RW, Mansfield JW (2011) *Pseudomonas syringae* pv. *phaseolicola*: from "has bean" to supermodel. Mol Plant Pathol 12:617–622

Badel JL, Shimizu R, Oh HS, Collmer A (2006) A *Pseudomonas syringae* pv. *tomato avrE1/hopM1* mutant is severely reduced in growth and lesion formation in tomato. Mol Plant Microbe Interact 19: 99–111

Baltrus DA, Nishimura MT, Dougherty KM, Biswas S, Muhktar S, Vicente JG et al (2012) The molecular basis of host specialization in bean pathovars of *Pseudomonas syringae*. Mol Plant Microbe Interact 25:877–888

Baltrus DA, Nishimura MT, Romanchuk A, Chang JH, Mukhtar MS, Cherkis K et al (2011) Dynamic evolution of pathogenicity revealed by sequencing and comparative genomics of 19 *Pseudomonas syringae* isolates. PLoS Pathog 7:e1002132

Barreteau H, Bouhss A, Fourgeaud M, Mainardi JL, Touze T, Gerard F et al (2009) Human- and plant-pathogenic *Pseudomonas* species produce bacteriocins exhibiting colicin M-like hydrolase activity towards peptidoglycan precursors. J Bacteriol 191:3657–3664

Bender CL, Alarcón-Chaidez F, Gross DC (1999) *Pseudomonas syringae* phytotoxins: mode of action,

regulation, and biosynthesis by peptide and polyketide synthetases. Microbiol Mol Biol Rev 63:266–292

Berti AD, Thomas MG (2009) Analysis of achromobactin biosynthesis by *Pseudomonas syringae* pv. *syringae* B728a. J Bacteriol 191:4594–4604

Block A, Alfano JR (2011) Plant targets for *Pseudomonas syringae* type III effectors: virulence targets or guarded decoys? Curr Opin Microbiol 14:39–46

Bodilis J, Ghysels B, Osayande J, Matthijs S, Pirnay J-P, Denayer S et al (2009) Distribution and evolution of ferripyoverdine receptors in *Pseudomonas aeruginosa*. Environ Microbiol 11:2123–2135

Budzikiewicz H (2004) Siderophores of the *Pseudomonadaceae* sensu stricto (fluorescent and non-fluorescent *Pseudomonas* spp.). Fortschr Chem Org Naturst 87:81–237

Buell CR, Joardar V, Lindeberg M, Selengut J, Paulsen IT, Gwinn ML et al (2003) The complete genome sequence of the *Arabidopsis* and tomato pathogen *Pseudomonas syringae* pv. *tomato* DC3000. Pr Proc Natl Acad Sci USA 100:10181–10186

Bultreys A, Gheysen I (2000) Production and comparison of peptide siderophores from strains of distantly related pathovars of *Pseudomonas syringae* and *Pseudomonas viridiflava* LMG 2352. Appl Environ Microbiol 66:325–331

Bultreys A, Gheysen I, Maraite H, de Hoffmann E (2001) Characterization of fluorescent and nonfluorescent peptide siderophores produced by *Pseudomonas syringae* strains and their potential use in strain identification. Appl Environ Microbiol 67:1718–1727

Bultreys A, Gheysen I, Wathelet B, Maraite H, de Hoffmann E (2003) High-performance liquid chromatography analyses of pyoverdin siderophores differentiate among phytopathogenic fluorescent *Pseudomonas* species. Appl Environ Microbiol 69:1143–1153

Burch AY, Shimada BK, Mullin SWA, Dunlap CA, Bowman MJ, Lindow SE (2012) *Pseudomonas syringae* coordinates production of a motility-enabling surfactant with flagellar assembly. J Bacteriol 194:1287–1298

Cai R, Lewis J, Yan S, Liu H, Clarke CR, Campanile F et al (2011) The plant pathogen *Pseudomonas syringae* pv. *tomato* is genetically monomorphic and under strong selection to evade tomato immunity. PLoS Pathog 7:e1002130

Carrion VJ, Gutierrez-Barranquero JA, Arrebola E, Bardaji L, Codina JC, de Vicente A et al (2013) The mangotoxin biosynthetic operon (mbo) is specifically distributed within *Pseudomonas syringae* genomospecies 1 and was acquired only once during evolution. Appl Environ Microbiol 79:756–767

Chien C-F, Mathieu J, Hsu C-H, Boyle P, Martin GB, Lin N-C (2013) Nonhost resistance of tomato to the bean pathogen *Pseudomonas syringae* pv. *syringae* B728a Is due to a defective E3 ubiquitin ligase domain in AvrPtoB B728a. Mol Plant Microbe Interact 26:387–397

Clarke CR, Cai R, Studholme DJ, Guttman DS, Vinatzer BA (2010) *Pseudomonas syringae* strains naturally lacking the classical *P. syringae hrp/hrc* locus are common leaf colonizers equipped with an atypical type III secretion system. Mol Plant Microbe Interact 23:198–210

Cody YS, Gross DC (1987) Outer membrane protein mediating iron uptake via pyoverdin$_{pss}$, the fluorescent siderophore produced by *Pseudomonas syringae* pv. *syringae*. J Bacteriol 169:2207–2214

Cornelis P, Bodilis J (2009) A survey of TonB-dependent receptors in fluorescent pseudomonads. Environ Microbiol Reports 1:256–262

Cornelis P, Matthijs S (2002) Diversity of siderophore-mediated iron uptake systems in fluorescent pseudomonads: not only pyoverdines. Environ Microbiol 4:787–798

Cornelis P, Hohnadel D, Meyer JM (1989) Evidence for different pyoverdine-mediated iron uptake systems among *Pseudomonas aeruginosa* strains. Infect Immun 57:3491–3497

Cunnac S, Chakravarthy S, Kvitko BH, Russell AB, Martin GB, Collmer A (2011) Genetic disassembly and combinatorial reassembly identify a minimal functional repertoire of type III effectors in *Pseudomonas syringae*. Proc Natl Acad Sci USA 108:2975–2980

Davies JK, Reeves P (1975) Genetics of resistance to colicins in *Escherichia coli* K-12: cross-resistance among colicins of group A. J Bacteriol 123:102–117

Denayer S, Matthijs S, Cornelis P (2007) Pyocin S2 (Sa) kills *Pseudomonas aeruginosa* strains via the FpvA type I Ferripyoverdine receptor. J Bacteriol 189:7663–7668

Diallo MD, Monteil CL, Vinatzer BA, Clarke CR, Glaux C, Guilbaud C, Desbiez CEC, Morris CE (2012a) *Pseudomonas syringae* naturally lacking the canonical type III secretion system are ubiquitous in nonagricultural habitats, are phylogenetically diverse and can be pathogenic. ISME J 6:1–11

Diallo MD, Monteil CL, Vinatzer BA, Clarke CR, Glaux C, Guilbaud C, Desbiez CEC, Morris CE (2012b) *Pseudomonas syringae* naturally lacking the canonical type III secretion system are ubiquitous in nonagricultural habitats, are phylogenetically diverse and can be pathogenic. ISME J 6:1325–1335

Expert D, Toussaint A (1985) Bacteriocin-resistant mutants of *Erwinia chrysanthemi*: possible involvement of iron acquisition in phytopathogenicity. J Bacteriol 163:221–227

Fan J, Crooks C, Creissen G, Hill L, Fairhurst S, Doerner P, Lamb C (2011) *Pseudomonas sax* genes overcome aliphatic isothiocyanate-mediated non-host resistance in *Arabidopsis*. Science 331:1185–1188

Feil H, Feil WS, Chain P, Larimer F, DiBartolo G, Copeland A et al (2005) Comparison of the complete genome sequences of *Pseudomonas syringae* pv. *syringae* B728a and pv. *tomato* DC3000. Proc Natl Acad Sci USA 102:11064–11069

Ferrante P, Clarke CR, Cavanaugh KA, Michelmore RW, Buonaurio R, Vinatzer BA (2009) Contributions of the effector gene *hopQ1-1* to differences in host range between *Pseudomonas syringae* pv. *phaseolicola* and *P.syringae* pv. *tabaci*. Mol Plant Pathol 10:837–842

Filiatrault MJ, Stodghill PV, Bronstein PA, Moll S, Lindeberg M, Grills G et al (2010) Transcriptome analysis of *Pseudomonas syringae* identifies new genes, noncoding RNAs, and antisense activity. J Bacteriol 192:2359–2372

Filiatrault MJ, Stodghill PV, Myers CR, Bronstein PA, Butcher BG, Lam H et al (2011) Genome-wide identification of transcriptional start sites in the plant pathogen *Pseudomonas syringae* pv. *tomato* str. DC3000. PLoS One 6:e29335

Fones H, Davis CAR, Rico A, Fang F, Smith JAC, Preston GM (2010) Metal hyperaccumulation armors plants against disease. PLoS Pathog 6:e1001093

Gardan L, Shafik H, Belouin S, Broch R, Grimont F, Grimont PA (1999) DNA relatedness among the pathovars of *Pseudomonas syringae* and description of *Pseudomonas tremae* sp. nov. and *Pseudomonas cannabina* sp. nov. (ex Sutic and Dowson 1959). Int J Syst Bacteriol 49:469–478

Garrett CM, Crosse JE (1963) Observations on lysogeny in the plant pathogens *Pseudomonas morsprunorum* and *P. syringae*. J Appl Microbiol 26:27–34

Garrett CM, Panagopoulos CG, Crosse JE (1966) Comparison of plant pathogenic pseudomonads from fruit trees. J Appl Microbiol 29:342–356

Geng X, Cheng J, Gangadharan A, Mackey D (2012) The coronatine toxin of *Pseudomonas syringae* is a multifunctional suppressor of Arabidopsis defense. Plant Cell 24:4763–4774

Ghequire MGK, Li W, Proost P, Loris R, De Mot R (2012) Plant lectin-like antibacterial proteins from phytopathogens *Pseudomonas syringae* and *Xanthomonas citri*. Environ Microbiol Reports 4:373–380

Green S, Studholme DJ, Laue BE, Dorati F, Lovell H, Arnold D et al (2010) Comparative genome analysis provides insights into the evolution and adaptation of *Pseudomonas syringae* pv. *aesculi* on *Aesculus hippocastanum*. PLoS One 5:e10224

Grinter R, Milner J, Walker D (2012) Bacteriocins active against plant pathogenic bacteria. Biochem Soc Trans 40:1498–1502

Groll M, Schellenberg B, Bachmann AS, Archer CR, Huber R, Powell TK et al (2008) A plant pathogen virulence factor inhibits the eukaryotic proteasome by a novel mechanism. Nature 452:755–758

Gross H, Loper JE (2009) Genomics of secondary metabolite production by *Pseudomonas* spp. Nat Prod Rep 26:1408–1446

Haapalainen M, Mosorin H, Dorati F, Wu RF, Roine E, Taira S et al (2012) Hcp2, a secreted protein of the phytopathogen *Pseudomonas syringae* pv. *tomato* DC3000, is required for competitive fitness against bacteria and yeasts. J Bacteriol 194:4810–4822

Harrison PW, Lower RPJ, Kim NKD, Young JPW (2010) Introducing the bacterial "chromid": not a chromosome, not a plasmid. Trends Microbiol 18:141–148

Hattermann DR, Ries SM (1989) Motility of *Pseudomonas syringae* pv. *glycinea* and its role in infection. Phytopathology 79:284–289

Hirano S, Upper C (2000) Bacteria in the leaf ecosystem with emphasis on *Pseudomonas syringae*—a pathogen, ice nucleus, and epiphyte. Microbiol Mol Biol Rev 64:624

Hockett KL, Burch AY, Lindow SE (2013) Thermoregulation of genes mediating motility and plant interactions in *Pseudomonas syringae*. PLoS One 8:e59850

Holtsmark I, Eijsink VGH, Brurberg MB (2008) Bacteriocins from plant pathogenic bacteria. FEMS Microbiol Lett 280:1–7

Hwang MSH, Morgan RL, Sarkar SF, Wang PW, Guttman DS (2005) Phylogenetic characterization of virulence and resistance phenotypes of *Pseudomonas syringae*. Appl Environ Microbiol 71:5182–5191

Jackson RW, Vinatzer B, Arnold DL, Dorus S, Murillo J (2011) The influence of the accessory genome on bacterial pathogen evolution. Mob Genet Elements 1:55–65

Joardar V, Lindeberg M, Jackson RW, Selengut J, Dodson R, Brinkac LM et al (2005) Whole-genome sequence analysis of *Pseudomonas syringae* pv. *phaseolicola* 1448A reveals divergence among pathovars in genes involved in virulence and transposition. J Bacteriol 187:6488–6498

Jones AM, Wildermuth MC (2011) The phytopathogen *Pseudomonas syringae* pv. *tomato* DC3000 has three high-affinity iron-scavenging systems functional under iron limitation conditions but dispensable for pathogenesis. J Bacteriol 193:2767–2775

Jones JDG, Dangl JL (2006) The plant immune system. Nature 444:323–329

Joyner DC, Lindow SE (2000) Heterogeneity of iron bioavailability on plants assessed with a whole-cell GFP-based bacterial biosensor. Microbiology 146:2435–2445

Jülich M, Taraz K, Budzikiewicz H, Geoffroy V, Meyer JM, Gardan L (2001) The structure of the pyoverdin isolated from various *Pseudomonas syringae* pathovars. Z Naturforsch C 56:687–694

Karamanoli K, Bouligaraki P, Constantinidou HIA, Lindow SE (2011) Polyphenolic compounds on leaves limit iron availability and affect growth of epiphytic bacteria. Ann Appl Biol 159:99–108

Kim JJ, Sundin GW (2000) Regulation of the rulAB mutagenic DNA repair operon of *Pseudomonas syringae* by UV-B (290 to 320 Nanometers) radiation and analysis of *rulAB*-mediated mutability in vitro and in planta. J Bacteriol 182:6137–6144

Kniskern JM, Barrett LG, Bergelson J (2010) Maladaptation in wild populations of the generalist plant pathogen *Pseudomonas syringae*. Evolution 65:818–830

Koskella B, Thompson JN, Preston GM, Buckling A (2011) Local biotic environment shapes the spatial scale of bacteriophage adaptation to bacteria. Am Nat 177:440–451

Kvitko BH, Park DH, Velásquez AC, Wei C-F, Russell AB, Martin GB et al (2009) Deletions in the repertoire of *Pseudomonas syringae* pv. *tomato* DC3000 Type III secretion effector genes reveal functional overlap among effectors. PLoS Pathog 5:e1000388

Lacombe SEV, Rougon-Cardoso A, Sherwood E, Peeters N, Dahlbeck D, van Esse HP et al (2010) Interfamily transfer of a plant pattern-recognition receptor confers broad-spectrum bacterial resistance. Nat Biotechnol 28:365–369

Laue H (2006) Contribution of alginate and levan production to biofilm formation by *Pseudomonas syringae*. Microbiology 152:2909–2918

Lavermicocca P, Lonigro SL, Evidente A, Andolfi A (1999) Bacteriocin production by *Pseudomonas syringae* pv. *ciccaronei* NCPPB2355. Isolation and partial characterization of the antimicrobial compound. J Appl Microbiol 86:257–265

Lavermicocca P, Lonigro SL, Valerio F, Evidente A, Visconti A (2002) Reduction of olive knot disease by a bacteriocin from *Pseudomonas syringae* pv. *ciccaronei*. Appl Environ Microbiol 68:1403–1407

Lin N-C, Martin GB (2005) An *avrpto/avrptoB* mutant of *Pseudomonas syringae* pv. *tomato* DC3000 does not elicit Pto-mediated resistance and is less virulent on tomato. Mol Plant Microbe Interact 18:43–51

Lin N-C, Martin GB (2007) Pto- and Prf-mediated recognition of AvrPto and AvrPtoB restricts the ability of diverse *Pseudomonas syringae* pathovars to infect tomato. Mol Plant Microbe Interact 20:806–815

Lindeberg M, Cunnac S, Collmer A (2012) *Pseudomonas syringae* type III effector repertoires: last words in endless arguments. Trends Microbiol 20:199–208

Lindeberg M, Myers CR, Collmer A, Schneider DJ (2008) Roadmap to new virulence determinants in *Pseudomonas syringae*: insights from comparative genomics and genome organization. Mol Plant Microbe Interact 21:685–700

Lindow SE, Andersen G, Beattie GA (1993) Characteristics of insertional mutants of *Pseudomonas syringae* with reduced epiphytic fitness. Appl Environ Microbiol 59:1593–1601

Lovell HC, Mansfield JW, Godfrey SAC, Jackson RW, Hancock JT, Arnold DL (2009) Bacterial evolution by genomic island transfer occurs via DNA transformation In Planta. Curr Biol 19:1586–1590

Ma Z, Smith JJ, Zhao Y, Jackson RW, Arnold DL, Murillo J, Sundin GW (2007) Phylogenetic analysis of the pPT23A plasmid family of *Pseudomonas syringae*. Appl Environ Microbiol 73:1287–1295

Marcelletti S, Ferrante P, Petriccione M, Firrao G, Scortichini M (2011) *Pseudomonas syringae* pv. *actinidiae* draft genomes comparison reveal strain-specific features involved in adaptation and virulence to *Actinidia* species. PLoS One 6:e27297

McCann HC, Nahal H, Thakur S, Guttman DS (2012) Identification of innate immunity elicitors using molecular signatures of natural selection. Proc Natl Acad Sci USA 109:4215–4220

Melotto M, Underwood W, Koczan J, Nomura K, He SY (2006) Plant stomata function in innate immunity against bacterial invasion. Cell 126:969–980

Meyer JM, Neely A, Stintzi A, Georges C, Holder IA (1996) Pyoverdin is essential for virulence of *Pseudomonas aeruginosa*. Infect Immun 64:518–523

Meyer J-M (2000) Pyoverdines: pigments, siderophores and potential taxonomic markers of fluorescent *Pseudomonas* species. Arch Microbiol 174:135–142

Minor SM, Nordeen RO, Pachall R (1996) Partial characterization of bacteriophages of *Pseudomonas syringae* pv. *tomato*. Proc Ark Acad Sci 50:137–140

Mohr TJ, Liu H, Yan S, Morris CE, Castillo JA, Jelenska J, Vinatzer BA (2008) Naturally occurring nonpathogenic isolates of the plant pathogen *Pseudomonas syringae* lack a type III secretion system and effector gene orthologues. J Bacteriol 190:2858–2870

Morris CE, Sands DC, Vanneste JL, Montarry J, Oakley B, Guilbaud C, Glaux C (2010) Inferring the evolutionary history of the plant pathogen *Pseudomonas syringae* from its biogeography in headwaters of rivers in North America, Europe, and New Zealand. mBio 1: e00107-10-e00107-20

Morris CE, Sands DC, Vinatzer BA, Glaux C, Guilbaud C, Buffière A et al (2008) The life history of the plant pathogen *Pseudomonas syringae* is linked to the water cycle. ISME J 2:321–334

Nadarasah G, Stavrinides J (2011) Insects as alternative hosts for phytopathogenic bacteria. FEMS Microbiol Rev 35:555–575

Nicaise V, Roux M, Zipfel C (2009) Recent advances in PAMP-triggered immunity against bacteria: pattern recognition receptors watch over and raise the alarm. Plant Physiol 150:1638–1647

Nordeen RO, Morgan MK, Currier TC (1983) Isolation and partial characterization of bacteriophages of the phytopathogen *Pseudomonas syringae*. Appl Environ Microbiol 45:1890–1898

Oguiza JA, Kiil K, Ussery DW (2005) Extracytoplasmic function sigma factors in *Pseudomonas syringae*. Trends Microbiol 13:565–568

Owen JG, Ackerley DF (2011) Characterization of pyoverdine and achromobactin in *Pseudomonas syringae* pv. *phaseolicola* 1448a. BMC Microbiol 11:218

O'Brien HE, Thakur S, Guttman DS (2011) Evolution of plant pathogenesis in *Pseudomonas syringae*: a genomics perspective. Ann Rev Phytopathol 49:269–289

O'Brien HE, Thakur S, Gong Y, Fung P, Zhang J, Yuan L et al (2012) Extensive remodeling of the *Pseudomonas syringae* pv. *avellanae* type III secretome associated with two independent host shifts onto hazelnut. BMC Microbiol 12:141

Panopoulos NJ, Schroth MN (1974) Role of flagellar motility in the invasion of bean leaves by *Pseudomonas phaseolicola*. Phytopathology 64:1389–1397

Parret AHA, De Mot R (2002) Bacteria killing their own kind: novel bacteriocins of *Pseudomonas* and other gamma-proteobacteria. Trends Microbiol 10: 107–112

Preston GM (2000) *Pseudomonas syringae* pv. *tomato*: the right pathogen, of the right plant, at the right time. Mol Plant Pathol 1:263–275

Prior ES, Andrews JA, Nordeen OR (2007) Characterization of bacteriophages of *Pseudomonas syringae* pv. *tomato*. Proc Ark Acad Sci 61:84–90

Quigley NB, Gross DC (1994) Syringomycin production among strains of *Pseudomonas syringae* pv. *syringae*: conservation of the *syrB* and *syrD* genes and activation of phytotoxin production by plant signal molecules. Mol Plant Microbe Interact 7:78–90

Quiñones B, Dulla G, Lindow SE (2005) Quorum sensing regulates exopolysaccharide production, motility, and virulence in *Pseudomonas syringae*. Mol Plant Microbe Interact 18:682–693

Ramos C, Matas IM, Bardaji L, Aragón IM, Murillo J (2012) *Pseudomonas savastanoi* pv. *savastanoi*: some like it knot. Mol Plant Pathol 13:998–1009

Records AR, Gross DC (2010) Sensor kinases RetS and LadS regulate *Pseudomonas syringae* type VI secretion and virulence factors. J Bacteriol 192:3584–3596

Rico A, Preston GM (2008) *Pseudomonas syringae* pv. *tomato* DC3000 uses constitutive and apoplast-induced nutrient assimilation pathways to catabolize nutrients that are abundant in the tomato apoplast. Mol Plant Microbe Interact 21:269–282

Rodríguez-Palenzuela P, Matas IM, Murillo J, López-Solanilla E, Bardaji L, Pérez-Martínez I, Rodríguez-Moskera ME, Penyalver R, López MM, Quesada JM, Biehl BS, Perna NT, Glasner JD, Cabot EL, Neeno-Eckwall E, Ramos C (2010) Annotation and overview of the *Pseudomonas savastanoi* pv. *savastanoi* NCPPB 3335 draft genome reveals the virulence gene complement of a tumour-inducing pathogen of woody hosts. Environ Microbiol 12:1–17

Sarris PF, Skandalis N, Kokkinidis M, Panopoulos NJ (2010) In silico analysis reveals multiple putative type VI secretion systems and effector proteins in *Pseudomonas syringae* pathovars. Mol Plant Pathol 11:795–804

Sarris PF, Trantas EA, Baltrus DA, Bull CT, Wechter WP, Yan S et al (2013) Comparative genomics of multiple strains of *Pseudomonas cannabina* pv. *alisalensis*, a potential model pathogen of both monocots and dicots. PLoS One 8:e59366

Sato M (1983) Phage-induction from lysogenic strains of *Pseudomonas syringae* pv. *mori* by the extract from mulberry leaves. Ann Phytopathol Soc Japan 49:259–261

Schellenberg B, Ramel C, Dudler R (2010) *Pseudomonas syringae* virulence factor syringolin A counteracts stomatal immunity by proteasome inhibition. Mol Plant Microbe Interact 23:1287–1293

Scholz-Schroeder BK, Hutchison ML, Grgurina I, Gross DC (2001) The contribution of syringopeptin and syringomycin to virulence of *Pseudomonas syringae* pv. *syringae* strain B301D on the basis of *sypA* and *syrB1* biosynthesis mutant analysis. Mol Plant Microbe Interact 14:336–348

Schulze-Lefert P, Panstruga R (2011) A molecular evolutionary concept connecting nonhost resistance, pathogen host range, and pathogen speciation. Trends Plant Sci 16:117–125

Scortichini M, Marcelletti S, Ferrante P, Petriccione M, Firrao G (2012) *Pseudomonas syringae* pv. *actinidiae*: a re-emerging, multi-faceted, pandemic pathogen. Mol Plant Pathol 13:631–640

Segonzac C, Zipfel C (2011) Activation of plant pattern-recognition receptors by bacteria. Curr Opin Microbiol 14:54–61

Smith EE, Sims EH, Spencer DH, Kaul R, Olson MV (2005) Evidence for diversifying selection at the pyoverdine locus of *Pseudomonas aeruginosa*. J Bacteriol 187:2138–2147

Spencer DH, Kas A, Smith EE, Raymond CK, Sims EH, Hastings M et al (2003) Whole-genome sequence variation among multiple isolates of *Pseudomonas aeruginosa*. J Bacteriol 185:1316–1325

Stavrinides J, McCloskey JK, Ochman H (2009) Pea aphid as both host and vector for the phytopathogenic bacterium *Pseudomonas syringae*. Appl Environ Microbiol 75:2230–2235

Studholme DJ (2011) Application of high-throughput genome sequencing to intrapathovar variation in *Pseudomonas syringae*. Mol Plant Pathol 12:829–838

Sundin GW (2007) Genomic insights into the contribution of phytopathogenic bacterial plasmids to the evolutionary history of their hosts. Ann Rev Phytopathol 45:129–151

Taguchi F, Suzuki T, Inagaki Y, Toyoda K, Shiraishi T, Ichinose Y (2010) The siderophore pyoverdine of *Pseudomonas syringae* pv. *tabaci* 6605 is an intrinsic virulence factor in host tobacco infection. J Bacteriol 192:117–126

Takase H, Nitanai H, Hoshino K, Otani T (2000) Impact of siderophore production on *Pseudomonas aeruginosa* infections in immunosuppressed mice. Infect Immun 68:1834–1839

Tampakaki AP, Skandalis N, Gazi AD, Bastaki MN, Panagiotis FS, Charova SN et al (2010) Playing the "Harp": evolution of our understanding of *hrp/hrc* genes. Ann Rev Phytopathol 48:347–370

Taylor J, Teverson D, Allen D (1996) Identification and origin of races of *Pseudomonas syringae* pv. *phaseolicola* from Africa and other bean growing areas. Plant Pathol 45:469–478

Thakur PB, Vaughn-Diaz VL, Greenwald JW, Gross DC (2013) Characterization of five ECF sigma factors in the genome of *Pseudomonas syringae* pv. *syringae* B728a. PLoS ONE 8:e58846

Thomma BPHJ, Nurnberger T, Joosten MHAJ (2011) Of PAMPs and effectors: the blurred PTI-ETI dichotomy. Plant Cell 23:4–15

Ude S, Arnold DL, Moon CD, Timms-Wilson T, Spiers AJ (2006) Biofilm formation and cellulose expression among diverse environmental *Pseudomonas* isolates. Environ Microbiol 8:1997–2011

Verdier V, Triplett LR, Hummel AW, Corral R, Cernadas RA, Schmidt CL et al (2012) Transcription activator-like (TAL) effectors targeting OsSWEET genes enhance virulence on diverse rice (*Oryza sativa*) varieties when expressed individually in a TAL effector-deficient strain of *Xanthomonas oryzae*. New Phytol 196:1197–1207

Vidaver AK, Mathys ML, Thomas ME, Schuster ML (1972) Bacteriocins of the phytopathogens *Pseudomonas syringae*, *P. glycinea*, and *P. phaseolicola*. Can J Microbiol 18(6):705–713

Visca P, Imperi F, Lamont IL (2007) Pyoverdine siderophores: from biogenesis to biosignificance. Trends Microbiol 15:22–30

Vivian A, Murillo J, Jackson RW (2001) The roles of plasmids in phytopathogenic bacteria: mobile arsenals? Microbiology 147:763–780

Wei C-F, Kvitko BH, Shimizu R, Crabill E, Alfano JR, Lin N-C et al (2007) A *Pseudomonas syringae* pv. *tomato* DC3000 mutant lacking the type III effector HopQ1-1 is able to cause disease in the model plant *Nicotiana benthamiana*. Plant J 51:32–46

Wensing A, Braun SD, Buttner P, Expert D, Völksch B, Ullrich MS, Weingart H (2010) Impact of siderophore production by *Pseudomonas syringae* pv. *syringae* 22d/93 on epiphytic fitness and biocontrol activity against *Pseudomonas syringae* pv. *glycinea* 1a/96. Appl Environ Microbiol 76:2704–2711

Wichmann G, Sun J, Dementhon K, Glass NL, Lindow SE (2008) A novel gene, *phcA* from *Pseudomonas syringae* induces programmed cell death in the filamentous fungus *Neurospora crassa*. Mol Microbiol 68:672–689

Wilson M, Lindow SE (1994) Coexistence among epiphytic bacterial populations mediated through nutritional resource partitioning. Appl Environ Microbiol 60:4468–4477

Wroblewski T, Caldwell KS, Piskurewicz U, Cavanaugh KA, Xu H, Kozik A et al (2009) Comparative large-scale analysis of interactions between several crop species and the effector repertoires from multiple pathovars of *Pseudomonas* and *Ralstonia*. Plant Physiol 150:1733–1749

Young JM (2010) Taxonomy of *Pseudomonas syringae*. J. Plant Pathol 92(1): S1-5–14

Yu X, Lund SP, Scott RA, Greenwald JW, Records AH, Nettleton D et al (2013) Transcriptional responses of *Pseudomonas syringae* to growth in epiphytic versus apoplastic leaf sites. Proc Natl Acad Sci USA 110:E425–E434

Zhang S, Sundin GW (2004) Long-term effect of mutagenic DNA repair on accumulation of mutations in *Pseudomonas syringae* B86-17. J Bacteriol 186:7807–7810

4 *Pseudomonas syringae* Genomics: From Comparative Genomics of Individual Crop Pathogen Strains Toward Population Genomics

Boris A. Vinatzer and Caroline L. Monteil

4.1 Who Is *Pseudomonas syringae*

Pseudomonas syringae sensu lato is a species that includes a genetically and phenotypically diverse group of Gammaproteobacteria, whose exact taxonomic status is still in flux. This group includes a few dozen causative agents of economically important crop diseases and a large number of genetically distinct strains isolated from wild plants, leaf litter, and compartments of the water cycle (e.g., clouds, precipitation, snowpack, lakes, and rivers) (Morris et al. 2008, 2010, 2013; Monteil et al. 2012). The named species most closely related to *P. syringae* are *P. cichorii* (Swingle 1925) and *P. viridiflava* (Burkholder 1930), which were originally described as species separate from *P. syringae sensu stricto* (Van Hall 1904) because of distinguishable phenotypic characteristics compared to *P. syringae sensu stricto*. Interestingly, even considering only crop pathogens, strains related to the *P. syringae*-type strain could be assigned to nine different species based on DNA similarity (or better dissimilarity). However, since no consistent phenotypic characteristics distinguish these nine groups, most of these groups could not be described as named species and are thus referred to as numbered "genomospecies" 1 through 9 (Gardan et al. 1999). It was recently shown that each genomospecies corresponds to a phylogenetic group (Bull et al. 2011) based on multilocus sequence typing (MLST) (Maiden et al. 1998).

Each genomospecies of *P. syringae* gathers different crop pathogens that are described based on their host range and the type of disease symptoms they cause. They are referred to as "pathovars" (Dye et al. 1980), and each pathovar is represented by a pathotype strain similar to type strains of named species. However, many genetic lineages of *P. syringae* have not been assigned to pathovars because either they did not cause disease on any tested plant species (Clarke et al. 2010) or simply because they were isolated from non-plant substrates and have not been tested for host range. They are currently simply assigned to phylogenetic groups based on MLST building on the MLST scheme originally developed by Sarkar and Guttman (2004).

While Chap. 3 describes and compares genomes of relatively distantly related strains belonging to different pathovars—or even different genomospecies—in this chapter, we will focus on how genome sequencing of very closely related strains belonging to the same phylogenetic group—or even the same pathovar—has started to give insight into various aspects of pathogen emergence, evolution, molecular

B. A. Vinatzer (✉) · C. L. Monteil
Department of Plant Pathology, Physiology, and Weed Science, Virginia Tech, Latham Hall, Ag Quad Lane, Blacksburg, VA, USA
e-mail: vinatzer@vt.edu

C. L. Monteil
INRA, UR0407 Pathologie Végétale, 84143, Montfavet cedex, France

D. C. Gross et al. (eds.), *Genomics of Plant-Associated Bacteria*,
DOI: 10.1007/978-3-642-55378-3_4, © Springer-Verlag Berlin Heidelberg 2014

host–microbe interactions, and ecology. We will also show how genome sequencing has the potential to transform plant disease diagnostics and unravel geographic routes of transmission. In the last part of the chapter, we will give our opinion on how we expect population genomics to give new insight into the life history and life cycle of *P. syringae* as long as we deeply sample from all environments in which *P. syringae* populations have been found to reside.

4.2 *Pseudomonas syringae*: One Population or Many?

Because strains that belong to *P. syringae* could actually be assigned to at least nine different species based on DNA similarity (Gardan et al. 1999), the question is how homogenous these strains are in regard to ecological niches they occupy, their interaction with other organisms, and their population structure. Importantly, is there one *P. syringae* population or are there many? In other words, does each pathovar represent a separate population or does each genomospecies or phylogenetic group represent a population or is there just one large *P. syringae* population?

This is not an easy question to answer since it is not even obvious what a bacterial population is. Sexual reproducing organisms belong to the same population if they interbreed. Different populations are isolated from each other, but occasional migration of individuals between populations may occur. In the case of bacteria, the boundaries of populations are not well defined. Although bacteria belonging to the same population can exchange DNA by horizontal gene transfer through site-directed recombination of genomic islands or homologous recombination of any genomic region, the relative contribution of recombination compared to mutation is different in different bacterial species. Absurdly, in bacteria, we may even have species corresponding to a "clonal" population, i.e., a population within which there is no recombination, which takes the original definition of "population" *ad absurdum*.

After performing a population genetic analysis of a collection of crop pathogens belonging to *P. syringae,* Sarkar and Guttman (2004) concluded that *P. syringae* consists of a population, in which members rarely recombine. Whole-genome sequencing (WGS) of multiple isolates of *P. syringae* pv. *aesculi* (*Pae*), the causative agent of bleeding canker of horse chestnut (Green et al. 2010), WGS of multiple isolates of *P. syringae* pv. *tomato* (*Pto*), the pathogen that causes bacterial speck disease around the world (Cai et al. 2011a), WGS of multiple isolates of *P. syringae* pv. *actinidiae* (*Psa*) from kiwifruit plants with bacterial canker disease (Mazzaglia et al. 2012), and WGS of multiple isolates of *Pseudomonas cannabina* pv. *alisalensis* (*Pal*) of diseased crucifers, tomato, or monocots (Sarris et al. 2013), confirmed that each of these pathovars consists of only one or a small number of clonal lineages. Each of these pathogens shows very little genetic variation between isolates. These pathogens could thus be defined as "genetically monomorphic." "Genetically monomorphic" pathogens were described by Mark Achtman as pathogens with very little genetic variation between isolates, suggesting that these pathogens only recently emerged, in the order of dozens, hundreds, or maybe thousands of years (Achtman 2008).

The interesting question is do genetically monomorphic pathogens represent clonal populations? We do not believe so. Yes, if a pathogen is genetically monomorphic, it has not significantly recombined with other bacteria since its emergence and, in that sense, it appears to have a clonal population structure. However, genetically monomorphic pathogens only very recently emerged. Therefore, it is not appropriate to compare a genetically monomorphic pathogen with a genetically diverse pathogen that emerged much earlier. In the first case, we look at a time scale of a few hundred or thousand years and in the second case at a time scale of a few hundreds of thousands or millions of years. Therefore, genetically monomorphic pathogens may be clonal populations when we choose to look at them individually over a window of hundreds of years, but they may be

part of recombining populations if we choose to look at them as a group of populations over a window of hundreds of thousands of years. In fact, a whole-genome phylogeny of isolates belonging to the genetically monomorphic *Pto* pathogen T1 and the *Pto* pathogens JL1065 and DC3000 suggests that the ancestors of these lineages recombined (our unpublished data). A very similar picture emerged when constructing a whole-genome phylogeny of *Psa* (McCann et al. 2013). In fact, the authors inferred that 10 % of each sequenced *Psa* genome is derived from recombination.

When performing MLST of *Pto* and of closely related crop pathogens in combination with isolates from compartments of the water cycle, it became evident that ancestors of crop pathogens and ancestors of environmental *P. syringae* isolates recombined (Monteil et al. 2013). One of the genes that was found to have recently recombined between crop strains and environmental isolates was the *fliC* gene encoding bacterial flagellin, which has an important role in plant–microbe interactions (Clarke et al. 2013).

Furthermore, since there appears to be relatively little recombination between the different phylogenetic groups of *P. syringae* based on MLST (Sarkar and Guttman 2004) but considerable recombination within groups (Cai et al. 2011b; Yan et al. 2008), we may conclude that the different phylogenetic groups within *P. syringae* each represents a separate population and that there is no overall *P. syringae* population.

Taking all these results together, we propose that the different *P. syringae* phylogenetic groups/genomospecies each represents separate recombining populations that consist of crop pathogens and related environmental strains. We further believe that genetically monomorphic crop pathogens occasionally emerge from these populations and then acquire an apparent "short-term" clonal population structure while spreading around the world from crop field to crop field [see also Monteil et al. (2013)]. Smith et al. (1993) defined such a combination of recombining and clonal populations as "epidemic population structure." However, to confirm this hypothesis, a full-fledged population genomic analysis of crop pathogens and environmental isolates will be necessary (see below).

4.3 Geographic Origin of *Pseudomonas syringae* Crop Pathogens and Their Routes of Transmission

For some fungal and oomycete pathogens, strong evidence has been found for their geographic origin, for example, *Phytophthora infestans* in the Andes of South America (Gomez-Alpizar et al. 2007; Stukenbrock and McDonald 2008). From South America, *P. infestans* seems to have migrated to North America and from there to Europe and, finally, to Ireland where it caused the Irish potato famine. Also for some bacterial pathogens, like *Xylella fastidiosa* subsp. *fastidiosa*, convincing evidence has been found for an origin in Central America and introduction into the USA has been hypothesized to have occurred with the importation of coffee plants from Costa Rica to California around 1880 (Nunney et al. 2010). Also for *Xanthomonas axonopodis* pv. *manihotis,* a South American origin seems likely (Bart et al. 2012). As expected, for some genetically monomorphic bacterial human pathogens, origin and international routes of transmission have been reconstructed in great detail. The most impressive example is the reconstruction of the international spread of *Yersinia pestis*, the causal agent of plague, out of China based on whole-genome sequences of hundreds of strains (Morelli et al. 2010; Cui et al. 2013). The conceptual basis behind such studies is relatively simple and is mainly based on one population genetic principle: The geographic origin of a species (or a population) is inferred to be where the highest genetic diversity is found (Stukenbrock and McDonald 2008). In fact, members of a population accumulate mutations over time. Therefore, where the pathogen has existed for the longest time, i.e., where it originated, pathogen isolates can be expected to be much more different from each other than in those geographic

areas where the pathogen was recently introduced. This is also due to the fact that the genetic diversity in the new area is derived from the usually small number of bacteria that migrated from the geographic origin of the pathogen (for example on a single lot of seeds). This phenomenon is called a "population bottleneck effect."

Another aspect to consider is that populations at the geographic origin diverged less from their ancestor than the pathogen populations in the geographic areas into which the pathogen was recently introduced. In fact, members of the original population will accumulate mutations, but as long as these mutations do not confer a strong selective advantage, these mutations will stay at a relatively low frequency in this population. In contrast, when a very small number of bacteria of a population are introduced into a new geographic area, the mutations present in these few bacteria will now be present in their entire progeny, that is, in all members of the population in this new geographic area. In other words, the introduced bacteria become the founders of the new population in the new geographic area. This process is known as the "founder effect," and the genetic differences between the population at the center of origin and the populations in the new geographic area are said to be due to "genetic drift" (instead of selection). Therefore, each time a few members of a population are introduced from one geographic area into the next, the genetic distance from the original population can be assumed to have increased. Importantly, the mutations present in each founder of each geographic transmission are preserved during the following geographic transmissions, making it possible to reconstruct the geographic routes of transmission by correlating phylogeny of isolates with their geographic origin.

Since recent estimates of yearly mutation rates for bacterial human pathogens are in the order of 1–10 mutation per million base pairs per year (for example, Morelli et al. 2010; Nuebel et al. 2010), it is obvious that if we assume a similar mutation rate for plant pathogens, the geographic origin of recently emerged *P. syringae* crop pathogens can only be reconstructed by sequencing whole genomes. It is thus exciting that sequencing of a large number of genomes has become affordable using next-generation sequencing with prices as low as $100/genome. Therefore, studies of the geographic origin and routes of transmission can now also be performed for *P. syringae* crop pathogens. However, there is another challenge that needs to be overcome before performing such studies: Populations from different geographic areas can only be compared, and phylogeny can only be correlated with geography (in what is referred to as "phylogeography") if a sufficient number of isolates are available from each geographic area to be representative of the populations in these areas. Here, we will show how limited availability of isolates has in fact precluded in-depth phylogeographic studies of *P. syringae* crop pathogens so far.

The first study that compared multiple strains of a recently emerged *P. syringae* pathogen focused on *Pae*, the causative agent of bleeding canker disease of European horse chestnut in northwestern Europe (Green et al. 2010). This disease was first described in 2002/2003 and spread from England within a few years all the way to Scotland. The genomes of three British isolates of the pathogen from 2006 and 2008 were found to be almost indistinguishable from each other (3 mutations in 3 million base pairs) but to be significantly different (1,613 mutations in 3 million base pairs) from a *P. syringae* pathogen that was isolated in 1969 in India and that is the causative agent of a foliar disease of horse chestnut. The bleeding canker disease of horse chestnut has not been observed anywhere else in the world, and the *Pae* strain from India that only causes foliar symptoms is the closest known relative. Also, no environmental isolate of *P. syringae* closely related to *Pae* has yet been identified in Europe. Therefore, the only conclusion that can be made is that the UK outbreak of this disease is due to a single clone that started spreading on horse chestnut very recently. However, nothing can be said about the geographic origin of this clone. Did it preexist in an environmental reservoir, for example, in

compartments of the water cycle, or on wild plants in Europe and jumped onto horse chestnut? Or was it introduced from another geographic area where it caused bleeding canker in the past? Maybe the disease has simply not been observed in the geographic area where it first emerged because other factors in that area limit the severity of the disease, for example, unfavorable climate, tolerance of horse chestnut genotypes that grow in that area, and/or absence of insect vectors that appear to contribute to the spread of *Pae* in the UK (Green et al. 2010).

Bacterial speck disease of tomato is caused by *Pto* and was first described in the USA by Bryan (1933). Interestingly, reports of the disease popped up around the world in the 1970s (Goode and Sasser 1980). It is not known whether this was due to a real increase in disease incidence or whether it was the awareness of the disease that increased and led to an increase in disease reports. Interestingly, analyzing over 100 *Pto* isolates from the 1940s until 2008 from around the world, it was found that two separate lineages of *Pto* (called T1 and JL1065) were common in Europe and North America in the 1970s, while a third lineage (DC3000) was only found in the 1940s in Canada and 1960s in the UK (Cai et al. 2011a) with one more isolate of this lineage from Japan (Sarkar and Guttman 2004). In the 1980s, the T1 lineage almost completely replaced the JL1065 lineage in Europe and North America. The few available *Pto* isolates from Africa and Australia belonged to either T1 or JL1065, but isolates from South America all belonged to T1. Since the T1 lineage was found on every continent from which isolates were available, any of these continents could be the geographic origin of the T1 lineage. Unfortunately, genomes of only five representative isolates of the T1 lineage from Europe and North America were sequenced, while all other isolates were only analyzed with a small number of genome-derived molecular markers. Moreover, the number of available *Pto* isolates from outside of Europe and North America was very limited. Therefore, it was impossible to determine which continent had the population with the highest diversity. Interestingly though, three out of eight genotypes of the *Pto* T1 population in Europe and North America were found to be present in both continents, showing that T1 strains have been exchanged several times between these two continents during recent years. Some of the same genotypes were found also in Australia. However, T1 isolates from Colombia in South America all had at least one mutation absent from all other T1 populations around the world and absent from the other *Pto* lineages and all related *P. syringae* pathovars. This suggests that the *Pto* population in South America may be relatively isolated compared to the *Pto* population of the Northern Hemisphere.

The mode of intercontinental transfer of *Pto* is impossible to know at this point. It could be by seed since *Pto* was found to be transmitted by seed (McCarter et al. 1983), but it could also be atmospheric movement. In fact, strains indistinguishable from *Pto* DC3000 were isolated in a creek upstream of any agricultural activity in New Zealand (Monteil et al. 2013), suggesting that at least bacteria belonging to the *Pto* DC3000 lineage can also travel through the water cycle and possibly travel long distance through the atmosphere. Therefore, it is possible that *Pto* T1 is exchanged between continents via atmospheric movement.

A relatively more detailed phylogeography has been obtained for *Psa*, the causative agent of bacterial canker disease of kiwifruit. Bacterial canker of kiwifruit was first described in Japan and China in the 1980s and then in Korea in the early 1990s (Scortichini et al. 2012). In 2008, a very severe epidemic started spreading throughout Italy and then the rest of Europe and Turkey. In 2010, the disease was found in New Zealand and in Chile. Three genome-sequenced isolates from Japan and Korea collected during the outbreak in the 1980s and 1990s were found to be clearly distinct from the 2008 epidemic and represent a separate population with strains that differ by more than 2,000 mutations/million base pairs from strains of the 2008 epidemic (Mazzaglia et al. 2012). This high number of mutations could not have accumulated in less than 30 years, and therefore, the Korean and Japanese populations can be excluded as the

parent population from which the 2008 population emerged. Two strains from China (isolated from the same location in the same year) were instead found to be highly similar to four strains isolated in Europe with fewer than 2 mutations per million base pairs (of vertically inherited core genome) distinguishing the strains from the two continents (Mazzaglia et al. 2012). Later, a small number of genome sequences of *Psa* strains from New Zealand and Chile were compared and found to be also extremely closely related to the European and Chinese strains (Butler et al. 2013). Interestingly though, one additional strain from China was found to be more divergent (1 mutations/16,000 base pairs) compared to all 2008 outbreak strains. While many more *Psa* strains isolated in New Zealand were sequenced later (McCann et al. 2013), the number of strains from China, Europe, and Chile is still very limited and any conclusions made from these data must be considered preliminary and are based on the pure assumption that the isolates from the different countries/continents are representative of the diversity in these countries. This is likely to be the case for Europe, New Zealand, and Chile since the disease emerged so recently in these areas. However, China, Japan, and Korea may harbor many more genotypes than currently known. Nonetheless, the finding that China harbors at least one genotype that is closely related to all the 2008 isolates, but is clearly more divergent, supports the conclusion that China is the geographic origin (Butler et al. 2013). Moreover, there is strong circumstantial evidence pointing to China as geographic origin of the current epidemic: (1) The center of diversification of the kiwifruit genus *Actinidia* is in China, making it likely that *P. syringae* strains adapted to *Actinidia* species possibly coevolving with *Actinidia* over hundreds of thousands of years; (2) the disease broke out in China years before it broke out in Europe, New Zealand, and Chile; and (3), plant material is known to have been imported from China to Europe and possibly to New Zealand.

The next important question is what were the international routes of transmission of *Psa*? Since the disease broke out first in Italy and then in New Zealand and isolates from these countries could initially not be distinguished, it was suggested that Italy was the source of the New Zealand outbreak. It was thought that possibly strains from China were imported with contaminated plant material into Italy and from there to New Zealand. However, several results point to a direct import of *Psa* from China to New Zealand: (1) Some mutations are shared between Chinese and New Zealand isolates but not with European isolates and (2) the same variant of a genomic island is present in one of the sequenced Chinese strains and all of the sequenced New Zealand strains, but this variant of the genomic island is absent in the European strains (Butler et al. 2013). Interestingly, a different Chinese strain has the identical variant of the genomic island present in the European strains, while the Chilean strains carry yet another version of the same island (Butler et al. 2013). This suggests (but does not prove) that the European, the New Zealand, and the Chilean *Psa* populations are each derived from separate independent importations of *Psa* from China.

4.4 Comparison of Closely Related Genomes Provides Insight into Molecular Plant: Microbe Interactions for Crop Improvement

Comparison of genomes of closely related strains with different phenotypes is one avenue to identify the genetic differences at the basis of the observed phenotypic differences. In the case of plant pathogens, the most relevant phenotypic differences are differences in host range and differences in virulence. The identification of genes responsible for these differences is facilitated by the fact that many *P. syringae* genes that determine virulence and, to some degree, host range are already described. It is known that conserved microbial-associated molecular patterns (MAMPs) trigger immunity in most plants but that some MAMP alleles evade recognition. For example, most flagellin alleles trigger a plant immune response, but flagellin of *Ralstonia*

solanacearum and *Agrobacterium tumefaciens* do not (Pfund et al. 2004; Felix et al. 1999). It is also known that *P. syringae* translocates via a type III secretion system (T3SS) so-called effector proteins into plant cells that suppress MAMP-triggered immunity in host plants, while some of them trigger an immune response (Effector-triggered immunity (ETI)) in non-host plants (Jones and Dangl 2006).

When comparing two closely related strains, whereby strain A causes disease on a certain plant species and strain B does not cause disease on this plant species, two hypotheses can be experimentally tested: (1) Strain A contains effectors necessary to suppress the immune system of the plant species, and strain B is missing those effectors and (2) strain B contains effectors or MAMP alleles that trigger an immune response in the plant species, and strain A is missing those effectors or has different MAMP alleles. Either hypothesis may be correct, or a combination of both hypotheses may apply. Moreover, other genes necessary for successful invasion of—or interaction with—the host species may be missing in strain B.

Two groups of closely related *P. syringae* strains have so far been used for this kind of comparison: (1) *Pto* DC3000 that causes disease on tomato and *Arabidopsis thaliana* and the closely related *Pto* T1 strain that causes disease on tomato but not on *A. thaliana* (Almeida et al. 2009) and (2) *P. syringae* pv. *phaseolicola* (*Pph*) strains that are pathogens of bean (*Phaseoulus vulgaris*) and mung bean (*Vigna radiata*) and the soybean pathogen *P. syringae* pv. *glycinea* (*Pgl*) (Baltrus et al. 2012). While there is one known case where deleting a single effector gene (*hopQ1*) from a strain (*Pto* DC3000) expanded the host range of the strain to an additional species (*Nicotiana benthamiana*) (Wei et al. 2007), genome comparisons and follow-up genetic manipulations of *Pto* T1 and *Pto* DC3000 and *Pph* and *Pgl* showed that host range differences may more often be the result of multiple genes (Sohn et al. 2012; Baltrus et al. 2012).

In the case of *Pto* T1, it was found that at least two of its effectors, AvrRpt2 and HopAS1, trigger an immune response in *A. thaliana* (Almeida et al. 2009; Sohn et al. 2012). Deleting these two effectors significantly increased growth of *Pto* T1 on *A. thaliana*. However, the double-deletion mutant still grows significantly less than *Pto* DC3000 and does not cause any disease symptoms at the minimum inoculum dose at which *Pto* DC3000 causes disease symptoms on *A. thaliana* (Sohn et al. 2012). Expressing individual *Pto* DC3000 effectors in the double-deletion mutant only marginally increased growth of *Pto* T1 in *A. thaliana* (our unpublished data). Therefore, either multiple *Pto* DC3000 effectors need to be expressed in the *avrRpt2/hopAS1* deletion mutant of *Pto* T1 at the same time to allow it to reach the virulence of *Pto* DC3000 or T3SS-independent genes necessary for full virulence on *A. thaliana* are also missing from *Pto* T1 but are present in *Pto* DC3000.

In the case of the *Pph* and *Pgl* comparison, expression of the *Pgl* effectors *hopC1* and *hopM1* in one of the *Pph* strains only slightly reduced growth on bean. On the other hand, expression of the *Pph* effector *avrB2* in *Pgl* only slightly increased growth of *Pgl* on bean. Therefore, adaptation of *Pph* to bean appears to be due to multiple effector differences (Baltrus et al. 2012). Only by deleting and expressing multiple effectors in the same strain in order for a *Pph* strain to acquire the same effector repertoire as the *Pgl* strain (and/or the *Pgl* strain to obtain the same effector repertoire of a *Pph* strain) will it be possible to determine whether effector differences alone can explain the host range differences between these strains. Possibly, allelic differences in effector sequences or expression level differences of effectors also contribute to the observed host range differences. Importantly, since T3SS mutants of the *Pph* and *Pgl* strains grew to the same population density, the differences in their host range appear to be limited to effectors and, possibly, allelic differences in the structural components of their T3SSs.

While comparison of *Pto* T1 with *Pto* DC3000 and of *Pph* strains with *Pgl* revealed differences in effectors, but differences in MAMPs were not noted, comparison of multiple isolates all belonging to the *Pto* T1 lineage

revealed unexpected differences in MAMPs within flagellin (Cai et al. 2011a). First, a group of strains from Colombia in South America had an amino acid substitution in the flagellin epitope flg22, a known MAMP (Felix et al. 1999), although these strains were otherwise indistinguishable from other *Pto* T1 isolates. Although it was known that different pathogens have different flagellin alleles and different strains of *Xanthomonas campestris* have different flagellin alleles (Sun et al. 2006), this difference among isolates belonging to the same genetic lineage was surprising. Moreover, all strains from Colombia had a non-synonymous mutation in another region of flagellin downstream of the flg22 epitope and all strains from Europe and North America isolated after 1980 had another non-synonymous mutation only three codons away from it (Cai et al. 2011a). This strongly suggested that the region of flagellin containing these two mutations (called flgII from now on) represented a new MAMP and *Pto* T1 adapted independently in Colombia and in North America/Europe to tomato by evading recognition of this MAMP through allelic variation. This hypothesis was experimentally confirmed: A 28-amino-acid-long peptide corresponding to the flgII allele typical of European isolates before 1980 triggered a stronger immune response in tomato than peptides corresponding to the two mutated alleles from Colombia and from North America/Europe after 1980 (Cai et al. 2011a). Therefore, comparison of almost identical isolates of the same genetic lineage of pathovar *Pto* allowed identification of a previously unknown MAMP. Moreover, allelic variation at this MAMP influences bacterial fitness *in planta* depending on the plant genotype suggesting that different plants have different alleles of the receptor that recognize alleles of flgII with different affinity (Clarke et al. 2013).

The flgII receptor is different from the flg22 receptor and not present outside of the Solanaceae family (Clarke et al. 2013). Therefore, cloning different alleles of the flgII receptor and expression of these receptor alleles in crops that either do not have any native flgII receptor or have a flgII receptor with low affinity for the flgII alleles of their most important pathogens could be used to improve crop disease resistance (or at least significantly decrease pathogen growth) similar to what was proposed in regard to the MAMP receptor elongation factor tu receptor (EFR) (Lacombe et al. 2010).

Identifying differences in effectors or MAMPs is not the only approach to find new targets for crop improvement for disease resistance based on genome comparisons. Identifying the conserved core repertoire of effectors present in every single strain of a pathogen is another promising approach. The underlying hypothesis is that the effectors that are present in every single strain of a pathogen are the effectors that are most important for virulence and that cannot be easily lost by the pathogen without a reduction in virulence. This approach has been proposed for *Xanthomonas manihotis* (Bart et al. 2012) but can easily be applied to *P. syringae* pathogens. For example, the genomes of five isolates of *P. cannabina* pv. *alisalensis* (*Pcal*) have been sequenced and the core effector repertoire identified (Sarris et al. 2013). The individual *Pcal* effectors could be cloned and tested for triggering a defense response in a panel of plant species to identify putative resistance genes to these effectors. The identified resistance genes could then be either bred into *Pcal* hosts or cloned and transferred into high-yielding commercial *Pcal* hosts by genetic engineering. The same could be done with the identified core effectors of *Psa* (McCann et al. 2013).

4.5 Present and Future of Genome-Based Diagnostics and Epidemiology

In plant disease diagnostics today, polymerase chain reaction (PCR) and real-time PCR (RT-PCR) have become routine and complement—or even replace—diagnostic techniques based on phenotype, for example, Biolog® or fatty acid analysis. Some of the PCR and RT-PCR primers have been designed based on individual whole-genome sequences or based on comparison of whole-genome sequences of the pathogen of

interest with genome sequences of other related pathogens, for example, for *Xanthomonas carotae* (Kimbrel et al. 2011) or *R. solanacearum* race 3 biovar 2 (Guidot et al. 2009). Even more precise diagnostics are possible with markers based on the comparison of genomes of different strains of the same pathogen. For example, Studholme and colleagues have developed simple genotyping assays based on diversity revealed by WGS in *Xanthomonas musae* (Wasukira et al. 2013) and Balestra et al. (2013) have designed primers that can distinguish between different clones of *Psa*.

Going one-step further, MLST is starting to be used for precise identification of pathogens. For example, the causative agent of a bacterial leaf spot outbreak of parsley in Ohio (USA) was recently identified as *P. syringae* pv. *coriandricola* (Xu and Miller 2013). The use of MLST for *P. syringae* and other bacterial crop pathogens is possible because a dedicated MLST database was established to simply compare sequences of individual loci with a collection of characterized pathogen strains or to perform a full-fledged MLST study of a pathogen sequencing all loci that were included in published MLST analyses of pathotype strains (Bull et al. 2011; Young et al. 2008).

While MLST has sufficient resolution for precise identification of a *P. syringae* pathovar, it is not always sufficient for differentiating between different lineages within the same pathovar. For example, MLST can distinguish between the Korean/Japanese population of *Psa* and the isolates of the 2008 outbreak in Europe, but MLST cannot distinguish between *Pae* causing bleeding canker of horse chestnut in Europe and the Indian strain causing only leaf spotting (Green et al. 2010). Also, in regard to epidemiology, MLST usually does not provide the resolution to identify the source of a disease outbreak. Therefore, the question is how can WGS be translated into the diagnostic practice to improve strain identification and determine the source of disease outbreaks?

The potential of WGS for improving diagnostics and epidemiology for bacterial human pathogens is very similar to the potential we see for improving diagnostics of plant pathogens. It was recently proposed that WGS could replace multiple separate steps and tests in the diagnosis of human bacterial pathogens and in epidemiological investigations (Didelot et al. 2012). In fact, not only could WGS precisely identify a plant pathogen to the pathovar level or beyond, it could even provide a list of genes present in an outbreak strain encoding antibiotics resistance and thus help choose the best disease control strategy to deploy. Moreover, WGS has the power to give a precise view of the effector repertoire of an outbreak strain and suggest which crop cultivars may be resistant to the outbreak strain. Therefore, WGS could inform growers in regard to the best choice of cultivars that are possibly resistant to a new outbreak strain and that could be planted the following year.

As with MLST, WGS can only be successful if appropriate databases and tools for genome analysis are developed and become accessible in plant disease clinics. All research results in regard to effector repertoires and the corresponding resistance genes known to recognize them need to be included in such a database. Moreover, to identify the source of a disease outbreak, all isolates that are identified by any diagnostic clinic need to be automatically added to a central database. In this way, if a new pathogen or a new pathogen lineage emerges in one geographic area, it will be possible to automatically follow its spread to other areas as long as outbreak isolates are routinely sequenced. Currently, routine genome sequencing is still too expensive in order to implement such a strategy, but we can expect prices to decrease further and we believe that it is only a question of time until such an a strategy will become reality.

While sequencing the genome of isolated bacteria is an effective tool in identification and epidemiology, it would be even better if pathogens could be precisely identified directly from a plant sample without pathogen isolation in a metagenomic approach. This could be done by extracting DNA from a plant sample, or it could be done by extracting RNA from a plant sample. The advantage of RNA over DNA is that by extracting RNA and then reverse transcribing all

RNA into DNA, RNA viruses would also be detected. The resolution lost by only sequencing transcribed genes would be minimal for diagnostics and may be negligible for epidemiological purposes since, for example, most of the single-nucleotide polymorphisms (SNPs) that were identified between *Pto* isolates were intragenic (Cai et al. 2011a). All genes may not be transcribed *in planta*, but even if only half the genes of a bacterium were transcribed *in planta*, this would probably still provide the necessary resolution for source identification. Another advantage of RNA sequencing would be that even physiological problems, like mineral deficiencies, might be identifiable if the sample is fresh or flash-frozen in liquid nitrogen immediately after collection. Of course, as for sequencing isolated bacteria directly, such a culture-independent sequencing approach would require an even more comprehensive database and collection of bioinformatics tools to interpret the massive amount of data that would be obtained.

4.6 From Individual Genomes Analysis to Population Genomics: The Next Step to Infer *Pseudomonas syringae* Crop Adaptation and Evolutionary History

We expect the next key findings in ecology and evolution of *P. syringae* to be revealed by "population genomics," an approach that has already demonstrated its power in clinical microbiology. While WGS is extremely useful for disease diagnostics and to determine the molecular determinants of phenotypic traits or to reconstruct international routes of transmission (as described above), it cannot correctly answer these and other more fundamental questions about ecological differentiation or evolutionary history unless these questions are addressed in the context of a population, i.e., applying population genomics. The field of population genomics has existed for dozens of years in its early form whereby genomic divergence within and between populations was assessed using a small number of genomic loci (Nosil and Buerkle 2010). Today, WGS allows extending this approach to whole genomes and thereby dramatically increases the quantity of information that we can analyze and provides access to genomic variation that was previously undetectable. In particular, we can pinpoint all genomic regions that have diverged between individuals and estimate gene flow allowing the inference of population structure, adaptation, and evolutionary history of organisms with an accuracy previously unimaginable (Nosil and Buerkle 2010; Nosil et al. 2009).

When genomes are compared between individuals, the individuals are expected to be representative of the population or metapopulation for which we want to determine genomic variation. Sampling over space, niche, and time is thus the foundation of any population genomic study, because it drives the interpretation of the results we obtain. The less the sample reflects the population, the less accurate our inferences are. This biggest limitation is due to the fact that our observations are always an approximation of the reality depending on the sampling representativeness. Paraphrasing Hunt et al. (2008), "in most ecological sampling, the true habitats or niches are unknown and can only be observed as projections onto the sampling dimensions ('projected habitats')." In the case of *P. syringae* and other plant pathogens, the question is how to get the most representative sample a priori while we do not know the population structure.

Therefore, carefully choosing the sampling strategy prior to sampling is indispensable and requires considering several factors. The choice of individuals will determine the success of the study in answering the questions we pose. In fact, sampling is what constitutes the difference between studies of population genomics and those of comparative genomics. Comparative genomics provides information about differences in gene content and allelic differences between genomes to test and develop hypotheses on how these differences determine phenotypic differences between organisms. Comparative genomics also permits development of lines of work to investigate ecology and epidemiology of plant

pathogens (Sarkar et al. 2006; Potnis et al. 2011; Baltrus et al. 2011; Mann et al. 2013). However, comparative genomics cannot extrapolate the findings to the population scale and cannot distinguish the effect of randomness from real selection by the environment. It is true that when whole-genome sequencing was first applied to epidemiological or evolutionary studies, the sample size was necessarily small because of sequencing costs and time required for gene annotation. Ten genomes were sufficient to obtain first genome-wide phylogenetic trees or to compare genomes of different pathogens or pathogen lineages. However, one individual can never be representative of a population whatever the criterion is (e.g., niche, biology, and lineage). Traits of that individual can be variable within the population, and this variability must be taken into account in the analysis. From a statistical point of view, the higher the number of individuals is, the more powerful the analysis is. However, even considering the low cost of sequencing today, we still cannot sequence as many genomes as we would like to and we still need to make a choice of which strains to sequence. To this end, it is important to select strains from a diverse set of samples that are representative of the environments that the pathogen occupies. In fact, even a large number of strains from a restricted number of samples may not be representative of the pathogen population. Fortunately, sequencing costs are still decreasing and this will allow sequencing more and more strains and make strain selection easier and easier (Didelot et al. 2012).

However, even if we were able to sequence as many strains as we want, choosing the best sampling strategy would still be important. Choice of sampling location, host, environment, or date, the relative number of different samples, and the choice of how many strains to sequence from each sample will always strongly affect how representative the sequenced individuals are of the population. Therefore, to accurately infer the processes that shape genomic variation, samples and strains must be selected properly. Moreover, that choice has to be tailored to the question and to the evolutionary scale at which the effects of the investigated processes are visible. Accordingly, year of isolation, geographic origin, genetic relatedness among individuals, and strain phenotypes do not always have the same importance. For example, when the objective of a study is to investigate early events in ecological differentiation, as in Shapiro et al. (2012), the investigated processes occur over a short time scale in a limited space compared to the evolutionary history of a bacterial species. Therefore, closely related bacteria isolated over a short time and occupying different niches at the same geographic location were chosen. A similar example is the investigation of the microevolution of *Staphphyloccoccus aureus* within a single host by Young et al. (2012) for which the authors selected dozens of isolates over a 13-month period from the same patient. For studies of biogeography, instead, we need to maximize space and time of sampling. For example, inference of the historical transmission routes of *Y. pestis* over continents (Morelli et al. 2010; Cui et al. 2013) required this kind of sampling. Finally, when the interest concerns pathogen–host specificity, studies may maximize the number of samples from different hosts (Fitzgerald et al. 2001; Sheppard et al. 2013a, b).

4.7 *Pseudomonas syringae* in the Footsteps of Human and Animal Bacterial Pathogens: What Have We Learned from Clinical Population Genomic Studies?

Population genomic studies of bacterial human pathogens have already dramatically improved our vision of pathogen ecology and evolution and are revolutionizing medical diagnostics and disease epidemiology. Development of statistical models, databases, and software, like the bacterial isolate genome sequence database (BIGSdb) pipeline (Jolley and Maiden 2010), have made it possible to handle genomic data of hundreds of strains of the same pathogen species. Such tools make it possible, for example, to streamline association-mapping methods to

determine the genomic basis of adaptation (Sheppard et al. 2013b). Additionally, several Bayesian modeling approaches have been developed to infer population structure, clonal relationships, and genomic fluxes from large populations and gene sets (Falush et al. 2003; Didelot et al. 2009, 2010; Marttinen et al. 2012; Corander et al. 2008; Shapiro et al. 2012).

Numerous processes associated with genomic divergence between populations and evolutionary history of human pathogens have been revealed, bringing to light new research perspectives. For example, population genomic studies applied to *Salmonella enterica* have given insights into population structure and the role of recombination far beyond the limitations of classical approaches (den Bakker et al. 2011; Desai et al. 2013; Didelot et al. 2011; Zhou et al. 2013). Classification into serovars and even phylogenies and recombination analysis based on MLST still missed important genetic information (Achtman et al. 2012). However, using a population genomic analysis of 10 % of the core genome of 114 isolates significantly refined our knowledge of the relationships between serovars, identified subpopulations, pinpointed donors and recipients of recombination events, and obtained insight into emergence of genetic lineages and estimated the age of these lineages (Didelot et al. 2011). By a similar approach, Joseph et al. (2012) unraveled the population structure and genomic fluxes within the obligate pathogen *Chlamydia trachomatis*.

In order to contain disease spread, it is necessary to understand how the bacterial variant that is causing a disease outbreak emerged. It is especially important to assess the importance of species introgression (acquisition of genomic regions from a different bacterial species) and to determine the environment in which it occurred, because this may allow identifying hot spots of bacterial diversification and prevention of future emergence of new variants. Mechanisms of introgression and the regions of the genome that are affected have been recognized and better understood, thanks to population genomics. One of the best examples is the genome-wide introgression that occurred in the zoonotic pathogen *Campylobacter coli*. Based on MLST data, Sheppard et al. (2011) had previously observed extensive DNA acquisition from another related species *Campylobacter jejuni* into one lineage of *C. coli*. Both species cause gastrointestinal symptoms in humans characterized by different host ranges in agricultural and non-agricultural environments. Through WGS of 30 strains of the two species, they were able to determine that gene flow occurred most frequently in those regions of the genome that are most similar between the two species (Sheppard et al. 2013a). These results also suggest that farming has played an important role in the diversification of *Campylobacter* spp. by enhancing physical opportunities for genetic exchange between the two species. Thus, agriculture appears to provide the conditions that are conducive to the emergence of new adapted hybrids and their proliferation. Sheppard et al. (2013b) then went further in the study of host adaptation applying for the first time whole-genome association mapping to a bacterial model and identified a genomic region significantly associated with isolates from cattle but frequently absent from genomes of strains that infect birds.

The high resolution of the information provided by whole-genome sequences facilitates the study of small genetic changes undetectable with classical genetic approaches. With population genetics, the assumptions we make about gene organization and dynamics in bacterial genomes leave variation at many loci undetected. We can thus not study small microevolutionary events associated with a small numbers of mutations per genome. With population genomics instead, these barriers disappear, permitting the study of microevolutionary events associated with niche adaptation. This can lead to new fundamental knowledge in microbial ecology. Whole-genome comparisons at the population scale have even revealed the processes involved in the early phases of ecological differentiation and speciation of bacteria (Shapiro et al. 2012).

The more we are interested in subtle changes and short time scales, the less powerful classical genetic population analysis is. This is especially well illustrated by the within-host

microevolutionary study of methicillin-sensitive *Staphylococcus aureus* causing fatal bloodstream infections (Young et al. 2012). The genome-wide study of Wilson and his partners (2012) enabled identification of a small number of mutations separating disease-causing variants from commensal bacteria. Following the pathogen population within a patient over a year, they demonstrated that the initial population accumulated 30 SNPs. Just eight of these SNPs affected protein function including a transcriptional regulator and were associated with the emergence of pathogenic variants inside the nasal carrier population. The high resolution of the approach thus brought new insight into the evolutionary dynamics of bacterial pathogens by demonstrating how small evolutionary events can lead to the emergence of virulent variants from nonpathogenic populations within the same host.

These few examples show the many possibilities of how population genomics could be applied to *P. syringae* and could help answer many unresolved questions about its ecology and evolution. Similar to what was shown for *S. aureus*, small evolutionary events might occur *in planta* in a few weeks or months from seed germination to maturity. The microevolutionary dynamics unraveled by population genomics within the plant host could also reveal regulatory and structural changes that lead to pathogenicity from a single cell via mutation or via acquisition or loss of genomic islands as experimentally revealed previously (Pitman et al. 2005; Lovell et al. 2009).

4.8 Looking Beyond Agriculture to Better Infer the Ecology and Evolution of *Pseudomonas syringae* Crop Pathogens

Previously, we discussed the importance of sampling in population genomic studies. We pointed out how we sample based on our current knowledge of a pathogen's ecology; in particular, we sample those environments in which we assume the pathogen population to reside. Therefore, our view of the plant–pathogen system is going to determine our sampling strategy and because our sampling strategy will determine the results of the population genomic analysis that we apply, our view of the plant pathogens system will ultimately condition the results we obtain and their interpretation. In plant pathology, the selection of isolates is mostly based on virulence traits, host range, or epidemic history. This is why studies of plant pathogens are often incomplete. Most of the time, strain choice is biased because we impose our current perspective of a plant pathogen's life history on that choice. Didelot et al. (2011) point out that when only highly virulent isolates are chosen, it is not possible to find the true evolutionary history of a pathogen because we only consider a small portion of the pathogen population and we cannot correctly infer gene flow and population structure.

For most pathogens that can survive outside of their hosts, limiting sampling to hosts is a problem. Sampling from diseased plants is only sufficient when we are interested in the evolution of a clonal lineage or its phylogeography. However, it is not sufficient if we want to understand how the pathogenic lineage relates to the wider diversity and how it originally emerged. The range of ecosystems, where non-obligate plant pathogens can evolve and adapt, have been unexplored and underestimated. A series of investigations of the ecology of *P. syringae* outside the agricultural context and, more specifically, in alpine ecosystems, have highlighted the various facets of *P. syringae* lifestyles. *P. syringae* population dynamics have been found to be inextricably associated with the water cycle (Morris et al. 2008). From the clouds through precipitation to river water, snowpack, and leaf litter, *P. syringae* is widely present and abundant (Diallo et al. 2012; Monteil et al. 2012; Morris et al. 2008, 2010). Based on these new insights into *P. syringae* life history, Morris et al. (2013) proposed scenarios of the role of the Earth's processes in *P. syringae* ecology and evolutionary history. Evidence was gathered that population dynamics of crop pathogens participate in wider metapopulation dynamics through the global freshwater ecosystem. These

environments are not only routes of dissemination for crop pathogens, but they represent reservoirs of genetic diversity and may be motors of diversification. While populations of alpine meadows or surface water are not directly impacted by cropping areas, most of them harbor virulence traits and genes typical of crop pathogens (Monteil et al. 2013). *P. syringae* studies carried out outside of the agricultural context are still exceptions among bacterial plant pathogens, but clues from others species, such as *Erwinia* spp. or *Pantoea* spp., suggest they could have similar life histories (Morris et al. 2009). *Phytophthora* species are another example of plant pathogens for which recent studies showed the importance of other ecosystems beyond crops and diseased trees. In the past decade, new evidence emerged showing abundance and high diversity of *Phytophthora* species in forest soils and forest streams. *Phytophthora* species present in natural habitats could lead to emergence of cryptic species potentially able to cause emerging destructive epidemics (Hansen et al. 2012).

Undoubtedly, determination of genomic fluxes between populations residing in natural environments and those causing disease on crops is indispensable to infer the evolutionary history and adaptation to a pathogenic lifestyle. Recently, a recombination analysis coupled to Bayesian coalescent analysis was performed with *P. syringae* strains isolated from snowpack and alpine streams that are closely related to the tomato pathogen *Pto*. This study revealed that ancestors of environmental populations and those of monomorphic crop pathogens recombined (Monteil et al. 2013). Moreover, virulence gene repertoires coding for T3SS effectors in *Pto* lineages are present in environmental strains. Environmental strains induced slightly less severe symptoms on tomato but had a wider host range. These results suggest that crop pathogens evolved through a small number of evolutionary events from an environmental population of less aggressive ancestors. Environmental pools and crop pathogen populations thus appear to interact, and environmental populations could be a source of novel genes. Similar observations led to the same conclusions for the causal agent of Legionnaire's disease: *Legionella pneumophila*. Interestingly, Coscolla and Gonzalez-Candelas (2007) found recombination rates to be high within environmental populations that are abundant in freshwater. When clinical isolates were included in the analysis, conflicting signals within trees based on sequences of individual loci strongly suggested recombination between environmental and clinical isolates (Coscolla and Gonzalez-Candelas 2009).

These observations show how important it is to enlarge the population framework outside of hosts to have a representative sample for population genomic studies. The environment may select genes that can be involved in plant–microbe interaction, virulence, and host range (Morris et al. 2009). In all environments, trophic interactions within the microbial community (e.g., predation/prey relationships, competition for resources, and mutualism) shape bacterial genomes, which act both as sink and donor of genetic information. Genetic exchanges and mutations lead to the formation and selection of new bacterial variants. Since the maintenance of genes in a genome has a cost, the presence of virulence genes in environmental strains suggests that these genes might have another function outside the agricultural context (Martinez 2013). These traits may then become virulence traits if they confer an adaptive advantage for *P. syringae* fitness *in planta*. Although this has not formally been demonstrated for any gene in *P. syringae*, some of the phytotoxins produced by some *P. syringae* strains (Bender et al. 1999) are known to have antimicrobial activity and may have evolved under selection pressure during competition with other microbes in the environment. Several examples of putative dual-use traits related to pathogenic and environmental fitness have been demonstrated for human pathogens (Morris et al. 2009). Those traits are involved in the formation of biofilms, resistance to predation by nematodes or protists, iron sequestration, oxidation, and resistance to antibiotic compounds. For example, *Pseudomonas aeruginosa* and *Vibrio* spp. are abundant in soil and aquatic substrates known to be

environmental reservoirs (Vezzulli et al. 2010; Selezska et al. 2012). They possess T3SSs similar to *P. syringae* to deliver effectors to eukaryotic cells and infect their human hosts. Evidence supports the hypothesis that their T3SSs were originally associated with bacterial survival in soil and water as defense against bacteriovorous amoebae (Matz et al. 2008, 2011). But grazing is not the only selection pressure at the origin of traits involved in virulence. For example, seawater is a reservoir of *Vibrio cholerae* and *Vibrio* sp. use the same factors associated with human intestinal colonization for attachment to chitin shells of crustaceans (Vezzulli et al. 2010; Pruzzo et al. 2008). Also, part of the resistome of clinical bacteria is shared with the soil microbiome. A metagenomic approach showed the presence of resistance genes to five classes of antibiotics in nonpathogenic soil-dwelling proteobacteria that were identical to antibiotic resistance genes of human pathogens (e.g., *P. aeruginosa*, *Klebsellia pneumonia*, and *Acinetobacter baumanii*) (Forsberg et al. 2012).

An example of the importance of the non-host environment for an opportunistic human pathogen is *Staphylococcus aureus.* Genomic diversity, evolutionary relationships, and methicillin resistance trait were investigated by Fitzgerald et al. (2001) on the basis of 36 clinical isolates. The authors showed that the methicillin resistance gene had been horizontally acquired several times independently. Where these events occurred is still unclear, but since *Staphylococcus aureus* is present in the rhizosphere (Berg et al. 2005), it may very well be that they occurred in the soil.

Therefore, conclusions about the evolutionary pressures shaping pathogen genomes are biased if they are made on the sole basis of pathogenic populations isolated from the host, while the species has a life style that includes non-host environments. Because this is the case for *P. syringae*, interpretation of genetic changes underlying virulence or host range based on strains isolated only from crop plants might be misleading, like in the studies of Baltrus et al. (2011) or Cai et al. (2011b).

The water cycle does not only consist of water from hydrological networks, but also includes the biosphere and consequently wild plants. *P. syringae* populations on wild plants or non-diseased crops have usually not been taken into account in population genetics and population genomics studies of *P. syringae* so far. Yet, *P. syringae* is an epiphyte and pathogenic lineages can be present in the phyllosphere of non-hosts and hosts at high population density without causing disease (Morris et al. 2008; Hirano and Upper 2000). No population genomic study has yet been performed to determine how genetically diverse epiphytic *P. syringae* populations are and how different or similar their virulence traits are compared to pathogen populations on diseased hosts. Including wild plants in studies of *P. syringae* evolution may thus also give new insight into plant adaptation. Indeed, as *P. syringae* crop pathogens have close relatives in the environment, crop plants have wild relatives outside of cropping areas. We can assume that the pool of wild relatives of crops plays an important role in host adaptation of environmental *P. syringae* population. The emergence of a *P. syringae* crop pathogen might possibly involve a progressive increase in fitness on wild relatives and lead to a population adapted to relatives of a certain crop from which a clonal crop pathogen may occasionally emerge, as was suggested for *Psa* (McCann et al. 2013). Importantly, most of the plant families of agricultural interest are present in alpine ecosystems like Rosaceae, Cruciferae, Fabaceae, Brassicaceae, Fagaceae, or Solanaceae (Sherman et al. 2008; Taberlet et al. 2012). Since Monteil et al. (2012) estimated that about 10^8 cfu of *P. syringae* may reside in one square meter of alpine vegetation, wild plants of alpine ecosystems should be included when investigating the adaptation of *P. syringae* to crop plants.

4.9 Conclusion

So far, population genomic studies based on whole-genome sequences of *P. syringae* pathogens are relatively few and relatively small in

regard to the number of strains that were analyzed. However, because WGS allows us to detect even a single mutation in a whole genome, population genomic approaches show great promise for describing evolutionary processes at a resolution that was inconceivable just a few years ago. Studies of human pathogens have pioneered the field and will facilitate similar studies on bacterial plant pathogens like *P. syringae*. In the past 10 years, population genomic studies of human pathogens brought new insights into how populations are structured, how genes are exchanged, and how populations adapt to a pathogenic lifestyle. By identifying the processes leading to the emergence of virulent isolates, host specialization, and pathogen movement, we expect to be able to better prevent and control crop diseases in the future. We may even be able to develop advanced warning systems alerting us of new pathogens or pathogen variants allowing more time to develop resistant cultivars.

Sampling remains the most important issue to address in population genomic studies in order to obtain results that reflect the actual pathogen population. Depending on the question, the factors to consider in the choice of strains in regard to time, location, host, and/or environment of isolation will be different. Importantly though, evolutionary history reconstruction of non-obligate pathogens like *P. syringae* will always need to take into account more than the diversity within the diseased crop host. They need to extend to the pathogen population residing on *non-diseased* hosts and non-host plants and non-plant environments. We expect that following these guidelines, population genomic studies of *P. syringae* will significantly deepen and broaden our view of plant pathogen ecology, evolution, epidemiology, and molecular plant–microbe interactions.

Reference

Achtman M (2008) Evolution, population structure, and phylogeography of genetically monomorphic bacterial pathogens. Annu Rev Microbiolo 62:53–70

Achtman M, Wain J, Weill F-X, Nair S, Zhou Z, Sangal V, Krauland MG, Hale JL, Harbottle H, Uesbeck A, Dougan G, Harrison LH, Brisse S, Grp SEMS (2012) Multilocus sequence typing as a replacement for serotyping in *Salmonella enterica*. PLoS Pathog 8(6):e1002776. doi:10.1371/journal.ppat.1002776

Almeida NF, Yan S, Lindeberg M, Studholme DJ, Schneider DJ, Condon B, Liu H, Viana CJ, Warren A, Evans C, Kemen E, Maclean D, Angot A, Martin GB, Jones JD, Collmer A, Setubal JC, Vinatzer BA (2009) A draft genome sequence of *Pseudomonas syringae* pv. *tomato* T1 reveals a type III effector repertoire significantly divergent from that of *Pseudomonas syringae* pv. *tomato* DC3000. Mol Plant Microbe Interact 22(1):52–62. doi:10.1094/MPMI-22-1-0052

Balestra GM, Taratufolo MC, Vinatzer BA, Mazzaglia A (2013) A multiplex PCR assay for detection of *Pseudomonas syringae* pv. *actinidiae* and differentiation of populations with different geographic origin. Plant Dis 97(4):472–478. doi:10.1094/pdis-06-12-0590-re

Baltrus DA, Nishimura MT, Dougherty KM, Biswas S, Mukhtar MS, Vicente J, Holub EB, Dangl JL (2012) The molecular basis of host specialization in bean pathovars of *Pseudomonas syringae*. Mol Plant Microbe Interact 25(7):877–888. doi:10.1094/mpm1-08-11-0218

Baltrus DA, Nishimura MT, Romanchuk A, Chang JH, Shahid Mukhtar M, Cherkis K, Roach J, Grant SR, Jones CD, Dangl JL (2011) Dynamic evolution of pathogenicity revealed by sequencing and comparative genomics of 19 *Pseudomonas syringae* isolates. PLoS Pathog 7(7):e1002132

Bart R, Cohn M, Kassen A, McCallum EJ, Shybut M, Petriello A, Krasileva K, Dahlbeck D, Medina C, Alicai T, Kumar L, Moreira LM, Rodrigues Neto J, Verdier V, Angelica Santana M, Kositcharoenkul N, Vanderschuren H, Gruissem W, Bernal A, Staskawicz BJ (2012) High-throughput genomic sequencing of cassava bacterial blight strains identifies conserved effectors to target for durable resistance. Proc Natl Acad Sci USA 109(28):E1972–E1979. doi:10.1073/pnas.1208003109

Bender CL, Alarcon-Chaidez F, Gross DC (1999) *Pseudomonas syringae* phytotoxins: mode of action, regulation, and biosynthesis by peptide and polyketide synthetases. Microbiol Mol Biol Rev 63(2):266–292

Berg G, Eberl L, Hartmann A (2005) The rhizosphere as a reservoir for opportunistic human pathogenic bacteria. Environ Microbiol 7(11):1673–1685. doi:10.1111/j.1462-2920.2005.00891.x

Bryan MK (1933) Bacterial speck of tomatoes. Phytopathology 23:897–904

Bull CT, Clarke CR, Cai R, Vinatzer BA, Jardini TM, Koike ST (2011) Multilocus sequence typing of *Pseudomonas syringae* sensu lato confirms previously described genomospecies and permits rapid identification of *P. syringae* pv. *coriandricola* and *P.*

syringae pv. *apii* causing bacterial leaf spot on parsley. Phytopathology 101(7):847–858. doi:10.1094/PHYTO-11-10-0318

Butler MI, Stockwell PA, Black MA, Day RC, Lamont IL, Poulter RTM (2013) *Pseudomonas syringae* pv. *actinidiae* from recent outbreaks of kiwifruit bacterial canker belong to different clones that originated in China. PLoS One 8(2):e57464. doi:10.1371/journal.pone.0057464

Cai R, Lewis J, Yan S, Liu H, Clarke CR, Campanile F, Almeida NF, Studholme DJ, Lindeberg M, Schneider D, Zaccardelli M, Setubal JC, Morales-Lizcano NP, Bernal A, Coaker G, Baker C, Bender CL, Leman S, Vinatzer BA (2011a) The plant pathogen *Pseudomonas syringae* pv. *tomato* is genetically monomorphic and under strong selection to evade tomato immunity. PLoS Pathog 7(8):e1002130. doi:10.1371/journal.ppat.1002130

Cai R, Yan S, Liu H, Leman S, Vinatzer BA (2011b) Reconstructing host range evolution of bacterial plant pathogens using *Pseudomonas syringae* pv. *tomato* and its close relatives as a model. Infect Genet Evol 11(7):1738–1751. doi:10.1016/j.meegid.2011.07.012

Clarke CR, Cai R, Studholme DJ, Guttman DS, Vinatzer BA (2010) *Pseudomonas syringae* strains naturally lacking the classical *P. syringae hrp/hrc* locus are common leaf colonizers equipped with an atypical type III secretion system. Mol Plant Microbe Interact 23(2):198–210. doi:10.1094/MPMI-23-2-0198

Clarke CR, Chinchilla D, Hind SR, Taguchi F, Miki Y, Martin GB, Leman S, Felix G, Vinatzer BA (2013) Allelic variation in two distinct *Pseudomonas syringae* flagellin epitopes modulates the strength of plant immune responses but not bacterial motility. New Phytol 200(3):847–860. doi:10.1111/nph.12408

Corander J, Marttinen P, Siren J, Tang J (2008) Enhanced Bayesian modelling in BAPS software for learning genetic structures of populations. Bmc Bioinform 9(1):539. doi:10.1186/1471-2105-9-539

Coscolla M, Gonzalez-Candelas F (2007) Population structure and recombination in environmental isolates of *Legionella pneumophila*. Environ Microbiol 9(3):643–656. doi:10.1111/j.1462-2920.2006.01184.x

Coscolla M, Gonzalez-Candelas F (2009) Comparison of clinical and environmental samples of *Legionella pneumophila* at the nucleotide sequence level. Infect Genet Evol 9(5):882–888. doi:10.1016/j.meegid.2009.05.013

Cui Y, Yu C, Yan Y, Li D, Li Y, Jombart T, Weinert LA, Wang Z, Guo Z, Xu L, Zhang Y, Zheng H, Qin N, Xiao X, Wu M, Wang X, Zhou D, Qi Z, Du Z, Wu H, Yang X, Cao H, Wang H, Wang J, Yao S, Rakin A, Li Y, Falush D, Balloux F, Achtman M, Song Y, Wang J, Yang R (2013) Historical variations in mutation rate in an epidemic pathogen, *Yersinia pestis*. Proc Natl Acad Sci USA 110(2):577–582. doi:10.1073/pnas.1205750110

den Bakker HC, Switt AIM, Cummings CA, Hoelzer K, Degoricija L, Rodriguez-Rivera LD, Wright EM, Fang R, Davis M, Root T, Schoonmaker-Bopp D, Musser KA, Villamil E, Waechter H, Kornstein L, Furtado MR, Wiedmann M (2011) A whole-genome single nucleotide polymorphism-based approach to trace and identify outbreaks linked to a common *Salmonella enterica* subsp enterica Serovar Montevideo pulsed-field gel electrophoresis type. Appl Environ Microbiol 77(24):8648–8655. doi:10.1128/aem.06538-11

Desai PT, Porwollik S, Long F, Cheng P, Wollam A, Clifton SW, Weinstock GM, McClelland M (2013) Evolutionary genomics of *Salmonella enterica* subspecies. MBio 4(2): e00579–12. doi:10.1128/mBio.00579-12

Diallo MD, Monteil CL, Vinatzer BA, Clarke CR, Glaux C, Guilbaud C, Desbiez C, Morris CE (2012) *Pseudomonas syringae* naturally lacking the canonical type III secretion system are ubiquitous in nonagricultural habitats, are phylogenetically diverse and can be pathogenic. ISME J 6(7):1325–1335. doi:10.1038/ismej.2011.202

Didelot X, Bowden R, Street T, Golubchik T, Spencer C, McVean G, Sangal V, Anjum MF, Achtman M, Falush D, Donnelly P (2011) Recombination and population structure in *Salmonella enterica*. Plos Genet 7(7):e1002191. doi:10.1371/journal.pgen.1002191

Didelot X, Bowden R, Wilson DJ, Peto TEA, Crook DW (2012) Transforming clinical microbiology with bacterial genome sequencing. Nat Rev Genet 13(9):601–612. doi:10.1038/nrg3226

Didelot X, Darling A, Falush D (2009) Inferring genomic flux in bacteria. Genome Res 19(2):306–317. doi:10.1101/gr.082263.108

Didelot X, Lawson D, Darling A, Falush D (2010) Inference of homologous recombination in bacteria using whole-genome sequences. Genetics 186(4):U1435–U1567. doi:10.1534/genetics.110.120121

Dye DW, Bradbury JF, Goto M, Hayward AC, Lelliott RA, Schroth MN (1980) International standards for naming pathovars of phytopathogenic bacteria and a list of pathovar names and pathotype strains. Rev Plant Pathol 59(4):153–168

Falush D, Stephens M, Pritchard JK (2003) Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. Genetics 164(4):1567–1587

Felix G, Duran JD, Volko S, Boller T (1999) Plants have a sensitive perception system for the most conserved domain of bacterial flagellin. Plant J Cell Mol Biol 18(3):265–276

Fitzgerald JR, Sturdevant DE, Mackie SM, Gill SR, Musser JM (2001) Evolutionary genomics of *Staphylococcus aureus*: insights into the origin of methicillin-resistant strains and the toxic shock syndrome epidemic. Proc Natl Acad Sci USA 98(15):8821–8826. doi:10.1073/pnas.161098098

Forsberg KJ, Reyes A, Wang B, Selleck EM, Sommer MOA, Dantas G (2012) The shared antibiotic resistome of soil bacteria and human pathogens. Science 337(6098):1107–1111. doi:10.1126/science.1220761

Gardan L, Shafik H, Belouin S, Broch R, Grimont F, Grimont PAD (1999) DNA relatedness among the pathovars of *Pseudomonas syringae* and description of *Pseudomonas tremae* sp. nov. and *Pseudomonas cannabina* sp. nov. (ex Sutic and Dowson 1959). Int J Syst Bacteriol 49:469–478

Gomez-Alpizar L, Carbone I, Ristaino JB (2007) An Andean origin of *Phytophthora infestans* inferred from mitochondrial and nuclear gene genealogies. Proc Natl Acad Sci USA 104(9):3306–3311. doi:10.1073/pnas.0611479104

Goode MJ, Sasser M (1980) Prevention—the key to controlling bacterial spot and bacterial speck of tomato. Plant Dis 64(9):831–834

Green S, Studholme DJ, Laue BE, Dorati F, Lovell H, Arnold D, Cottrell JE, Bridgett S, Blaxter M, Huitema E, Thwaites R, Sharp PM, Jackson RW, Kamoun S (2010) Comparative genome analysis provides insights into the evolution and adaptation of *Pseudomonas syringae* pv. *aesculi* on *Aesculus hippocastanum*. PLoS One 5(4):e10224. doi:10.1371/journal.pone.0010224

Guidot A, Elbaz M, Carrere S, Siri MI, Pianzzola MJ, Prior P, Boucher C (2009) Specific genes from the potato brown rot strains of *Ralstonia solanacearum* and their potential use for strain detection. Phytopathology 99(9):1105–1112. doi:10.1094/phyto-99-9-1105

Hansen EM, Reeser PW, Sutton S (2012) *Phytophthora* beyond agriculture. Annu Rev Phytopathol. doi:10.1146/annurev-phyto-081211-172946

Hirano SS, Upper CD (2000) Bacteria in the leaf ecosystem with emphasis on *Pseudomonas syringae*—a pathogen, ice nucleus, and epiphyte. Microbiol Mol Biol Rev 64(3):624–653

Hunt DE, David LA, Gevers D, Preheim SP, Alm EJ, Polz MF (2008) Resource partitioning and sympatric differentiation among closely related bacterioplankton. Science 320(5879):1081–1085. doi:10.1126/science.1157890

Jolley KA, Maiden MCJ (2010) BIGSdb: scalable analysis of bacterial genome variation at the population level. Bmc Bioinform 11(1):595. doi:10.1186/1471-2105-11-595

Jones JDG, Dangl JL (2006) The plant immune system. Nature 444(7117):323–329. doi:10.1038/nature05286

Joseph SJ, Didelot X, Rothschild J, de Vries HJC, Morre SA, Read TD, Dean D (2012) Population genomics of *Chlamydia trachomatis*: insights on drift, selection, recombination, and population structure. Mol Biol Evol 29(12):3933–3946. doi:10.1093/molbev/mss198

Kimbrel JA, Givan SA, Temple TN, Johnson KB, Chang JH (2011) Genome sequencing and comparative analysis of the carrot bacterial blight pathogen, *Xanthomonas hortorum* pv. *carotae* M081, for insights into pathogenicity and applications in molecular diagnostics. Mol Plant Pathol 12(6):580–594. doi:10.1111/j.1364-3703.2010.00694.x

Lacombe S, Rougon-Cardoso A, Sherwood E, Peeters N, Dahlbeck D, van Esse HP, Smoker M, Rallapalli G, Thomma BPHJ, Staskawicz B, Jones JDG, Zipfel C (2010) Interfamily transfer of a plant pattern-recognition receptor confers broad-spectrum bacterial resistance. Nat Biotech 28(4):365–369. http://www.nature.com/nbt/journal/v28/n4/abs/nbt.1613.html#supplementary-information

Lovell HC, Mansfield JW, Godfrey SAC, Jackson RW, Hancock JT, Arnold DL (2009) Bacterial evolution by genomic Island transfer occurs via DNA transformation in planta. Curr Biol 19(18):1586–1590. doi:10.1016/j.cub.2009.08.018

Maiden MCJ, Bygraves JA, Feil E, Morelli G, Russell JE, Urwin R, Zhang Q, Zhou JJ, Zurth K, Caugant DA, Feavers IM, Achtman M, Spratt BG (1998) Multilocus sequence typing: a portable approach to the identification of clones within populations of pathogenic microorganisms. Proc Natl Acad Sci USA 95(6):3140–3145. doi:10.1073/pnas.95.6.3140

Mann RA, Smits THM, Buhlmann A, Blom J, Goesmann A, Frey JE, Plummer KM, Beer SV, Luck J, Duffy B, Rodoni B (2013) Comparative genomics of 12 strains of *Erwinia amylovora* identifies a pan-genome with a large conserved core. PLoS One 8(2):e55644. doi:10.1371/journal.pone.0055644

Martinez JL (2013) Bacterial pathogens: from natural ecosystems to human hosts. Environ Microbiol 15(2):325–333. doi:10.1111/j.1462-2920.2012.02837.x

Marttinen P, Hanage WP, Croucher NJ, Connor TR, Harris SR, Bentley SD, Corander J (2012) Detection of recombination events in bacterial genomes from large population samples. Nucleic Acids Res 40(1):e6. doi:10.1093/nar/gkr928

Matz C, Moreno AM, Alhede M, Manefield M, Hauser AR, Givskov M, Kjelleberg S (2008) *Pseudomonas aeruginosa* uses type III secretion system to kill biofilm-associated amoebae. ISME J 2(8):843–852. doi:10.1038/ismej.2008.47

Matz C, Nouri B, McCarter L, Martinez-Urtaza J (2011) Acquired type III secretion system determines environmental fitness of epidemic *Vibrio parahaemolyticus* in the interaction with bacterivorous protists. PLoS One 6(5):e20275

Mazzaglia A, Studholme DJ, Taratufolo MC, Cai R, Almeida NF, Goodman T, Guttman DS, Vinatzer BA, Balestra GM (2012) *Pseudomonas syringae* pv. *actinidiae* (PSA) isolates from recent bacterial canker of kiwifruit outbreaks belong to the same genetic lineage. PLoS One 7(5): e36518. doi:10.1371/journal.pone.0036518

McCann HC, Rikkerink EHA, Bertels F, Fiers M, Lu A, Rees-George J, Andersen MT, Gleave AP, Haubold B, Wohlers MW, Guttman DS, Wang PW, Straub C, Vanneste J, Rainey PB, Templeton MD (2013) Genomic analysis of the kiwifruit pathogen *Pseudomonas syringae* pv. *actinidiae* provides insight into the origins of an emergent plant disease. PLoS Pathog 9(7):e1003503. doi:10.1371/journal.ppat.1003503

McCarter SM, Jones JB, Gitaitis RD, Smitley DR (1983) Survival of *Pseudomonas syringae* pv *tomato* in

association with tomato seed, soil, host tissue, and epiphytic weed hosts in Georgia. Phytopathology 73(10):1393–1398

Monteil CL, Cai R, Liu H, Mechan Llontop ME, Leman S, Studholme DJ, Morris CE, Vinatzer BA (2013) Non-agricultural reservoirs contribute to emergence and evolution of *Pseudomonas syringae* crop pathogens. New Phytol. doi:10.1111/nph.12316

Monteil CL, Guilbaud C, Glaux C, Lafolie F, Soubeyrand S, Morris CE (2012) Emigration of the plant pathogen *Pseudomonas syringae* from leaf litter contributes to its population dynamics in alpine snowpack. Environ Microbiol 14(8):2099–2112. doi:10.1111/j.1462-2920.2011.02680.x

Morelli G, Song Y, Mazzoni CJ, Eppinger M, Roumagnac P, Wagner DM, Feldkamp M, Kusecek B, Vogler AJ, Li Y, Cui Y, Thomson NR, Jombart T, Leblois R, Lichtner P, Rahalison L, Petersen JM, Balloux F, Keim P, Wirth T, Ravel J, Yang R, Carniel E, Achtman M (2010) *Yersinia pestis* genome sequencing identifies patterns of global phylogenetic diversity. Nat Genet 42(12):1140–1143. doi:10.1038/ng.705

Morris CE, Bardin M, Kinkel LL, Moury B, Nicot PC, Sands DC (2009) Expanding the paradigms of plant pathogen life history and evolution of parasitic fitness beyond agricultural boundaries. PLoS Pathog 5(12):e1000693

Morris CE, Monteil CL, Berge O (2013) The life history of *Pseudomonas syringae*: linking agriculture and Earth system processes. Annu Rev Phytopathol 51:85–104

Morris CE, Sands DC, Vanneste JL, Montarry J, Oakley B, Guilbaud C, Glaux C (2010) Inferring the evolutionary history of the plant pathogen *Pseudomonas syringae* from its biogeography in headwaters of rivers in North America, Europe, and New Zealand. mBio 1(3):1–10

Morris CE, Sands DC, Vinatzer BA, Glaux C, Guilbaud C, Buffiere A, Yan SC, Dominguez H, Thompson BM (2008) The life history of the plant pathogen *Pseudomonas syringae* is linked to the water cycle. ISME J 2(3):321–334. doi:10.1038/ismej.2007.113

Nosil P, Buerkle A (2010) Population genomics. Nat Educ Knowl 3(10):8

Nosil P, Funk DJ, Ortiz-Barrientos D (2009) Divergent selection and heterogeneous genomic divergence. Mol Ecol 18(3):375–402. doi:10.1111/j.1365-294X.2008.03946.x

Nuebel U, Dordel J, Kurt K, Strommenger B, Westh H, Shukla SK, Zemlickova H, Leblois R, Wirth T, Jombart T, Balloux F, Witte W (2010) A timescale for evolution, population expansion, and spatial spread of an emerging clone of methicillin-resistant *Staphylococcus aureus*. PLoS Pathog 6(4):e1000855. doi:10.1371/journal.ppat.1000855

Nunney L, Yuan X, Bromley R, Hartung J, Montero-Astua M, Moreira L, Ortiz B, Stouthamer R (2010) Population genomic analysis of a bacterial plant pathogen: novel insight into the origin of pierce's disease of grapevine in the US. PLoS One 5(11):e15488. doi:10.1371/journal.pone.0015488

Pfund C, Tans-Kersten J, Dunning FM, Alonso JM, Ecker JR, Allen C, Bent AF (2004) Flagellin is not a major defense elicitor in *Ralstonia solanacearum* cells or extracts applied to *Arabidopsis thaliana*. Mol Plant Microbe Interact 17(6):696–706. doi:10.1094/MPMI.2004.17.6.696

Pitman AR, Jackson RW, Mansfield JW, Kaitell V, Thwaites R, Arnold DL (2005) Exposure to host resistance mechanisms drives evolution of bacterial virulence in plants. Current biology CB 15(24):2230–2235. doi:10.1016/j.cub.2005.10.074

Potnis N, Krasileva K, Chow V, Almeida NF, Patil PB, Ryan RP, Sharlach M, Behlau F, Dow JM, Momol MT, White FF, Preston JF, Vinatzer BA, Koebnik R, Setubal JC, Norman DJ, Staskawicz BJ, Jones JB (2011) Comparative genomics reveals diversity among xanthomonads infecting tomato and pepper. BMC Genomics 12(1):146. doi:10.1186/1471-2164-12-146

Pruzzo C, Vezzulli L, Colwell RR (2008) Global impact of *Vibrio cholerae* interactions with chitin. Environ Microbiol 10(6):1400–1410

Sarkar SF, Gordon JS, Martin GB, Guttman DS (2006) Comparative genomics of host-specific virulence in *Pseudomonas syringae*. Genetics 174(2):1041

Sarkar SF, Guttman DS (2004) Evolution of the core genome of *Pseudomonas syringae*, a highly clonal, endemic plant pathogen. Appl Environ Microbiol 70(4):1999–2012. doi:10.1128/aem.70.4.1999-2012.2004

Sarris PF, Trantas EA, Baltrus DA, Bull CT, Wechter WP, Yan S, Ververidis F, Almeida NF, Jones CD, Dangl JL, Panopoulos NJ, Vinatzer BA, Goumas DE (2013) Comparative genomics of multiple strains of *Pseudomonas cannabina* pv. *alisalensis*, a potential model pathogen of both Monocots and Dicots. PLoS One 8(3):e59366. doi:10.1371/journal.pone.0059366

Scortichini M, Marcelletti S, Ferrante P, Petriccione M, Firrao G (2012) *Pseudomonas syringae* pv. *actinidiae*: a re-emerging, multi-faceted, pandemic pathogen. Mol Plant Pathol 13(7):631–640. doi:10.1111/j.1364-3703.2012.00788.x

Selezska K, Kazmierczak M, Müskeb M, Garbe J, Schobert M, Häussler S, Wiehlmann L, Rhode C, Sikorski J (2012) *Pseudomonas aeruginosa* population structure revisited under environmental focus: impact of water quality and phage pressure. Env Microbiol 14(8):1952–1967

Shapiro BJ, Friedman J, Cordero OX, Preheim SP, Timberlake SC, Szabo G, Polz MF, Alm EJ (2012) Population genomics of early events in the ecological differentiation of bacteria. Science 336(6077):48–51. doi:10.1126/science.1218198

Sheppard SK, Colles FM, McCarthy ND, Strachan NJC, Ogden ID, Forbes KJ, Dallas JF, Maiden MCJ (2011) Niche segregation and genetic structure of *Campylobacter jejuni* populations from wild and agricultural host species. Mol Ecol 20(16):3484–3490. doi:10.1111/j.1365-294X.2011.05179.x

Sheppard SK, Didelot X, Jolley KA, Darling AE, Pascoe B, Meric G, Kelly DJ, Cody A, Colles FM, Strachan NJC, Ogden ID, Forbes K, French NP, Carter P, Miller WG, McCarthy ND, Owen R, Litrup E, Egholm M, Affourtit JP, Bentley SD, Parkhill J, Maiden MCJ, Falush D (2013a) Progressive genome-wide introgression in agricultural *Campylobacter coli*. Mol Ecol 22(4):1051–1064. doi:10.1111/mec.12162

Sheppard SK, Didelot X, Meric G, Torralbo A, Jolley KA, Kelly DJ, Bentley SD, Maiden MCJ, Parkhill J, Falush D (2013b) Genome-wide association study identifies vitamin B-5 biosynthesis as a host specificity factor in Campylobacter. Proc Natl Acad Sci USA 110(29):11923–11927

Sherman R, Mullen R, Haomin L, Zhendong F, Yi W (2008) Spatial patterns of plant diversity and communities in Alpine ecosystems of the Hengduan Mountains, Northwest Yunnan China. J Plant Ecol 1(2):117–136. doi:10.1093/jpe/rtn012

Smith JM, Smith NH, Orourke M, Spratt BG (1993) How clonal bacteria are? Proc Natl Acad Sci USA 90(10):4384–4388. doi:10.1073/pnas.90.10.4384

Sohn KH, Saucet SB, Clarke CR, Vinatzer BA, O'Brien HE, Guttman DS, Jones JDG (2012) Hop AS1 recognition significantly contributes to Arabidopsis non-host resistance to *Pseudomonas syringae* pathogens. New Phytol 193(1):58–66. doi:10.1111/j.1469-8137.2011.03950.x

Stukenbrock EH, McDonald BA (2008) The origins of plant pathogens in agro-ecosystems. Annu Rev Phytopathol 46:75–100. doi:10.1146/annurev.phyto.010708.154114

Sun WX, Dunning FM, Pfund C, Weingarten R, Bent AF (2006) Within-species flagellin polymorphism in *Xanthomonas campestris* pv *campestris* and its impact on elicitation of Arabidopsis FLAGELLIN SENSING2-dependent defenses. Plant Cell 18(3):764–779. doi:10.1105/tpc.105.037648

Taberlet P, Zimmermann NE, Englisch T, Tribsch A, Holderegger R, Alvarez N, Niklfeld H, Coldea G, Mirek Z, Moilanen A, Ahlmer W, Marsan PA, Bona E, Bovio M, Choler P, Cieślak E, Colli L, Cristea V, Dalmas J-P, Frajman B, Garraud L, Gaudeul M, Gielly L, Gutermann W, Jogan N, Kagalo AA, Korbecka G, Küpfer P, Lequette B, Letz DR, Manel S, Mansion G, Marhold K, Martini F, Negrini R, Niño F, Paun O, Pellecchia M, Perico G, Piękoś-Mirkowa H, Prosser F, Puşcaş M, Ronikier M, Scheuerer M, Schneeweiss GM, Schönswetter P, Schratt-Ehrendorfer L, Schüpfer F, Selvaggi A, Steinmann K, Thiel-Egenter C, van Loo M, Winkler M, Wohlgemuth T, Wraber T, Gugerli F, IntraBioDiv C (2012) Genetic diversity in widespread species is not congruent with species richness in alpine plant communities. Ecol Lett 15(12):1439–1448. doi:10.1111/ele.12004

Vezzulli L, Pruzzo C, Huq A, Colwell RR (2010) Environmental reservoirs of *Vibrio cholerae* and their role in cholera. Environ Microbiol Rep 2(1):27–33. doi:10.1111/j.1758-2229.2009.00128.x

Wasukira A, Tayebwa J, Thwaites R, Paszkiewicz K, Aritua V, Kubiriba J, Smith J, Grant M, Studholme DJ (2013) Genome-wide sequencing reveals two major sub-lineages in the genetically monomorphic pathogen *Xanthomonas campestris* pv. *musacearum*. Genes. doi:10.3390/genes3030361

Wei C-F, Kvitko BH, Shimizu R, Crabill E, Alfano JR, Lin N-C, Martin GB, Huang H-C, Collmer A (2007) A *Pseudomonas syringae* pv. *tomato* DC3000 mutant lacking the type III effector HopQ1-1 is able to cause disease in the model plant *Nicotiana benthamiana*. Plant J 51(1):32–46. doi:10.1111/j.1365-313X.2007.03126.x

Xu X, Miller SA (2013) First report of bacterial leaf spot of parsley caused by *Pseudomonas syringae* pv. *coriandricola* in Ohio. Plant Dis 97(7):988. doi:10.1094/pdis-12-12-1181-pdn

Yan S, Liu H, Mohr TJ, Jenrette J, Chiodini R, Zaccardelli M, Setubal JC, Vinatzer BA (2008) Role of recombination in the evolution of the model plant pathogen *Pseudomonas syringae* pv. *tomato* DC3000, a very atypical tomato strain. Appl Environ Microbiol 74(10):3171–3181. doi:10.1128/aem.00180-08

Young BC, Golubchik T, Batty EM, Fung R, Larner-Svensson H, Votintseva AA, Miller RR, Godwin H, Knox K, Everitt RG, Iqbal Z, Rimmer AJ, Cule M, Ip CLC, Didelot X, Harding RM, Donnelly P, Peto TE, Crook DW, Bowden R, Wilson DJ (2012) Evolutionary dynamics of *Staphylococcus aureus* during progression from carriage to disease. Proc Natl Acad Sci USA 109(12):4550–4555. doi:10.1073/pnas.1113219109

Young JM, Park DC, Shearman HM, Fargier E (2008) A multilocus sequence analysis of the genus *Xanthomonas*. Syst Appl Microbiol 31(5):366–377. doi:10.1016/j.syapm.2008.06.004

Zhou Z, McCann A, Litrup E, Murphy R, Cormican M, Fanning S, Brown D, Guttman DS, Brisse S, Achtman M (2013) Neutral genomic microevolution of a recently emerged pathogen, *Salmonella enterica* serovar Agona. PLoS Genet 9(4):e1003471. doi:10.1371/journal.pgen.1003471

Genetics and Functional Genomics of the *Pseudomonas fluorescens* Group

5

Sarah Craven Seaton and Mark W. Silby

5.1 Introduction

5.1.1 General Characteristics

The fluorescent pseudomonads are of tremendous interest in diverse areas including pathogenicity, food spoilage, bioremediation, and biological control. In particular, members of the *Pseudomonas fluorescens* group are ubiquitous, environmental saprophytes that are found in soil, foliage, freshwater, and seawater (Cornelis 2008). Because most *P. fluorescens* strains are psychrotrophic, unable to grow at temperatures above 32 °C, generally the species is considered to pose no health threat; however, in rare cases, *P. fluorescens* has been linked to disease (Gershman et al. 2008; Hsueh et al. 1998).

Members of the *P. fluorescens* group are Gram-negative rod-shaped chemoheterotrophs that are motile by means of polar flagella and belong to the γ subclass of proteobacteria. They have simple nutritional requirements and can utilize an array of small organic molecules as sources of carbon and energy (Palleroni 1984). *P. fluorescens* strains are strictly respiratory; however, some can utilize nitrate as an electron acceptor in place of O_2, and all strains can use nitrate as a nitrogen source. *P. fluorescens* strains produce and secrete siderophores to scavenge iron when in iron-deplete environments, the major example of which is pyoverdine, the yellow–green fluorescent pigment lending these organisms the characteristic fluorescence for which they are named.

The ubiquity of members of the *P. fluorescens* group in various environments (up to 10^6 in a gram of soil) is, in part, a consequence of the simple nutritional requirements of the organisms, the range of carbon sources they utilize, as well as their genetic and metabolic adaptability. Soil-dwelling *P. fluorescens* are distributed throughout the world and have evolved the metabolic capacity to subsist on or degrade an array of plant-derived phenolics. Thus, the species in this group contribute greatly to the turnover of organic matter and to global carbon cycles. Such inherent metabolic diversity has also been exploited for biodegradation of natural and xenobiotic pollutants including styrene (Baggi et al. 1983; Beltrametti et al. 1997), naphthalene and other polycyclic aromatic hydrocarbons (Chauhan et al. 2008; Foght and Westlake 1996), and chlorinated hydrocarbons (Vandenbergh and Kunka 1988), among others. Genetic and genomic tools continue to be applied to deciphering the mechanisms underlying these metabolic capabilities, with a view to

S. C. Seaton
Center for Adaptation Genetics and Drug Resistance, Department of Molecular Microbiology, Tufts University School of Medicine, Boston, USA
e-mail: sarah.craven@tufts.edu

M. W. Silby (✉)
Department of Biology, University of Massachusetts Dartmouth, Dartmouth, USA
e-mail: mark.silby@umassd.edu

D. C. Gross et al. (eds.), *Genomics of Plant-Associated Bacteria*,
DOI: 10.1007/978-3-642-55378-3_5,

development of improved bioproducts which can be deployed to remediate contaminated environments.

The relationship between bacteria and the plants and soil they inhabit is well known, but on a functional level, it is still poorly understood. Many pseudomonads live in a commensal relationship with plants, utilizing plant-exuded nutrients and occupying sites provided by the architecture of the plant. Such commensal species, in turn, can have profound effects on plant health by suppressing phytopathogens, enhancing local access to nutrients, degrading environmental pollutants, or inducing systemic resistance in the plant host. Of particular interest is the capacity of *Pseudomonas* spp. to produce a variety of bioactive secondary metabolites, including antibiotics that thwart plant pathogens and hormones which promote plant growth. While the genetics underlying some of these metabolic processes is well established, big questions remain. The advent of genomic approaches leveraging sequence and expression data has begun to reveal the complexity and details of processes required for interactions of *P. fluorescens* with its environment. Furthermore, genome comparisons and genome mining approaches have identified new features which are likely critical to the success of these organisms, but had remained undetected using other approaches. Genetic and functional genomic analysis of *P. fluorescens* living in natural environments is giving us new insight into the way of life of these bacteria outside of the laboratory.

5.1.1.1 Taxonomy

The genus *Pseudomonas* is large and diverse, currently comprised of more than 200 named species that are phenotypically and genotypically well defined. The genus was first described by Migula (1894), based almost exclusively on the morphological characteristics of its members and thus for many years served as a repository for any number of Gram-negative, aerobic non-sporulating rods with polar flagella, despite the fact that such characteristics are shared by bacterial genera that are only distantly related from a phylogenetic standpoint. The taxonomy of the genus has aptly been described as tumultuous, and one estimate suggests that in the first decades of the twentieth century, the number of named species approached the one thousand mark (Cornelis 2008). Increasingly sensitive methods in molecular taxonomy have redistributed many of the former *Pseudomonas* species into other genera, while species designations continue to see reorganization.

In particular, heterogeneity within the *P. fluorescens* species has long been a topic for investigation. In the 1960s, Stanier et al. (1966) first introduced the division of *P. fluorescens* into five distinct biovars, defined by exhaustive phenotypic characterization. However, even this early delineation was described by Stanier as tenuous, highlighting the poor cohesion within the *fluorescens* complex. Sensitive methods that rely on DNA sequence comparisons, alone or in addition to phenotypic characterization, provide a clearer, but continually evolving portrait of *P. fluorescens* phylogeny. Classification based on 16S rRNA gene sequence permits the ascription of an organism to the *Pseudomonas* genus, but provides little resolution at the species level. Polyphasic investigations that assign phylogeny based on the sequence of multiple genes (for example, 16S rRNA, *gyrB*, *rpoB*, and *rpoD*) provide better resolution, but the genes chosen for analysis impact the final result. Recent reports relying on multilocus sequence typing divide the *P. fluorescens* complex into nine distinct subgroups (Mulet et al. 2012).

5.1.1.2 Phylogenomics

Whole-genome sequences allow for a comprehensive appraisal of relationships and unequivocal assignments into groups based on overall relationships. This can be based on the average nucleotide identity (ANI), average amino acid identity (AAI) among the predicted proteins of conserved genes, or a phylogenomic approach which builds phylogenetic trees based on similarity of all genes rather than just a handful as in MLST. The availability of complete genome

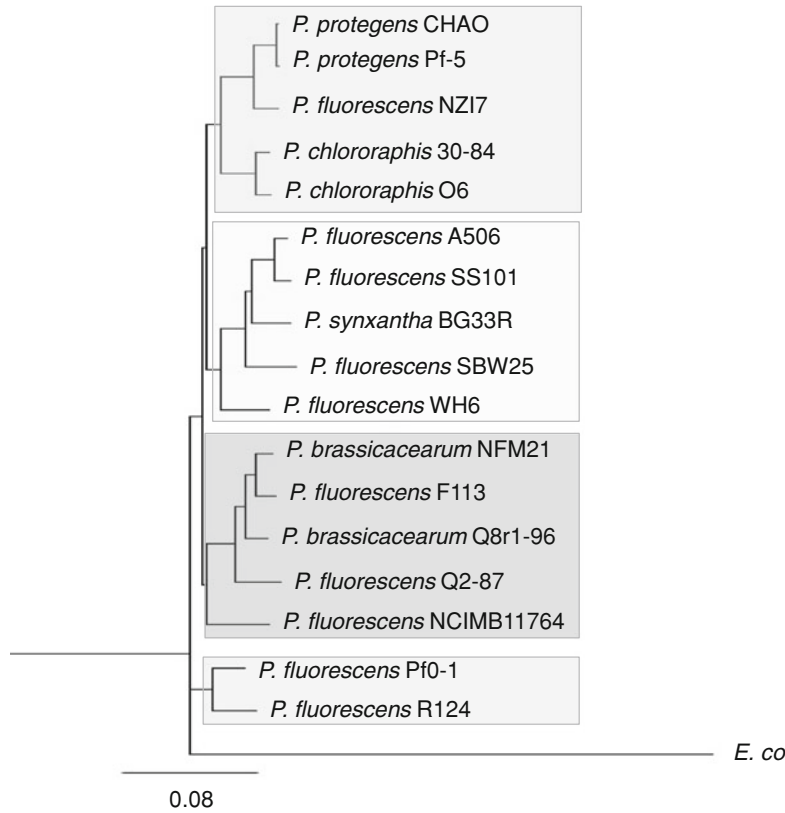

Fig. 5.1 Phylogenetic tree depicting the relationships of sequenced members of the *Pseudomonas fluorescens* group. The neighbor-joining tree is built based on concatenated alignments of four housekeeping genes, using Geneious Tree Builder with default parameters (Geneious version 6.1.6): *rpoD*, *gyrB*, *gltA*, and *gapA*, with *E. coli* as the out-group. *Boxed* regions indicate the four subgroups as defined by whole-genome phylogenetic analysis described by Marchi et al. (2013)

sequence information for members of the *P. fluorescens* complex has led to the reorganization of the *P. fluorescens* into a group comprised of several species. Initial comparative genomic analysis of the first three complete *P. fluorescens* genome sequences (Pf0-1, SBW25, and Pf-5) highlighted a large number of strain-specific genes (Silby et al. 2009), suggesting that the sequenced strains belong to a complex rather than to a single species. In addition, the AAI scores between the three strains were well below what has been deemed the cutoff in terms of defining a species. This genomic analysis clearly showed that the three *P. fluorescens* strains were more closely related to each other than to other species in the genus, but not close enough to each other to meet current AAI-based definition of a species. Even with the benefit of phylogenomic analysis, the estimation of the structure of the *P. fluorescens* group is highly dependent on the number and identity of the strains included. Loper et al. (2012) suggested three main clades from analysis of 10 strains (Loper et al. 2012), while a more recent phylogenomic appraisal of 50 strains yielded five subgroups (Redondo-Nieto et al. 2013). The genomic variation among the *P. fluorescens* group means that new species (e.g., *Pseudomonas protegens*) are frequently introduced, leading to a state of flux with regard to the number of species in any given subgroup. What continues to be apparent is the considerable diversity and the ongoing need to clarify the taxonomy of this complex. For researchers who isolate putative new members of the group, it may not be clear how to determine their relative phylogenetic position, given that some of the recent analyses utilize whole-genome phylogenies. Such a method is not always available to laboratories which do not routinely carry out bioinformatic analyses. A simpler and more accessible procedure is to generate a tree using housekeeping genes. Figure 5.1 shows a phylogenetic tree which was generated from concatenation of four housekeeping genes (*rpoD*, *gltA*,

gyrB, and *gapA*). The topology is consistent with that generated by Marchi et al. (2013) using the core genome (Marchi et al. 2013), indicating that new isolates can be placed into the phylogeny based on housekeeping genes with a relatively high degree of confidence. In this chapter, we will discuss multiple species within the "*Pseudomonas fluorescens* complex" including the recently renamed *P. protegens* Pf-5 and members of the closely related *P. brassicacearum*, *P. synxantha*, and *P. chlororaphis*. We additionally recognize that restructuring of the *P. fluorescens* complex is ongoing and anticipate that many of the current species designations may change as whole-genome sequences become available.

5.2 *P. fluorescens* Genomes

5.2.1 Genome Organization

The genetic repertoire within the *P. fluorescens* complex reflects the diverse lifestyles of these ubiquitous bacteria. Colonization of diverse habitats, from soil and water to plant and animal hosts, correlates with the ability to exploit many different nutrition sources and a high potential for adaptation to new and changing environmental conditions. Such versatility necessitates the maintenance of relatively large genomes, ranging from 5.51 Mbp to over 7 Mbp, in the sequenced strains, each with approximately 6,000 CDSs (Table 5.1).

Given the broad range of isolates in the *P. fluorescens* group, it is not surprising that there is considerable variability in the genome content (discussed below). An additional striking feature of the genomes is the overall organization of genes, at the level of individual genes and on a genome-wide scale. Several functional genomics studies have indicated that in at least two strains of *P. fluorescens* as well as *Pseudomonas putida*, divergent overlapping sense/antisense pairs of genes may be common features (Fernandez et al. 2013; Silby et al. 2009; Silby and Levy 2008). These were identified by IVET screening (see below), taking advantage of the non-biased nature of the method in terms of sequences that can be identified. Studies in some other species also indicate that these may be normal genomic features, but functional studies are somewhat limited. In SBW25, at least 42 of these "antisense" genes are activated during growth on plant surfaces (Gal et al. 2003; Rainey and Preston 2000; Silby et al. 2009). The presence of these in Pf0-1 and SBW25 suggests that similar organization occurs in other *P. fluorescens* genomes, although this has yet to be demonstrated. Something as seemingly simple as estimating the number of genes in a bacterial genome has turned out to be far more challenging than expected in the early days of microbial genomics.

When examined as a whole, two interesting and possibly related features of *P. fluorescens* genomes are apparent. At least two studies, comparing different isolates, have shown that the genomes of *P. fluorescens* can be considered in two major regions. Around the origin of replication, there is a relatively high degree of synteny among isolates. In contrast, toward the replication terminus, synteny drops dramatically. The genomes appear compartmentalized. Most of the conserved (core) genes in strains in the *P. fluorescens* group are found in the two-thirds of the genome flanking the origin. In contrast, there is relatively low conservation in the one-third of the genome around the replication terminus (Loper et al. 2012; Silby et al. 2009) (Fig. 5.2). In this region, there are many strain-specific genes, suggesting the possibility of a mechanism by which strain-specific genes are concentrated in this area. Such strain-specific genes may represent flexible genome content in the group, which confers niche-specific capabilities on isolates from different environments.

A second feature of the genomes of *P. fluorescens* is apparent extensive rearrangement by reciprocal recombination around the replication terminus (Fig. 5.2). This has been observed in comparisons among *P. fluorescens* Pf0-1 and SBW25 and *P. protegens* Pf-5 (Silby et al. 2009) and when Pf0-1 was compared with *P. fluorescens*

Table 5.1 Features of selected strains of the *Pseudomonas fluorescens* group

Strain	Source	Chromosome size (Mbp)	G + C (%)	Protein-coding sequences (CDSs)
P. fluorescens				
Pf0-1	Agricultural loam soil, Massachusetts, USA	6.44	60.5	5,722
SBW25	Sugar beet phyllosphere, Oxfordshire, UK	6.72	60.5	5,921
A506	Pear phyllosphere, California, USA	6.02	59.9	5,334
F113	Sugar beet rhizosphere, Ireland	6.85	60.8	5,862
WH6	Grass rhizosphere, Oregon, USA	6.27	60.6	5,860
SS101	Wheat rhizosphere, the Netherlands	6.18	60.0	5,372
R124	Cave isolate	6.30	60.3	5,224
Q2-87	Wheat rhizosphere, Washington, USA	6.37	60.6	5,596
NCIMB11764	Enrichment from River Tees sediment, UK	7.02	59.0	6,416
P. protegens				
Pf-5	Soil, Texas, USA	7.07	63.3	6,108
CHAO	Tobacco rhizosphere, Morens, Switzerland	6.84	60.8	6,095
P. brassicacearum				
NFM421	*Arabidopsis* rhizosphere, Australia	6.84	60.8	6,095
Q8r1-96[a]	Wheat rhizosphere, Washington, USA	6.6	61.0	5,715
P. chlororaphis				
O6	Soil, Utah, USA	6.98	62.9	6,223
30–84	Wheat rhizosphere, Kansas, USA	6.67	62.9	5,848
P. synxantha				
BG33R	Peach tree rhizosphere, South Carolina, USA	6.3	59.7	5,509

[a] Listed in GenBank as *P. fluorescens,* but categorized by Loper et al. (2012) as *P. brassicacearum*

F113 and *P. brassicacearum* NFM421 (Redondo-Nieto et al. 2013). The mechanism underlying the recombination has not been determined. Examination of Pf0-1, SBW25, and Pf-5 genomes did not reveal any significant homology correlated with regions which had been translocated, indicating that *recA*-dependent homologous recombination is unlikely to be responsible. The functional significance, if any, also remains unknown.

5.2.1.1 Core and Pan-Genome

As of July 15, 2013, the NCBI hosts details of four completely sequenced and closed *P. fluorescens* genomes (Table 5.1: Pf0-1, SBW25, F113, and A506), along with sequence data for nearly 20 other *P. fluorescens* draft genomes which are assembled into more than one contig. When considering the broader *P. fluorescens* group, the depth of sequence data becomes far more pronounced, including fully sequenced

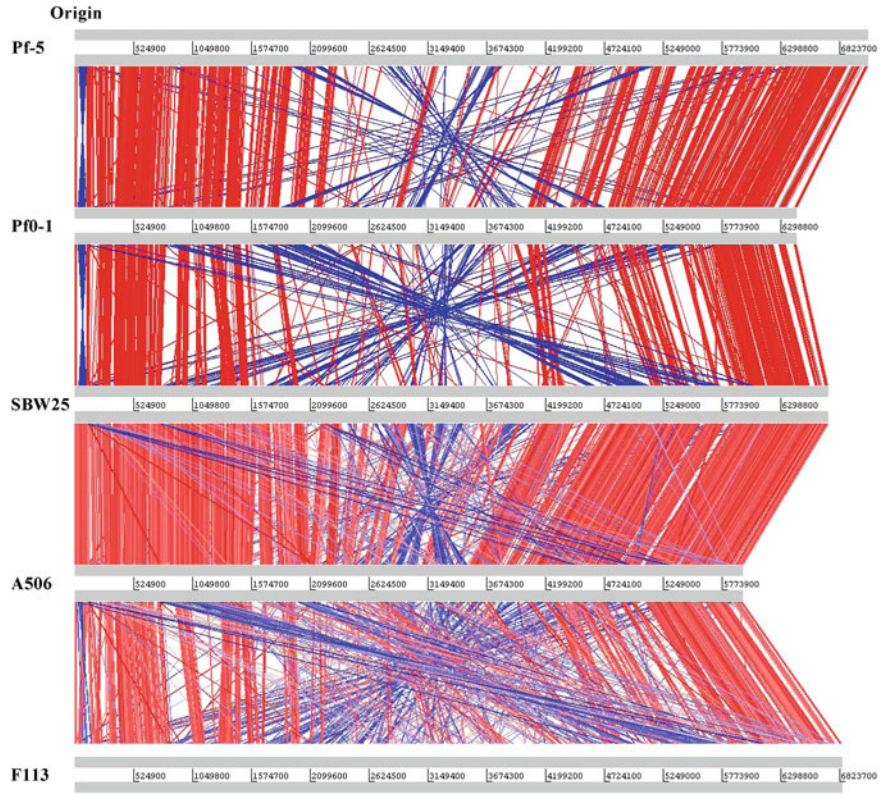

Fig. 5.2 Comparison of five closed genomes of *P. fluorescens* group members. Amino acid matches between 6-frame translations are shown. Comparisons were generated using TBlastX at the Double ACT Website (http://www.hpa-bioinfotools.org.uk/pise/double_act.html) and visualized using the Artemis Comparison Tool (*ACT*) (Carver et al. 2005). For each genome, forward and reverse strands are shown as *gray lines*, with nucleotide coordinates included. *Red bars* between genomes indicate regions of synteny, while *blue bars* show inverted matches. Notable is the generally higher level of conservation and synteny near the origin (*left* and *right* ends) and the sequences which have been inverted and translocated around the terminus

members of the *P. protegens*, *P. brassicacearum*, *P. synxantha*, and *P. chlororaphis* species (Table 5.1).

Two recent studies have comprehensively examined genome diversity within the *P. fluorescens* complex, highlighting a relatively small core genome and a large number of strain-specific genes (Loper et al. 2012; Redondo-Nieto et al. 2013). Loper et al. (2012) analyzed a total of 10 strains, in which they describe a core genome of just 2,789 predicted protein-coding genes, among a pan-genome of 13,872 predicted protein-coding genes. The small core genome, representing only 45–52 % of the total proteome of each strain, highlights the extreme diversity among the strains tested and combined with phylogenomic analyses has led to the reassignment of several of the strains to other species. Here, we define a more stringent *P. fluorescens* core genome—that is, the set of orthologous genes common to strains that currently retain the *P. fluorescens* designation—in a multiway BLAST analysis via the GView server (Petkau et al. 2010). Using e-value and percent identity cutoffs of 1 e^{-10} and 80 %, respectively, a core genome of 2,396 predicted CDSs was defined (Fig. 5.3a). Genes conserved in the core of all strains encode proteins that contribute mainly to fundamental housekeeping functions, including those proteins involved in protein and nucleic acid synthesis and central metabolism. Furthermore, core genes are found in the highest

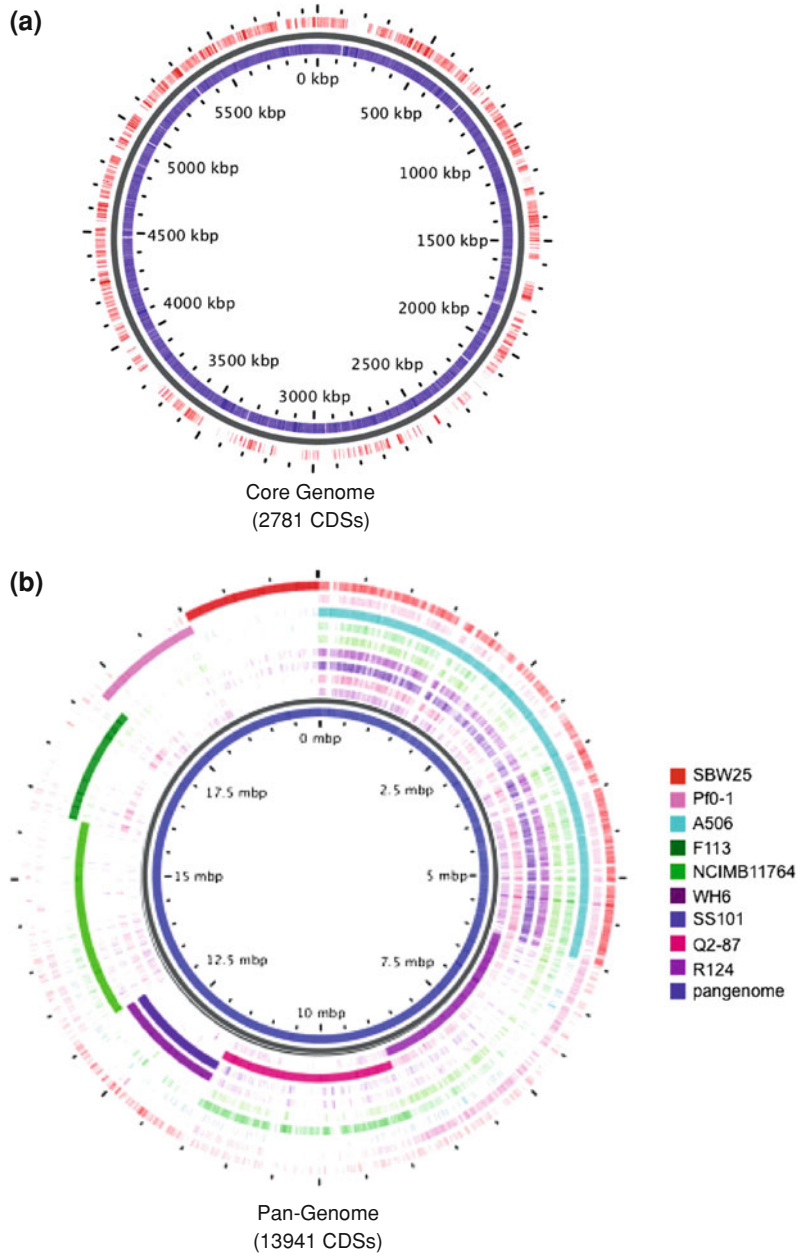

Fig. 5.3 Core genome and pan-genome analyses of selected *P. fluorescens* genomes. **a** The core genome (the set of genes shared by all *P. fluorescens* strains under analysis). The *inner circle* (*blue*) shows all of the predicted CDSs in the genome of strain A506, the smallest of the sequenced genomes. The *outermost circle* (*red*) indicates only those CDSs for which a BLAST hit (>60 % identity) was present between the A506 reference and all of the genomes in the query (all *P. fluorescens* strains listed in Table 5.1). Of note is the large proportion of core genes clustered around the replication origin, with less conservation around the terminus, as discussed in the text. **b** The full complement of genes, or pan-genome, of the nine *P. fluorescens* strains analyzed. The pan-genome, represented by the *innermost circle* (*blue*), was constructed by concatenating all the unique regions among each of the nine query genomes. Subsequently, a BLAST atlas was created for each query genome, as shown in the *outer circles*. The strains appear in the following order, from *outer circle* to *inner*: SBW25, Pf0-1, A506, F113, NCIMB11764, WH6, SS101, Q2-87, and R124. Analysis and images done using the GView server (Petkau et al. 2010)

proportion around the origin of replication (Fig. 5.3a) and strain-specific genes highest in the third of the genome surrounding the terminus, consistent with previous observations that levels of synteny decrease with distance from the origin (Loper et al. 2012; Silby et al. 2009).

The remarkable diversity of habitat and ecology of *P. fluorescens* group is further reflected in the large pan-genome of the sequenced members of the species. The pan-genome of the strains sequenced by Loper et al. (2012) consists of 13,872 putative protein-coding genes, a substantially larger pan-genome that has been estimated for *Pseudomonas aeruginosa* and *Pseudomonas syringae*. A similarly sized pan-genome is apparent in our analysis of the nine sequenced strains that retain the *P. fluorescens* designation (13,914 CDSs; Fig. 5.3b). Furthermore, of the CDSs comprising the *P. fluorescens* pan-genome, 5,798 do not have orthologs in the genomes of other *Pseudomonas* spp., reflecting high-level differentiation of genes in the group and likely a high frequency of horizontal gene acquisition from other taxa (Loper et al. 2012). Both the large size of the pan-genome and the restricted group of core genes are likely reflective of the varied lifestyles of members of the group and the distinct circumstances they must deal with in their different niches. Indeed, the inventory of genes present in *P. fluorescens* isolates is likely to continue to increase as new genome sequences become available.

5.2.1.2 Repetitive and Mobile Genetic Elements

The large inventory of genes in the *P. fluorescens* complex is due, in part, to the presence of horizontally acquired mobile genetic elements (MGEs) including prophages, plasmids, and transposons. Such elements are common in the genomes of environmental bacteria and are often niche specific. While transient, those elements harboring genes that confer a fitness advantage are selected and become a part of the flexible gene pool of a particular species or group.

Prophages. Analysis of *Pseudomonas* genome sequences reveals that most contain prophages formed when temperate bacteriophages integrate into the host genome (Loper et al. 2012; Mavrodi et al. 2009; Silby et al. 2009). These genetic elements not only encode phage-related functions, but can also harbor accessory, or cargo, genes that alter the phenotype and fitness of the host and thus allow survival in new ecological niches. For instance, phage-related genes are induced in *P. fluorescens* strains confronted with a bacterial competitor (Garbeva et al. 2011) and under environmental stress such as desiccation (Silby, unpublished), indicating a potential ecological role for these genetic elements. Alternatively, temperate phages that remain capable of development through the lytic pathway ultimately result in phage-induced host cell lysis. Thus, maintenance of prophages can adversely affect populations of rhizosphere-inhabiting *P. fluorescens* (Keel et al. 2002) and can also provide a mechanism for exchange of genes, leading to the generation of increased diversity among related strains.

Comparative analysis of genomes of the *P. fluorescens* group reveals a number of phage elements, with each genome encoding one to six prophage regions (Loper et al. 2012). The genome of *P. protegens* Pf-5, for instance, encodes six prophage elements, collectively spanning over 130 kbp of the genome. Notable are the prophages (designated prophage 01, in each genome) inserted into a common position between *mutS* and *cinA* in most of the sequenced strains (Loper et al. 2012; Mavrodi et al. 2009). These putative prophages display mosaic structure and appear to have undergone widespread recombination. Given the absence of some essential phage genes for head development, it is possible that these all encode putative F-type, R-type, or hybrid F/R-type pyocins. Pyocins, which are proteinaceous, narrow-spectrum bacteriocins, can kill sensitive strains of *Pseudomonas* and thus may provide a selective advantage in natural habitats. Indeed, widespread distribution of F-type and R-type pyocins among *P. fluorescens* soil isolates (Mavrodi

et al. 2009) indicates positive selection for the maintenance of these elements. Unlike the prophage 01 elements, many of the prophage regions carry a set of cargo genes that are distinct from those in prophages of other strains. Putative proteins encoded by phage cargo genes include additional bacteriocins, UV resistance proteins, putative fimbrial subunits, adenine- and cytosine-specific DNA methyltransferases, and an array of conserved hypothetical proteins whose functions remain unknown (Loper et al. 2012). Prophage regions represent a portion of the flexible gene pool of *P. fluorescens* strains, but still only account for a fraction of the strain-to-strain variability that is evident in the *P. fluorescens* complex.

Repetitive extragenic palindromic (REP) elements. REP elements—short (approximately 20 bp) extragenic sequences with palindromic structure—are the predominant class of repetitive DNA in bacterial genomes. REP elements are abundant in the intergenic regions of *Pseudomonas* genomes, and while the function of these sequences remains in question, roles in mRNA stability, insertion of genetic elements, binding of proteins such as IHF, DNA polymerase I, and DNA gyrase have been suggested. An elegant study of the occurrence and distribution of REP elements in the genome of *P. fluorescens* SBW25 (Bertels and Rainey 2011a) revealed three distinct sequence groups of REPs (GI, GII, and GIII) and provided the first evidence that individual REP elements are in fact part of a larger mobile element consisting of a REP doublet in inverted orientation or REP doublet forming hairpin (REPIN). Population sequencing data revealed that the REP doublets, or REPINs, are excised from the chromosome, suggesting that these elements are active MGEs. REPINs, however, are too short to encode their own transposition machinery, and thus, their movement requires activity encoded by some other fully autonomous element or an orphan transposase gene. Indeed, proximal to the REPIN elements are often found transposases that are distantly related to those of the IS*200* family of insertion sequences. It is these so-called RAYTs, REP-associated tyrosine transposases, that likely mediate transposition of the REPINs.

The possibility that REPINs are transposable elements is intriguing given their widespread distribution in *P. fluorescens* genomes. At least one type of REP element, occurring at least 250 times, is observed in each of the completely sequenced genomes of the *P. fluorescens* group, with the exception of strain Pf0-1. Across all of the REP-containing strains, a primary conserved REP sequence (REPa) can be found; however, the REPa sequence is not identical among the strains and many strains encode more than one distinct REP class. In addition to the primary REPa element, secondary REP elements are present in a number of genomes: REPb is present in *P. chlororaphis* strains 30-84 and O6, and *P. protegens* Pf-5; REPc and REPd display scattered distribution among the sequenced strains; REPe is only found in the genome of SBW25. Distinct RAYT proteins were identified as associated with each REP class, with the exception of REPc. As described above and in (Bertels and Rainey 2011a, b), the majority of REP sequences were found as oppositely oriented pairs separated by a uniform distance, lending further support to the REPIN as the minimal mobilizable unit. Differences in the distribution and abundance of REP elements among closely related *Pseudomonas* strains suggest that the evolution of such elements is independent of the core genome. However, this does not appear to be the case for the associated RAYT-encoding genes, as phylogenetic analysis of RAYT sequences closely resembles that seen in the MLSA tree of the *Pseudomonas* strains, suggesting that the RAYT genes may have been a stable part of the genomes since their divergence (Mavrodi et al. 2009). While the potential influence on genome evolution is considerable, it is important to keep in mind that experimental evidence for mobility of REPINs is limited (Bertels and Rainey 2011b). Interestingly in strain Pf-5, the RAYT associated with REPa sequences contains a nonsense mutation likely rendering the RAYT non-functional. Nonetheless, the Pf-5 genome contains 999 copies of REPa, indicating that these REP sequences were

dispersed prior to the RAYT mutation or that the REPINs may be mobilized through the activity of some other transposase. Much about the origin, maintenance and dissemination of these elements, as well as their possible function, remains to be determined.

Plasmids. The genomes of plant-associated *Pseudomonas* species frequently contain a significant component of extrachromosomal DNA in the form of one or more plasmids (Vivian et al. 2001). Because plasmids often encode context-dependent benefits that facilitate survival and colonization in a specialized niche or habitat, the transfer of these plasmids within and between bacterial populations is thought to facilitate adaptation to novel environments (Gal et al. 2003).

Long-term field studies of the phytosphere microflora of sugar beet at a site in Oxford, UK, revealed an abundance of self-transmissible plasmids, collectively called the pQBR plasmids, that transfer freely among the resident *Pseudomonas* community (Lilley and Bailey 1997b). At least five genetically unrelated plasmid replicons were present at the site, representing five distinct sequence groups (groups I–V) as determined using RFLP analysis. The identified plasmids were selected based on the ability to confer mercury resistance, and all harbor a Tn*5041*-like transposon encoding the *mer* operon. However, the large size of the plasmids (ranging from 60 to over 400 kb) and the persistence of the plasmids during repeated sampling over several years suggested that additional environment traits, perhaps those for adaptation to or survival in the terrestrial plant environment, may select for maintenance of these plasmids in the field.

A substantial body of work has genetically (Lilley and Bailey 1997b; Tett et al. 2007; Turner et al. 2002; Zhang et al. 2004b) and ecologically (Lilley and Bailey 1997a, 2002; Zhang et al. 2004a, b) characterized plasmids of the pQBR group. pQBR103, isolated from *P. fluorescens* SBW25, was the first representative to be fully sequenced and at 425 kb is one of the largest self-transmissible plasmids yet sequenced from the phytosphere microbial community (Tett et al. 2007). Sequence analysis confirmed that pQBR103 carries a near-perfect copy of the Tn*5042* Hg^R transposon, a common characteristic of environmental plasmids. RulAB homologues were also identified, which explains the enhanced UV resistance the plasmid confers upon *P. fluorescens* SBW25 (Zhang et al. 2004b), a factor known to influence survival of pseudomonads in the light-exposed phytosphere environment. Strikingly, pQBR103 lacks genes often found in other large environmental plasmids, including those involved in nutrient uptake and utilization (Gal et al. 2003). In fact, of the nearly 500 predicted coding sequences housed on pQBR103, only 20 % show similarity to known proteins or functional domains, making it difficult to predict the functional role of the plasmid in the phytosphere. However, carriage of the plasmid in *P. fluorescens* SBW25 enhances fitness of bacteria colonizing mature sugar beet plants (Lilley and Bailey 1997a) and the plasmid persists at the same field site over multiple years, strongly suggesting that pQBR103-encoded traits confer some selectable advantage in the terrestrial plant environment. pQBR103-encoded genes that are specifically induced on the plant surface included three helicase-like genes (*helA*, *helB*, and *helC*), an oligoribonuclease gene (*orn*), an AlgZ-like transcriptional regulator, a response regulator protein, and the three Tn*5042* transposase subunits, among others. Interestingly, expression of a greater proportion of pQBR103 genes than SBW25 chromosomal genes is induced while growing on plant surfaces (Zhang et al. 2004b), indicating that the coding density of ecologically important traits is quite high.

The majority of the fully sequenced *P. fluorescens* strains, many of which were propagated in the laboratory for years prior to sequencing, do not harbor plasmids; however, it is likely that pseudomonads in natural environments commonly acquire and share plasmids that are stably or transiently maintained. Indeed, genome sequencing of *P. fluorescens* A506 revealed the presence of an approximately 57-kbp conjugative plasmid, whose sequence features and ecological function have recently been characterized (Stockwell et al. 2013). The plasmid, designated

pA506, is a member of the well-characterized pPT23A plasmid family and contains 67 predicted coding sequences arranged in a highly mosaic structure composed of blocks of genes of variable G + C content, likely correlating with plasmid regions from distinct origins. Among the predicted genes are those involved in plasmid replication and partitioning, as well as elements typical of certain integrative conjugative elements (ICEs) found in the chromosomes of *Pseudomonas* species, suggesting a widespread mechanism for the transfer of traits among *Pseudomonas* populations. Genes related to ecological fitness of the host bacterium were also identified, including those encoding a putative pyocin immunity protein, which may protect strain A506 from pyocins produced by other bacteria in the environment. Like pQBR103, *rulAB* genes, which are involved in tolerance to UV irradiation, are present. Indeed, strain A506 is significantly more tolerant to UV irradiation than a plasmid-cured derivative. Also like pQBR103, a large percentage of the gene inventory of pA506 encodes proteins of unknown function, and thus, ecological benefits of carrying pA506, beyond resistance to UV, are difficult to predict. Interestingly, pA506 is mobilizable not only to other pseudomonads, but also to members of the *Enterobacteriaceae* and thus may represent a vehicle for genetic exchange between diverse bacteria in soil and on plant surfaces (Stockwell et al. 2013). In addition, because the plant epiphytic A506 is available as a commercial biocontrol agent (marketed under the name BlightBan), the role of plasmid pA506 in epiphytic fitness and biocontrol will be of continued interest.

Genomic Islands. Members of the *P. fluorescens* complex carry a number of genomic islands, including the so-called ICEs, which are plasmid-like in their replication and conjugation machinery but additionally encode phage-like integrase genes to allow for site-specific integration into the host chromosome. In addition to genes required for replication and transfer, ICEs often carry unique "cargo" genes that reflect the specific lifestyle of their hosts and may include pathogenicity factors and metabolic, regulatory, and transport genes. In *P. aeruginosa* and *P. syringae*, ICEs encode pathogenicity factors that enable host colonization, while an ICE in *Pseudomonas* sp. strain B13 encodes genes that allow the host to metabolize chlorinated aromatic compounds. The *P. protegens* Pf-5 genome contains two ICEs (Mavrodi et al. 2009). The 115-kb genomic island 01 (PFGI-1) is integrated into one of two $tRNA^{Lys}$ genes, in a region of conserved synteny likely representing an integration "hot spot" for ICEs in *Pseudomonas* spp. The first two-thirds of PFGI-1 encode plasmid replication, partitioning, and conjugation functions; however, it is not known whether PFGI-1 exists as an episome in strain Pf-5 or any other *Pseudomonas* host. It is not clear what role, if any, PFGI-1 may play in the survival of *P. protegens* Pf-5 in the rhizosphere. However, the element contains a unique 35-kb DNA segment that encodes genes not related to plasmid maintenance or transfer, many of which do not have homologs elsewhere in the Pf-5 genome or in the genome of related strains including *P. fluorescens* Pf0-1. These PFGI-1 cargo genes encode functions that may enhance rhizosphere adaptation, including a non-heme catalase which may protect the host from oxidative stress, both the kinase and response regulators of a two-component system that may enable the host to sense and respond to changes in the environment and other genes with putative roles in protection from environmental stress.

The function of the second genomic island in *P. protegens*, PFGI-2, is even more cryptic. PFGI-2 is similar to ICE-like elements in Pf0-1 and SBW25 (Silby et al. 2009). Approximately half of the small (16.8 kb) PFGI-2 element contains a gene cluster found in the exchangeable effector locus (EEL) of a tripartite type III secretion pathogenicity island (T-PAI) from the plant pathogen *P. viridiflava*. However, strain Pf-5 does not encode a type III protein secretion pathway, suggesting that PFGI-2 is inactive and simply represents a T-PAI remnant anchored in the chromosome (Mavrodi et al. 2009). Similar ICEs are also present in other sequenced members of the *P. fluorescens* group. Strain BG33R, for instance, encodes a 154-kb genomic island, with

similarity to the PFGI-1 island in Pf-5. However, unlike PFGI-1, the BG33R island contains genes encoding uptake and catabolism of indole-3-acetic acid, quinolones and haloaromatic compounds, as well as a MexCD drug efflux pump and other transporters, a toxin-like subunit protein and several transposases (Loper et al. 2012). Overall, individual members of the *P. fluorescens* complex carry from two to seven genomic islands, and these elements contribute to the heterogeneity of the strains and likely provide functions specific to the habitat of the host strain. Whether some or all of them are capable of excision from their genomic location is unclear. Functional studies similar to those done for islands in *P. aeruginosa* (Mathee et al. 2008) and *P. syringae* (Godfrey et al. 2013) are required if we are to fully understand the data generated by DNA sequencing.

5.2.2 Genetic Capacity for Secondary Metabolite Production and Biocontrol

From an application point of view, critical features of *P. fluorescens* genomes are the genes for biosynthesis of antifungal and antibacterial metabolites and the genetic circuits required for their regulation. Genetic approaches have revealed much about the pathways to make a range of such compounds. Although there was intense genetic research over many years, the rise of genomic tools has enabled the discovery of hitherto unknown compounds, expanding completeness of knowledge on the repertoire of antimicrobial compounds made. Genomic approaches also allow rapid appraisal of the variation in biosynthetic gene content for phenazines, which reflects the nature of the end product, and has given insight into the evolutionary history of the genes responsible for 2,4-DAPG production.

5.2.2.1 Repertoire of Secondary Metabolites

Pseudomonas strains synthesize an array of chemically diverse metabolites with antibacterial, antifungal, and antihelminthic activity. The production of such compounds is irrefutably linked to biocontrol efficacy, and *Pseudomonas*-produced compounds that exhibit activity against agriculturally relevant pathogens include 2,4-diacetylphloroglucinol (Landa et al. 2002), hydrogen cyanide (Voisard et al. 1989), phenazines (Hofte and Altier 2010; Huang et al. 2004), pyrrolnitrin (Hammer et al. 1999), pyoluteorin (Howell and Stipanovic 1980), viscosinamide (Nielsen et al. 1999), and coronatin (Bender et al. 1998), among others. The molecular basis for some, such as phenazines and 2,4-DAPG, was characterized with "traditional" molecular genetic techniques, but other important compounds such as orfamide remained unknown prior to whole-genome analysis (Gross et al. 2007). Genomic analysis has also been instrumental in extending the understanding of the evolution of *phl* genes and the production of phenazine variants (discussed below). Whole-genome surveys additionally reveal the patchy distribution of many of the secondary metabolite gene clusters, with single strains harboring quite different repertoires of secondary metabolite biosynthetic loci (Fig. 5.4a). Of the gene clusters linked to bioactive metabolite production, only the genes specifying pyoverdine synthesis are present in all sequenced *P. fluorescens* strains. Other biosynthetic loci are present only in a subset of strains and can often be found in different genetic contexts, indicating a complex pattern of inheritance punctuated by acquisition and/or loss of gene clusters throughout the evolution of these strains. An example of such complex inheritance is highlighted by the distribution of 2,4-DAPG biosynthesis in members of the *P. fluorescens* group (Fig. 5.4b).

The distribution of 2,4-diacetylphloroglucinol biosynthesis. The low molecular weight phenolic 2,4-diacetylphloroglucinol (DAPG) is implicated in myriad ecological functions and contributes to the biocontrol of plant disease due to its toxicity to a wide range of plant pathogenic fungi, bacteria, and helminthes (Haas and Defago 2005; Laville et al. 1992; Raaijmakers et al. 1997, 2002). DAPG additionally can affect the plant host by promoting amino acid exudation from plant roots (Phillips et al. 2004; Brazelton et al.

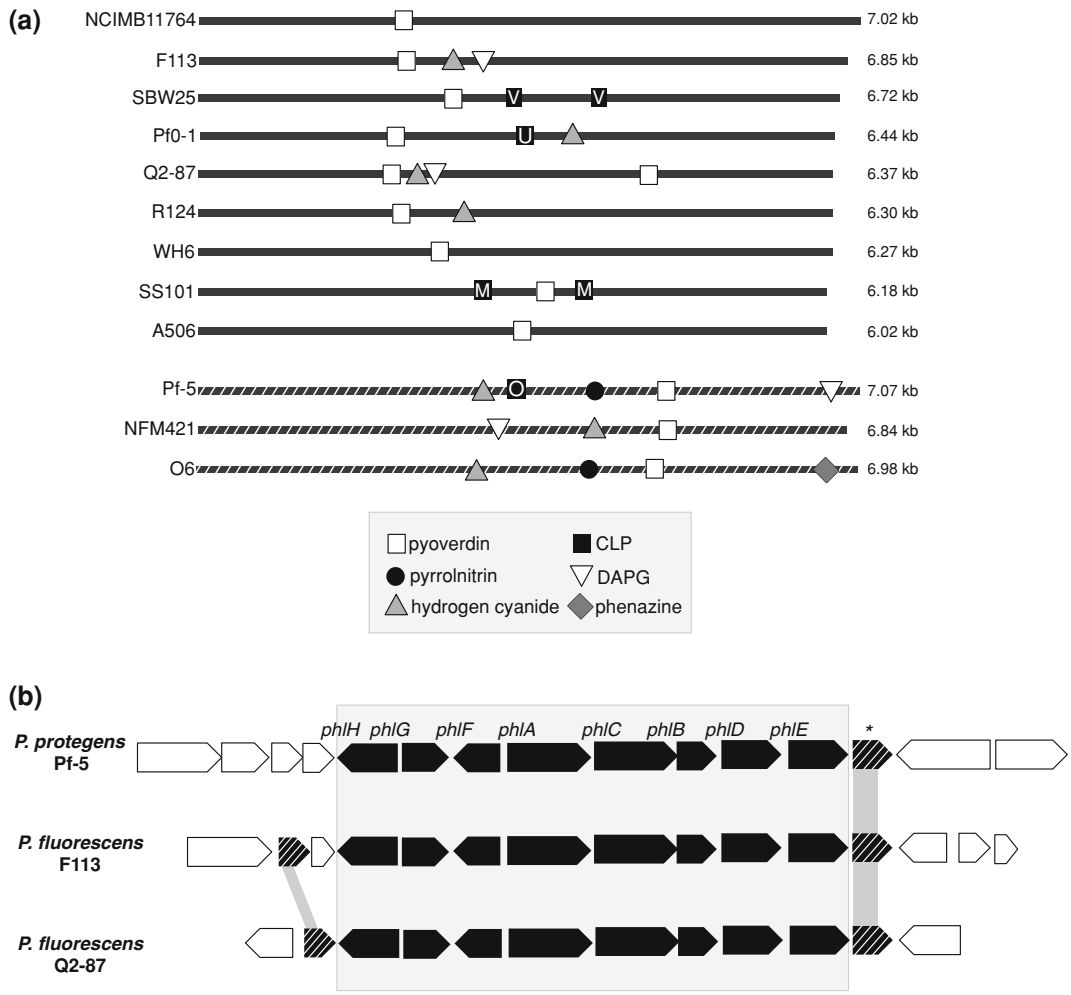

Fig. 5.4 Genomic distribution and context of secondary metabolite genes in sequenced strains of the *P. fluorescens* group. **a** The distribution of secondary metabolite gene clusters is analyzed in the nine sequenced strains designated as *P. fluorescens* (*solid gray*) and in representative members of the *P. protegens*, *P. brassicacearum*, and *P. chlororaphis* species (*dashed*). The presence of metabolite biosynthetic genes, along with the approximate location, is shown along the length of the genome. Selected secondary metabolite gene clusters are as follows: pyoverdin (*open square*), pyrrolnitrin (*closed circle*), hydrogen cyanide (*gray triangle*), cyclic lipopeptide (*closed square*), 2,4-diacetylphloroglucinol (*inverted triangle*), and phenazine (*gray diamond*). The identity of a particular CLP is represented by a letter designation: *O* orfamide A; *M* massetolide; *V* viscosin; or *U* unknown. Note that for the pyoverdin biosynthetic genes, the genomic location is representative of the locus encoding *pvdIJ*, and the remaining genes are spread throughout different locations in the respective genomes (for example, 7 different loci in SBW25 and 3 different loci in Pf0-1 and Pf-5). Similarly, the CLP biosynthetic genes are spread across two loci in strains SS101 and SBW25, as shown by distinct boxes. **b** The *phl* gene cluster is located in three different genetic contexts in the sequenced *P. fluorescens* genomes. The sequence and organization of the *phlHGFACBDE* cluster are conserved in all DAPG-producing *P. fluorescens* strains (*black arrows* within *boxed region*). The sequence immediately 3′ of the *phl* cluster is homologous among the main genomic location groups (dashed arrows). In the 5′ flanking region, strains F113 and Q2-87 share homology, but are distinct from that of Pf-5. The presence of a gene in the region following *phlE* in all DAPG producers (*asterisk*), but absent in any other *Pseudomonas* strain, suggests that this may constitute an additional *phl* gene, although its function has yet to be determined

2008) and by triggering systemic resistance leading to enhanced plant protection, as has been shown for *Arabidopsis* against fungal and bacterial pathogens (Iavicoli et al. 2003). Despite the diverse activities and apparent ecological benefits of DAPG, its production is limited to only a subset of *P. fluorescens* strains. First discovered in *P. fluorescens* F113 (Fenton et al. 1992), the 8-kb *phlHGFACBDE* biosynthetic locus is shared by all DAPG producers and comprises genes for biosynthesis (*phlACBD*) and degradation (*phlG*) of DAPG, with additional genes important for the regulation (*phlF* and *phlH*) of DAPG production and its transport (*phlE*). An additional gene, *phlI*, lacks significant homology to genes in the databases and was only recently linked to DAPG production because of its presence in alignments of the *phl* locus in multiple strains (Moynihan et al. 2009) (Fig. 5.4b).

While the *phl* locus is conserved at the sequence and organizational levels in DAPG producers, the capacity to produce DAPG is restricted to only a minority of *P. fluorescens* strains with no other bacteria having the capacity to produce this compound (Troppens et al. 2013). Also striking is the observation that the genomic context of the *phl* operon is variable, as a comparison of the flanking genes reveals at least three different genetic loci (Fig. 5.4b) in sequenced members of the *P. fluorescens* complex (Moynihan et al. 2009). There is a clear phylogenetic divide between producers and non-producers. Among the producers, there are three distinct phylogenetic clusters (Moynihan et al. 2009). In each cluster, the *phl* genes are at a specific location, and this location varies between the different clusters. The relatively low frequency of phloroglucinol production in strains of the *P. fluorescens* complex combined with the presence of the *phl* cluster in different genetic contexts is reminiscent of horizontally transferred genetic elements. Indeed, recent reports highlight the complex evolution of the *phl* locus and suggest that *phlD* and the *phlACB* operon have distinct evolutionary origins and likely were acquired in two separate horizontal gene transfer events (Troppens et al. 2013). Initial theories that *phlD* was acquired from plants, which were based on the similarity of PhlD to plant chalcone synthases, have been shown to be incorrect. Instead, the *phlD* gene in *Pseudomonas* appears to have its origins in the *Actinobacteria*. The *phlACB* operon, however, appears to have been acquired independently, and the observation that homologs of these three genes are found in multiple archaeal groups suggests that the *phlACB* cluster was acquired from an archaeon. *phlD* was likely acquired first, as no lineages lacking *phlD* but carrying *phlACB* have been identified. Also striking is the observation that *phlD* and *phlACB* are clustered in all DAPG-producing strains, indicating selective pressure for their linkage, as is often the case for biosynthetic and catabolic gene clusters in bacteria (Reams and Neidle 2004).

Further complicating the issue is the apparently widespread loss of the *phl* locus, some estimates suggesting that as few as 1 % of *P. fluorescens* isolates are capable of DAPG production (Moynihan et al. 2009; Troppens et al. 2013). Given the presumed ecological functions of DAPG in antibiosis and cell-to-cell signaling, it is difficult to imagine why the locus has been retained in so few lineages. Studies of the frequency and geographic distribution of DAPG producers offer some insight into the ecological context that selects for *phl* maintenance. For instance, studies from diverse geographic locations and using soils of varied physiochemical composition indicate that DAPG producers are highly enriched in soils suppressive to take-all decline of wheat (de Souza et al. 2003; Landa et al. 2002; Raaijmakers et al. 1997). The proportion of *phl*+ and *phl*− populations is also affected by the resident plant population, as specific plant cultivars exhibit a differential capacity to select for specific resident populations and long-term crop monoculture leads to an enrichment of 2,4-DAPG producers (McSpadden Gardener et al. 2005; Landa et al. 2006). Overall, genomic analysis of *Pseudomonas* spp. from complex environments reveals how lineages become specialized to a particular niche, necessitating maintenance and loss of genes required for fitness under specific conditions.

The genetics of phenazine production. The bacterial phenazines include more than 50 known nitrogen-containing heterocyclic pigments that contribute to the ecological fitness and biocontrol activity of the producing strain. Phenazines are produced by diverse bacterial genera, including *Pseudomonas*, *Burkholderia*, *Streptomyces*, *Pectobacterium*, and *Brevibacterium*, as well as by the archaeon *Methanosarcina mazei*. Fluorescent pseudomonads remain the best-studied phenazine producers, and while *P. fluorescens* strains were previously thought to produce only the yellow compound phenazine-1-carboxylic acid (PCA), recent genomic evidence reveals that the spectrum of phenazines produced is more broad.

Genetic and enzymatic details of PCA biosynthesis have been described by others in detail (Ahuja et al. 2008; Blankenfeldt et al. 2004; Mavrodi et al. 2001, 2004, 2006; Thomashow et al. 1990). In addition to PCA, which is the phenazine product of the core biosynthetic pathway coded for by the operon *phzABCDEFG*, *P. chlororaphis* 30-84 produces 2-hydroxy-PCA by hydroxylation of PCA mediated by the enzyme PhzO (Maddula et al. 2008). The genomic location of *phzO* is adjacent to the core PCA genes, as is also the case with genes responsible for core modification in other pseudomonads. For example, in *P. aeruginosa* PAO1, the *phz1* operon is flanked by *phzM* (upstream and divergently encoded) and *phzS* (downstream). Both PhzM and PhzS are involved in the modification of the PCA core; PhzS converts PCA to 1-OH-phenazine, while the combined action of PhzS and PhzM leads to the production of pyocyanin. When *phz* genes are identified by genome sequencing or genetic analysis, examination of surrounding sequence can yield genes with novel roles in the production of modified phenazine compounds (Mavrodi et al. 2001).

Taking advantage of whole-genome sequences enables rapid searching for genes related to those in other organisms, but which are not found in the same genomic context. Again using the PAO1 example, a *phzH* gene similar to that in *P. chlororaphis* PCL1391 was identified using BLAST. In *P. chlororaphis* PCL1391, *phzH* is downstream of the *phz* core genes and is responsible for conversion of PCA to phenazine-1-carboxamide (Chin et al. 2001). In PAO1, *phzH* is not clustered with the other *phz* genes and is most easily identified by genomics. Similar approaches are likely to lead to an expanded view of phenazine derivatives made by members of the *P. fluorescens* family (Mavrodi et al. 2001).

Recent work aimed at the ecology of phenazine production has given insight into the distribution of phenazine$^+$ pseudomonads in natural environments. In a study of over 80 dryland wheat fields in the inland Pacific Northwest, an abundance of phenazine$^+$ pseudomonads was detected in the rhizosphere of commercially grown cereals (Mavrodi et al. 2012a, b), where their presence was positively correlated with the transient accumulation of high levels of PCA. Surprisingly, phenazine$^+$ populations seem to flourish under arid conditions and population densities decreased in neighboring fields subject to irrigation, suggesting that phenazine$^+$ strains are uniquely adapted to the plant rhizosphere under conditions of water stress. These data illustrate the importance of the environment in selection of the different genotypes in the *P. fluorescens* group, a parameter which we suggest should be included when organizing members of the group.

Genomics and the discovery of new biocontrol metabolites. Examination of the genome sequences of *P. fluorescens* reveals the potential for the synthesis of numerous toxic metabolites, the production of which has not yet been detected. A recently published complete genome of a *P. fluorescens* isolate is that of F113 (Redondo-Nieto et al. 2012). In addition to loci known to be required for 2,4-DAPG and HCN production, Redondo-Nieto et al. (2012) discovered genes with similarity to lankacidin synthesis genes from *Streptomyces rochei*. The 15 genes, including polyketide synthase and non-ribosomal peptide synthetase modules, are clustered in *S. rochei* but separated into two clusters in F113. Comparative genome analysis showed that no other *Pseudomonas* genome

carries these genes, adding to the picture of broad and varied genome content among the *P. fluorescens* group. Further examination of the F113 genome indicates the potential to synthesize an insecticidal toxin which correlates well with the ability to kill insects, and numerous genes predicted to specify hemolysin/hemagglutinins and adhesin or agglutination proteins (Redondo-Nieto et al. 2012, 2013).

P. protegens Pf-5 has been intensively studied for its potential in biological control of plant pathogens for many years. Despite this detailed work, the genome sequence revealed that Pf-5 might make antimicrobial compounds that had not been detected previously. Indeed, using a "genomisotopic" approach, the cyclic lipopeptide (CLP) orfamide A was discovered (Gross et al. 2007). Orfamide A has a role in motility of Pf-5 and shows antimicrobial activity, indicating that it is likely an important compound for Pf-5 in its natural environment. The potential for CLP biosynthesis, which is mediated by non-ribosomal peptide synthetases (NRPSs), is quite straightforward to identify based on DNA sequence data. These synthetases consist of conserved modules which mediate amino acid activation (adenylation by the A-domain), thioesterification of the activated amino acid (by the T-domain), and peptide bond formation (by the condensation C-domain) which adds amino acids to the growing chain. The modules are generally in an order which corresponds with the amino acids they add to the chain, which aids in the prediction of the primary amino acid sequence, which can be accomplished using Web-based tools, for instance NRPS predictor 2 (Rottig et al. 2011). Thus, the CLP sequence, which varies with respect to the length and composition of the peptide moiety even among closely related *Pseudomonas* strains, can be predicted solely from DNA sequence. Whether the non-ribosomal peptide is cyclic can be predicted by the examination of the TE domain which releases the peptide from the NRPS. In the case of CLPs, the TEs (known as peptide cyclases) have characteristic amino acid sequence features distinct from those TEs which release linear peptides (Sieber and Marahiel 2005). This technique has been used, for instance, to predict the structures of as yet uncharacterized CLPs from Pf0-1 and BG33R, which include an 11- and 9-amino acid peptide, respectively (Loper et al. 2012).

Genome mining has also led to the identification of a large number of orphan gene clusters—that is, loci with sequence characteristic of secondary metabolite genes but without known biosynthetic products. Work on one such orphan NRPS/PKS cluster combined both genomic analysis and metabolic profiling and ultimately led to the discovery of five rhizoxin analogs with activity against the phytopathogens *Botrytis cinerea* and *Phytophthora ramorum* and toxicity toward plants and human cell lines (Loper et al. 2008). Additional orphan clusters are predicted to encode the biosynthesis of novel siderophores, based on sequence similarity to siderophore biosynthetic loci in other bacteria. Soil pseudomonads likely produce a repertoire of natural products that have yet to be identified, and these examples illustrate the power of genomic analysis for the discovery of novel systems.

5.2.3 Genetic Tools and Tractability

To gain the maximum benefit from the availability of genome sequences, and to increase our understanding of these sequences, experimental studies are required. Critical to these are genetic experiments which seek to ascribe function to predicted genes and add functional data to models built with genome and transcriptome data.

Pseudomonas sp. are generally straightforward to manipulate genetically. DNA can be introduced by conjugation, and for several species, electroporation is also established (Itoh et al. 1994; Smith and Iglewski 1989). Combined with the many important traits, the tractability has likely driven much of the use of pseudomonads as model systems in research areas such as experimental evolution and biofilm development. The ability to do genetic, and more recently functional genomic, experiments

makes these bacteria popular choices when considering pathogenesis (*P. aeruginosa*, *P. syringae*) and environmental research (*P. fluorescens*, *P. putida*).

In recent years, a considerable collection of resources have been applied with success in *P. fluorescens*. The workhorses of genetics are replicative plasmids, suicide plasmids, and vectors for introducing transposons for mutagenesis. Several examples of each exist. Shuttle vectors based on the plasmids pUC18/19 for moving cloned DNA from *E. coli* to *Pseudomonas* were constructed in the 1990s (Schweizer 1991; West et al. 1994). These pUCP18/19 derivatives have been widely used, along with other vectors such as the pBBRMCS series (Kovach et al. 1994, 1995), which have the added advantage of being mobilizable by conjugation. Tools have also been developed which permit the straightforward generation of random and site-specific mutations. Several Tn*5* derivatives have been made which facilitate the creation of random mutants carrying a range of markers and with promoter fusions to reporters such as *lacZ*, *phoA*, and *luxCDABE* and the gene for GFP (de Lorenzo et al. 1990; Tang et al. 1999). For the construction of targeted mutants, suicide plasmids such as pUIC3 (Rainey 1999), pSR47s (Andrews et al. 1998), and pKO3 (Monds et al. 2007) have all been used successfully in allele exchange mutagenesis of *P. fluorescens* strains. Of note, the splicing by overlap extension approach (Horton et al. 1989) is relatively simple in *P. fluorescens* Pf0-1, allowing the construction of precise unmarked deletion mutants (Silby and Levy 2004). Recently, several useful Tn*7* constructs have become available. The Tn*7* generally integrates at a specific site downstream of the gene *glmS* in many bacteria, and this insertion is considered to be neutral in terms of effect on growth. This means that genes for complementation or overexpression can be cloned in Tn*7* and delivered to a predictable site in the genome (Choi et al. 2005). Tn*7* elements have been made which differ only in the antibiotic resistance gene (Monds et al. 2007; Silby and Levy 2008), which we have used to create essentially isogenic but differently marked versions of strains of Pf0-1. Other useful features in Tn*7* constructs include a P_{BAD} promoter enabling inducible expression of genes downstream (Damron et al. 2013) and cloned *luxCDABE* genes which can be used to report activity of cloned promoters or for the localization of bacteria by luminescence (Damron et al. 2013). Many other plasmids have been constructed for use in *Pseudomonas*. The few highlighted here demonstrate part of a broad range of tools available which help make *Pseudomonas* sp. attractive model systems.

5.3 *P. fluorescens* Functional Genomics

The increasing number of *P. fluorescens* genomes and the ever-improving annotation and comparative genomic tools have combined to generate a reasonably comprehensive picture of the genetic complement of members of this species. The great challenge is to understand this information and to do so in a context that is as meaningful as possible. Leveraging genomic sequence data and genome-wide expression technologies allows genomic activity to be viewed as a whole, complementing many years of genetic work in which genes and gene networks have been examined one at a time. For an organism like *P. fluorescens*, which dwells primarily in soils and on plants, the ultimate goal is to understand the organism in these natural conditions. The utilization of genome-wide investigations has begun to reveal the secrets of *P. fluorescens* strains living under non-standard culture conditions. Functional genomic studies are even starting to show us how *P. fluorescens* adapts to stress in its natural environment (Lim et al. 2012, 2013a; Varivarn et al. 2013) and how it behaves during encounters with other microbial species (Garbeva et al. 2011) and in response to host plants (Jousset et al. 2011; Kidarsa et al. 2012; Matilla et al. 2007).

5.3.1 Gene Expression Under Nutrient Limitation

In a number of microarray-based studies, laboratory culture conditions have been manipulated to address the transcriptional response of *Pseudomonas* spp. to nutritional conditions that the organism is likely to encounter in natural environments, including limited bioavailability of the micronutrients iron (Lim et al. 2012) and zinc (Lim et al. 2013a) and the presence of plant-secreted products such as tannins (Lim et al. 2013b). We generally consider such conditions as "extreme" because we are used to working under laboratory conditions. However, for the bacterium, conditions such as low iron availability are routinely encountered. Genome-wide studies can give us insight into the activities of *P. fluorescens* under environmentally relevant conditions and can begin to ascribe functional significance to the numerous "hypothetical" genes identified in genome annotation studies. These genomic studies should not be viewed as esoteric. Indeed, understanding *P. fluorescens* under conditions that get close to nature will allow better prediction of success and failure of applications associated with these bacteria. If we know some key environment-related traits which favor success, we know what to look for in new isolates being considered for applications such as biocontrol.

Iron, which acts as a cofactor for a number of essential enzymes, is a micronutrient necessary for bacterial growth. While often abundant in soil, iron exists primarily in the insoluble ferric oxide form and is thus not available for use by bacteria under aerobic conditions. *P. fluorescens* strains overcome the limited bioavailability of iron through a variety of mechanisms—secreting siderophores, encoding a barrage of TonB-dependent receptors that "steal" siderophores produced by other microorganisms, through haem uptake systems, by direct uptake of ferrous iron and use of iron storage proteins such as bacterioferritins (Andrews et al. 2003; Cornelis and Matthijs 2002; Hartney et al. 2011). Despite the importance of iron and its general low bioavailability in soil, how members of the *P. fluorescens* group respond to iron limitation has been unknown until a recent study investigated the effect of iron limitation on global gene and protein expression profiles of *P. protegens* Pf-5 (Lim et al. 2012). The transcription of over 300 genes was significantly altered (180 increased and 121 decreased) in iron-starved cells versus those supplemented with $FeCl_2$, indicating a broad genetic and physiological response to iron stress. As expected, genes encoding iron homeostasis were upregulated under iron-deplete conditions. These include genes for siderophore biosynthesis, membrane proteins to facilitate uptake of iron(II) and iron(III), and a bacterioferritin-associated ferredoxin that mobilizes iron stored in bacterioferritin B. Such genome-wide studies are revealing unexpected consequences of environmental conditions, such as the reduction in flagellar gene expression and presumably motility, as well as increased expression of a ribosomal protein and genes predicted to encode refractile inclusion bodies (R bodies). Inclusion of a source of ferric iron in the experiment rather than an iron-limited condition would take the setup a step closer to natural conditions and provide a picture of the whole transcriptome activity under a nutrient condition more relevant to the bacterium's "normal" habitat.

A complementary study addressed the genetic and physiological response of *P. protegens* to zinc limitation (Lim et al. 2013a). Zinc is a particularly important micronutrient, estimated to act as a cofactor in as many as 5 % of the proteins found in bacteria. While zinc is essential for many cellular functions, excess levels can be toxic, and thus, intracellular zinc concentrations must be finely modulated. Transcriptional profiles of *P. protegens* Pf-5 under zinc limitation reveal 73 and 28 genes that are significantly up- and downregulated, respectively. Genes with altered expression were from a variety of functional classes including, as expected, upregulation of the global regulatory gene, *zur*, involved in zinc homeostasis as well as a number of zinc-specific ABC transporters

and TonB-dependent receptors predicted to facilitate zinc uptake. Additional responses include the upregulation of the *xseA* DNA repair gene and reduced expression of a gene encoding the CopZ copper chaperone. Due to the requirement for zinc cofactors in various DNA replication and repair enzymes, increased expression of *xseA* may be a response to compensate for increased DNA replication errors during zinc limitation. Altered expression of *copZ* suggests interplay between copper and zinc homeostasis in *P. fluorescens*; however, the mechanism remains unclear. Again, application of genome-wide analytical tools has generated a picture of the impact of zinc on cellular physiology that is far broader than would be predicted, which most likely would not be captured by genetic or biochemical studies.

In addition to the role of micronutrients on normal cellular physiology, iron and zinc are linked to the expression of biocontrol phenotypes in some *Pseudomonas* species. In the case of *P. protegens* Pf-5, iron-deplete conditions resulted in altered expression of a number of genes involved in bioactive metabolite synthesis, including downregulation of hydrogen cyanide production (Blumer and Haas 2000; Lim et al. 2012) and upregulation of genes involved in the production of DAPG, pyrrolnitrin, and orfamide A (Lim et al. 2013a). On the contrary, downregulation of several genes involved in the DAPG production was observed in the transcriptional profile of Pf-5 under zinc depletion, supporting the observation that DAPG production can be enhanced by exogenous zinc amendment (Duffy and Defago 1997, 1999). Successful application of *Pseudomonas* spp. to biocontrol requires an increased understanding of how these organisms modulate bioactive metabolite synthesis in the field and during interaction with host plants. It is clear that external factors including local concentrations of micronutrients alter bioactive compound production, as does the presence of plant-exuded compounds (described below).

5.3.2 Gene Expression During Interaction with Plants and Plant-Exuded Compounds

Plants release into the rhizosphere a wide variety of low molecular weight compounds including polysaccharides, organic acids, phenolics, and amino acids from their roots, selecting specific microbial communities and modulating microbial activity. The link between plant-exuded compounds and changes in bacterial gene expression has been assessed, for instance, using reporter fusions to specific genes involved in bioactive metabolite synthesis (de Werra et al. 2011; Jousset et al. 2011). Given the high abundance of plant-derived compounds in soil, it is important to understand how such compounds may alter the overall physiology of rhizosphere pseudomonads. Recent studies, for example, those by Lim et al. (2013b), Matilla et al. (2007), and Kidarsa et al. (2012) assess global transcriptional profiles of *Pseudomonas* spp. to plant-derived compounds, roots, and seeds, respectively. Taken together, such studies provide evidence for the differential expression of hundreds of genes in response to plant-exuded compounds and during interaction with plant or seed surfaces. Genes with altered expression encode functions including biofilm formation, motility, type VI secretion, antibiotic and exoenzyme production, and iron homeostasis, among others. During *Pseudomonas* interaction with plant roots, Matilla et al. (2007) showed upregulation of several regulatory systems and genes for metabolism and transport, consistent with the expectation that changes in gene expression would reflect adaptation to different nutrient sources such as root exudates.

Microarray studies have proven extremely useful in improving understanding of gene expression of *P. fluorescens* under conditions which are moving toward mimicking natural environments. However, there is still some way to go. These studies represent a significant step and will inform the next generation of

experiments as environmental conditions begin to be approximated with greater accuracy. Even more powerful, however, is the ability to assess global bacterial gene expression directly in a complex environment. While challenging, this has been attempted using in vivo expression technology and is becoming more accessible with the use of transcriptome sequencing. Soon, analysis of gene expression in the environment will be limited only by a researcher's ability to isolate a sufficient quantity and quality of mRNA from even the most complex microbial habitat.

5.3.3 Gene Expression in Natural Environments: Functional Genomics In Situ

As recently as thirty years ago, microbial activity in natural environments such as the rhizosphere remained somewhat of an unknown. Study of rhizosphere bacteria was restricted to pure cultures studied in the laboratory, limiting knowledge to single, cultivable species under relatively unnatural conditions. Advances in methodology including fluorescent in situ hybridization (FISH) and stable isotope probing (SIP) provided for the first time a means to detect and link functional activity to uncultured bacteria in the rhizosphere or other natural environments; however, detail at the genetic level remained absent.

An approach to identify and understand the function of genes in complex environments is the promoter-trapping strategy IVET (in vivo expression technology) (Mahan et al. 1993). Broadly, IVET allows the selection of bacterial promoters that are active in a specified niche. Such promoter-trapping strategies rely on positive selection of niche-induced promoters by using genetic complementation of a conditionally lethal mutation. Promoters active in the environment under study, but inactive under laboratory conditions, are isolated on a genome-wide scale independent of whether the loss of those sequences is lethal and thus allowing the recovery of essential and nonessential genes that contribute to ecological success. IVET was first applied to identify *Xanthomonas campestris* genes induced during infection of turnips (Osbourn et al. 1987) and has since been adapted to examine genes induced in pathogens during infection of hosts (Slauch et al. 1994), during *Rhizobium*–legume symbiosis (Oke and Long 1999), and when bacteria interact with fungi (Lee and Cooksey 2000) and plants (Boch et al. 2002).

Various IVET strategies have been successfully applied to study the adaptation of *Pseudomonas* species to different environments, including plant surfaces, the rhizosphere, and bulk soil (Gal et al. 2003; Rainey 1999; Silby and Levy 2004). Complementary IVET strategies, based on *panB* and *dapB*, were used to identify *P. fluorescens* SBW25 genes that are specifically induced in the sugar beet rhizosphere (Rainey 1999). In agreement with the saprophytic lifestyle of soil pseudomonads, one quarter of the identified genes were related to nutrient acquisition and metabolism. These include metabolism of organic acids and xylose as well as genes involved in amino acid and nucleotide uptake and transport, suggesting that *P. fluorescens* can utilize plant-exuded carbon sources and may scavenge amino acids and nucleotides as a source of carbon or nitrogen. An additional suite of genes with elevated expression in the rhizosphere were those involved in the oxidative stress response, suggesting that efficient elimination of the harmful effects of oxidative stress is an important factor in rhizosphere competence.

A similar *dapB*-based IVET approach identified 22 *P. fluorescens* Pf0-1 promoters with elevated expression in bulk soil (Silby and Levy 2004). Genes with predicted environmental relevance were again identified including those involved in nutrient acquisition and metabolism, nucleotide transport, and stress response. Of particular interest, however, was the surprising number of fusions in which the sequence driving expression of *dapB* had no similarity to known genes or proteins. Nearly half of the in vitro expressed sequences were these so-called

cryptic fusions, transcribed on the antisense strand opposite a known open reading frame. Individual antisense transcripts have been shown to function as antisense RNA regulators, while others encode a polypeptide, and it is clear that this class of genes plays a role in the ecological success of *P. fluorescens* Pf0-1 (Silby and Levy 2008). IVET experiments in other *Pseudomonas* spp. have identified the existence of additional sense/antisense transcriptional pairs (Silby and Levy 2004; Fernandez et al. 2013), suggesting an importance of this class of transcripts to the ecology of pseudomonads in a variety of niches. It is important to note that identification of antisense or non-predicted transcripts is not possible with traditional microarray studies described in the previous section, as these sequences would be missing from the array.

Transcriptome sequencing combines the genome-wide scope of microarrays with IVET's unbiased ability to discover previously non-predicted genes, using the high throughput of new-generation sequencing technology to sequence cDNA derived from all transcripts in a given sample. Often called RNA-seq, this approach allows quantification of gene expression, providing a snapshot of all transcripts under conditions of interest. Comparative approaches can then reveal the network of genes specifically important in each environment. For optimal utility, transcriptome sequence data should be mapped onto a complete genome sequence for which a high-quality annotation exists, although efficient de novo assembly of transcriptome data into contigs of expressed sequences means that even genotypes for which genome sequence data are not available are amenable to the approach.

To date, transcriptome sequencing of *P. fluorescens* strains has been limited to only a few environments. Our recent work focuses on understanding gene expression patterns during survival and persistence of *P. fluorescens* strain Pf0-1 in soil, under varying moisture content. To survive in nature, bacteria need to adapt to rapid and sometimes severe perturbations in their environment, including drastic swings in moisture content of soil, which varies naturally as a consequence of weather and varies in agricultural settings because of irrigation. Some studies have examined the response of *Pseudomonas* sp. to water stress in vitro, but these have relied on simulated conditions to reduce water availability (Chang et al. 2007; Gulez et al. 2012). We have completed a study of the genome-wide transcriptional response of Pf0-1 to dehydration, while the bacterium is growing in soil, providing the first insight into factors that are important for adaptation to dehydration in situ. When soil colonized by Pf0-1 was allowed to lose 40 % of its moisture, expression of >200 genes was elevated at least 5-fold as a rapid response (1 h) to dehydration and also after five days reflecting longer-term adaptation strategies. There was only limited overlap between rapid and long-term responses to dehydration, indicating a distinction between immediate stress response and broader stress tolerance. Genes specifying production of alginate and another extracellular polysaccharide were upregulated during water stress, suggesting a protective role for these products in dry soil. While alginate expression was induced immediately and remained high, genes for the second EPS were only induced after a longer dehydration time. Our data indicate that Pf0-1 uses different polysaccharides for specific purposes in adaptation to water stress. Our continued transcriptome analysis of Pf0-1 under a variety of environment conditions, including gene expression profiles in complex multispecies bacterial communities, moves forward the knowledge of pseudomonas behavior in natural settings and will improve our understanding of what features may enhance the biocontrol potential of these organisms.

5.4 Future Perspectives

Genome sequencing is now a routine task for many research groups. Computational tools for the assembly and preliminary annotation of new genome sequences are freely available, bringing the possibility of sequencing all new *P. fluorescens* isolates within reach. We expect many novel and useful traits to be identified by careful

examination of the genome sequences, followed by experimental studies. The pending deluge of new *P. fluorescens* genome sequences seems likely to increase the pan-genome estimate and decrease the core genome size. Given the variability in the group, we anticipate that it will take many sequences for the core and pan-genome numbers to reach a stable plateau. In addition, these new genome sequences will inevitably lead to more revisions of the species and of the *P. fluorescens* group. With the ease of generating sequence data, our expectation is that genome sequences will form a platform for sophisticated new functional and comparative studies using high-throughput genomic methodologies.

Potential and limits to biocontrol. A major long-term goal of comparative and functional genomics of *P. fluorescens* is the improvement in biological control strategies to combat plant pathogens. The major features of an effective biocontrol strain are the ability to compete and thrive in the target environment and the ability to inhibit the growth of the target pathogens. Genomic studies coupled with other approaches have begun to shed light on mechanisms for persistence in soil and on plant surfaces and have led to an expanded catalog of antimicrobial compounds that can be produced by *P. fluorescens*. It may be that none of the *P. fluorescens* isolates currently being investigated is the "complete package" for application in biocontrol; in fact, it is likely that many will be needed for use in different agricultural environments. The genomic information gleaned from the isolates already sequenced provides a catalog of features known to favor biocontrol, and because of the relatively low cost of genome sequencing, at present these traits can be readily identified in new isolates by comparative genomics. Those isolating these organisms can search for the ability to produce compounds known to be effective in combatting a particular pathogen and can determine the presence of a range of genes of importance in environmental performance. Importantly, the tools are available to permit the detection of novel genes for antimicrobial production.

Genomic investigations have provided a baseline from which to extend the repertoire of biological control agents and show promise for improvement of *Pseudomonas*-mediated biocontrol. Current attention, for example, makes use of sequence information and functional genomic studies to direct strain engineering and environmental augmentation for improved biocontrol. Genetic engineering strategies often focus on secondary metabolite biosynthesis. For instance, one study to improve strain efficacy involved insertion of PCA biosynthetic operon from *P. fluorescens* 2-79 into the chromosome of strain SBW25, a robust rhizosphere colonizer known to control a number of fungal phytopathogens but that does not produce PCA. The *phz* operon in the engineered SBW25 permitted strong constitutive expression of PCA production independent of cell density and significantly improved the ability of SBW25 to reduce damping-off disease of pea seedlings caused by *Pythium ultimum* (Timms-Wilson et al. 2000). Using a similar strategy, PCA production was engineered in *P. fluorescens* strain Q8r1-96, a superior root colonizer that produces the unrelated antibiotic DAPG and is known to control take-all disease of wheat. Recombinants expressing the PCA cassette not only produced high levels of PCA, but also showed increased production of DAPG, and were effective at controlling both take-all disease and *Rhizoctonia* root rot even when inoculated at doses one to two orders of magnitude lower than the control (Huang et al. 2004). The engineered derivatives additionally showed increased efficacy in field trials spanning three years, as wheat treated with the recombinant strains showed yields up to 20 % greater than those from treatments with the wild type (Mavrodi et al. 2006). Such studies confirm that the in situ performance of an already effective biological control agent can be significantly enhanced by the insertion of novel functional traits.

While it is clear that many *Pseudomonas* strains have the genetic capacity to produce an array of biocontrol traits, the presence of relevant genes is simply a first step along the path toward effective exploitation of these traits.

Understanding how such genes are regulated in complex environments, as well as an understanding of how to ensure survival and fitness of desirable populations in situ, is equally important. As such, studies of microbial community dynamics in natural environments are of increasing importance. For instance, in the case of some *P. fluorescens* strains, the presence of other, competing soil bacteria induces the production of bioactive metabolites and thus may enhance biocontrol potential (de Boer et al. 2007; Garbeva and de Boer 2009; Garbeva et al. 2011). Such interspecies bacteria–bacteria interactions may be a driving force for bioactive compound production in complex microbial communities, where this so-called chemical warfare may offer a competitive advantage to the producing strain. Other environmental factors may be equally important, as it is clear that nutrient availability (Lim et al. 2012, 2013a), soil moisture (Varivarn et al. 2013), and temperature (Arana et al. 2010) can have a drastic effect on both the survival of biocontrol organisms and expression of biocontrol traits.

We see the next generation of genomic analysis as being a combination of genome sequencing and high-throughput automated annotation, with transcriptome sequencing. The latter adds a layer of functional information, which can be integrated into metabolic models and combined with high-throughput metabolomic and proteomic studies to produce a system-wide view of the functioning of cells and populations. Critical to this approach is the choice of experimental frameworks which move closer toward "real-world" situations, ensuring that the models built have the potential to reflect life in natural environments rather than life as a monoculture in a flask of laboratory growth medium. Of equal importance is the realization in the post-genomic era that genomic analysis gives a view of the potential of an organism, but only with continued genetic, physiological, metabolic, and environmental experiments can the potential revealed in sequence and transcriptomic studies be fully characterized and understood.

References

Ahuja EG, Janning P, Mentel M, Graebsch A, Breinbauer R, Hiller W, Costisella B, Thomashow LS, Mavrodi DV, Blankenfeldt W (2008) PhzA/B catalyzes the formation of the tricycle in phenazine biosynthesis. J Am Chem Soc 130(50):17053–17061

Andrews HL, Vogel JP, Isberg RR (1998) Identification of linked *Legionella pneumophila* genes essential for intracellular growth and evasion of the endocytic pathway. Infect Immun 66(3):950–958

Andrews SC, Robinson AK, Rodriguez-Quinones F (2003) Bacterial iron homeostasis. FEMS Microbiol Rev 27(2–3):215–237

Arana I, Muela A, Orruno M, Seco C, Garaizabal I, Barcina I (2010) Effect of temperature and starvation upon survival strategies of *Pseudomonas fluorescens* CHA0: comparison with Escherichia coli. FEMS Microbiol Ecol 74(3):500–509

Baggi G, Boga MM, Catelani D, Galli E, Treccani V (1983) Styrene catabolism by a strain of *Pseudomonas fluorescens*. Syst Appl Microbiol 4(1):141–147

Beltrametti F, Marconi AM, Bestetti G, Colombo C, Galli E, Ruzzi M, Zennaro E (1997) Sequencing and functional analysis of styrene catabolism genes from *Pseudomonas fluorescens* ST. Appl Environ Microbiol 63(6):2232–2239

Bender CL, Palmer DA, Penaloza-Vazquez A, Rangaswamy V, Ullrich M (1998) Biosynthesis and regulation of coronatine, a non-host-specific phytotoxin produced by *Pseudomonas syringae*. Subcell Biochem 29:321–341

Bertels F, Rainey PB (2011a) Curiosities of REPINs and RAYTs. Mob Genet Elements 1(4):262–268

Bertels F, Rainey PB (2011b) Within-genome evolution of REPINs: a new family of miniature mobile DNA in bacteria. PLoS Genet 7(6):e1002132

Blankenfeldt W, Kuzin AP, Skarina T, Korniyenko Y, Tong L, Bayer P, Janning P, Thomashow LS, Mavrodi DV (2004) Structure and function of the phenazine biosynthetic protein PhzF from *Pseudomonas fluorescens*. Proc Natl Acad Sci USA 101(47):16431–16436

Blumer C, Haas D (2000) Iron regulation of the *hcnABC* genes encoding hydrogen cyanide synthase depends on the anaerobic regulator ANR rather than on the global activator GacA in *Pseudomonas fluorescens* CHA0. Microbiology 146(Pt 10):2417–2424

Boch J, Joardar V, Gao L, Robertson TL, Lim M, Kunkel BN (2002) Identification of *Pseudomonas syringae* pv. *tomato* genes induced during infection of *Arabidopsis thaliana*. Mol Microbiol 44(1):73–88

Brazelton JN, Pfeufer EE, Sweat TA, Gardener BB, Coenen C (2008) 2,4-diacetylphloroglucinol alters plant root development. Mol Plant Microbe Interact 21(10):1349–1358

Carver TJ, Rutherford KM, Berriman M, Rajandream MA, Barrell BG, Parkhill J (2005) ACT: the Artemis comparison tool. Bioinformatics 21(16):3422–3423

Chang WS, van de Mortel M, Nielsen L, de Guzman GN, Li X, Halverson LJ (2007) Alginate production by *Pseudomonas putida* creates a hydrated microenvironment and contributes to biofilm architecture and stress tolerance under water-limiting conditions. J Bacteriol 189(22):8290–8299

Chauhan A, Fazlurrahman, Oakeshott JG, Jain RK (2008) Bacterial metabolism of polycyclic aromatic hydrocarbons: strategies for bioremediation. Indian J Microbiol 48(1):95–113

Chin AWTF, Thomas-Oates JE, Lugtenberg BJ, Bloemberg GV (2001) Introduction of the *phzH* gene of *Pseudomonas chlororaphis* PCL1391 extends the range of biocontrol ability of phenazine-1-carboxylic acid-producing *Pseudomonas* spp. strains. Mol Plant Microbe Interact 14(8):1006–1015

Choi KH, Gaynor JB, White KG, Lopez C, Bosio CM, Karkhoff-Schweizer RR, Schweizer HP (2005) A Tn7-based broad-range bacterial cloning and expression system. Nat Methods 2(6):443–448 (nmeth765)

Cornelis P (ed) (2008) *Pseudomonas*: genomics and molecular biology. Caister Academic Press, Norfolk

Cornelis P, Matthijs S (2002) Diversity of siderophore-mediated iron uptake systems in fluorescent pseudomonads: not only pyoverdines. Environ Microbiol 4(12):787–798

Damron FH, McKenney ES, Barbier M, Liechti GW, Schweizer HP, Goldberg JB (2013) Construction of mobilizable mini-Tn7 vectors for bioluminescent detection and single copy promoter lux reporter analysis in Gram-negative bacteria. Appl Environ Microbiol 79(13):4149–4153

de Boer W, Wagenaar AM, Klein Gunnewiek PJ, van Veen JA (2007) In vitro suppression of fungi caused by combinations of apparently non-antagonistic soil bacteria. FEMS Microbiol Ecol 59(1):177–185

de Lorenzo V, Herrero M, Jakubzik U, Timmis KN (1990) Mini-Tn5 transposon derivatives for insertion mutagenesis, promoter probing, and chromosomal insertion of cloned DNA in gram-negative eubacteria. J Bacteriol 172(11):6568–6572

de Souza JT, Weller DM, Raaijmakers JM (2003) Frequency, diversity, and activity of 2,4-diacetylphloroglucinol-producing fluorescent *Pseudomonas* spp. in Dutch take-all decline soils. Phytopathology 93(1):54–63

de Werra P, Huser A, Tabacchi R, Keel C, Maurhofer M (2011) Plant- and microbe-derived compounds affect the expression of genes encoding antifungal compounds in a pseudomonad with biocontrol activity. Appl Environ Microbiol 77(8):2807–2812

Duffy BK, Defago G (1997) Zinc improves biocontrol of *Fusarium* crown and root rot of tomato by *Pseudomonas fluorescens* and represses the production of pathogen metabolites inhibitory to bacterial antibiotic biosynthesis. Phytopathology 87(12):1250–1257

Duffy BK, Defago G (1999) Environmental factors modulating antibiotic and siderophore biosynthesis by *Pseudomonas fluorescens* biocontrol strains. Appl Environ Microbiol 65(6):2429–2438

Fenton AM, Stephens PM, Crowley J, O'Callaghan M, O'Gara F (1992) Exploitation of gene(s) involved in 2,4-diacetylphloroglucinol biosynthesis to confer a new biocontrol capability to a *Pseudomonas* strain. Appl Environ Microbiol 58(12):3873–3878

Fernandez M, Conde S, Duque E, Ramos JL (2013) In vivo gene expression of *Pseudomonas putida* KT2440 in the rhizosphere of different plants. Microb Biotechnol 6:307–313

Foght JM, Westlake DW (1996) Transposon and spontaneous deletion mutants of plasmid-borne genes encoding polycyclic aromatic hydrocarbon degradation by a strain of *Pseudomonas fluorescens*. Biodegradation 7(4):353–366

Gal M, Preston GM, Massey RC, Spiers AJ, Rainey PB (2003) Genes encoding a cellulosic polymer contribute toward the ecological success of *Pseudomonas fluorescens* SBW25 on plant surfaces. Mol Ecol 12(11):3109–3121

Garbeva P, de Boer W (2009) Inter-specific interactions between carbon-limited soil bacteria affect behavior and gene expression. Microb Ecol 58(1):36–46

Garbeva P, Silby MW, Raaijmakers JM, Levy SB, Boer W (2011) Transcriptional and antagonistic responses of *Pseudomonas fluorescens* Pf0-1 to phylogenetically different bacterial competitors. ISME J 5(6):973–985

Gershman MD, Kennedy DJ, Noble-Wang J, Kim C, Gullion J, Kacica M, Jensen B, Pascoe N, Saiman L, McHale J, Wilkins M, Schoonmaker-Bopp D, Clayton J, Arduino M, Srinivasan A (2008) Multistate outbreak of *Pseudomonas fluorescens* bloodstream infection after exposure to contaminated heparinized saline flush prepared by a compounding pharmacy. Clin Infect Dis 47(11):1372–1379

Godfrey SA, Lovell HC, Mansfield JW, Corry DS, Jackson RW, Arnold DL (2013) The stealth episome: suppression of gene expression on the excised genomic island PPHGI-1 from *Pseudomonas syringae* pv. *phaseolicola*. PLoS Pathog 7(3):e1002010

Gross H, Stockwell VO, Henkels MD, Nowak-Thompson B, Loper JE, Gerwick WH (2007) The genomisotopic approach: a systematic method to isolate products of orphan biosynthetic gene clusters. Chem Biol 14(1):53–63

Gulez G, Dechesne A, Workman CT, Smets BF (2012) Transcriptome dynamics of *Pseudomonas putida* KT2440 under water stress. Appl Environ Microbiol 78(3):676–683

Haas D, Defago G (2005) Biological control of soilborne pathogens by fluorescent pseudomonads. Nat Rev Microbiol 3(4):307–319

Hammer PE, Burd W, Hill DS, Ligon JM, van Pee K (1999) Conservation of the pyrrolnitrin biosynthetic gene cluster among six pyrrolnitrin-producing strains. FEMS Microbiol Lett 180(1):39–44

Hartney SL, Mazurier S, Kidarsa TA, Quecine MC, Lemanceau P, Loper JE (2011) TonB-dependent outer-membrane proteins and siderophore utilization in *Pseudomonas fluorescens* Pf-5. Biometals 24(2):193–213

Hofte M, Altier N (2010) Fluorescent pseudomonads as biocontrol agents for sustainable agricultural systems. Res Microbiol 161(6):464–471

Horton RM, Hunt HD, Ho SN, Pullen JK, Pease LR (1989) Engineering hybrid genes without the use of restriction enzymes: gene splicing by overlap extension. Gene 77(1):61–68

Howell C, Stipanovic R (1980) Suppression of *Pythium ultimum*-induced damping off of cotton seedlings by *Pseudomonas fluorescens* and its antibiotic, pyoluteorin. Phytopathology 70:712–715

Hsueh PR, Teng LJ, Pan HJ, Chen YC, Sun CC, Ho SW, Luh KT (1998) Outbreak of *Pseudomonas fluorescens* bacteremia among oncology patients. J Clin Microbiol 36(10):2914–2917

Huang Z, Bonsall RF, Mavrodi DV, Weller DM, Thomashow LS (2004) Transformation of *Pseudomonas fluorescens* with genes for biosynthesis of phenazine-1-carboxylic acid improves biocontrol of rhizoctonia root rot and in situ antibiotic production. FEMS Microbiol Ecol 49(2):243–251

Iavicoli A, Boutet E, Buchala A, Metraux JP (2003) Induced systemic resistance in *Arabidopsis thaliana* in response to root inoculation with *Pseudomonas fluorescens* CHA0. Mol Plant Microbe Interact 16(10):851–858

Itoh N, Kouzai T, Koide Y (1994) Efficient transformation of *Pseudomonas* strains with pNI vectors by electroporation. Biosci Biotechnol Biochem 58(7):1306–1308

Jousset A, Rochat L, Lanoue A, Bonkowski M, Keel C, Scheu S (2011) Plants respond to pathogen infection by enhancing the antifungal gene expression of root-associated bacteria. Mol Plant Microbe Interact 24(3):352–358

Keel C, Ucurum Z, Michaux P, Adrian M, Haas D (2002) Deleterious impact of a virulent bacteriophage on survival and biocontrol activity of *Pseudomonas fluorescens* strain CHAO in natural soil. Mol Plant Microbe Interact 15(6):567–576

Kidarsa TA, Shaffer BT, Goebel NC, Roberts DP, Buyer JS, Johnson A, Kobayashi DY, Zabriskie TM, Paulsen IT, Loper JE (2012) Genes expressed by the biological control bacterium *Pseudomonas protegens* Pf-5 on seed surfaces under the control of the global regulators GacA and RpoS. Environ Microbiol 3:716–735

Kovach ME, Elzer PH, Hill DS, Robertson GT, Farris MA, Roop RM 2nd, Peterson KM (1995) Four new derivatives of the broad-host-range cloning vector pBBR1MCS, carrying different antibiotic-resistance cassettes. Gene 166(1):175–176

Kovach ME, Phillips RW, Elzer PH, Roop RM 2nd, Peterson KM (1994) pBBR1MCS: a broad-host-range cloning vector. Biotechniques 16(5):800–802

Landa BB, Mavrodi OV, Raaijmakers JM, Gardener BBM, Thomashow LS, Weller DM (2002) Differential ability of genotypes of 2,4-diacetylphloroglucinol-producing *Pseudomonas fluorescens* strains to colonize the roots of pea plants. Appl Environ Microbiol 68(7):3226–3237

Landa BB, Mavrodi OV, Schroeder KL, Allende-Molar R, Weller DM (2006) Enrichment and genotypic diversity of *phlD*-containing fluorescent *Pseudomonas* spp. in two soils after a century of wheat and flax monoculture. FEMS Microbiol Ecol 55(3):351–368

Laville J, Voisard C, Keel C, Maurhofer M, Defago G, Haas D (1992) Global control in *Pseudomonas fluorescens* mediating antibiotic synthesis and suppression of black root rot of tobacco. Proc Natl Acad Sci USA 89(5):1562–1566

Lee SW, Cooksey DA (2000) Genes expressed in *Pseudomonas putida* during colonization of a plant-pathogenic fungus. Appl Environ Microbiol 66(7):2764–2772

Lilley AK, Bailey MJ (1997a) Impact of plasmid pQBR103 acquisition and carriage on the phytosphere fitness of *Pseudomonas fluorescens* SBW25: burden and benefit. Appl Environ Microbiol 63(4): 1584–1587

Lilley AK, Bailey MJ (1997b) The acquisition of indigenous plasmids by a genetically marked pseudomonad population colonizing the sugar beet phytosphere is related to local environmental conditions. Appl Environ Microbiol 63(4):1577–1583

Lilley AK, Bailey MJ (2002) The transfer dynamics of *Pseudomonas* sp. plasmid pQBR11 in biofilms. FEMS Microbiol Ecol 42(2):243–250

Lim CK, Hassan KA, Penesyan A, Loper JE, Paulsen IT (2013a) The effect of zinc limitation on the transcriptome of *Pseudomonas protegens* Pf-5. Environ Microbiol 15:702–715

Lim CK, Hassan KA, Tetu SG, Loper JE, Paulsen IT (2012) The effect of iron limitation on the transcriptome and proteome of *Pseudomonas fluorescens* Pf-5. PLoS ONE 7(6):e39139

Lim CK, Penesyan A, Hassan KA, Loper JE, Paulsen IT (2013b) Effect of tannic acid on the transcriptome of the soil bacterium *Pseudomonas protegens* Pf-5. Appl Environ Microbiol 79:3141–3145

Loper JE, Hassan KA, Mavrodi DV, Davis EW 2nd, Lim CK, Shaffer BT, Elbourne LD, Stockwell VO, Hartney SL, Breakwell K, Henkels MD, Tetu SG, Rangel LI, Kidarsa TA, Wilson NL, van de Mortel JE, Song C, Blumhagen R, Radune D, Hostetler JB, Brinkac LM, Durkin AS, Kluepfel DA, Wechter WP, Anderson AJ, Kim YC, Pierson LS 3rd, Pierson EA, Lindow SE, Kobayashi DY, Raaijmakers JM, Weller DM, Thomashow LS, Allen AE, Paulsen IT (2012) Comparative genomics of plant-associated *Pseudomonas* spp.: insights into diversity and inheritance of traits involved in multitrophic interactions. PLoS Genet 8(7):e1002784

Loper JE, Henkels MD, Shaffer BT, Valeriote FA, Gross H (2008) Isolation and identification of rhizoxin analogs from *Pseudomonas fluorescens* Pf-5 by using a genomic mining strategy. Appl Environ Microbiol 74(10):3085–3093

Maddula VS, Pierson EA, Pierson LS 3rd (2008) Altering the ratio of phenazines in *Pseudomonas chlororaphis* (*aureofaciens*) strain 30-84: effects on biofilm

formation and pathogen inhibition. J Bacteriol 190(8): 2759–2766

Mahan MJ, Slauch JM, Mekalanos JJ (1993) Selection of bacterial virulence genes that are specifically induced in host tissues. Science 259(5095):686–688

Marchi M, Boutin M, Gazengel K, Rispe C, Gauthier JP, Guillerm-Erckelboudt AY, Lebreton L, Barret M, Daval S, Sarniguet A (2013) Genomic analysis of the biocontrol strain *Pseudomonas fluorescens* Pf29Arp with evidence of T3SS and T6SS gene expression on plant roots. Environ Microbiol Rep 5(3):393–403

Mathee K, Narasimhan G, Valdes C, Qiu X, Matewish JM, Koehrsen M, Rokas A, Yandava CN, Engels R, Zeng E, Olavarietta R, Doud M, Smith RS, Montgomery P, White JR, Godfrey PA, Kodira C, Birren B, Galagan JE, Lory S (2008) Dynamics of *Pseudomonas aeruginosa* genome evolution. Proc Natl Acad Sci USA 105(8):3100–3105

Matilla MA, Espinosa-Urgel M, Rodriguez-Herva JJ, Ramos JL, Ramos-Gonzalez MI (2007) Genomic analysis reveals the major driving forces of bacterial life in the rhizosphere. Genome Biol 8(9):R179

Mavrodi DV, Blankenfeldt W, Thomashow LS (2006) Phenazine compounds in fluorescent *Pseudomonas* spp. biosynthesis and regulation. Annu Rev Phytopathol 44:417–445

Mavrodi DV, Bleimling N, Thomashow LS, Blankenfeldt W (2004) The purification, crystallization and preliminary structural characterization of PhzF, a key enzyme in the phenazine-biosynthesis pathway from *Pseudomonas fluorescens* 2-79. Acta Crystallogr D Biol Crystallogr 60(Pt 1):184–186

Mavrodi DV, Bonsall RF, Delaney SM, Soule MJ, Phillips G, Thomashow LS (2001) Functional analysis of genes for biosynthesis of pyocyanin and phenazine-1-carboxamide from *Pseudomonas aeruginosa* PAO1. J Bacteriol 183(21):6454–6465

Mavrodi DV, Loper JE, Paulsen IT, Thomashow LS (2009) Mobile genetic elements in the genome of the beneficial rhizobacterium *Pseudomonas fluorescens* Pf-5. BMC Microbiol 9:8

Mavrodi DV, Mavrodi OV, Parejko JA, Bonsall RF, Kwak YS, Paulitz TC, Thomashow LS, Weller DM (2012a) Accumulation of the antibiotic phenazine-1-carboxylic acid in the rhizosphere of dryland cereals. Appl Environ Microbiol 78(3):804–812

Mavrodi OV, Mavrodi DV, Parejko JA, Thomashow LS, Weller DM (2012b) Irrigation differentially impacts populations of indigenous antibiotic-producing *Pseudomonas* spp. in the rhizosphere of wheat. Appl Environ Microbiol 78(9):3214–3220

McSpadden Gardener BB, Gutierrez LJ, Joshi R, Edema R, Lutton E (2005) Distribution and biocontrol potential of *phlD*(+) pseudomonads in corn and soybean fields. Phytopathology 95(6):715–724

Migula W (1894) Uber ein neues System der Bakterien. Arb Bakteriol Inst Karlsruhe 1:235–238

Monds RD, Newell PD, Gross RH, O'Toole GA (2007) Phosphate-dependent modulation of c-di-GMP levels regulates *Pseudomonas fluorescens* Pf0-1 biofilm formation by controlling secretion of the adhesin LapA. Mol Microbiol 63(3):656–679

Moynihan JA, Morrissey JP, Coppoolse ER, Stiekema WJ, O'Gara F, Boyd EF (2009) Evolutionary history of the *phl* gene cluster in the plant-associated bacterium *Pseudomonas fluorescens*. Appl Environ Microbiol 75(7):2122–2131

Mulet M, Gomila M, Scotta C, Sanchez D, Lalucat J, Garcia-Valdes E (2012) Concordance between whole-cell matrix-assisted laser-desorption/ionization time-of-flight mass spectrometry and multilocus sequence analysis approaches in species discrimination within the genus *Pseudomonas*. Syst Appl Microbiol 35(7):455–464

Nielsen TH, Christophersen C, Anthoni U, Sorensen J (1999) Viscosinamide, a new cyclic depsipeptide with surfactant and antifungal properties produced by *Pseudomonas fluorescens* DR54. J Appl Microbiol 87(1):80–90

Oke V, Long SR (1999) Bacterial genes induced within the nodule during the *Rhizobium*-legume symbiosis. Mol Microbiol 32(4):837–849

Osbourn AE, Barber CE, Daniels MJ (1987) Identification of plant-induced genes of the bacterial pathogen *Xanthomonas campestris* pathovar campestris using a promoter-probe plasmid. EMBO J 6(1):23–28

Palleroni NJ (1984) *Pseudomonadaceae*. In: Kreig NR, Holt JG (eds) Bergey's manual of systematic biology. Williams and Wilkins, Baltimore, pp 141–199

Petkau A, Stuart-Edwards M, Stothard P, Van Domselaar G (2010) Interactive microbial genome visualization with GView. Bioinformatics 26(24):3125–3126

Phillips DA, Fox TC, King MD, Bhuvaneswari TV, Teuber LR (2004) Microbial products trigger amino acid exudation from plant roots. Plant Physiol 136(1):2887–2894

Raaijmakers JM, Vlami M, de Souza JT (2002) Antibiotic production by bacterial biocontrol agents. Antonie Van Leeuwenhoek 81(1–4):537–547

Raaijmakers JM, Weller DM, Thomashow LS (1997) Frequency of antibiotic-producing *Pseudomonas* spp. in natural environments. Appl Environ Microbiol 63(3):881–887

Rainey PB (1999) Adaptation of *Pseudomonas fluorescens* to the plant rhizosphere. Environ Microbiol 1(3):243–257

Rainey PB, Preston GM (2000) In vivo expression technology strategies: valuable tools for biotechnology. Curr Opin Biotechnol 11(5):440–444

Reams AB, Neidle EL (2004) Selection for gene clustering by tandem duplication. Annu Rev Microbiol 58:119–142

Redondo-Nieto M, Barret M, Morrisey JP, Germaine K, Martinez-Granero F, Barahona E, Navazo A, Sanchez-Contreras M, Moynihan JA, Giddens SR, Coppoolse ER, Muriel C, Stiekema WJ, Rainey PB, Dowling D, O'Gara F, Martin M, Rivilla R (2012) Genome sequence of the biocontrol strain *Pseudomonas fluorescens* F113. J Bacteriol 194(5): 1273–1274

Redondo-Nieto M, Barret M, Morrissey J, Germaine K, Martinez-Granero F, Barahona E, Navazo A, Sanchez-Contreras M, Moynihan JA, Muriel C, Dowling D, O'Gara F, Martin M, Rivilla R (2013) Genome sequence reveals that *Pseudomonas fluorescens* F113 possesses a large and diverse array of systems for rhizosphere function and host interaction. BMC Genom 14(1):54

Rottig M, Medema MH, Blin K, Weber T, Rausch C, Kohlbacher O (2011) NRPSpredictor2–a web server for predicting NRPS adenylation domain specificity. Nucleic Acids Res 39(Web Server issue):W362–W367

Schweizer HP (1991) *Escherichia-Pseudomonas* shuttle vectors derived from pUC18/19. Gene 97(1):109–121

Sieber SA, Marahiel MA (2005) Molecular mechanisms underlying nonribosomal peptide synthesis: approaches to new antibiotics. Chem Rev 105(2): 715–738

Silby MW, Cerdeno-Tarraga AM, Vernikos GS, Giddens SR, Jackson RW, Preston GM, Zhang XX, Moon CD, Gehrig SM, Godfrey SA, Knight CG, Malone JG, Robinson Z, Spiers AJ, Harris S, Challis GL, Yaxley AM, Harris D, Seeger K, Murphy L, Rutter S, Squares R, Quail MA, Saunders E, Mavromatis K, Brettin TS, Bentley SD, Hothersall J, Stephens E, Thomas CM, Parkhill J, Levy SB, Rainey PB, Thomson NR (2009) Genomic and genetic analyses of diversity and plant interactions of *Pseudomonas fluorescens*. Genome Biol 10(5):R51

Silby MW, Levy SB (2004) Use of in vivo expression technology to identify genes important in growth and survival of *Pseudomonas fluorescens* Pf0-1 in soil: discovery of expressed sequences with novel genetic organization. J Bacteriol 186(21):7411–7419

Silby MW, Levy SB (2008) Overlapping protein-encoding genes in *Pseudomonas fluorescens* Pf0-1. PLoS Genet 4(6):e1000094

Slauch JM, Mahan MJ, Mekalanos JJ (1994) In vivo expression technology for selection of bacterial genes specifically induced in host tissues. Methods Enzymol 235:481–492

Smith AW, Iglewski BH (1989) Transformation of *Pseudomonas aeruginosa* by electroporation. Nucleic Acids Res 17(24):10509

Stanier RY, Palleroni NJ, Doudoroff M (1966) The aerobic pseudomonads: a taxonomic study. J Gen Microbiol 43(2):159–271

Stockwell VO, Davis EW 2nd, Carey A, Shaffer BT, Mavrodi DV, Hassan KA, Hockett K, Thomashow LS, Paulsen IT, Loper JE (2013) pA506: A conjugative plasmid of the plant epiphyte *Pseudomonas fluorescens* A506. Appl Environ Microbiol 79(17):5272–5282

Tang X, Lu BF, Pan SQ (1999) A bifunctional transposon mini-Tn5GFP-Km which can be used to select for promoter fusions and report gene expression levels in *Agrobacterium tumefaciens*. FEMS Microbiol Lett 179(1):37–42

Tett A, Spiers AJ, Crossman LC, Ager D, Ciric L, Dow JM, Fry JC, Harris D, Lilley A, Oliver A, Parkhill J, Quail MA, Rainey PB, Saunders NJ, Seeger K, Snyder LA, Squares R, Thomas CM, Turner SL, Zhang XX, Field D, Bailey MJ (2007) Sequence-based analysis of pQBR103; a representative of a unique, transfer-proficient mega plasmid resident in the microbial community of sugar beet. ISME J 1(4):331–340

Thomashow LS, Weller DM, Bonsall RF, Pierson LS (1990) Production of the antibiotic phenazine-1-carboxylic acid by fluorescent *Pseudomonas* species in the rhizosphere of wheat. Appl Environ Microbiol 56(4):908–912

Timms-Wilson TM, Ellis RJ, Renwick A, Rhodes DJ, Mavrodi DV, Weller DM, Thomashow LS, Bailey MJ (2000) Chromosomal insertion of phenazine-1-carboxylic acid biosynthetic pathway enhances efficacy of damping-off disease control by *Pseudomonas fluorescens*. Mol Plant Microbe Interact 13(12): 1293–1300

Troppens DM, Moynihan JA, Barret M, O'Gara F, Morrissey JP (2013) Genetics and evolution of 2,4-diacetylphloroglucinol synthesis in *Pseudomonas fluorescens*. In: de Bruijn FJ (ed) Molecular microbial ecology of the rhizosphere, vol 2. Wiley, New York, pp 593–605

Turner SL, Lilley AK, Bailey MJ (2002) Two *dnaB* genes are associated with the origin of replication of pQBR55, an exogenously isolated plasmid from the rhizosphere of sugar beet. FEMS Microbiol Ecol 42(2):209–215

Vandenbergh PA, Kunka BS (1988) Metabolism of volatile chlorinated aliphatic hydrocarbons by *Pseudomonas fluorescens*. Appl Environ Microbiol 54(10):2578–2579

Varivarn K, Champa LA, Silby MW, Robleto EA (2013) Colonization strategies of *Pseudomonas fluorescens* Pf0-1: activation of soil-specific genes important for diverse and specific environments. BMC Microbiol 13:92

Vivian A, Murillo J, Jackson RW (2001) The roles of plasmids in phytopathogenic bacteria: mobile arsenals? Microbiology 147(Pt 4):763–780

Voisard C, Keel C, Haas D, Defago G (1989) Cyanide production by *Pseudomonas fluorescens* helps suppress black root rot of tobacco under gnotobiotic conditions. EMBO J 8(2):351–358

West SE, Schweizer HP, Dall C, Sample AK, Runyen-Janecky LJ (1994) Construction of improved *Escherichia-Pseudomonas* shuttle vectors derived from pUC18/19 and sequence of the region required for their replication in *Pseudomonas aeruginosa*. Gene 148(1):81–86

Zhang XX, Lilley AK, Bailey MJ, Rainey PB (2004a) Functional and phylogenetic analysis of a plant-inducible oligoribonuclease (*orn*) gene from an indigenous *Pseudomonas* plasmid. Microbiology 150(Pt 9):2889–2898

Zhang XX, Lilley AK, Bailey MJ, Rainey PB (2004b) The indigenous *Pseudomonas* plasmid pQBR103 encodes plant-inducible genes, including three putative helicases. FEMS Microbiol Ecol 51(1):9–17

The Genomics of *Xanthomonas oryzae*

6

Lindsay Triplett, Ralf Koebnik, Valerie Verdier, and Jan E. Leach

6.1 Introduction

Few bacterial plant pathogens, particularly those with narrow host ranges, have the global impact of *Xanthomonas oryzae* (*Xo*) pathovars. The host they attack, rice (*Oryza sativa*), is the staple food crop for more than half of the world's population. Currently, *Xo* has two recognized pathovars, *X. oryzae* pv. *oryzae* (*Xoo*) and *X. oryzae* pv. *oryzicola* (*Xoc*), which differ in the symptoms they cause on rice. *Xoo* causes bacterial blight (BB) of rice, characterized by spreading marginal lesions, while *Xoc* causes bacterial leaf streak (BLS).

Under disease-favorable conditions and in rice hosts with ineffective resistance, BB can cause yield losses up to 70 % (Reddy et al. 1979; Mew et al. 1993), although more typical reports ranged from 20 to 50 % (Ou 1985). BLS disease, caused by *Xoc*, is less widespread, occurring in tropical and subtropical regions of Asia, Africa, and Australia (Gonzalez et al. 2007), and is less severe than BB, with losses usually ranging between 10 and 20 %. Control methods for the diseases caused by *Xo* are limited. Although chemical means for control are available (Chaudhary et al. 2012; Devadath 1989), their use and effectiveness is limited by cost, high variability in susceptibility among strains, and the impracticality of applying treatments during heavy rains characteristic of tropical growing seasons (Gnanamanickam et al. 1999; Devadath 1989; Gu et al. 2005; Yoshimura et al. 1998). Cultural controls such as drainage, plant spacing, and fertilizer management are also recommended (Leung et al. 2003). Regulatory controls aimed at prevention of *Xo* movement into unaffected areas restrict the movement of rice seed internationally. Importation and interstate movement of these pathogens within the USA has long been regulated by the USDA plant protection and quarantine (PPQ) program within the animal and plant health inspection service (APHIS). Now, all *Xos* are designated as select agents (http://www.aphis.usda.gov/programs/ag_selectagent/) according to the Public Health Security and Bioterrorism Preparedness and Response Act of 2002 (Public Law 107–188; June 12, 2002).

The most reliable method for control of *Xoo* is varietal resistance. Qualitative resistance has proved highly useful in controlling BB; in areas

L. Triplett · V. Verdier · J. E. Leach (✉)
Department of Bioagricultural Sciences and Pest Management and Program in Plant Molecular Biology, Colorado State University, Fort Collins, CO 80523-1177, USA
e-mail: Jan.Leach@colostate.edu

L. Triplett
e-mail: Lindsay.Triplett@colostate.edu

V. Verdier
e-mail: Valerie.Verdier@ird.fr

R. Koebnik · V. Verdier
Institut de Recherche pour le Développement, UMR Résistance des Plantes aux Bioagresseurs, IRD-CIRAD-UM2, 911 Avenue Agropolis, BP 6450134394, Montpellier Cedex 5, France
e-mail: Ralf.Koebnik@ird.fr

D. C. Gross et al. (eds.), *Genomics of Plant-Associated Bacteria*,
DOI: 10.1007/978-3-642-55378-3_6,

where effective resistance is deployed, BB causes yield losses of less than 1 % (Savary et al. 2000a, b; Leung et al. 2003). More than 30 BB R genes, named *Xa1* to *Xa38*, have been described (Bhasin et al. 2012; Verdier et al. 2012b), and the development of robust molecular markers for several of these genes is making them more available to breeders (Verdier et al. 2012c). It is interesting that, while more than 80 R genes have been identified and 21 have been cloned conferring resistance to rice blast, a fungal disease of rice caused by *Magnaporthe oryzae* (Ballini et al. 2008; Lv et al. 2013), the sequences of only eight R genes conferring resistance to BB are known [*Xa1*, *Xa3/Xa26*, *xa5*, *xa13*, *Xa21*, *Xa21D*, *Xa24*, and *Xa27* (Yoshimura et al. 1998; Gu et al. 2005; Song et al. 1995; Iyer and McCouch 2004; Blair et al. 2003; Wang et al. 1998; Xiang et al. 2006; Ogawa et al. 1987; Khush and Angeles 1999)]. This may be related to the atypical nature of BB resistance mechanisms; while rice blast resistance genes fall into the category of dominant NB-LRR-type R genes, most of the BB R genes cloned are not in this category (Verdier et al. 2012b), and 14 are recessive genes (Zhang and Wang 2013).

Oddly, given the successful identification of R genes controlling BB, single R genes controlling BLS have not been found in cultivated rice, despite concerted screening efforts (Chen et al. 2006). Currently, the most promising sources of resistance to BLS that are identified from rice are based on quantitative genetic resistance (Tang et al. 2000; Zheng et al. 2005; Han et al. 2008; Chen et al. 2006; Raymundo et al. 1999). Two novel sources of single-gene resistance to BLS have been found by searching outside of *Oryza sativa*. One R gene, *Rxo1*, was cloned from maize and shown to confer resistance to BLS when introduced into rice (Zhao et al. 2004b, 2005). More recently, a second recessive R gene, *bls1*, was localized to chromosome 6 of *Oryza rufipogon* (He et al. 2012).

Despite the regulatory, cultural, and genetic controls employed against *Xo*, the species remain a significant threat to rice yields and the global food supply. Given the importance of these two diseases and of rice, it is not surprising that the pathogens and their interactions with rice have been the focus of significant research efforts. This chapter will focus on the contributions of genomic sequencing projects to the understanding of *Xo* biology and diversity and the future questions that genomics will help address.

6.1.1 The Biology of *Xanthomonas oryzae*

The *Xo* pathovars are particularly fascinating because, despite the close relatedness of *Xoo* and *Xoc*, the two enter and establish relationships with different host tissues, and the disease symptoms they cause are distinct. *Xoo* is a vascular pathogen and is particularly devastating when it infects rice leaves at the seedling stage, when the plants are especially vulnerable. The pathogen can enter the vessels directly through wounds generated during transplanting or by the wind-driven rains during typhoons. Movement between leaves and plants occurs as bacterial exudates are blown by wind or splashing rains, or as leaves rub against one another. Alternatively, *Xoo* can gain access to vessels by moving with guttation fluids through natural openings called hydathode water pores located on the edges of rice leaves (Guo and Leach 1989; Mew et al. 1984; Tabei 1977). Once in the water pore, the bacteria multiply in the epithem and enter the vessels through the vascular pass (Tabei 1977). In early stages of infection, the rice leaves turn grayish green and curl. As disease progresses, the leaves turn yellow to straw color, and if infected at the seedling stage, the plant wilts, dries up, and dies. If leaves of older plants are infected, the pathogen moves down the leaf in the vessels, killing the entire leaf.

Xoc is an intercellular pathogen that can enter plants either through wounds or by invading the stomata (Mew 1993). As the bacteria multiply between the mesophyll parenchyma cells, the disease symptoms first manifest as small, water-soaked streaks between leaf veins. The streaks, which appear translucent when backlit, then turn

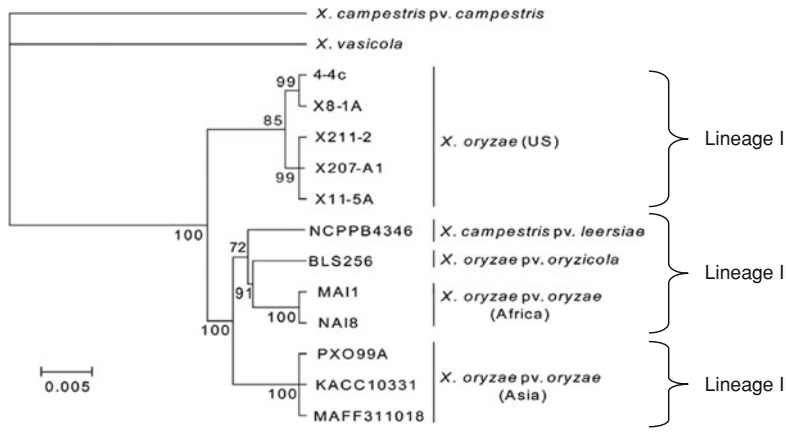

Fig. 6.1 Phylogenetic relationships of *Xanthomonas oryzae* strains. *Xanthomonas oryzae* forms three major lineages; the US *Xo* group branches separately from the Asian *Xoo* lineage and the lineage containing African *Xoo*, *Xoc*, and a native grass pathogen named *Xanthomonas campestris* pv. *leersiae*. Figure, reprinted from Triplett et al. (2011), is based on Bayesian phylogenetic analysis of three concatenated housekeeping genes

light brown to yellowish gray. As disease progresses, the entire leaf becomes brown and dies. As with BB, in high humidity, yellow droplets containing masses of bacteria can be observed on the leaf surface. At later stage of infection, it is difficult to distinguish between leaf blight and leaf streak in the field.

6.1.2 Distribution and Taxonomic Diversity

Xanthomonas oryzae pathovars are widely distributed. BB disease causes losses in both temperate and tropical environments, particularly in irrigated or rainfed lowland areas (Mew et al. 1993; Verdier et al. 2012b). *Xoo* has been reported in almost all major rice-growing areas, except for the Mediterranean Basin (Spain, Turkey, and Egypt). In addition to reports in most East and Southeast Asian countries, *Xoo* has also been reported in Australia (Aldrick et al. 1973), Iran (Ghasemie et al. 2008), and in several sub-Saharan African countries (Awoderu and John 1984). There are single reports of *Xoo* in several other rice-producing areas, including North and Latin America, but *Xoo* has not been confirmed to be endemic in these areas (Jones et al. 1989; Lozano 1977). *Xoc* is found in Asia and has also been reported in Madagascar and more recently in West Africa (Buddenhagen 1985; Wonni et al. 2011). In recent years, the disease has been observed with increasing frequency in Asia and Africa, likely due to the planting of susceptible varieties (Xie et al. 1990; Tang et al. 2000; Gonzalez et al. 2007; Wonni et al. 2011) and possibly as a result of a changing environment (Llano 1999).

Taxonomic studies have shown that *Xo* forms a distinct genetic lineage that is not more closely related to other *Xanthomonas* pathogens infecting monocots than to species infecting dicots. While differentiation according to observed symptoms classified *X. oryzae* into two pathovars (pvs. *oryzae* and *oryzicola*), analysis of strains from the USA and Africa has challenged the view that these two subgroups represent two genetically distinct lineages. Multi-locus sequence analysis points to three major genetic *Xo* lineages: (1) Asian *Xoo,* (2) Asian and African *Xoc* and African *Xoo*, and (3) *Xo* isolates from the USA (Fig. 6.1). Additional phylogenetic analyses of *Xoo* strains concur that Asian and African strains belong to distinct genetic lineages and that African *Xoo* strains share several features with *Xoc* (Gonzalez et al. 2007; Soto-Suarez et al. 2010b; Hajri et al. 2012). The group of isolates from the USA, referred to as US*Xo,* causes weak blight-like symptoms and yield losses of less than 1 %

(Jones et al. 1989), but has no formally designated pathovar. The US*Xo* strains fall into a separate clade from all known Asian and African *Xoo* and *Xoc* strains, suggesting an extended period of geographic separation. One group of strains causing blight-like symptoms on the native grass *Leersia hexandra* was originally named *X. campestris* pv. *leersiae* or *X. leersiae*, but those strains with sequence information appear to be bona fide *X. oryzae* strains related to *Xoc* and African *Xoo* (Triplett et al. 2011).

6.2 Genetic Characterization of *X. oryzae* Prior to Genomic Sequencing

6.2.1 Genetic Races

Prior to the publication of the first *X. oryzae* genome, much research interest of *Xoo* genetics was focused on the identification of races and the mechanism of race specificity. (No race structure has been characterized in *Xoc* or US*Xo,* although US*Xo* elicits HR on many varieties). Races are groups of strains sharing a common phenotype of virulence or avirulence to a set of differential cultivars carrying different R genes. Near-isogenic lines (NILs) such as the IRBB lines developed at the International Rice Research Institute are commonly used to identify *Xoo* races (Ogawa et al. 1990, 1991; Ogawa and Yamamoto 1987) (Table 6.1). Each NIL carries one specific resistance gene (*Xa* gene), which was incorporated into the indica cultivar IR24 by traditional breeding techniques and/or using marker-assisted selection. Based on their reaction on the IRBB lines, eleven *Xoo* races (Philippine races 1–10) were identified in the Philippines, 14 in Sri Lanka (Ochiai et al. 2000), and three in Africa (Table 6.1, A1, A2, and A3) (Gonzalez et al. 2007). Interestingly, HR-inducing resistance to African A3 strains was observed in all IRBB lines tested including IR24. In China, successively 9, 61, and 2 races were described among *Xoo* strains collected from different provinces (Liu et al. 2007, 2009a; Yang et al. 2013), with the new race FXP1 strains carrying genetic determinants that induce resistance (incompatibility) on the NILs carrying *xa5* and *Xa7* (Table 6.1). The emergence of new races after the widespread deployment of *Xa* resistance genes is a common occurrence (Mew et al. 1992). For example, resistance to *Xa4* was overcome in more than one instance by the emergence of new races (Pandey et al. 1986; Mew et al. 1992; Vera Cruz et al. 2000).

In studies among *Xoo* strains in Japan, Nepal, Korea, and India, numerous other races have been named in addition to the 30 in Table 6.1. Because these races were reported based on phenotyping isolates on different sets of NILs or other cultivars, it is not known how many *Xoo* races have emerged (Jeung et al. 2006; Lee et al. 1999; Noda et al. 1996). A single mutation or gene transfer event can change pathogen race, and races cannot be correlated to phylogeny or geographic group (Ochiai et al. 2005).

6.2.2 Avirulence and Virulence Genes

Beginning in the mid-1980s, the implementation of techniques for genetic analysis of *Xanthomonas* conjugation (via bi- or triparental mating) (Choi and Leach 1994) and transformation (via electroporation) (White and Gonzalez 1995; Choi and Leach 1994) of DNA into *Xo* strains allowed transposon insertional mutagenesis, complementation, and screening of cosmid libraries in *Xo* to identify the *avr* genes conferring race specificity (Sun et al. 2006; Daniels and Leach 1993). After it was discovered that *avrBs3*, an *avr* gene cloned from *X. vesicatoria* (Bonas et al. 1989), was highly similar in sequence to other *Xanthomonas* avirulence proteins, hybridization with *avrBs3* followed by inoculation experiments was used to identify cosmid clones containing *avrBs3*-like *Xoo* effector genes *avrxa5, avrXa10,* and *avrXa7* (Hopkins et al. 1992). The *avrBs3*-family effectors, now called transcriptional activator-like (TAL) effectors, function by binding to effector-specific sites in the host genome and

Table 6.1 Races of *Xoo* identified in the Philippines, Africa, China, and Sri Lanka on 12 near-isogenic lines containing a single gene for resistance

NILs	*Xa* gene	Philippines											Africa			China		Sri Lanka												
		1	2	3B	3C	4	5	6	7	8	9a	10	A1	A2	A3	FXP1a	FXP2b	1	2	3	4	5	6	7	8	9	10	11	12	13	14
IR24	*Xa18*	S	S	S	S	S	S	S	S	S	S	S	S	S	R	S	S	–	–	–	–	–	–	–	–	–	–	–	–	–	–
IRBB1	*Xa1*	S	S	S	S	S	S	S	S	S	S	S	–	–	–	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S
IRBB2	*Xa2*	S	S	S	S	S	S	S	S	S	S	S	–	–	–	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S
IRBB3	*Xa3*	S	S	S	S	S	S	S	S	S	S	S	S	R	R	S	S	S	S	S	S	R	R	S	S	R	R	R	S	R	S
IRBB4	*Xa4*	R	S	S	S	MR	R	S	R	R	S	R	R	R	R	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S
IRBB5	*xa5*	R	R	R	R	S	R	S	R	R	R	R	R	R	R	R	R	S	S	S	R	S	R	S	R	S	R	R	R	R	R
IRBB7	*Xa7*	MS	R	R	R	S	R	S	R	R	S	R	R	R	R	R	R	S	S	S	S	S	S	S	S	S	S	S	R	R	R
IRBB8	*xa8*	S	S	S	S	S	MR	S	S	S	S	S	MS	R	R	S	MR	S	S	S	S	S	S	R	R	R	R	R	S	S	R
IRBB10	*Xa10*	S	R	S	S	S	R	S	R	S	S	S	S	MR	R	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S
IRBB11	*Xa11*	S	S	S	S	S	S	S	S	S	S	S	S	R	R	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S
IRBB13	*xa13*	S	S	S	S	S	S	R	S	S	S	S	S	MR	R	S	S	S	R	S	S	S	S	S	S	S	R	S	S	S	S
IRBB14	*Xa14*	S	S	S	S	S	R	S	S	R	S	S	S	MR	R	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S
IRBB21	*Xa21*	R	R	R	MR	R	R	MR	MR	MR	MR	S	MS	MR	R	S	S	S	R	R	R	R	R	R	R	R	R	R	R	R	R

Resistance and susceptibility reaction are expressed in lesion lengths measured 14 days after leaf clip inoculation on rice plants. Resistant (R) <5 cm, moderately resistant (MR) 5–10 cm, moderately susceptible (MS) 10–15 cm, and susceptible (S) >15 cm

activating expression of specific host genes. Some activate host genes that enhance susceptibility and thus contribute to virulence, while some also activate host resistance genes and thus function as avirulence factors [for review, see (Boch and Bonas 2010)]. For example, the TAL effector AvrXa7 from *Xoo* functions as a virulence determinant by activating expression of a sugar transporter, OsSWEET14 (Chen et al. 2012); if the plants contain the R gene *Xa7*, AvrXa7 induces resistance (Hopkins et al. 1992). Another TAL effector homolog, *avrXa27,* was cloned as a gene directly inducing the activation of *Xa27* (Gu et al. 2005). Microarray analysis of induced genes in rice helped to determine that the TAL effector, *pthXo1*, was the factor triggering *Xa13*-mediated susceptibility (Yang et al. 2006). Like many effector proteins, some TAL effectors from *Xoo* and *Xoc* are critical components of a pathogen's virulence (Bai et al. 2000; Verdier et al. 2012a; Yang and White 2004; Sugio et al. 2007; Yu et al. 2011), while others have no known contribution (Bai et al. 2000). A few *avr* genes are not TAL effectors; *avrRxo1* was cloned as an *Xoc* gene triggering HR in maize carrying the *Rxo1* gene (Zhao et al. 2004a). Several genes required for the AvrXa21–Xa21 interaction, called *rax*, have also been discovered (Lee et al. 2006).

Numerous other genes were implicated in the *Xo* infection process prior to the availability of the genome sequence. In addition to transposon mutagenic screens, studies in *X. campestris* often served as a guide for amplification and identification of homologous genes in *Xo* (Li et al. 2004). This work led to the identification, sequencing, and characterization of the virulence roles of the *hrp* cluster for type III secretion (Zhu et al. 2000), the *xps* cluster involved in type II secretion of enzymes (Ray et al. 2000), the *gum* cluster of exopolysaccharide synthesis genes (Dharmapuri and Sonti 1999), the *rpf* cluster of quorum sensing and biofilm formation genes (Tang et al. 1996; Chatterjee and Sonti 2002), and distinct clusters of lipopolysaccharide and O-antigen formation genes (Patil and Sonti 2004). Genes involved in the utilization of host resources including phytic acid (Chatterjee et al. 2003), purines (Chatterjee and Sonti 2005), xylan (Rajeshwari et al. 2005), and iron (Subramoni and Sonti 2005) were also required for *Xoo* virulence. Virulence factors characterized also encoded stress resistance factors including the *aroE* gene required for pigment production (Goel et al. 2001) and genes encoding catalase (*kat*) and peroxidase (*pox*) enzymes (Chamnongpol et al. 1995; Chittoor et al. 1997).

6.3 Insights from the Sequenced Genomes of *Xo* Strains

As of this writing, 12 complete *Xanthomonas* genomes have been published, four of which are in the species *X. oryzae*: strains KACC10331 (*Xoo*K; Korea), MAFF311018 (*Xoo*M; Japan), and PXO99A (*Xoo*P; Philippines) in the pathovar *oryzae* (Lee et al. 2005; Ochiai et al. 2005; Salzberg et al. 2008) and the strain BLS256 of pathovar *oryzicola* (Bogdanove et al. 2011). All four complete *Xo* genomes were sequenced by shotgun sequencing using the Sanger method, and the annotations were manually curated; the *Xoo*P genome annotation was later improved by additional analysis and validation (Lei et al. 2013). In addition, two draft genomes of *Xo* strains X11-5A and X8-1A were sequenced using Illumina technology, assembled into over 300 contigs, and automated annotation was performed (Triplett et al. 2011). Draft Illumina genome sequence has been generated but is not yet publicly available for multiple strains of *Xoo* and *Xoc* isolated from China and West and Central Africa (authors, unpublished data; White, pers. comm.); however, phylogenetic and gene content information based on analysis of these genomes has been published and will be discussed below (Triplett et al. 2011; Zhao et al. 2012a; Hajri et al. 2012). This section will focus on the insights gained from sequencing *Xo* genomes.

6.3.1 DNA Sequence Characteristics

Basic general characteristics of the four completed *Xo* genomes are shown in Table 6.2. *Xo*

Table 6.2 General characteristics of publicly available sequenced genomes of *Xanthomonas oryzae*

Strain	KACC10331	MAFF311018	PXO99A	BLS256	X8-1A	X11-5A
Pathovar	*oryzae*	*oryzae*	*oryzae*	*oryzicola*	Undesignated	Undesignated
Abbreviation	*Xoo*K	*Xoo*M	*Xoo*P	*Xoc*	US*Xo*	US*Xo*
Country of origin	Korea	Japan	Philippines	Philippines	United States	United States
Genome status	Complete	Complete	Complete	Complete	Draft	Draft
Length (bp)	4,941,439	4,940,217	5,240,075	4,831,739	4,679,331	4,641,765
GC content (%)	63.7	63.7	63.6	64.1	64.0	64.0
Annotated genes	4,637	4,372	5,083	4,480	4,886	4,655
Total IS elements	714	712	683	245	ND*	ND
Distinct IS elements	27	26	27	21	ND	ND
IS element families	8	7	8	6	ND	ND
TAL effector genes	15	17	19	26	0	0
Plasmids	0	0	0	0	1	0
Source	Lee et al. (2005)	Ochiai et al. (2005)	Salzberg et al. (2008)	Bogdanove et al. (2011)	Triplett et al. (2011)	Triplett et al. (2011)

*Accurate assembly of IS elements is not possible in draft genomes due to repetitive sequence collapse during assembly

genomes are similar in size, GC content, and gene number to other *Xanthomonas* genomes, but encode a much larger number of insertion sequence elements and TAL effectors (Meyer and Bogdanove 2009). The sequenced Asian strains of *Xo* do not contain plasmids, but the US*Xo* strain X8-1A genome contains a plasmid similar to pXAV38 from *X. axonopodis* pv. *vesicatoria* (Triplett et al. 2011).

6.3.2 The Gene Content of *X. Oryzae* Genomes

6.3.2.1 TAL Effectors

Genome sequencing revealed a wealth of information about the diversity, organization, and evolution of TAL effectors, formerly AvrBs3/PthA-family effectors, in the genomes of *Xo*. It was known from Southern hybridization experiments that unlike genomes of *X. axonopodis* pv. *citri*, *X. vesicatoria*, and *X. gardneri*, which contain single TAL effectors, *Xo* genomes contained many and varying copies of TAL effectors (Hopkins et al. 1992; Ryba-White et al. 1995; Leach et al. 1995), and *Xo* genome assemblies confirmed this: 15–19 are found in Asian *Xo* strains and 26 in *Xoc*. The genes are clustered in nine loci in the *Xoo*P genome, eight loci in the *Xoo*M genome, seven loci in the *Xoo*K genome, and 11 loci in the *Xoc* genome, with one to eight TAL effector genes in each individual locus. In both *Xoc* and *Xoo*, TAL coding sequences are not clustered in operons but are separated from each other by a conserved sequence containing a ~100-bp putative promoter sequence that is conserved among *Xoc* and *Xoo* (Bogdanove et al. 2011).

Xanthomonas oryzae genomes revealed interesting observations about the evolutionary history of TAL effectors in *Xo*. Phylogenetic analysis based on the N- and C-terminal portions of *Xo* TAL effectors shows that the effectors form *Xoc* group separately from those of *Xoo* (Bogdanove et al. 2011), while the TAL effectors from *Xoo*M and *Xoo*P do not form distinct clades. Later, TalC from the African *Xoo* strain BAI3 was shown to be part of the *Xoc* clade of TAL effectors (Yu et al. 2011). Together, these results suggest that a single TAL effector entered the genome of a progenitor of *Xoo* and *Xoc*, and amplification and diversification occurred after the separation of the Asian *Xoo* group from the *Xoc*/African *Xoo* group. Further amplification and shrinkage events then occurred after the separation of *Xoo*M and *Xoo*P strains. Interestingly, both draft genomes from the divergent clade of US*Xo* strains showed no evidence of full or partial TAL effector sequences (Triplett et al. 2011). This suggests that TAL effectors never entered the US*Xo* genomes and that the transfer of the initial TAL effector occurred after the divergence of the US*Xo* group from the *Xoo*/*Xoc* progenitor. The high degree of duplication of TAL effectors could have been mediated by transposition or recombination, conferring a selective advantage to strains with multiple TAL effectors. However, tracing the mechanism of TAL effector duplication is difficult due to the extensive level of rearrangement in the *Xo* genomes; while IS elements flank some TAL effector loci (Bogdanove et al. 2011), no specific family of IS element has been reported to be the predominant element associated with TAL effectors. Interestingly, Southern hybridization and unpublished sequence analysis indicates that African strains of *Xoo* contain fewer copies of TAL effectors than the Asian strains (Gonzalez et al. 2007; Hajri et al. 2012), as does the *Xo* strain known as *X. campestris* pv. *leersiae* (5 copies; authors, unpublished). Sequencing TAL effectors in additional strains from the *Xoc* clade, including African *Xoo* and *X. campestris* pv. *leersiae*, is needed to understand whether genomic factors may influence TAL duplication, whether TAL effectors contribute to tissue specificity in *Xo*, and whether there are additional TAL effector variants in the *Xo* species.

6.3.3 Other Type III Secreted Effectors

Type III secreted effectors other than the TAL have been predicted in *Xo* genomes on the basis of homology to known effectors, N-terminal

features characteristic of effectors, or the presence of conserved promoter elements indicative of a gene in the HrpX regulon, i.e., a plant-inducible promoter (PIP) box or a −10 box (Furutani et al. 2006). Ochiai et al. (2005) identified 37 putative secreted effectors in the *Xoo*M genome based on PIP-box presence; Furutani et al. (2009) later identified 60 putative effectors based on all three effector criteria. By fusing these 60 genes to the translocation reporter gene *cyaA* and measuring reporter activity in planta, the authors discovered that 16 of the predicted effectors were secreted into plant cells, including seven previously unidentified effectors (Furutani et al. 2009). White et al. (2009) proposed a Xop nomenclature for effectors in *Xanthomonas* species, identifying 22 putative effectors in the *Xoo*M genome (the 16 known translocated effectors plus six additional effector homologs not assayed in the CyaA study) and 21 in the unpublished genome of the African strain CFBP1947. Of this list, only XopC, XopG, XopL, XopAD, and XopAE have not yet been shown to be translocated by *Xoo*, although a later study demonstrated translocation of *Xoo* XopAE from *X. campestris* (Zhao et al. 2013).

Comparative analysis of the completed genomes demonstrates that *Xo* effector repertoires are strikingly homogenous for a species with a high degree of genomic plasticity. However, there is a small pool of variable effectors. Of 26 putative effector genes in the *Xoc* genome, only five are absent or disrupted in completed *Xoo* genomes (XopAF, XopAJ, XopAK, XopO, and XopI), making these potential candidates as mediators of the *hrp*-dependent suppression of the rice defense response to *Xoo* TAL effectors by *Xoc* (Bogdanove et al. 2011). Like *Xoc*, the draft genome of the African *Xoo* strain CFBP1947 is missing the effector XopT; however, this strain does not encode *Xoc* effectors XopAJ, XopAF, XopO, or XopAK (Hajri et al. 2012). The draft genomes of the US*Xo* group, which causes *Xoo*-like symptoms but is equally related to the *Xoo* and *Xoc* lineages, are predicted to contain intact copies of "*Xoc*-specific" effectors XopAK and XopI and "*Xoo*-specific" effector XopT in addition to all other *Xoo* effectors except XopU [(Triplett et al. 2011), unpublished data]. One effector present in all *X. oryzae* strains, XopC2, is specific to *Xo* among all *Xanthomonas* species (Bogdanove et al. 2011).

6.3.4 Virulence Factors

The major gene clusters associated with pathogenesis in *Xanthomonas* are the *hrp* (type III secretion), *xps* and *xcs* (type II secretion), *rpf* (quorum sensing and biofilm formation), *lps* (lipopolysaccharide), and *gum* (xanthan exopolysaccharide synthesis) clusters. The genes encoding the *hrp*, *xps*, and *gum* clusters are strongly conserved among *Xanthomonas* spp, varying only in the border regions and occasional presence of IS elements between genes (Lu et al. 2009). For example, in the *hrp* clusters of *Xoo*K and *Xoo*M, but not *Xoc*, a cluster of IS elements is inserted in region II between *xopF1* and *hrpF* (Lu et al. 2009). While the central *gum* cluster (*gumB–gumN*) is conserved among *Xanthomonas*, the genomes of *X. campestris* encode *gum* genes of unknown function downstream of *gumN* that are lacking in the *Xo* genomes. Mutagenic analysis of the *Xo gum* cluster genes showed that genes involved in exopolysaccharide synthesis and export are required for full *Xo* virulence, but genes that function in posttranslational modifications such as acetylation do not affect symptoms (Kim et al. 2009). This contrasts with studies in *X. campestris* pv. *campestris*, in which deletion of xanthan acetylation genes *gumF* and *gumG* cause a reduction in virulence (Katzen et al. 1998). The highly conserved *xps* type II secretion system is required for full virulence and for secretion of xylase, cellulase, cellobiosidase, and lipase enzymes in *Xoo* (Jha et al. 2007; Ray et al. 2000). *Xo* genomes lack the *xcs* type II secretion clusters of *X. campestris*, which have no known role in virulence.

In *Xanthomonas*, quorum sensing mediated by diffusible signal factor (DSF) is mediated by the putative DSF synthases RpfF and RpfB, the

sensor RpfC, and the response regulator RpfG. Mutagenesis of *rpfF* and of *rpfC* and G decreases virulence in *Xoo* and *Xoc*, respectively (Tang et al. 1996; Guo et al. 2012; Wang et al. 2007). The roles of additional *rpf* genes A, D, E, H, and I are not known. All the *rpf* genes are conserved among *Xo* genome sequences, with the exception of absent or disrupted *rpfH* and *rpfI* genes in some strains (Lu et al. 2009). Nucleotide sequence analysis determined that *rpfC* and *rpfF* have relatively divergent sequences among *Xanthomonas* and are under low levels of purifying selection (Lu et al. 2009), although it is not known how this sequence divergence affects the chemical makeup and perception of the group of compounds that compose DSF.

Of the virulence clusters of *Xo*, the greatest diversity is seen among the lipopolysaccharide (LPS) synthesis genes. LPS is an outer membrane component which acts as a PAMP to trigger plant defense responses. *Xoo* and *Xoc* have distinct clusters of seven and 15 LPS synthesis genes, respectively, with only two homologous genes shared between them (Lu et al. 2009). LPS synthesis genes are important for virulence in *Xoo* and *Xoc* (Guo et al. 2012; Wang et al. 2007, 2013a), although it is not known how LPS O-antigen diversity might impact host-range adaptation.

The BLS256 genome report includes tabular comparison of HD-GYP-, GGDEF-, and EAL-domain proteins, TonB-dependent receptors, adhesins, and two-component systems in the *Xo* genomes and other representative *Xanthomonas* sp. (Bogdanove et al. 2011), and these are further reviewed in Ryan et al. (2011). A variety of fimbrial and afimbrial adhesins contribute to *Xoo* colonization of the external surfaces of the leaf as well as internal colonization of vessels (Boch and Bonas 2010; Pradhan et al. 2012; Ray et al. 2002). Adhesins also contribute to *Xoc* virulence (Wang et al. 2007). Adhesin content is similar among *Xo* genomes, with the notable exception of *fhaB*, a large (>10 kb) afimbrial adhesin gene involved in biofilm formation in other *Xanthomonas* sp. (Li and Wang 2011). *fhaB* and associated genes are present in the genomes of *Xoo*P, *Xoc*, and US*Xo*, but absent in *Xoo*M and *Xoo*K. GGDEF- and HD-GYP/EAL-domain proteins are involved in the synthesis and degradation, respectively, of the signaling messenger cyclic di-GMP. Of the 35 *X. campestris* pv. *campestris* genes with a GGDEF or EAL domain, 31 are present in the *Xoc* genome and 26 are present in the genomes of *Xoo*M and *Xoo*P (Bogdanove et al. 2011). Two of the few two-component systems characterized in *Xoo*, RpfC/RpfG and PdeK/PdeR, are virulence factors and regulators of cyclic di-GMP (Yang et al. 2012; Tang et al. 1996).

TonB-dependent receptors (TBDRs) are transporters of iron–siderophore complexes, vitamins, and sugars and thus may play a crucial role in growth and virulence in the plant host. TBDRs are identified in the genome based on the presence of a TonB box sequence, an N-terminal signal sequence, and plug and β-barrel domains (Blanvillain et al. 2007). While most bacterial species encode fewer than 14 TBDRs, strains of *X. campestris* and *X. axonopodis* encode between 52 and 68, and up to 17 of these contributed to symptom development on plants in *X. campestris* pv. *campestris* (Blanvillain et al. 2007). The genomes of *Xoc*, *Xoo*M, *Xoo*K, and *Xoo*P each encode 36 complete TBDRs (Bogdanove et al. 2011; Blanvillain et al. 2007), although TBDR pools are not identical between the four genomes (Bogdanove et al. 2011); at least one TBDR has been implicated in *Xoo* virulence (Xu et al. 2012). The reduced numbers of cyclic di-GMP and TBDR proteins in *Xo* compared with other *Xanthomonas* species may be a consequence of adaptation to a narrow ecological niche. In addition, loss of some TBDRs might be compensated for by the acquisition of TAL effectors that increase availability of sugars in the plant environment.

Both US*Xo* draft genomes harbor large numbers of predicted genes with similarity to non-ribosomal peptide synthase–polyketide synthase (NRPS-PKS) clusters of the sugarcane pathogen *X. albilineans* (Triplett et al. 2011). In *X. albilineans*, one NRPS-PKS cluster produces the peptide toxin albicidin, a GyrB-targeting protein toxic against plants and against other bacteria, and several other NRPS-PKS clusters

with unknown function are encoded. NRPS genes have been identified in the African *Xoo* strains as well (NCBI accession JQ348075). The NRPS genes identified in African *Xoo* and US*Xo* are highly conserved on the amino acid level, but they share only slightly less than 80 % identity with the *X. albilineans* NRPS genes (unpublished data). NRPS clusters similar to the syringomycin synthesis genes of *Pseudomonas syringae* are reported in the *Xoc* genome (Ryan et al. 2011).

6.3.5 IS Elements

Insertion sequence (IS) elements are mobile genetic elements that can constitute a significant component of eubacterial and archaeal genomes (Siguier et al. 2006). With several hundred complete and partial IS elements per genome representing six distinct families (Table 6.2), *Xo* genomes contain a much higher number and diversity of IS elements than other *Xanthomonas* species (Fig. 6.2) (Bogdanove et al. 2011). IS element expansion often follows restriction to a single host, and *Xo* does occupy a narrow ecological niche, but the number of IS elements in *Xo* is extreme even among specialized pathogens (Mira et al. 2006). In a mathematical modeling study of IS element distribution, the three *Xoo* genomes had the highest copy numbers of IS5 out of 525 proteobacterial genomes studied in 183 genera, suggesting a low fitness cost for IS5 in *Xoo* (Bichsel et al. 2013). The high number of IS elements compared with other *Xanthomonas* could also stem from unusual activity levels of an *Xo* gene that modulates IS element activity, such as a recombinase, DNA chaperone, or methylase (Mahillon and Chandler 1998). By generating insertions, deletions, and duplications, and by modulating activity of nearby genes, IS elements are drivers of genomic diversity and strain adaptation (Boch et al. 2009), potentially benefiting *Xo* fitness through the generation of duplicated beneficial genes such as TAL effectors. Insertion sequence elements are associated with seven of ten major chromosomal rearrangements in the *Xoo*P genome compared with other *Xoo* strains (Salzberg et al. 2008), and whole-genome alignments show a striking level of genome rearrangement among the four *Xo* genomes compared with the rearrangements between four *X. campestris* genomes (Bogdanove et al. 2011).

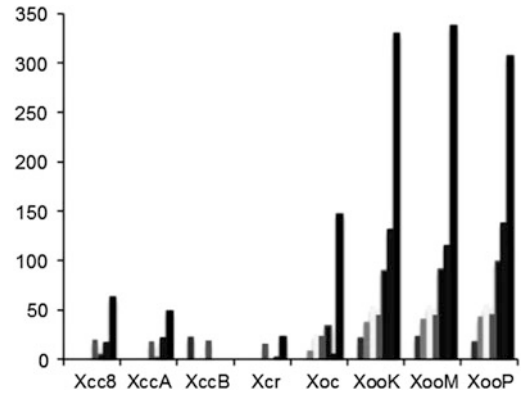

Fig. 6.2 Comparison of selected IS elements in the four completed *Xo* genomes (four *right* columns) with four strains of *Xanthomonas campestris* (*left* columns): *Xcc8* (*Xanthomonas campestris* pv. *campestris* strain 8004), *XccA* (*Xanthomonas campestris* pv. *campestris* strain ATCC33913), *XccB* (*Xanthomonas campestris* pv. *campestris* strain B100), and *Xcr* (*Xanthomonas campestris* pv. *raphani*). Bars representing IS elements in each genome are, left to right, ISL3, IS630, IS30, IS3, ISCNY, IS4, and IS5. Figure reprinted from (Bogdanove et al. 2011)

6.4 Applications of the *Xo* Genomes

Gene function studies The availability of *Xo* genome sequences facilitated an expansion of single gene function studies involving targeted mutagenesis and cloning, library screening, and proteomics, yielding information about the roles of diverse genes. For example, *Xoo* genomes have contributed to the characterization of individual virulence factors including siderophores (Subramoni et al. 2012; Pandey and Sonti 2010), adhesins (Pradhan et al. 2012), a manganese efflux system (Li et al. 2011), and TonB-dependent receptors (Xu et al. 2012). Virulence and regulatory roles were characterized for the master regulator *rsmA* and two-component

systems (Yang et al. 2012; Zhu et al. 2011). Other *Xoo* genes characterized contribute to general pathogen biology and stress responses, such as SAM salvage pathway components and SOS response genes (Sukchawalit et al. 2006). A screen of transposon mutants in *Xoc* revealed that type III secretion genes and type IV pilins for attachment and motility are important in virulence (Wang et al. 2007). Other novel virulence factors identified in *Xoc* using genome-guided cloning and gene disruption or analysis of transposon-induced mutations include a novel extracellular protease (Zou et al. 2012), an aspartate metabolism gene (Qian et al. 2012a), quorum sensing-regulated genes (Zhao et al. 2012b), and several genes of unknown function (Shen et al. 2012; Wang et al. 2013a).

Genomes also enable systematic mutagenesis and cloning studies to elucidate the roles of every individual member of a gene family. Several studies have used systematic approaches to identify the true repertoires of secreted effectors in *Xoo* from among dozens of candidates, leading to the identification of the first non-TAL effectors that contribute to *Xoo* virulence. In the first systematic translocation and mutagenesis study of *Xoo* effectors [(Furutani et al. 2009), discussed above], the effector XopR was found to inhibit virulence in rice and suppress the innate immune response in *Arabidopsis* (Akimoto-Tomiyama et al. 2012; Furutani et al. 2009). In another study, the *Xoo*M genome sequence was used to amplify and fuse 41 candidate effectors in *Xoo* strain 13751 to an N-terminal truncation of the *X. campestris* pv. *campestris* gene *avrBs1*; fusions were introduced into an *avrBs1* deletion mutant of *X. campestris* and tested in an HR-based translocation assay on pepper (Zhao et al. 2013). Nine proteins were translocated, including one not identified among the 16 previously found in *Xoo*M. Mutagenesis of the nine effectors confirmed XopR as a virulence determinant in strain 13751 (Zhao et al. 2013). It is possible that the *avrBs1* fusion assay is less sensitive than the *cyaA* assay as a test of protein translocation; either way, these studies show that a relatively small proportion of the *Xoo* genes with characteristics of type III effectors (homology to other known effectors and presence of *cis* regulatory promoter elements) are actually translocated into plant cells in these assays. A small proportion of these contribute to virulence in the standard scissor-clip assay (Kauffman et al. 1973) used for rice. In addition to XopR, mutagenesis studies using other strains or infection conditions reported virulence roles for XopZ and XopN (Song and Yang 2010; Cheong et al. 2013).

Although few individual secreted proteins have measurable impacts on symptom development, they may have roles in combination, on different plant hosts, or in different environmental contexts. Transgenic overexpression studies in plants, an alternate strategy to study the individual contributions of Xops, uncovered putative roles of five additional *Xoo* effectors in suppressing PAMP-triggered immunity (Sinha et al. 2013; Yamaguchi et al. 2013a, b). Pathogen-associated molecular patterns (PAMPs) betray the pathogen to the plant and trigger defense responses (Block and Alfano 2011). Xoo2875 (also called XopAA), a homolog of the early chlorosis factor from *X. campestris* pv. *vesicatoria* (Morales et al. 2005), was shown to promote colonization by a *Xoo* type III secretion mutant of rice plants overexpressing this effector, possibly by binding to OsBAK1 and/or OsBiSERK1 and interfering with PAMP signaling (Yamaguchi et al. 2013a). Another effector, Xop1488 (or XopY), was shown to interact with the receptor-like cytoplasmic kinase, OsRLCK185, thus suppressing its OsCERK1-mediated phosphorylation and consequently inhibiting peptidoglycan- and chitin-induced immunity (Yamaguchi et al. 2013b). Interactors of XopN include the resistance-related transcription factor OsVOZ2, which is required for rice resistance to *Xoo* (Cheong et al. 2013). Marker-free strategies have been developed for engineering double- and triple-knockout mutants in *Xo* (Zou et al. 2011). A systematic study of effector subsets in an effector-depleted strain, such as demonstrated for the plant pathogen *Pseudomonas syringae* (Cunnac et al. 2012b), may further illuminate the collective contribution of type III effectors to pathogenicity. Systematic analyses have also extended beyond *Xoo*

effectors; bioinformatic prediction and mutagenesis of genes for *Xoc*-specific secreted proteins led to the identification of the novel quorum sensing-regulated virulence factor Epv (Qian et al. 2012b).

6.4.1 Proteomics and Gene Expression Studies

While the majority of *Xanthomonas* functional genomics studies have been performed in *X. campestris* (Ryan et al. 2011), the *Xo* genome sequences have enabled microarray-, proteomics-, or reporter-based studies of gene expression in vitro and *in planta*. A systematic analysis of reporter fusions was used to identify nine novel Hrp-regulated promoters from a pool of 21 putative PIP-box promoters identified in the genome, including several controlling genes of unknown function (Furutani et al. 2006). An *Xoo–Xoc* dual genomics microarray determined genes upregulated in the apoplast-mimicking minimal medium XOM2, showing that many genes are only upregulated in this medium in *Xoo* (Seo et al. 2008). This demonstrated that there are key differences in the regulation of virulence genes in *Xoo* and *Xoc,* which could contribute to their tissue specificity differences. Other studies have used microarray analysis to identify differences in expression in *Xoo* strains with deleted master regulatory genes. A microarray was used to identify 330 genes differentially regulated by the LuxR-type response regulator OryR (González et al. 2012), and 214 genes were found to be differentially regulated in *Xoo* deficient in the PhoPQ two-component system, including virulence and regulatory genes (Sriariyanun et al. 2012).

In a third microarray study, 139 genes were differentially regulated in a *Xoo*Δ*rpfF* strain incapable of producing the quorum-sensing signal diffusible signal factor (DSF); attachment genes were downregulated, and genes for motility and extracellular enzymes were upregulated in the absence of the autoinducer (Rai et al. 2012). Expression findings were validated with functional analysis showing that the *rpfF* deletion mutant exhibits decreased aggregation and exopolysaccharide production and increased motility. The results suggest that *Xoo* DSF triggers a transition from a free-swimming invasion stage to biofilm formation in the xylem. This contrasts with the model of DSF function in the non-vascular pathogen *X. campestris*, in which DSF decreases attachment and increases motility, and parallels the model in the vascular pathogen *Xylella fastidiosa*, in which DSF increases attachment and suppresses motility (Rai et al. 2012). Intriguingly, study of an *rpfF* deletion mutant of *Xoc* exhibited increased aggregation and decreased expression of flagellar motility genes, suggesting that unlike *Xoo*, *Xoc* may follow the non-vascular, *X. campestris* model of DSF function (Zhao et al. 2011). In a follow-up study, 33 DSF-regulated extracellular proteins were identified using a similar experimental setup and four of them, including three type II-secreted hydrolases, were shown to contribute to virulence (Qian et al. 2013). Further expression analyses are needed to determine whether *Xoo* and *Xoc* respond differently to cell density signals in the host, or whether the observed differences stem from different in vitro experimental conditions.

Custom microarrays were also developed to identify genes expressed during infection in plants; 147 genes were differentially expressed in the African *Xoo* strain MAI1 during infection on rice, including Xop and TAL effectors, TonB-dependent receptors, IS elements, and adhesins (Soto-Suarez et al. 2010a). A microarray approach was also used to identify differentially expressed genes in *Xoo* strain KACC10331 at 1, 3, and 7 days post-infection (Zhang et al. 2009). Finally, *Xo* genomes have also aided the identification of proteins identified in proteomic analysis. Mass spectrometry of xylem fluid from infected rice identified 64 proteins secreted by *Xoo* during infection, including novel virulence factors validated by

mutagenesis (Gonzalez et al. 2012). A comparison of proteins from media-grown *Xoo* with proteins extracted from infected leaf tissue demonstrated production of Hrp proteins, adhesins, and quorum-sensing regulators during infection (Wang et al. 2013b). When cells of *Xoo* ectopically expressing the key regulatory virulence gene *hrpX* were compared to wild-type cells grown in vitro, seven proteins with different abundances in both samples were identified. Among these were two different peptidyl–prolyl cis–trans isomerases (Robin et al. 2013), pointing to a role of such chaperones in virulence, as had been shown before for *X. campestris* pv. *campestris* (Zang et al. 2007).

6.4.2 In Silico and Structural Genomics Techniques

The *Xoo* genomes have also been used to test novel in silico genomics techniques. The genome of MAFF311018 was used to perform in vitro and in silico analysis of restriction landmark genome scanning analysis (RLGS) to identify specific areas methylated in the genome, demonstrating the validity of this technique for pinpointing methylation-rich regions (Ichida et al. 2007). The sequence of *N*-acetylglucosamine-1-phosphate uridyltransferase (GlmU) gene from the *Xoo*M genome was used to create a structural model of the protein based on alignment with crystallized GlmU homologs, and this model was used to perform in silico screening of potential antimicrobial small molecules that would bind to GlmU and suppress *Xo* growth (Min et al. 2012). Three compounds identified, including the flavonoid luteolin, inhibited the activity of GlmU and had whole-cell antimicrobial properties against *Xoo*. This study demonstrates the usefulness of genome-derived information in the search for novel chemical controls against *Xo*. The RpfF protein, the synthase of diffusible signal factor quorum-sensing molecule, has also been modeled in silico for potential application in inhibitor screening (Reddy et al. 2012).

6.4.3 Development of TAL Effector-Based Applications

Determination of how TAL effectors target specific sequences provided an important clue to enable discovery of TAL targets in the host genome. For years, it was known that TAL specificity resided in the central repeat region, and specifically at the highly variable residues at positions 12 and 13 of each 33–34 aa repeat (Leach and White 1996). However, the clever deciphering of the code did not occur until 2009, when two groups independently showed that the amino acids at positions 12 and 13, called the repeat-variable diresidue (RVD), specify a nucleotide in the target; thus, the number and composition of RVDs in the TAL effector define the length and nucleotide sequence of the target (Boch et al. 2009; Moscou and Bogdanove 2009). Cross-comparison of predicted TAL sequences from *Xo* genomes with expression data from infected rice provided key validation for the TAL binding code, showing that TAL effectors were likely upregulating expression of predicted targets (Moscou and Bogdanove 2009). Synthetic TAL effectors are now being applied to introduce a variety of site-directed alterations in diverse organisms, a technology with wide-ranging implications for research and therapeutics (reviewed in (Doyle et al. 2013). Discoveries characterizing the binding specificity of major TAL effectors of *X. oryzae* to specific sequences in plant promoters (Moscou and Bogdanove 2009; Boch et al. 2009) and the computational prediction of the binding sequences (Doyle et al. 2012; Grau et al. 2013) have presented new avenues for resistance to bacterial blight and streak in rice. By engineering multiple TAL effector binding elements (EBEs) upstream of a known resistance gene, researchers have developed transgenic rice with qualitative resistance to multiple strains of *Xoo* and *Xoc* (Hummel et al. 2012). EBE editing has also been used to inactivate a TAL-targeted susceptibility gene to *Xoo* (Li et al. 2012), another avenue to bacterial blight resistance.

Designer TAL effectors (dTALEs), synthetic effectors designed to activate specific genes in the rice genome, have been developed as a tool to help elucidate the role and mechanism of plant disease resistance genes (Li et al. 2013).

The understanding of TAL effectors from the genomes can be applicable toward conventional breeding techniques. Sustainable control measures for *Xo* depend on understanding and characterizing genes involved in rice–*Xo* interactions, particularly those that determine race specificity. However, functional overlap among dozens of effectors in *Xo* genomes has made functional characterization of TAL effectors difficult, as well as correlation of genetic races with specific effectors. The absence of TAL effectors in the US*Xo* draft genome led to the development of *USXo* strain X11-5A as a tool for screening germplasm for resistance and susceptibility to individual TAL effectors (Verdier et al. 2012a). In a study of the effect of three individual TAL effectors on X11-5A virulence on 21 varieties of rice, novel sources of *Xa7* activity were revealed. In addition, TAL effector-mediated susceptibility was suppressed in some genetic backgrounds, suggesting that rice genetic loci other than known targets could moderate the effects of TAL effectors through an unknown mechanism (Verdier et al. 2012a).

6.4.4 Strain Typing and Diagnostics

Understanding strain diversity, movement, and population dynamics through strain typing is especially important for quarantine pathogens like *Xo*. Genome sequence availability has led to a technological shift in bacterial strain typing from fingerprinting approaches (AFLP, RFLP, and rep-PCR) to sequence and repeat-based techniques (Li et al. 2009b). Complete genome sequences are mined to identify short, hypermutable repetitive elements known as microsatellites, simple sequence repeats (SSR), or variable-number tandem repeats (VNTR). Differences in repeat number within these elements can be used to distinguish strains with a high degree of specificity. Repeat-based typing is performed with much greater repeatability than fingerprinting methods, yielding results that are easy to score and digitize. A microsatellite-based typing scheme has been developed for *Xoc* (Zhao et al. 2012a), and markers specific to additional *Xo* lineages are under development (authors, unpublished).

Clustered regularly interspersed short palindromic repeats (CRISPR) represent another repetitive sequence useful for population structure determination. CRISPR loci, part of a system that protects bacteria from foreign DNA, consist of repeated sequence separated by spacers that vary in length (Bogdanove et al. 2010). CRISPR-based typing, also called spoligotyping or spacer oligotyping, is well-established in mycobacteria (Driscoll 2009), and CRISPR locus typing in *Xoo* has been developed to differentiate strains originating from different lineages and geographic areas (Salzberg et al. 2008; Semenova et al. 2009). However, the genomes of *Xoc* and African *Xoo* and US*Xo* lack CRISPR loci, thus limiting this typing method to Asian strains of *Xoo* where CRISPR loci appear to be omnipresent [(Triplett et al. 2011), unpublished results]. Finally, typing and detection of *Xo* has been performed through partial sequencing of two housekeeping genes, *gyrB* and *rpoB*, and of the 16S-23S intergenic spacer (Goncalves and Rosato 2002; Parkinson et al. 2009; Ferreira-Tonin et al. 2011). Genomics-based analysis was used to develop pathovar-specific detection assays to target the 16S-23S rDNA, a membrane fusion protein gene, and an *rhs* family gene (Adachi and Oku 2000; Kang et al. 2008; Cho et al. 2011).

Accurate detection and differentiation of *Xoo* and *Xoc* is critical for diagnostic and regulatory purposes, but most DNA-based diagnostic tests published prior to genome sequencing lacked pathovar-level specificity or sensitivity to globally diverse strains. Lang et al. (2010) used a computational approach to design primers based on numerous genomic loci specific to (and conserved among) *Xoc* or *Xoo* and then developed a multiplex PCR assay to simultaneously detect and distinguish the two pathovars. The multiplex assay was validated on a diverse

collection of strains, and a similar approach was later used to develop primers specific to *Xo* isolates from the USA (Triplett et al. 2011). The pathovar-specific loci are being used to develop loop-mediated isothermal amplification (LAMP) PCR protocols to enable rapid field-level diagnostics of *Xo* pathovars (Lang et al. 2014).

6.5 Future Perspectives

The genome sequences of *X. oryzae* have yielded a wealth of information about the content of virulence genes, IS elements, and taxonomic differences among *Xoo* and *Xoc*. The genomes have served as a framework for studies of *Xoo* and *Xoc* gene function and expression, comparative analyses of pathovar differences, and development of improved typing and diagnostic tools. The resulting findings have introduced many new questions. We are only beginning to understand the complex regulatory networks involving quorum sensing, two-component systems, and c-di-GMP signaling and how these affect bacterial responses to the environment. The basis of tissue specificity differences between the closely related *Xoo* and *Xoc* is still not known. No virulence function or biochemical target has been attributed to most Xop effectors, and rice target genes for many TAL effectors have only been predicted computationally. The majority of *avr* and R genes mediating race specificity have still not been identified. Because many resistance-triggering genes are unknown or are highly homologous, difficult-to-differentiate TAL effectors, race determination of field strains of *X. oryzae* is still limited to labor-intensive, low-throughput bioassays.

While bench experimentation will be necessary to completely answer most of these questions, the generation and analysis of additional genome and transcriptome data will provide important insight and direction. Transcriptomic studies will improve genome annotations and lead to better understanding of gene regulation not only in vitro but also in vivo (i.e., during host colonization), as demonstrated for the plant pathogen *Ralstonia solanacearum* (Jacobs et al. 2012). The continuous drop in sequencing costs and improvement in new techniques, such as laser microdissection and dual pathogen–host RNAseq, will ultimately allow the study of transcriptomic changes of tissue-specific subpopulations during the infection process. Sequencing of additional diverse *Xo* genomes, including those from Africa, Australia, and the Middle East, will provide an improved perspective of pathogen evolution and spread, allow development of clade-specific diagnostics, and provide a more comprehensive core and pangenome for comparative analysis of gene content. African *Xoo* are closely related to *Xoc*, even though the strains cause different diseases and inhabit different tissues; comparative analysis of African *Xoo* with *Xoc* could provide better clues toward the basis of tissue specificity. Interestingly, the *leersiae* group of *Xo*, equally genetically related to African *Xoo* and *Xoc* (Fig. 6.1), was originally reported to cause leaf streak symptoms like *Xoc* (Fang and Ren 1959); sequencing of this genome could also help illuminate the basis of *Xo* tissue specificity.

The future of *Xo* genomics will not be limited to a few representative strains. High-throughput genome sequencing is becoming a routine technique in studies characterizing bacterial populations involved in human outbreaks (Chan et al. 2012) and, to a lesser extent, plant disease epidemics (Bart et al. 2012; Baltrus et al. 2012). Population genomics can be used to diagnose diseases, reveal probable means of pathogen spread, and track the origin and dissemination of toxins and antibiotic resistance genes through complex communities (Chan et al. 2012). In *Xo*, identifying the comprehensive global pool of TAL effectors would be one primary goal in population-level *Xo* sequencing projects. The majority of known native TAL effectors are encoded by the *Xo* species, and genomic study of global *Xoo* and *Xoc* TAL effectors will be critical for understanding how these genes duplicate and evolve in nature, what the most common and conserved TAL effectors are, and how their function is affected by the host genome. Both field and laboratory studies have

correlated the durability of plant R genes with the fitness cost of pathogen loss of certain TAL effectors [reviewed in (Leach et al. 2001)]. With population genomics, the probability of R genes functioning in a specific rice-growing region could be predicted prior to deployment. In short, genomics-enabled diagnostics of a specific area's *Xo* race diversity would be an extremely useful resource for plant breeders to target their efforts.

The read length limitations of second-generation sequencing technologies have made repeat-rich *Xo* genomes, and particularly the large and repetitive TAL effectors, difficult to assemble. However, sequencing approaches being developed could soon make it feasible to survey TAL effector variation on a population scale (Weber et al. 2011; Briggs et al. 2012; Reyon et al. 2012). Pacific biosciences RS (PacBio) offers a single-molecule sequencing technology yielding long reads without GC-bias (Korlach 2012), and Oxford Nanopore promises low-cost sequencing with lengths of 50 kb or more from its nanopore sequencing technology (Loman et al. 2012). Researchers have recently begun to apply the PacBio sequencer toward characterizing bacterial genomic diversity, including analysis of DNA methylation patterns that could affect pathogenicity (Chin et al. 2011; Korlach 2012). As error rates and costs decrease, there will be tremendous potential for using new technologies to capture the long repetitive regions specific to TAL effectors in *Xo* and study genomic variability on a population scale.

References

Adachi N, Oku T (2000) PCR-mediated detection of *Xanthomonas oryzae* pv. *oryzae* by amplification of the 16S-23S rDNA spacer region sequence. J Gen Plant Pathol 66:303–309

Akimoto-Tomiyama C, Furutani A, Tsuge S, Washington EJ, Nishizawa Y, Minami E, Ochiai H (2012) XopR, a type III effector secreted by *Xanthomonas oryzae* pv. *oryzae*, suppresses microbe-associated molecular pattern-triggered immunity in *Arabidopsis thaliana*. Mol Plant Microbe Interact 25:505–514

Aldrick SJ, Buddenhagen IW, Reddy APK (1973) The occurrence of bacterial leaf blight in wild and cultivated rice in Northern Australia. Crop Pasture Sci 24:219–227

Awoderu VA, John VT (1984) Occurrence of bacterial leaf blight on rice in four Sahelian countries: Senegal, Mali, Niger and Upper Volta. WARDA Technol Newsl 5:36–39

Bai J, Choi SH, Ponciano G, Leung H, Leach JE (2000) *Xanthomonas oryzae* pv. *oryzae* avirulence genes contribute differently and specifically to pathogen aggressiveness. Mol Plant Microbe Interact 13:1322–1329

Ballini E, Morel JB, Droc G, Price A, Courtois B, Notteghem JL, Tharreau D (2008) A genome-wide meta-analysis of rice blast resistance genes and quantitative trait loci provides new insights into partial and complete resistance. Mol Plant Microbe Interact 21:859–868

Baltrus DA, Nishimura MT, Romanchuk A, Chang JH, Mukhtar MS, Cherkis K, Roach J, Grant SR, Jones CD, Dangl JL (2012) Dynamic evolution of pathogenicity revealed by sequencing and comparative genomics of 19 *Pseudomonas syringae* isolates. PLoS Path 7(7):e1002132

Bart R, Cohn M, Kassen A, McCallum EJ, Shybut M, Petriello A, Krasileva K, Dahlbeck D, Medina C, Alicai T, Kumar L, Moreira LM, Neto JR, Verdier V, Santana MA, Kositcharoenkul N, Vanderschuren H, Gruissem W, Bernal A, Staskawicz BJ (2012) High-throughput genomic sequencing of cassava bacterial blight strains identifies conserved effectors to target for durable resistance. Proc Natl Acad Sci USA 109(28):E1972–E1979. doi:10.1073/pnas.1208003109

Bhasin H, Bhatia D, Raghuvanshi S, Lore JS, Sahi GK, Kaur B, Vikal Y, Singh K (2012) New PCR-based sequence-tagged site marker for bacterial blight resistance gene *Xa38* of rice. Mol Breeding 30:607–611

Bichsel M, Barbour AD, Wagner A (2013) Estimating the fitness effect of an insertion sequence. J Math Biol 66:95–114

Blair MW, Garris AJ, Iyer AS, Chapman B, Kresovich S, McCouch SR (2003) High resolution genetic mapping and candidate gene identification at the *xa5* locus for bacterial blight resistance in rice (*Oryza sativa* L.). Theor Appl Genet 107(1):62–73

Blanvillain S, Meyer D, Boulanger A, Lautier M, Guynet C, Denance N, Vasse J, Lauber E, Arlat M (2007) Plant carbohydrate scavenging through TonB-dependent receptors: a feature shared by phytopathogenic and aquatic bacteria. PLoS ONE 2:e224

Block A, Alfano JR (2011) Plant targets for *Pseudomonas syringae* type III effectors: virulence targets or guarded decoys? Curr Opin Microbiol 14(1):39–46. doi:10.1016/j.mib.2010.12.011

Boch J, Bonas U (2010) *Xanthomonas* AvrBs3 family-type III effectors: discovery and function. Annu Rev Phytopathol 48:419–436. doi:10.1146/annurev-phyto-080508-081936

Boch J, Scholze H, Schornack S, Landgraf A, Hahn S, Kay S, Lahaye T, Nickstadt A, Bonas U (2009)

Breaking the code of DNA binding specificity of TAL-type III effectors. Science 326:1509–1512. doi:10.1126/science.1178811

Bogdanove AJ, Koebnik R, Lu H, Furutani A, Angiuoli SV, Patil PB, Van Sluys MA, Ryan RP, Meyer DF, Han SW, Aparna G, Rajaram M, Delcher AL, Phillippy AM, Puiu D, Schatz MC, Shumway M, Sommer DD, Trapnell C, Benahmed F, Dimitrov G, Madupu R, Radune D, Sullivan S, Jha G, Ishihara H, Lee SW, Pandey A, Sharma V, Sriariyanun M, Szurek B, Vera-Cruz CM, Dorman KS, Ronald PC, Verdier V, Dow JM, Sonti RV, Tsuge S, Brendel VP, Rabinowicz PD, Leach JE, White FF, Salzberg SL (2011) Two new complete genome sequences offer insight into host and tissue specificity of plant pathogenic *Xanthomonas* spp. J Bacteriol 193(19):5450–5464. doi:10.1128/JB.05262-11

Bogdanove AJ, Schornack S, Lahaye T (2010) TAL effectors: finding plant genes for disease and defense. Curr Opin Plant Biol 13:394–401. doi:10.1016/j.pbi.2010.04.010

Bonas U, Stall RE, Staskawicz B (1989) Genetic and structural characterization of the avirulence gene *avrBs3* from *Xanthomonas campestris* pv. *vesicatoria*. Mol Gen Genet 218:127–136

Briggs AW, Rios X, Chari R, Yang L, Zhang F, Mali P, Church GM (2012) Iterative capped assembly: rapid and scalable synthesis of repeat-module DNA such as TAL effectors from individual monomers. Nucleic Acids Res 40:e117

Buddenhagen IW (1985) Rice disease evaluation in Madagascar. Int Rice Commiss Newsl 34:74–78

Chamnongpol S, Vattanaviboon P, Loprasert S, Mongkolsuk S (1995) Atypical oxidative stress regulation of a *Xanthomonas oryzae* pv. *oryzae* monofunctional catalase. Can J Microbiol 41:541–547

Chan JZM, Pallen MJ, Oppenheim B, Constantinidou C (2012) Genome sequencing in clinical microbiology. Nat Biotechnol 30 (11):1068

Chatterjee S, Sankaranarayanan R, Sonti RV (2003) PhyA, a secreted protein of *Xanthomonas oryzae* pv. *oryzae*, is required for optimum virulence and growth on phytic acid as a sole phosphate source. Mol Plant Microbe Interact 16(11):973–982

Chatterjee S, Sonti RV (2002) rpfF mutants of *Xanthomonas oryzae* pv. *oryzae* are deficient for virulence and growth under low iron conditions. Mol Plant Microbe Interact 15:463–471

Chatterjee S, Sonti RV (2005) Virulence deficiency caused by a transposon insertion in the *purH* gene of *Xanthomonas oryzae* pv. *oryzae*. Can J Microbiol 51(7):575–581

Chaudhary SU, Iqbal J, Hussain M (2012) Effectiveness of different fungicides and antibiotics against bacterial leaf blight in rice. J Agric Res 50:109–117

Chen C, Zheng W, Huang X, Zhang D, Lin XQ (2006) Major QTL conferring resistance to bacterial leaf streak. Agric Sci China 5:216–220

Chen LQ, Qu XQ, Hou BH, Sosso D, Osorio S, Fernie AR, Frommer WB (2012) Sucrose efflux mediated by SWEET proteins as a key step for phloem transport. Science 335:207–211

Cheong H, Kim C-Y, Jeon J-S, Lee B-M, Moon JS, Hwang I (2013) *Xanthomonas oryzae* pv. *oryzae* type III effector XopN targets OsVOZ2 and a putative thiamine synthase as a virulence factor in rice. PLoS ONE 8(9):e73346

Chin C-S, Sorenson J, Harris JB, Robins WP, Charles RC, Jean-Charles RR, Bullard J, Webster DR, Kasarskis A, Peluso P, Paxinos EE, Yamaichi Y, Calderwood SB, Mekalanos JJ, Schadt EE, Waldor MK (2011) The origin of the Haitian cholera outbreak strain. New Eng J Med 364(1):33–42. doi:10.1056/NEJMoa1012928

Chittoor JM, Leach JE, White FF (1997) Differential induction of a peroxidase gene family during infection of rice by *Xanthomonas oryzae* pv. *oryzae*. Mol Plant Microbe Interact 10:861–871

Cho MS, Kang MJ, Kim CK, Seol Y-J, Hahn JH, Park SC, Hwang DJ, Ahn T-Y, Park DH, Lim CK (2011) Sensitive and specific detection of *Xanthomonas oryzae* pv. *oryzae* by real-time bio-PCR using pathovar-specific primers based on an rhs family gene. Plant Dis 95:589–594

Choi SH, Leach JE (1994) Genetic manipulation of *Xanthomonas oryzae* pv. *oryzae*. Int Rice Res Notes 19:31–32

Cunnac S, Chakravarthy S, Kvitko BH, Russell AB, Martin GB, Collmer A (2011) Genetic disassembly and combinatorial reassembly identify a minimal functional repertoire of type III effectors in *Pseudomonas syringae*. Proc Natl Acad Sci USA 108:2975–2980. doi:10.1073/pnas.1013031108 |PII:031108

Daniels MD, Leach JE (1993) Genetics of *Xanthomonas*. In: Swings JG, Civerolo EL (eds) *Xanthomonas*. Chapman and Hall, London, pp 301–339

Devadath S (1989) Chemical control of bacterial blight of rice. Bacterial blight of rice International Rice Research Institute. Manila, Philippines, pp 89–98

Dharmapuri S, Sonti RV (1999) A transposon insertion in the *gumG* homologue of *Xanthomonas oryzae* pv. *oryzae* causes loss of extracellular polysaccharide production and virulence. FEMS Microbiol Lett 179:53–59

Doyle EL, Booher NJ, Standage DS, Voytas DF, Brendel VP, VanDyk JK, Bogdanove AJ (2012) TAL effector-nucleotide targeter (TALE-NT) 2.0: tools for TAL effector design and target prediction. Nucleic Acids Res 40:W117–W122

Doyle EL, Stoddard BL, Voytas DF, Bogdanove AJ (2013) TAL effectors: highly adaptable phytobacterial virulence factors and readily engineered DNA-targeting proteins. Trends Cell Biol. doi:10.1016/j.tcb.2013.04.003

Driscoll JR (2009) Spoligotyping for molecular epidemiology of the *Mycobacterium tuberculosis* complex. Methods Mol Biol 551:117–128

Fang CT, Ren HC (1959) Further studies on the relationship of *Leersia spp.* with the bacterial leaf

blight and the bacterial leaf streak disease of rice. Acta Phytopathol Sinica 2:004

Ferreira-Tonin M, Rodrigues-Neto J, Harakava R, Destefano SAL (2011) Phylogenetic analysis of *Xanthomonas* based on partial *rpoB* gene sequences and species differentiation by PCR-RFLP. Int J Syst Evol Microbiol 62:1419–1424

Furutani A, Nakayama T, Ochiai H, Kaku H, Kubo Y, Tsuge S (2006) Identification of novel HrpXo regulons preceded by two cis-acting elements, a plant-inducible promoter box and a -10 box-like sequence, from the genome database of *Xanthomonas oryzae* pv. *oryzae*. FEMS Microbiol Lett 259:133–141

Furutani A, Takaoka M, Sanada H, Noguchi Y, Oku T, Tsuno K, Ochiai H, Tsuge S (2009) Identification of novel type III secretion effectors in *Xanthomonas oryzae* pv. *oryzae*. Mol Plant Microbe Interact 22:96–106

Ghasemie E, Kazempour MN, Padasht F (2008) Isolation and identification of *Xathomonas oryzae* pv. *oryzae* the causal agent of bacterial blight of rice in Iran. J Plant Protect Res 48:53–62

Gnanamanickam SS, Priyadarisini VB, Narayanan NN, Vasudevan P, Kavitha S (1999) An overview of bacterial blight disease of rice and strategies for its management. Curr Sci 77(11):1435–1444

Goel AK, Rajagopal L, Sonti RV (2001) Pigment and virulence deficiencies associated with mutations in the *aroE* gene of *Xanthomonas oryzae* pv. *oryzae*. Appl Environ Microbiol 67:245–250

Goncalves ER, Rosato YB (2002) Phylogenetic analysis of *Xanthomonas* species based upon 16S-23S rDNA intergenic spacer sequences. Int J Syst Bacteriol 52:355–361

Gonzalez C, Szurek B, Manceau C, Mathieu T, Sere Y, Verdier V (2007) Molecular and pathotypic characterization of new *Xanthomonas oryzae* strains from West Africa. Mol Plant Microbe Interact 20:534–546. doi:10.1094/MPMI-20-5-0534

Gonzalez JF, Degrassi G, Devescovi G, De Vleesschauwer D, Hafte M, Myers MP, Venturi V (2012) A proteomic study of *Xanthomonas oryzae* pv. oryzae in rice xylem sap. J Proteomics doi:10.1016/j.jprot.2012.07.019

González JF, Myers MP, Venturi V (2012) The inter-kingdom solo OryR regulator of *Xanthomonas oryzae* is important for motility. Mol Plant Pathol 14:211–221

Grau J, Wolf A, Reschke M, Bonas U, Posch S, Boch J (2013) Computational predictions provide insights into the biology of TAL effector target sites. PLoS Comput Biol 9:e1002962

Gu K, Yang B, Tian D, Wu L, Wang D, Sreekala C, Yang F, Chu Z, Wang GL, White FF, Yin Z (2005) R gene expression induced by a type-III effector triggers disease resistance in rice. Nature 435(7045):1122–1125

Guo A, Leach JE (1989) Examination of rice hydathode water pores exposed to *Xanthomonas campestris* pv. *oryzae*. Phytopathology 79(4):433–436

Guo W, Cui Y-P, Li Y-R, Che Y-Z, Yuan L, Zou L-F, Zou H-S, Chen G-Y (2012) Identification of seven *Xanthomonas oryzae* pv. *oryzicola* genes potentially involved in pathogenesis in rice. Microbiology 158:505–518. doi:10.1099/mic.0.050419-0

Hajri A, Brin C, Zhao S, David P, Feng JX, Koebnik R, Szurek B, Verdier V, Boureau T, Poussier S (2012) Multilocus sequence analysis and type III effector repertoire mining provide new insights into the evolutionary history and virulence of *Xanthomonas oryzae*. Mol Plant Pathol 13(3):288–302. doi:10.1111/j.1364-3703.2011.00745.x

Han QD, Chen ZW, Deng Y, Lan T, Guan HZ, Duan YL, Zhou YC, Lin MC, Wu WR (2008) Fine mapping of *qBlsr5a*, a QTL controlling resistance to bacterial leaf streak in rice. Acta Agron Sin 34:587–590

He WA, Huang DH, Li RB, Qiu YF, Song JD, Yang HN, Zheng JX, Huang YY, Li XQ, Liu C, Zhang YX, Ma ZF, Yang Y (2012) Identification of a resistance gene *bls1* to bacterial leaf streak in wild rice *Oryza rufipogon* Griff. J Integr Ag 11:962–969

Hopkins CM, White FF, Choi SH, Guo A, Leach JE (1992) Identification of a family of avirulence genes from *Xanthomonas oryzae* pv. *oryzae*. Mol Plant Microbe Interact 5:451–459

Hummel AW, Doyle EL, Bogdanove AJ (2012) Addition of transcription activator-like effector binding sites to a pathogen strain-specific rice bacterial blight resistance gene makes it effective against additional strains and against bacterial leaf streak. New Phytol 195:883–893

Ichida H, Maeda K, Ichise H, Matsuyama T, Abe T, Yoneyama K, Koba T (2007) *In silco* restriction landmark genome scanning analysis of *Xanthomonas oryzae* pathovar *oryzae* MAFF 311018. Biochem Biophys Res Commun 363:852–856

Iyer AS, McCouch SR (2004) The rice bacterial blight resistance gene *xa5* encodes a novel form of disease resistance. Mol Plant Microbe Interact 17(12):1348–1354

Jacobs JM, Babujee L, Meng F, Milling A, Allen C (2012) The in planta transcriptome of *Ralstonia solanacearum*: conserved physiological and virulence strategies during bacterial wilt of tomato. mBio 3:e00114–00112. doi:10.1128/mBio.00114-12

Jeung JU, Heu SG, Shin MS, Vera Cruz CM, Jena KK (2006) Dynamics of *Xanthomonas oryzae* pv. *oryzae* populations in Korea and their relationship to known bacterial blight resistance genes. Phytopathology 96:867–875. doi:10.1094/PHYTO-96-0867

Jha G, Rajeshwari R, Sonti RV (2007) Functional interplay between two *Xanthomonas oryzae* pv. *oryzae* secretion systems in modulating virulence on rice. Mol Plant Microbe Interact 20(1):31–40

Jones RK, Barnes LW, Gonzalez CF, Leach JE, Alvarez AM, Benedict AA (1989) Identification of low virulence strains of *Xanthomonas campestris* pv. *oryzae* from rice in the United States. Phytopathology 79:984–990

Kang MJ, Shim JK, Cho MS, Seol YJ, Hahn JH, Hwang DJ, Park DS (2008) Specific detection of *Xanthomonas oryzae* pv. *oryzicola* in infected rice plant by use of PCR assay targeting a membrane fusion protein gene. J Microbiol Biotechnol 18:1492

Katzen F, Ferreiro DU, Oddo CG, Lelmini MV, Becker A, Pauhler A, Lelpi L (1998) *Xanthomonas campestris* pv. *campestris gum* mutants: Effects on xanthan biosynthesis and plant virulence. J Bacteriol 180:1607–1617

Kauffman H, Reddy A, Hsiek S, Merca S (1973) An improved technique for evaluating resistance of rice varieties to *Xanthomonas oryzae*. Plant Dis Rep 57:537–541

Khush GS, Angeles ER (1999) A new gene for resistance to race 6 of bacterial blight in rice, *Oryza sativa* L. Rice Genet Newsl 16:92–93

Kim S, Kim J, Lee B, Cho J (2009) Mutational analysis of the *gum* gene cluster required for xanthan biosynthesis in *Xanthomonas oryzae* pv. *oryzae*. Biotech Lett 31:265–270

Korlach J (2012) Single molecule, Real-Time DNA sequencing yields a clearer picture of biological complexity: Helps scientists close genomes, study linked variants, and detect base modifications. Indust Biotechnol 8(6):333–336

Lang JM, Hamilton JP, Diaz MGQ, Van Sluys MA, Burgos MRG, Cruz CMV, Buell CR, Tisserat NA, Leach JE (2010) Genomics-based diagnostic marker development for *Xanthomonas oryzae* pv. *oryzae* and *X. oryzae* pv. *oryzicola*. Plant Dis 94:311–319. doi:10.1094/Pdis-94-3-0311

Lang JM, Langlois P, Nguyen MHR, Triplett LR, Purdie L, Holton T, Djikeng A, Vera Cruz CM, Verdier V, Leach JE (2014) Sensitive detection of *Xanthomonas oryzae* pv. *oryzae* and X. *oryzae* pv. *oryzicola* by loop mediated isothermal amplification. Appl Environ Microbiol Accepted

Leach JE, Leung H, Nelson RJ, Mew TW (1995) Population biology of *Xanthomonas oryzae* pv. oryzae and approaches to its control. Curr Opin Biotechnol 6:298–304

Leach JE, Vera-Cruz CM, Bai J, Leung H (2001) Pathogen fitness penalty as a predictor of durability of disease resistance genes. Annu Rev Phytopathol 39:187–224

Leach JE, White FF (1996) Bacterial avirulence genes. Annu Rev Phytopathol 34:153–179

Lee BM, Park YJ, Park DS, Kang HW, Kim JG, Song ES, Park IC, Yoon UH, Hahn JH, Koo BS, Lee GB, Kim H, Park HS, Yoon KO, Kim JH, Jung CH, Koh NH, Seo JS, Go SJ (2005) The genome sequence of *Xanthomonas oryzae* pathovar *oryzae* KACC10331, the bacterial blight pathogen of rice. Nucleic Acids Res 33:577–586

Lee SW, Han SW, Bartley LE, Ronald PC (2006) Unique characteristics of *Xanthomonas oryzae* pv. *oryzae* AvrXa21 and implications for plant innate immunity. Proc Natl Acad Sci USA 103:18395–18400

Lee SW, Choi SH, Han SS, Lee DG, Lee BY (1999) Distribution of *Xanthomonas oryzae* pv. *oryzae* strains virulent to *Xa21* in Korea. Phytopathology 89:928–933

Lei Y, Kang S, Gao J, Jia XS, Chen LL (2013) Improved annotation of a plant pathogen genome *Xanthomonas oryzae* pv. *oryzae* PXO99A. J Biomolec Struct Dyn 31:342–350. doi:10.1080/07391102.2012.698218

Leung H, Zhu Y, Revilla-Molina I, Fan JX, Chen H, Pangga I, Vera Cruz C, Mew TW (2003) Using genetic diversity to achieve sustainable rice disease management. Plant Dis 87:11561169

Li C, Tao J, Mao D, He C (2011) A novel manganese efflux system, YebN, is required for virulence by *Xanthomonas oryzae* pv. *oryzae*. PLoS ONE 6:e21983

Li G, Song CF, Pang XM, Yang Y, Wang JS (2009a) Analysis of pathotypic and genotypic diversity of *Xanthomonas oryzae* pv. *oryzae* in China. J Phytopathol 157:208–218. doi:10.1111/j.1439-0434.2008.01471.x

Li J, Wang N (2011) Genome-wide mutagenesis of *Xanthomonas axonopodis* pv. *citri* reveals novel genetic determinants and regulation mechanisms of biofilm formation. PLoS ONE 6(7):e21804

Li P, Lu X, Shao M, Long J, Wang J (2004) Genetic diversity of Harpins from *Xanthomonas oryzae* and their activity to induce hypersensitive response and disease resistance in tobacco. Sci China Series C: Life Sci 47:461–469

Li T, Huang S, Zhou J, Yang B (2013) Designer TAL effectors induce disease susceptibility and resistance to *Xanthomonas oryzae* pv. *oryzae* in rice. Mol Plant doi:10.1093/mp/sst034

Li T, Liu B, Spalding MH, Weeks DP, Yang B (2012) High-efficiency TALEN-based gene editing produces disease-resistant rice. Nat Biotechnol 30(5):390–392

Li W, Raoult D, Fournier PE (2009b) Bacterial strain typing in the genomic era. FEMS Micro Rev 33:892–916

Liu H, Yang W, Hu B, Liu F (2007) Virulence analysis and race classification of *Xanthomonas oryzae* pv. *oryzae* in China. J Phytopathol 155:129–135. doi:10.1111/j.1439-0434.2007.01197.x

Llano AI (1999) Factors affecting the development of bacterial leaf streak of rice (*Oryza sativa L.*) caused by *Xanthomonas oryzae* pv. *oryzicola* Swings et al.

Loman NJ, Constantinidou C, Chan JZM, Halachev M, Sergeant M, Penn CW, Robinson ER, Pallen MJ (2012) High-throughput bacterial genome sequencing: an embarrassment of choice, a world of opportunity. Nat Rev Microbiol 9:599–606

Lozano JC (1977) Identification of bacterial blight in rice, caused by *Xanthomonas oryzae*, in America. Plant Dis Rep 61:644–648

Lu H, Patil P, Van Sluys MA, White FF, Ryan RP, Dow JM, Rabinowicz P, Salzberg SL, Leach JE, Sonti R, Brendel V, Bogdanove A (2009) Acquisition and evolution of plant pathogenesis–associated gene clusters and candidate determinants of tissue-specificity in *Xanthomonas*. PLoS ONE 3(11):e3828

Lv Q, Xu X, Shang J, Jiang G, Pang Z, Zhou Z, Wang J, Liu Y, Li T, Li X (2013) Functional analysis of *Pid3-A4,* an ortholog of rice blast resistance gene *Pid3* revealed by allele mining in common wild rice. Phytopathology 103:594–599

Mahillon J, Chandler M (1998) Insertion sequences. Microbiol Mol Biol Rev 62(3):725–774

Mew TW (1993) *Xanthomonas oryzae* pathovars on rice: cause of bacterial blight and bacterial leaf streak. In: Swings JG, Civerolo EL (eds) *Xanthomonas*. Chapman and Hall, London, pp 30–40

Mew TW, Alvarez AM, Leach JE, Swings J (1993) Focus on bacterial blight of rice. Plant Dis 77:5–12

Mew TW, Mew IC, Huang J (1984) Scanning electron microscopy of virulent and avirulent strains of *Xanthomonas campestris* pv. *oryzae* on rice leaves. Phytopathology 74:635–641

Mew TW, Vera C, C. M., Medalla ES (1992) Changes in race frequency of *Xanthomonas oryzae* pv. *oryzae* in response to rice cultivars planted in the Philippines. Plant Dis 76:1029

Meyer DF, Bogdanove AJ (2009) Genomics-driven advances in *Xanthomonas* biology. In: Jackson RW (ed) Plant pathogenic bacteria: genomics and molecular biology. Horizon scientific press, UK, p 147–161

Min J, Lin D, Zhang Q, Zhang J, Yu Z (2012) Structure-based virtual screening of novel inhibitors of the uridyltransferase activity of *Xanthomonas oryzae* pv. *oryzae* GlmU. Eur J Med Chem 53:150–158

Mira A, Pushker R, Rodriguez-Valera F (2006) The Neolithic revolution of bacterial genomes. Trends Micro 14(5):200–206

Morales CQ, Posada J, Macneale E, Franklin D, Rivas I, Bravo M, Minsavage J, Stall RE, Whalen MC (2005) Functional analysis of the early chlorosis factor gene. Mol Plant Microbe Interact 18(5):477–486. doi:10.1094/MPMI-18-0477

Moscou MJ, Bogdanove AJ (2009) A simple cipher governs DNA recognition by TAL effectors. Science 326(5959):1501. doi:10.1126/science.1178817

Noda T, Yamamoto T, Kaku H, Horino O (1996) Geographical distribution of pathogenic races of *Xanthomonas oryzae* pv. *oryzae* in Japan in 1991 and 1993. Ann Phytopathol Soc Jpn 62:549–553

Ochiai H, Horino O, Miyajima K, Kaku H (2000) Genetic diversity of *Xanthomonas oryzae* pv. *oryzae* strains from Sri Lanka. Phytopathology 90(4):415–421. doi:10.1094/Phyto.2000.90.4.415

Ochiai H, Inoue V, Takeya M, Sasaki A, Kaku H (2005) Genome sequence of *Xanthomonas oryzae* pv. *oryzae* suggests contribution of large numbers of effector genes and insertion sequences to its race diversity. JARQ 39 (4):275–287

Ogawa T, Lin L, Tabien RE, Khush GS (1987) A new recessive gene for resistance to bacterial blight of rice. Rice Genet Newsl 4:98–100

Ogawa T, Tabien RE, Yamamoto T, Busto GA, Ikeda R (1990) Breeding for near-isogenic lines for resistance to bacterial blight in rice. Rice Genet Newsl 7(10)

Ogawa T, Yamamoto K, Khush G, Mew T (1991) Breeding of near-isogenic lines of rice with single genes for resistance to bacterial blight pathogen (*Xanthomonas campestris* pv. *oryzae*). Jpn J Breed 41:523–529

Ogawa T, Yamamoto T (1987) Selection of recurrent parents to develop near-isogenic lines resistant to bacterial leaf blight of rice. Jpn Agri Res 21:65–69

Ou SH (1985) Rice Diseases, 2nd edn. Association Applied Biology, Surrey

Pandey A, Sonti RV (2010) Role of the FeoB protein and siderophore in promoting virulence of *Xanthomonas oryzae* pv. *oryzae* on rice. J Bacteriol 192(12): 3187–3203. doi:10.1128/JB.01558-09

Pandey MP, Singh H, Mani SC (1986) Breakdown of *Xa4* gene for resistance to bacterial blight (BB) at Pantnagar, India. Int Rice Res News 11:19–20

Parkinson N, Cowie C, Heeney J, Stead D (2009) Phylogenetic structure of *Xanthomonas* determined by comparison of *gyrB* sequences. Int J Syst Evol Microbiol 59(2):264–274

Patil PB, Sonti RV (2004) Variation suggestive of horizontal gene transfer at a lipopolysaccharide (*lps*) biosynthetic locus in *Xanthomonas oryzae* pv. *oryzae*, the bacterial leaf blight pathogen of rice. BMC Microbiol 4:40

Pradhan BB, Ranjan M, Chatterjee S (2012) XadM, a novel adhesin of *Xanthomonas oryzae* pv. *oryzae*, exhibits similarity to Rhs family proteins and is required for optimum attachment, biofilm formation, and virulence. Mol Plant Microbe Interact 25(9):1157–1170

Qian G, Liu C, Wu G, Yin F, Zhao Y, Zhou Y, Zhang Y, Song Z, Fan J, Hu B (2012a) AsnB, regulated by diffusible signal factor and global regulator Clp, is involved in aspartate metabolism, resistance to oxidative stress and virulence in *Xanthomonas oryzae* pv. *oryzicola*. Mol Plant Pathol 14:145–157

Qian G, Zhang Y, Zhou Y, Liu C, Zhao Y, Song Z, Fan J, Hu B, Liu F (2012b) *epv*, encoding a hypothetical protein, is regulated by DSF-mediating quorum sensing as well as global regulator Clp and is required for optimal virulence in *Xanthomonas oryzae* pv. *oryzicola*. Phytopathology 102:841–847

Qian G, Zhou Y, Zhao Y, Song Z, Wang S, Fan J, Hu B, Venturi V, Liu F (2013) Proteomic analysis reveals novel extracellular virulence-associated proteins and functions regulated by the diffusible signal factor (DSF) in *Xanthomonas oryzae* pv. *oryzicola*. J Proteome Res doi:10.1021/pr4001543

Rai R, Ranjan M, Pradhan BB, Chatterjee S (2012) Atypical regulation of virulence-associated functions by a diffusible signal factor in *Xanthomonas oryzae* pv. *oryzae*. Mol Plant Microbe Interact 25:789–801

Rajeshwari R, Jha G, Sonti RV (2005) Role of an in planta-expressed xylanase of *Xanthomonas oryzae* pv. *oryzae* in promoting virulence on rice. Mol Plant Microbe Interact 18(8):830–837

Ray SK, Rajeshwari R, Sharma Y, Sonti RV (2002) A high molecular weight outer membrane protein of

Xanthomonas oryzae pv. *oryzae* exhibits similarity to non-fimbrial adhesins of animal pathogenic bacteria and is required for optimum virulence. Mol Microbiol 46(3):637–647

Ray SK, Rajeshwari R, Sonti RV (2000) Mutants of *Xanthomonas oryzae* pv. *oryzae* deficient in general secretory pathway are virulence deficient and unable to secrete xylanase. Mol Plant Microbe Interact 13:394–401

Raymundo AK, Briones AM, Ardales EY, Perez MT, Fernandez LC, Leach JE, Mew TW, Ynalvez MA, McLaren CG, Nelson RJ (1999) Analysis of DNA polymorphism and virulence in Philippine strains of *Xanthomonas oryzae* pv. *oryzicola*. Plant Dis 83(5): 434–440

Reddy APK, Mackenzie DR, Rouse DI, Rao AV (1979) Relationship of bacterial leaf-blight severity to grain-yield of rice. Phytopathology 69:967–969

Reddy V, Kumar Y, Raghavendra A, Sowjenya G, Kumar S, Ramyasree G, Reddy G (2012) *In silico* model of DSF synthase RpfF protein from *Xanthomonas oryzae* pv. *oryzae*: a novel target for bacterial blight of rice disease. Bioinformation 8:504

Reyon D, Tsai SQ, Khayter C, Foden JA, Sander JD, Joung JK (2012) FLASH assembly of TALENs for high-throughput genome editing. Nat Biotechnol 30:460–465

Robin GP, Ortiz E, Szurek B, Brizard JP, Koebnik R (2013) Comparative proteomics reveal new HrpX-regulated proteins of *Xanthomonas oryzae* pv. *oryzae*. J Proteomics doi:10.1016/j.jprot.2013.04.010

Ryan RP, Vorholter F-J, Potnis N, Jones JB, Van Sluys M-A, Bogdanove AJ, Dow JM (2011) Pathogenomics of *Xanthomonas*: understanding bacterium-plant interactions. Nat Rev Microbiol 9:344–355

Ryba-White M, Notteghem JL, Leach JE (1995) Comparison of *Xanthomonas oryzae* pv. *oryzae* strains from Africa, North America, and Asia by restriction fragment length polymorphism analysis. Int Rice Res News 20:25–26

Salzberg SL, Sommer DD, Schatz MC, Phillippy AM, Rabinowicz PD, Tsuge S, Furutani A, Ochiai H, Delcher AL, Kelley D, Madupu R, Puiu D, Radune D, Shumway M, Trapnell C, Aparna G, Jha G, Pandey A, Patil PB, Ishihara H, Meyer DF, Szurek B, Verdier V, Koebnik R, Dow JM, Ryan RP, Hirata H, Tsuyumu S, Won Lee S, Ronald PC, Sonti RV, Van Sluys MA, Leach JE, White FF, Bogdanove AJ (2008) Genome sequence and rapid evolution of the rice pathogen *Xanthomonas oryzae* pv. *oryzae* PXO99A. BMC Genomics 9:204

Savary S, Willocquet L, Elazegui FA, Castilla NP, Teng PS (2000a) Rice pest constraints in tropical Asia: quantification of yield losses due to rice pests in a range of production situations. Plant Dis 84:357–369

Savary S, Willocquet L, Elazegui FA, Teng PS, Du PV, Zhu D, Tang Q, Huang S, Lin Z, Singh HM, Srivastava RK (2000b) Rice pest constraints in tropical Asia: characterization of injury profiles in relation to production situations. Plant Dis 84:156–341

Semenova E, Nagornykh M, Pyatnitskiy M, Artamonova II, Severinov K (2009) Analysis of CRISPR system function in plant pathogen *Xanthomonas oryzae*. FEMS Microbiol Lett 296:110–116

Seo YS, Sriariyanun M, Wang L, Pfeiff J, Phetsom J, Lin Y, Jung KH, Chou HH, Bogdanove A, Ronald P (2008) A two-genome microarray for the rice pathogens *Xanthomonas oryzae* pv. *oryzae* and *X. oryzae* pv. *oryzicola* and its use in the discovery of a difference in their regulation of hrp genes. BMC Microbiol 8:99

Y-p Shen, L-f Zou, Y-r Li, H-s Zou, X-l Liu, G-y Chen (2012) Xoryp_08180 of *Xanthomonas oryzae* pv. *oryzicola*, encoding a hypothetical protein, is regulated by HrpG and HrpX and required for full virulence in rice. J Integr Ag 11:600–610

Siguier P, Filee J, Chandler M (2006) Insertion sequences in prokaryotic genomes. Curr Opin Microbiol 9:526–531

Sinha D, Gupta MK, Patel HK, Ranjan A, Sonti RV (2013) Cell Wall Degrading Enzyme Induced Rice Innate Immune Responses Are Suppressed by the Type 3 Secretion System Effectors XopN, XopQ, XopX and XopZ of *Xanthomonas oryzae* pv. *oryzae*. PLoS ONE 8(9):e75867

Song C, Yang B (2010) Mutagenesis of 18 Type III effectors reveals virulence function of XopZPXO99 in *Xanthomonas oryzae* pv. *oryzae*. Mol Plant Microbe Interact 23:893–902

Song W-Y, Wang G-L, Chen L-L, Kim HS, Pi LY, Holsten T, Gardner J, Wang B, Zhai W-X, Zhu L-H, Fauquet C, Ronald P (1995) A receptor kinase-like protein encoded by the rice disease resistance gene, *Xa21*. Science 270:1804–1806

Soto-Suarez M, Bernal D, Gonzalez C, Szurek B, Guyot R, Tohme J, Verdier V (2010a) In planta gene expression analysis of *Xanthomonas oryzae* pathovar *oryzae*, African strain MAI1. BMC Microbiol 10:170. doi:10.1186/1471-2180-10-170

Soto-Suarez M, Gonzalez C, Piegu B, Tohme J, Verdier V (2010b) Genomic comparison between *Xanthomonas oryzae* pv. *oryzae* and *Xanthomonas oryzae* pv. *oryzicola*, using suppression-subtractive hybridization. FEMS Microbiol Lett 308(1):16–23. doi:10.1111/j.1574-6968.2010.01985.x

Sriariyanun M, Seo YS, Phetsom J (2012) Identification of PhoP regulon members in *Xanthomonas oryzae* pv. *oryzae* PXO99A. In: 4th international conference on chemical, biological and environmental engineering, IACSIT Press, Singapore. doi:10.7763/IPCBEE

Subramoni S, Pandey A, Vishnupriya MR, Patel HK, Sonti RV (2012) The ColRS system of *Xanthomonas oryzae* pv. *oryzae* is required for virulence and growth in iron-limiting conditions. Mol Plant Pathol 13:690–703. doi:10.1111/j.1364-3703.2011.00777.x

Subramoni S, Sonti RV (2005) Growth deficiency of a *Xanthomonas oryzae* pv. *oryzae* fur mutant in rice leaves is rescued by ascorbic acid supplementation. Mol Plant Microbe Interact 18(7):644–651. doi:10.1094/MPMI-18-0644

Sugio A, Yang B, Zhu T, White FF (2007) Two type III effector genes of *Xanthomonas oryzae* pv. *oryzae* control the induction of the host genes *OsTFIIAγ1* and *OsTFX1* during bacterial blight of rice. Proc Natl Acad Sci 104:10720

Sukchawalit R, Vattanaviboon P, Utamapongchai S, Vaughn G, Mongkolsuk S (2006) Characterization of *Xanthomonas oryzae* pv. *oryzae recX*, a gene that is required for high level expression of *recA*. FEMS Microbiol Lett 205(1):83–89

Sun Q, Wu W, Qian W, Hu J, Fang R, He C (2006) High quality mutant libraries of *Xanthomonas oryzae* pv. *oryzae* and *X. campestris* pv. *campestris* generated by an efficient transposon mutagenesis system. FEMS Microbiol Lett 226:145–150

Tabei H (1977) Anatomical studies of rice plant affected with bacterial leaf blight, *Xanthomonas oryzae* (Uyeda et Ishiyama Dowson). Bull Kyushu Agri Expt Sta 19:193–257

Tang D, Wu W, Li W, Lu H, Worland AJ (2000) Mapping of QTLs conferring resistance to bacterial leaf streak in rice. Theor Appl Genet 101:286–291

Tang JL, Feng JX, Li QQ, Wen HX, Zhou DL, Wilson TJ, Dow JM, Ma QS, Daniels MJ (1996) Cloning and characterization of the *rpfC* gene of *Xanthomonas oryzae* pv. *oryzae*: involvement in exopolysaccharide production and virulence to rice. Mol Plant Microbe Interact 9:664–666

Triplett LR, Hamilton JP, Buell CR, Tisserat NA, Verdier V, Zink F, Leach JE (2011) Genomic analysis of *Xanthomonas oryzae* isolates from rice grown in the United States reveals substantial divergence from known *X. oryzae* pathovars. Appl Environ Microbiol 77(12):3930–3937. doi:10.1128/AEM.00028-11

Vera Cruz C, Bai J, Oña I, Leung H, Nelson R, Mew T, Leach JE (2000) Predicting durability of a disease resistance gene based on an assessment of the fitness loss and epidemiological consequences of avirulence gene mutation. Proc Natl Acad Sci USA 97: 13500–13505

Verdier V, Triplett LR, Hummel AW, Corral R, Cernadas RA, Schmidt CL, Bogdanove AJ, Leach JE (2012a) Transcription activator-like (TAL) effectors targeting *OsSWEET* genes enhance virulence on diverse rice (*Oryza sativa*) varieties when expressed individually in a TAL effector-deficient strain of *Xanthomonas oryzae*. New Phytol 196(4):1197–1207. doi:10.1111/j.1469-8137.2012.04367.x

Verdier V, Vera Cruz C, Leach JE (2012b) Controlling rice bacterial blight in Africa: needs and prospects. J Biotechnol 159:320–328. doi:10.1016/j.jbiotec.2011.09.020

Verdier V, Vera Cruz C, Leach JE (2012c) Controlling rice bacterial blight in Africa: needs and prospects. J Biotechnol 159(4):320–328. doi:10.1016/j.jbiotec.2011.09.020

Wang GL, Ruan DL, Song WY, Sideris S, Chen L, Pi LY, Zhang S, Zhang Z, Fauquet C, Gaut BS, Whalen MC, Ronald PC (1998) *Xa21D* encodes a receptor-like molecular with a leucine-rich repeat domain that determines race-specific recognition and is subject to adaptive evolution. Plant Cell 10:765–780

Wang L, Makino S, Subedee A, Bogdanove AJ (2007) Novel candidate virulence factors in rice pathogen *Xanthomonas oryzae* pv. *oryzicola* as revealed by mutational analysis. Appl Environ Microbiol 73(24): 8023–8027

Wang L, Vinogradov EV, Bogdanove AJ (2013a) Requirement of the lipopolysaccharide O-chain biosynthesis gene *wxocB* for type III secretion and virulence of *Xanthomonas oryzae* pv. *oryzicola* J Bacteriol doi:10.1128/JB.02299-12

Wang Y, Kim SG, Wu J, Huh HH, Lee SJ, Rakwal R, Agrawal GK, Park ZY, Kang KY, Kim ST (2013b) Secretome analysis of the rice bacterium Xanthomonas oryzae (*Xoo*) using in vitro and in planta systems. Proteomics:n/a-n/a

Weber E, Gruetzner R, Werner S, Engler C, Marillonnet S (2011) Assembly of designer TAL effectors by Golden Gate cloning. PLoS ONE 6:e19722

White FF, Potnis N, Jones JB, Koebnik R (2009) The type III effectors of *Xanthomonas*. Mol Plant Pathol 10(6):749–766

White TJ, Gonzalez CR (1995) Electroporation of *Xanthomonas*. Electroporation Protoc Microorganisms 47:135

Wonni I, Ouedraogo L, Verdier V (2011) First report of bacterial leaf streak caused by *Xanthomonas oryzae* pv. *oryzicola* on rice in Burkina faso. Plant Dis 95(1):72–73. doi:10.1094/Pdis-08-10-0566

Xiang Y, Cao Y, Xu C, Li X, Wang S (2006) *Xa3*, conferring resistance for rice bacterial blight and encoding a receptor kinase-like protein, is the same as Xa26. Theor Appl Genet 113(7):1347–1355

Xie G, Sun S, Chen J, Zhu X, Chen J, Ye Y, Feng Z, Liang M (1990) Studies on rice seed inspection of *Xanthomonas campestris* pv. *oryzicola*: Immunoradiometric assay. Chinese J Rice Sci 4:127–132

Xu J, Wu M, He C (2012) Identification and functional analysis of Tdrxoo, the member of TonB-dependent-receptor family proteins in *Xanthomonas oryzae* pv. *oryzae*. Acta Microbiol Sinica 50:155

Yamaguchi K, Nakamura Y, Ishikawa K, Yoshimura Y, Tsuge S, Kawasaki T (2013a) Suppression of rice immunity by *Xanthomonas oryzae* type III effector Xoo2875. Biosci Biotechnol Biochem 77(4):796–801

Yamaguchi K, Yamada K, Ishikawa K, Yoshimura S, Hayashi N, Uchihashi K, Ishihama N, Kishi-Kaboshi M, Takahashi A, Tsuge S, Ochiai H, Tada Y, Shimamoto K, Yoshioka H, Kawasaki T (2013b) A receptor-like cytoplasmic kinase targeted by a plant pathogen effector is directly phosphorylated by the chitin receptor and mediates rice immunity. Cell Host Microbe 13(3): 347–357. doi:10.1016/j.chom.2013.02.007

Yang B, Sugio A, White FF (2006) *Os8N3* is a host disease-susceptibility gene for bacterial blight of rice. Proc Natl Acad Sci USA 103:10503–10508. doi:10.1073/pnas.0604088103

Yang B, White FF (2004) Diverse members of the AvrBs3/PthA family of type III effectors are major

virulence determinants in bacterial blight disease of rice. Mol Plant Microbe Interact 17:1192–1200. doi:10.1094/MPMI.2004.17.11.1192

Yang F, Tian F, Sun L, Chen H, Wu M, Yang C-H, He C (2012) A novel two-component system PdeK/PdeR regulates c-di-GMP turnover and virulence of *Xanthomonas oryzae* pv. *oryzae*. Mol Plant Microbe Interact 25:1361–1369

Yang SQ, Liu SY, Zhao S, Yu Y-H, Li RB, Duan CJ, Tang JL, Feng JX (2013) Molecular and pathogenic characterization of new *Xanthomonas oryzae* pv. *oryzae* strains from the coastline region of Fangchenggang city in China. World J Microbiol Biotechnol 29:713–720. doi:10.1007/s11274-012-1227-7

Yoshimura S, Yamanouchi U, Katayose Y, Toki S, Wang ZX, Kono I, Yano M, Iwata N, Sasaki T (1998) Expression of *Xa1*, a bacterial blight-resistance gene in rice, is induced by bacterial inoculation. Proc Natl Acad Sci USA 95(4):1663–1668

Yu Y, Streubel J, Balzergue S, Champion A, Boch J, Koebnik R, Feng J, Verdier V, Szurek B (2011) Colonization of rice leaf blades by an African strain of *Xanthomonas oryzae* pv. *oryzae* depends on a new TAL effector that induces the rice nodulin-3 *Os11N3* gene. Mol Plant Microbe Interact 24(9):1102–1113. doi:10.1094/MPMI-11-10-0254

Zang N, Tang DJ, Wei ML, He YQ, Chen B, Feng JX, Xu J, Gan YQ, Jiang BL, Tang JL (2007) Requirement of a *mip*-like gene for virulence in the phytopathogenic bacterium *Xanthomonas campestris* pv. *campestris*. Mol Plant Microbe Interact 20(1):21–30. doi:10.1094/MPMI-20-0021

Zhang H, Wang S (2013) Rice versus *Xanthomonas oryzae* pv. *oryzae*: a unique pathosystem. Curr Opin Plant Biol doi:10.1016/j.pbi.2013.02.008

Zhang X, Gao S, Wu M, He C (2009) DNA microarray expression analysis of *Xanthomonas oryzae* pv. *oryzae* in rice leaves at early infection stages using selective bacterial transcript labeling with genome-directed primers. Scientia Agricultura Sinica 10:3501–3508

Zhao B, Ardales E, Raymundo A, Bai J, Trick HN, Leach JE, Hulbert S (2004a) The *avrRxo1* gene from the rice pathogen *Xanthomonas oryzae* pv. *oryzicola* confers a nonhost defense reaction on maize with resistance gene Rxo1. Mol Plant Microbe Interact 17:771–779

Zhao B, Lin X, Poland J, Trick H, Leach J, Hulbert S (2005) A maize resistance gene functions against bacterial streak disease in rice. Proc Natl Acad Sci USA 102(43):15383–15388

Zhao BY, Ardales E, Brasset E, Claflin LE, Leach JE, Hulbert SH (2004b) The Rxo1/ Rba1 locus of maize controls resistance reactions to pathogenic and non-host bacteria. Theor Appl Genet 109(1):71–79

Zhao S, Mo WL, Wu F, Tang W, Tang JL, Szurek B, Verdier V, Koebnik R, Feng JX (2013) Identification of non-TAL effectors in *Xanthomonas oryzae* pv. *oryzae* Chinese strain 13,751 and analysis of their role in the bacterial virulence. World J Microbiol Biotechnol 29:733–744. doi:10.1007/s11274-012-1229-5

Zhao S, Poulin L, Rodriguez RL, Serna NF, Liu SY, Wonni I, Szurek B, Verdier V, Leach JE, He YQ, Feng JX, Koebnik R (2012a) Development of a variable number of tandem repeats typing scheme for the bacterial rice pathogen *Xanthomonas oryzae* pv. *oryzicola*. Phytopathology 102:948–956. doi:10.1094/PHYTO-04-12-0078-R

Zhao Y, Qian G, Fan J, Yin F, Zhou Y, Liu C, Shen Q, Hu B, Liu F (2012b) Identification and characterization of a novel gene, *hshB*, in *Xanthomonas oryzae* pv. *oryzicola* co-regulated by quorum sensing and clp. Phytopathology 102:252–259

Zhao Y, Qian G, Yin F, Fan J, Zhai Z, Liu C, Hu B, Liu F (2011) Proteomic analysis of the regulatory function of DSF-dependent quorum sensing in *Xanthomonas oryzae* pv. *oryzicola*. Microb Pathogenesis 50:48–55

Zheng JS, Li YZ, Fang XJ (2005) Detection of QTL conferring resistance to bacterial leaf streak in rice chromosome 2 (*O. sativa* L. ssp. *indica*). Agri Sci China 38:1923–1925

Zhu PL, Zhao S, Tang JL, Feng JX (2011) The rsmA-like gene *rsmAXoo* of *Xanthomonas oryzae* pv. *oryzae* regulates bacterial virulence and production of diffusible signal factor. Mol Plant Pathol 12(3):227–237

Zhu W, MaGbanua MM, White FF (2000) Identification of two novel *hrp*-associated genes in the *hrp* gene cluster of *Xanthomonas oryzae* pv. *oryzae*. J Bacteriol 182:1844–1853

Zou HS, Song X, Zou LF, Yuan L, Li YR, Guo W, Che YZ, Zhao WX, Duan YP, Chen GY (2012) EcpA, an extracellular protease, is a specific virulence factor required by *Xanthomonas oryzae* pv. *oryzicola* but not by *X. oryzae* pv. *oryzae* in rice. Microbiology 158:2372–2383

Zou L, Li YR, Chen GY (2011) A non-marker mutagenesis strategy to generate poly-hrp gene mutants in the rice pathogen *Xanthomonas oryzae* pv. *oryzicola*. J Integr Ag 10:1139–1150

Genomics of *Xanthomonas citri* and Related Species

7

Neha Jalan, Qing Yan, Sunitha Kogenaru, Yinping Guo, Jeffrey B. Jones, James H. Graham, and Nian Wang

7.1 Introduction

Citrus is an important fruit for human health and nutrition. It is the most widely produced fruit and is grown in more than 80 countries. The citrus industry contributes enormously to the world economy with an estimation of more than 10 billion US$ annually (Ladaniya 2008). However, the citrus industry has been threatened by impending problems including pests, diseases, and climatic problems that have limited expansion of the industry and affected the quality and quantity of citrus production. Citrus bacterial canker is one of the most devastating citrus diseases (Graham et al. 2004). Citrus canker affects most commercially important citrus species and cultivars in many citrus-growing areas worldwide and severe infection due to the disease causes extensive damage to citrus harvest. There are no effective disease suppression strategies against citrus canker for the susceptible cultivars of citrus grown in wet, tropical, and subtropical areas (Schubert et al. 2001). A strictly enforced quarantine or regulatory program was employed in the countries or regions free of citrus canker to prohibit introduction of infected citrus fruit and plant materials.

7.2 Citrus Canker and Related Diseases Associated With Citrus Trees and the Corresponding Pathogens

Multiple types of citrus canker disease have been reported. Among them, Asiatic (A) type canker is the most widespread and destructive form of citrus canker and is caused by *Xanthomonas citri* subsp. *citri* (XccA). Phylogenetic analyses place XccA at the taxonomic position of Bacteria, Proteobacteria, Gammaproteobacteria, Xanthomonadales, *Xanthomonadaceae*, and *Xanthomonas*. *X. citri* is a straight rod-shaped, Gram-negative, obligately aerobic bacterium, which is motile by a single polar flagellum (Fig. 7.1). Bacterial colonies on many common bacteriological media usually display yellow pigment due to the production of xanthomonadin (Brunings and Gabriel 2003; Das 2003; da Silva et al. 2002).

N. Jalan · Q. Yan · S. Kogenaru · Y. Guo · N. Wang (✉)
Citrus Research and Education Center, Department of Microbiology and Cell Science, University of Florida, 700 Experiment Station Road, Lake Alfred, FL 33850, USA
e-mail: nianwang@ufl.edu

J. B. Jones
Department of Plant Pathology, University of Florida, Gainesville, FL 32611, USA

J. H. Graham
Citrus Research and Education Center, Department of Soil and Water Science, University of Florida, 700 Experiment Station Road, Lake Alfred, FL 33850, USA

D. C. Gross et al. (eds.), *Genomics of Plant-Associated Bacteria*,
DOI: 10.1007/978-3-642-55378-3_7,

Fig. 7.1 *Xanthomonas citri* subsp. *citri* causes canker lesion symptoms on citrus leaves (**a**), stem (**b**), and fruits (**c**). **d** The rod-shaped bacterium of *X. citri* subsp. *citri* contains a single polar flagellum. Enlarged view under an electron microscope

Xanthomonas fuscans subsp. *aurantifolii* (Xau) is also known to cause citrus canker but has a limited geographical distribution and limited host range. XauB is restricted to South America (Argentina, Uruguay, and Paraguay) and causes canker B. It mostly affects lemon (*C. limon*) and Mexican lime (*C. aurantifolia*), but is also found on sweet orange (*C. sinensis*) and grapefruit (Civerolo 1984). XauC is restricted to Brazil and causes canker C only on Mexican lime (Stall and Civerolo 1991).

Compared to Xau, XccA has a broad host range and affects most commercial citrus varieties within the Rutaceae family including grapefruit and Mexican lime (Sun et al. 2004). Two variants of XccA have also been identified. The variant designated as A* was found in Southeast Asia in the 1990s causing typical canker lesions on Mexican lime, but not grapefruit (Vernière et al. 1998; Sun et al. 2004). The second variant designated as A^w (Wellington strain) was isolated from Palm Beach County in southern Florida and described by Schubert et al. in Florida in late 1990s (Sun et al. 2004; Schubert et al. 2001). $XccA^w$ was found to be pathogenic to Mexican lime and alemow (*C. macrophyla*) plants, but not to grapefruit and orange. Both $XccA^w$ and XccA cause similar symptoms on Mexican lime, and the populations were similar in this susceptible host (Rybak et al. 2009).

In 1984, a disease similar to citrus canker was discovered in citrus nurseries and named as E type of citrus canker in central Florida and led to destruction of millions of seedlings (Schoulties and Miller 1985; Sun 1984). Leaf spots of this strain are irregular to round, 3–5 mm in diameter, flat, water-soaked, often necrotic in the center, and usually surrounded by a chlorotic halo. Water-soaked elongate lesions with necrotic centers are also observed on twigs but not on fruits (Cubero and Graham 2002). Unlike typical citrus canker, this disease does not show raised callus-like lesions. Further study indicated that this disease is different from citrus canker. This disease was renamed as citrus bacterial spot disease, and the pathogen was named as *X. axonopodis* pv. *citrumelo* (Gabriel et al. 1989).

In this review, we mainly focus on the A type of citrus canker due to its wide distribution and importance. The A type of citrus canker is hereinafter referred to as citrus canker in this review.

7.2.1 Citrus Canker Disease Origin and Its Current Distribution

Knowledge of the origin of disease is essential for understanding the evolution of the pathogen. Currently, citrus canker is reported to have originated in Southeast Asia, from where the disease was distributed to other citrus-growing areas. However, the exact geographical origin of this disease is still a matter of controversy (Das 2003). It was believed that citrus canker originated from Southern China, and the wild host plant was assumed to be *Fortunella hindsii* (Lee 1918). But a later investigation indicated that citrus canker lesions were detected on the oldest citrus herbaria that were collected from India in 1827–1831 and Indonesia in 1842–1844,

suggesting that citrus canker may have originated in India and Java (Jenkins 1933).

The geographical range of citrus canker has continued to expand to new citrus-growing areas since it was first observed in Southeast Asia. By the twentieth century, citrus canker was present in more than 30 counties in Asia, South and Central Africa, the Middle East, Australia, New Zealand, the Pacific Islands, South America, and southeastern USA. The disease has spread to all citrus-producing continents except Europe (Schubert et al. 2001; Das 2003).

Citrus canker has been introduced at least three times into the Florida in USA. The history of this disease in Florida can be traced back to 1910 when it was first introduced to North Florida and other Gulf States. The introduction of this disease is reported to have resulted from a shipment of infected nursery stock from Japan where citrus canker is endemic (Dopson 1964; Stall and Seymour 1983). However, citrus canker was not considered a bacterial disease until 1915 when a bacterium was identified as the causative agent of this disease (Hasse 1915). Quarantine was imposed immediately in 1915. Florida groves and orchards lost 250,000 trees and nurseries lost 3,100,000 plants to the citrus canker eradication program (CCEP) (Rhoads and DeBusk 1931), which was declared successful in 1933 (Dopson 1964). A second outbreak of citrus canker occurred in Tampa Bay area of West Central Florida in 1986, and successful eradication was declared in 1994. Around 88,000 commercial and 600 residential trees were removed and the eradication cost $27 million at that time. Only one year later, a third citrus canker outbreak was found on a residential tree near the Miami International Airport in 1995 (Schubert et al. 2001). It was determined that citrus canker had been there more than 2 years, and the disease had been estimated to spread to 12,950 ha (Schubert et al. 2001). The Florida CCEP, a joint effort of state and federal government, was managed by animal and plant health inspection service (APHIS) and the Florida Department of Agriculture and Consumer Services (FDACS), with the mission to destroy all infected trees. The joint eradication program has spent nearly $1 billion for surveys, eradication, and compensation over a 9-year campaign (Gottwald and Irey 2007). In spite of great efforts to eradicate citrus canker, the dis, the disease had spread extensively mainly due to the hurricanes in 2004 and 2005. Consequently, citrus canker had become endemic in Florida. The United States Department of Agriculture (USDA) determined that citrus canker eradication was not feasible in Florida on January 11, 2006 (USDA 2006). Eventually, the official decision to halt the CCEP was made by Florida House of Representatives on May 3, 2006 (Gottwald and Irey 2007).

7.2.2 Economic Losses and Costs of Citrus Canker Disease Management

Citrus canker poses a wide range of serious economic consequences to the citrus industry by affecting the production, marketing, and trade of commodities. Infection of citrus by XccA at early growth stage causes the fruit to crack or become malformed as they grow. Infection in the later growth stage causes canker lesions on leaves, stems, and fruit (Fig. 7.1). Although XccA does not affect human health, the scattered physical lesions on the canker-infected fruit damage the fruit surface and adversely affect the value of fruit that is suitable for fresh market. It was estimated that the fresh market accounts for nearly one quarter of the entire citrus production in the USA (Zansler 2004). Citrus canker also reduces the productivity and vigor of citrus trees that eventually results in reduction of fruit yield. It is estimated that as much as one-third of the pack-out rates for fresh fruit market will be reduced due to citrus canker disease (Zansler 2004). The severe infection of canker disease on fruit and leaves usually induces heavy defoliation, leaving only bare twigs (Graham et al. 2004).

Several formulations of copper bactericides, including copper oxychloride, copper sulfate, copper oxide, copper hydroxide, and ammonia-copper carbonate, that showed bactericidal activity have been used for controlling citrus

canker disease (Behlau et al. 2008, 2010; Schubert et al. 2001). More than one application is usually needed to maintain fresh fruit appearance standards. In addition to copper compounds, establishment of windbreaks significantly slows down wind speed, resulting in increased effectiveness of other control measures, and thus reducing the intensity of canker disease (Gottwald and Timmer 1995; Behlau et al. 2008; Leite and Mohan 1990).

The most serious economic problem due to citrus canker infection is the impact on commerce caused by quarantine restrictions imposed by canker-free citrus-growing areas and countries to interstate and international transport and sale of fruit and plant material from canker-infested areas (Das 2003; Canteros 2004). Movement of citrus plant materials has to be under strict restrictions within quarantine areas. Commercial citrus nursery sales are prohibited. Some processing plants and packinghouses refuse to accept fruit from quarantine areas (Gottwald et al. 2001). Only fresh fruit harvested from certified canker-free growing areas is allowed to be exported to European markets from Argentina, where citrus canker is endemic. As a result of strict inspections of canker disease being required before harvesting the fruit and during packing, higher costs are incurred (Muraro et al. 2001). It was estimated that the citrus industry in Florida would experience net revenue losses of $53 million and $2 million for red and white seedless grapefruit, respectively, by losing the European Union fresh market due to the quarantine issues of canker disease (Zansler 2004).

7.2.3 Disease Symptoms and Life Cycle

XccA can infect all young above-ground tissues of citrus. Maximum susceptibility of the citrus tissues is observed during the last half of the expansion phase of growth (Schubert et al. 2001). XccA enters host tissues through natural openings and wounds under favorable conditions. The earliest symptoms on leaves can be noticed with a hand lens, as tiny oily looking lesions on the abaxial surface about 4–7 days after inoculation (Schubert and Sun 1996; Stall and Seymour 1983). The primordial lesions may be confused with oil glands in the leaf, but can be distinguished by their aggregated distribution on the leaf surface. On leaves, stems, thorns, and fruit, the lesions become raised above the surface of host tissue and are rough to the touch. Lesions are usually visible on both sides of a leaf. The lesions continue to enlarge and may reach 10 mm in diameter and turn to a tan to brown color with a water-soaked margin. The water soaking may disappear as lesions age and is not prominent on resistant cultivars. A diagnostic yellow halo usually can be observed around the lesion and may fade or disappear as the lesion enlarges. Large number of lesions may form on a single fruit, leaf, or stem (Fig. 7.1). Older lesions on leaves and fruits tend to have more elevated margins and a sunken center, which on fruit result in a blemished appearance (Brunings and Gabriel 2003). Heavy infection eventually leads to defoliation, premature fruit abscission, and twig dieback on diseased plants.

XccA remains alive in lesions on leaves, stems, and fruits. Multiplication of XccA occurs mostly while the lesions are still expanding. The bacteria persist in the lesions on leaves and fruit until they fall and begin to decompose. Even after falling, XccA can be detected from buried citrus leaves with lesions up to 2–3 months (Hartung 1992; Graham et al. 1989). The bacterium could stay viable in lesions on woody branches up to a few years. Viable bacteria were detected from stem lesions on 5- to 7-year-old trunks of Mexican lime infected with the canker pathogen in Florida (Gottwald et al. 2002). Thus, the canker bacteria can remain viable as long as host cells in the vicinity of the lesions are alive, although the bacterial titer will drop considerably. Research also indicated that the canker bacteria may stay viable on non-host plants for several weeks under natural conditions (Gottwald et al. 2002). Bacterial cells ooze out when there is free moisture (e. g., rainwater) on the lesions. The bacterial concentration in the rainwater collected from foliage with lesions

could reach 10^5–10^8 CFU per ml (Stall et al. 1980). The bacterial cells could be dispersed to new growth and other healthy plants and serve as inoculum for further disease development. Windblown rain is the primary mechanism for short and medium distance dispersal of the pathogen. Long-distance spread usually happens by mechanical transmission or by human movement of diseased plant materials, or by tropic storms (Gottwald et al. 1997a, b).

7.3 Virulence Factors and Regulation of XccA

7.3.1 Protein Secretion Systems and Effectors

XccA harbors an extensive repertoire of virulence factors including surface polysaccharides, cell-wall-degrading enzymes, detoxification-related factors, effectors and their secretion systems, and a diffusible signal factor-mediated quorum-sensing (QS) system (da Silva et al. 2002). Intensive studies of virulence mechanisms of XccA were conducted, and numerous virulence-associated factors have been identified (Brunings and Gabriel 2003; Moreira et al. 2010a). In our recent effort to investigate virulence factors of XccA, an EZ-Tn5 mutant library containing 22,000 mutants of XccA strain 306 was used to screen virulence-deficient mutants in grapefruit. A total of 82 genes involved in the type II secretion system (T2SS), type III secretion system (T3SS) and effectors, quorum-sensing system, extracellular polysaccharide (EPS) and lipopolysaccharide (LPS) synthesis, regulatory factors, extracellular enzymes, adhesion, ABC transporters, plasmid stabilization, and general metabolic pathway were identified as being required for the full virulence *in planta* (Yan and Wang 2012).

Many Gram-negative plant pathogenic bacteria employ the T3SS to benefit the pathogen by injecting effector proteins directly into the host cells and thus manipulate the host cellular activities (Büttner and He 2009). XccA contains a *hrp* (hypersensitive response and pathogenicity) cluster of 26 genes from *hrpF* to *hpa2* which encodes the T3SS proteins. The *hrp* cluster genes of XccA are in a similar order with that of *X. campestris* pv. *campestris*, with the only exception of *hrpW*, which is located between *hrpF* and *hrpE* in *X. campestris* pv. *campestris*, but outside the *hrp* cluster in XccA (da Silva et al. 2002). Consistent with the critical role played by the T3SS in the virulence mechanism, mutation of *hrpB*, *hrpB4*, *hrcV,* and *hrcN* in XccA completely abolished the bacterial ability to cause citrus canker symptoms on citrus (Laia et al. 2009; Yan and Wang 2012).

Twenty-four known and putative T3SS effectors were found in the genome of XccA (Moreira et al. 2010a). Among them, PthA has been well studied and functions as a major pathogenicity determinant in the canker pathogen. PthA is necessary for XccA to cause citrus canker disease, and an exogenous insertion of *pthA* gene into *X. axonopodis* pv. *citrumelo*, which elicits citrus bacterial spot disease without the erumpent canker lesions, confers the ability to cause canker symptoms in citrus (Swarup et al. 1991). Importantly, transient expression of *pthA* itself in citrus plants is sufficient to induce hypertrophy and hyperplasia of host cells and cause canker lesions on citrus leaves (Duan et al. 1999). PthA is a member of the avirulence and pathogenicity (AvrBs3/PthA) or transcription activator-like (TAL) effector family. Members of this large family are widely distributed in phytopathogenic *Xanthomonas* species and play critical roles. AvrBs3/PthA effectors contain a near-identical, leucine-rich, tandem repeat region in their central part. Each repeat, in general, consists of 102 bp nucleotides encoding 34 amino acids. The number and nature of the central repeats may vary in different members of this effector family and thus contribute to the specific function of the AvrBs3/PthA effectors during the elicitation of resistance or virulence in the respective host–pathogen interactions (Szurek et al. 2001; Boch and Bonas 2010; Fujikawa et al. 2006; Yang and White 2004; White et al. 2009). It has been elucidated that AvrBs3 in *X. campestris* pv. *vesicatoria*, the causal agent of bacterial spot disease on pepper

and tomato, was translocated to the plant cell nucleus and induced the expression of *upa20* by direct binding to the promoter of *upa20* via the central repeat region (Kay et al. 2007). The *upa20* gene encodes a master regulator of plant cell size (Marois et al. 2002). Activation of *upa20* eventually resulted in hypertrophy of plant mesophyll cells that contributes to symptom development. Efforts have been made to define the target(s) of PthA4, a functional homolog of PthA, of XccA in host citrus. Interestingly, a very recent microarray analysis revealed that the wild-type XccA 306 but not the *pthA4* mutant strongly induced the expression of two host plant genes, which encode a lateral organ boundaries (LOB) domain protein and a nodulin MtN3 protein, respectively (Hu et al. 2012), which might be the direct targets of PthA4 and involved in citrus canker symptom development.

7.3.2 Exopolysaccharide and Lipopolysaccharide

Similar to other species of *Xanthomonas*, one characteristic feature of *X. citri* is the production of copious amounts of mucoid EPS (also named xanthan gum) (Subramoni et al. 2006). EPS of xanthomonads consists of β-1,4-linked D-glucose backbone with trisaccharide side chains containing mannose-(β-1,4)-glucuronic acid-(β-1,2)-mannose attached to alternate glucose residues in the backbone by α-1,3 linkages (Jansson et al. 1975) and can be released into its living environment when cultured in vitro or in vivo. EPS has long been recognized as an important virulence factor for many phytopathogenic bacteria (Leigh and Coplin 1992). In XccA, the *gum* cluster genes from *gumB* to *gumP* are responsible for the biosynthesis and exportation of EPS. Six *gum* genes, i.e., *gumCDEFJK*, were identified to be involved in biofilm formation in borosilicate glass tubes (Li and Wang 2011a). Biofilm refers to complex structures formed by communities of microorganisms that enhance attachment to biotic or abiotic surfaces. Biofilm formation protects microbes from diverse environmental stresses, contributes to the epiphytic fitness of microbes, and has been implicated in the virulence of many phytopathogenic bacteria (Danhorn and Fuqua 2007). Mutation of *gumB* gene in Xcc resulted in defective EPS production, biofilm formation, epiphytic survival, and reduced disease symptoms in lemons (Rigano et al. 2007), suggesting an important role of EPS in the early infection process of the canker pathogen. *gumF* and *gumK* were also reported to contribute to virulence of XccA (Yan and Wang 2012). Compared with the wild type, the *gumF* and *gumK* mutants showed a significant reduction of bacterial growth *in planta* after pressure inoculation into the intercellular spaces of Duncan grapefruit leaves. EPS also suppresses plant defense responses such as callose deposition in the plant cell wall, probably by chelation of divalent calcium ions that are present in the plant apoplast and are required for activation of plant defense responses (Aslam et al. 2008).

Another important virulence factor employed by the canker pathogen is LPS. LPS is a key component of the outer membrane of Gram-negative bacteria and plays multiple roles in plant–microbe interactions (Newman et al. 2007). It has been suggested that LPS can function as a protective wall to protect bacteria from variety of unfavorable environmental conditions such as antimicrobial compounds produced by host cells (Papo and Shai 2005; Dow et al. 1995). LPS is also a pathogen-associated molecular pattern (PAMP) that elicits or potentiates plant defense-related responses in plant–pathogen interactions (Zeidler et al. 2004). LPS is a tripartite amphipathic molecule composed of a lipid A moiety, a core oligosaccharide, and an O-antigen polysaccharide chain (Vorholter et al. 2001; Raetz and Whitfield 2002). In *Xanthomonas*, the *lps* gene cluster is flanked by highly conserved genes, *metB* and *etfA*, which are required for the biosynthesis and exportation of LPS (Patil and Sonti 2004). Variation in gene number and nucleotide sequence of the *lps* gene clusters have been found among different strains of *Xanthomonas* (Patil and Sonti 2004; Patil et al. 2007). For example, the *lps* cluster contains 6, 14, and

15 genes in *X. oryzae* pv. *oryzae*, XccA, and *X. campestris* pv. *campestris,* respectively. This variation was proposed to play a role in evading the host immune system in diverse host–pathogen interactions. The LPS structure of XccA has been defined recently (Casabuono et al. 2011). It is composed mainly of a penta- or tetra-acylated diglucosamine backbone attached to either two pyrophosphorylethanolamine (PP-EtNH2) groups or to one PP-EtNH2 group and one phosphorylethanolamine group. The core region consists of a branched oligosaccharide and two phosphate groups, whereas the O-antigen is composed of a rhamnose homo-oligosaccharide. Recent studies indicate that LPS is involved in multiple steps of the disease cycle of XccA. Mutation of *wxacO* (XAC3596) and *rfbC* (XAC3598) impaired LPS production as revealed by sodium dodecylsulphate–polyacrylamide gel electrophoresis (Li and Wang 2011b). The *wxacO* and *rfbC* mutants formed a compromised biofilm on glass or host plant leaves and were more sensitive to antimicrobial compounds such as polymyxin B and hydrogen peroxide. Both mutants also showed deficiency in virulence and growth on host leaves after spray inoculation. The mutants, which had intact flagella, had reduced motility on swimming and swarming assay plates (Li and Wang 2011b). Additionally, another six *lps* genes, i. e., *wzm*, *wzt*, XAC3591, XAC3593, XAC3595, and XAC3597, were reported to be involved in biofilm formation (Li and Wang 2011a). The *wzm* and *wzt* genes encode an ATP-binding cassette (ABC) transporter system responsible for export of the O-antigen polysaccharide of LPS (Cuthbertson et al. 2007). Mutation of *wzm* and *wzt* also severely affected the virulence of XccA and significantly reduced bacterial populations in the host plant (Yan and Wang 2012; Casabuono et al. 2011). Based on sequence analysis, two novel hypothetical genes (*orf3* and *orf5*) were reannotated in the *lps* cluster in XccA 306 (Patil et al. 2007). Gene *orf5* was renamed as *nlxA* (novel *lps* cluster gene of XccA), which is required for the LPS O-antigen biosynthesis by encoding a putative rhamnosyltransferase (Yan et al. 2012). The *nlxA* mutant caused much less canker symptoms on Duncan grapefruit leaves, and the bacterial population was significantly lower than the wild type in host plant (Yan et al. 2012; Yan and Wang 2012). Mutation of *orf3* (renamed as *nlxB*) also severely affected the virulence of XccA in Duncan grapefruit (Yan and Wang 2012). However, no detectable change was observed in the LPS pattern by mutation of *nlxB* gene.

7.3.3 Regulatory Systems

Xanthomonas citri, like other foliar bacterial pathogens, encounters diverse environmental conditions on leaf surface and the intercellular space during its infection cycle. *X. citri* has evolved elaborate regulatory systems to efficiently use its genetic resource to adapt to these environments.

7.3.3.1 Two-Component Regulatory System

The two-component regulatory system is one of the basic stimulus-response coupling mechanisms for bacteria to sense and respond to a variety of stimuli (Stock et al. 2000). A typical two-component system consists of a membrane-associated histidine kinase sensor and a cytoplasmic response regulator. Upon stimulation by a specific environmental signal, the sensor kinase is activated via autophosphorylation at a conserved histidine residue. The phosphoryl group is then transferred to the cognate response regulator, which results in a conformational change in the regulatory protein and activates its target genes.

Genome analysis indicates that XccA contains a large number of genes ($n = 114$) that belong to the two-component regulatory system (Qian et al. 2008; da Silva et al. 2002). Among them, HrpG is a two-component system response regulator responsible for regulation of the *hrp* genes in XccA. Similar to the typical response regulator, the 263-residue HrpG protein of XccA contains an N-terminal response regulator receiver domain that commonly receives a phosphoryl signal from sensor kinase. Yeast two-hybrid assay using HrpG as bait

revealed that HrpG can interact with itself and three proteins encoded by XAC0095, XAC1568, and XAC3683 (Alegria et al. 2004). Interestingly, XAC3683 encodes a putative histidine kinase sensor. This result implied that XAC3683 might serve as a cognate sensor kinase to activate the HrpG regulator. However, mutation of XAC3683 has no detectable effect on the bacterial virulence in Duncan grapefruit. How HrpG is activated in XccA remains to be addressed.

ColS/ColR is another two-component regulatory system that plays important role in regulation of the virulence genes. The ColS/ColR system was first identified in *Pseudomonas fluorescens* and contributes to its colonization of plant roots (Dekkers et al. 1998). ColS functions as the sensor kinase and ColR is the cognate response regulator. Further investigations of ColS/ColR in *P. putida* suggested that it plays multiple roles including transposition of a transposon (Rita Hõrak et al. 2004; Kivistik 2010), membrane function regulation, phenol tolerance (Kivistik et al. 2006), and heavy metal resistance (Hu and Zhao 2007).

ColS/ColR is a global regulatory system involved in various cellular processes, including virulence, hypersensitive response, and stress tolerance in *X. campestris* pv. *campestris* (Zhang et al. 2008). The ColS/ColR system also plays multiple roles in the virulence mechanism of XccA (Yan and Wang 2011). Mutation of *colS* and *colR* abolished the virulence of XccA *in planta*. The mutants formed a much reduced biofilm in glass tubes and on host leaf surfaces. Furthermore, the bacterial populations of the mutants *in planta* were significantly lower than that of the wild type. Quantitative reverse transcription-PCR (qRT-PCR) assay revealed that ColS/ColR positively regulated the expression of important virulence genes, including *hrpD6*, *hpaF*, the O-antigen LPS synthesis gene *rfbC*, and the catalase gene *katE*. Consistently, mutation of *colS* or *colR* altered the LPS production and impaired the catalase activity and tolerance of environmental stress, including phenol, copper, and hydrogen peroxide (Yan and Wang 2011).

7.3.3.2 Quorum-Sensing System

Quorum-Sensing is a widespread cell-to-cell communication mechanism utilized by many bacteria to monitor their population densities and to adapt their behaviors by regulating gene expression in response to fluctuations in cell population density (Miller and Bassler 2001; Keller and Surette 2006). It was first described as a regulatory mechanism that controls bioluminescence production in *Vibrio fischeri* (Nealson and Hastings 1979). QS has been identified in many Gram-negative and Gram-positive bacteria and plays important roles in a diverse array of physiological activities, including symbiosis, virulence, competence, conjugation, antibiotic production, motility, sporulation, and biofilm formation (González and Keshavan 2006; de Kievit and Iglewski 2000; Waters et al. 2008; Waters and Bassler 2005). Diverse QS systems that employ different signals including N-acyl-homoserine lactones (AHLs), LuxS-dependent autoinducer-2 (AI-2), and *Pseudomonas* quinolone signal (PQS), and the diffusible signal factor (DSF)have been found in Gram-negative pathogenic bacteria (Williams et al. 2007).

The DSF-mediated QS pathways have been found to be involved in virulence regulation in different strains of xanthomonads, such as *X. campestris* pv. *campestris*, *X. oryzae* pv. *oryzae*, *Xylella fastidiosa*, and XccA (He et al. 2010; Barber et al. 1997; Andrade et al. 2006; Chatterjee et al. 2008). The DSF structure from *X. campestris* pv. *campestris* was recently defined as an unsaturated fatty acid *cis*-11-methyldodecenoic acid (Wang et al. 2004). The regulation of pathogenicity factors (*rpf*) cluster of genes is required for the synthesis and perception of the DSF signal (Dow 2008). Gene *rpfF* encoding a putative enoyl-CoA hydratase functions as a key enzyme for DSF biosynthesis. The perception of DSF is dependent on a two-component system encoded by *rpfC* and *rpfG* (Barber et al. 1997; Slater et al. 2000). RpfC is a membrane-associated sensor kinase containing a CheY-like receiver motif and a C-terminal histidine phosphotransfer domain. RpfG is the response regulator with an HD-GYP

domain, which is involved in the hydrolysis of the bacterial second messenger cyclic-di-GMP (Slater et al. 2000; Ryan et al. 2010).

In XccA, the DSF-mediated QS system plays an important role in initial attachment of bacterial cells on host surfaces during the infection (Guo et al. 2012). Compared with the wild type, the *rpfF*, *rpfC*, and *rpfG* mutants formed less biofilm in polypropylene tubes. Green fluorescent protein (GFP)-labeled wild-type and mutant strains were used in a confocal laser scanning microscopy assay to determine the bacterial attachment to host leaves. The wild-type cells started to aggregate to form microcolonies on the leaf surfaces of Duncan grapefruit after one hour of incubation, but only a few individual cells of the *rpfF*, *rpfC*, and *rpfG* mutants attached to the leaves. Higher bacterial attachment and a biofilm matrix were observed on the leaf surfaces inoculated with wild type but not the *rpf* mutants at six hours after inoculation. A scanning electron microscope assay also showed that more wild-type cells occupied the depressions between epidermal cells and around stomata than the QS mutants. Consistently, after a spray inoculation which mimics the natural invasion conditions, the *rpfF*, *rpfC*, and *rpfG* mutants caused reduced canker lesions on grapefruit leaves than the wild type (Guo et al. 2012).

The DSF-mediated QS system is also required for the full virulence of XccA after entering the host plant. Compared with the wild-type XccA, deletion of the *rpfF*, *rpfC*, and *rpfG* genes showed no significant effect on the symptom development in Duncan grapefruit leaves when the strains were inoculated with a high bacterial concentration adjusted to 10^8 CFU/ml, but caused significantly fewer lesions on leaves when inoculated with a bacterial suspension adjusted to 10^4 CFU/ml (Guo et al. 2012). The QS mutants were also deficient in motility, and extracellular protease as was demonstrated in other *Xanthomonas* species (Barber et al. 1997; Ryan et al. 2010).

Yeast two-hybrid assays revealed that components of the DSF-mediated QS system can physically interact with a number of proteins including the NtrB/NtrC two-component regulatory system and proteins containing a GGDEF domain (Andrade et al. 2006). It has been reported that the GGDEF domain possesses diguanylate cyclase activity involved in the synthesis of bacterial second messenger cyclic-di-GMP (Tal et al. 1998; Paul et al. 2004). The cyclic-di-GMP can allosterically inhibit the DNA-binding activity of Clp (CRP [cyclic AMP receptor protein]-like protein) in XccA (Leduc and Roberts 2009), thus preventing it from regulating a subset of the DSF regulon genes (He et al. 2007). The HD-GYP domain of the RpfG protein can interact with nine GGDEF proteins. Among them, two were encoded by XAC0258 and XAC0424, respectively (Andrade et al. 2006). Their homolog proteins (XC_0249 and XC_0420) were found to control motility but not other DSF-mediated phenotypes in *X. campestris* pv. *campestris* (Ryan et al. 2010). The expression of XAC0424 was positively regulated by RpfG in the microarray analysis (Guo et al. 2012). Furthermore, the RpfC/RpfG two-component system interacts with the NtrB/NtrC two-component system with the specificity that RpfC interacts with NtrB and RpfG interacts with NtrC, respectively (Andrade et al. 2006). The NtrB/NtrC system has been shown to regulate σ^{54}-dependent transcription of nitrogen-regulated genes in many enteric bacteria (Ninfa et al. 1995). The sensor kinase NtrB possesses a C-terminal transmitter module that transfers the phosphoryl signal to an N-terminal receiver domain of the response regulator NtrC which contains a σ^{54}-activating domain and a DNA-binding domain. The NtrC-like σ^{54} activator (LuxO) and σ^{54} are involved in QS regulation network in *V. harveyi* and *V. cholera* (Lenz et al. 2004). Interestingly, mutation of XAC1969 (encoding σ^{54}, RpoN) impaired the biofilm formation (Li and Wang 2011a). The expression of XAC1969 was positively controlled by the QS in XccA (Guo et al. 2012). These results indicate that the NtrB/NtrC two-component system and σ^{54} could be another regulatory pathway involved in the QS-regulated biofilm formation in the canker bacteria.

7.3.4 Tools for Molecular Genetics Studies

As a model strain to study the plant–microbe interactions, diverse tools are suitable for the molecular genetics studies of XccA. The genome sequence of XccA 306 was published in 2002 and has greatly facilitated the molecular research on citrus canker (da Silva et al. 2002). Site-directed mutagenesis and random mutagenesis have been used to generate mutants. For example, EZ-Tn5 was used to create mutant libraries of XccA to identify virulence- and biofilm-related genes (Yan and Wang 2012; Li and Wang 2011a; Guo et al. 2010). Gene(s) of interest could be conveniently deleted using a homologous recombination strategy via a suicide plasmid pOK1 (Huguet et al. 1998). A subset of plasmids including pUFR053 and its derivatives could survive in XccA (El Yacoubi et al. 2007). Extra plasmids could be introduced into XccA cells by electroporation or by triparental mating method with the helper plasmid pRK2013 (Figurski and Helinski 1979).

7.4 Genome Sequencing of *X. citri* and Related Species

7.4.1 Genome Structure, General Features, and Distinguishing Characteristics

The complete genome sequence of XccA strain 306 was sequenced in 2002 (da Silva et al. 2002). Shotgun libraries with 46,462 clones were created using pUC18 to completely sequence the pathogen with 7X coverage, representing 98 % of the genome. The rest of the genome was decoded using a cosmid library (da Silva et al. 2002). The genome consists of a circular chromosome, 5.27 Mb in size and two plasmids 33 and 64 Kb in size. The annotated genome is publicly available for download from GenBank (accession number AE008923.1, AE008924.1, AE008925.1). With advances in sequencing technologies, other strains associated with citrus canker were also sequenced recently. Currently, draft sequences of XauB and XauC are available from GenBank (Moreira et al. 2010a). The complete genome of XccA^{w} strain 12879 has been sequenced (Jalan et al. 2013a), whereas XccA* strain 270 is currently being sequenced. Furthermore, the genome of another closely related citrus pathogen, *X. axonopodis* pv. *citrumelo* strain F1 that causes citrus bacterial spot, is also completed (Jalan et al. 2011). Multilocus sequence phylogenetic analysis of XccA and other *Xanthomonas* spp. using protein sequences of nine housekeeping genes (*uvrD*, *secA*, *carA*, *recA*, *groEL*, *dnaK*, *atpD*, *gyrB*, and *infB*) that are highly conserved in bacteria revealed that citrus canker pathogens form a closely related group as compared to other xanthomonads (Fig. 7.2). Overall, the availability of genome sequences of XccA along with related strains and comparative analyses have provided a foundation to elucidate the virulence mechanisms, genetics, and evolution of this important phytopathogen. Some of the key aspects of the genome are discussed further.

The sequenced genomes of XccA and related strains are similar in general characteristics (Table 7.1). Their sizes range from 4.87 to 5.39 million base pairs; G+C content is almost constant at 64.7–64.9 %. All the genomes encode for two ribosomal RNA (5S-16S-23S) operons and 51–54 transfer RNAs. The number of annotated genes ranges from 3,804 to 4,675 among the strains; however, this difference can be attributed to the size difference and the draft nature of genomes of XauB and XauC. Additionally, XccA contains plasmids, which introduce genetic variation as they encode for different virulence factors such as type III effectors and a type IV secretion system among others.

7.4.2 Genome Rearrangements

The gene content among XccA and related strains is largely conserved, but whole genome alignment of completely sequenced genomes

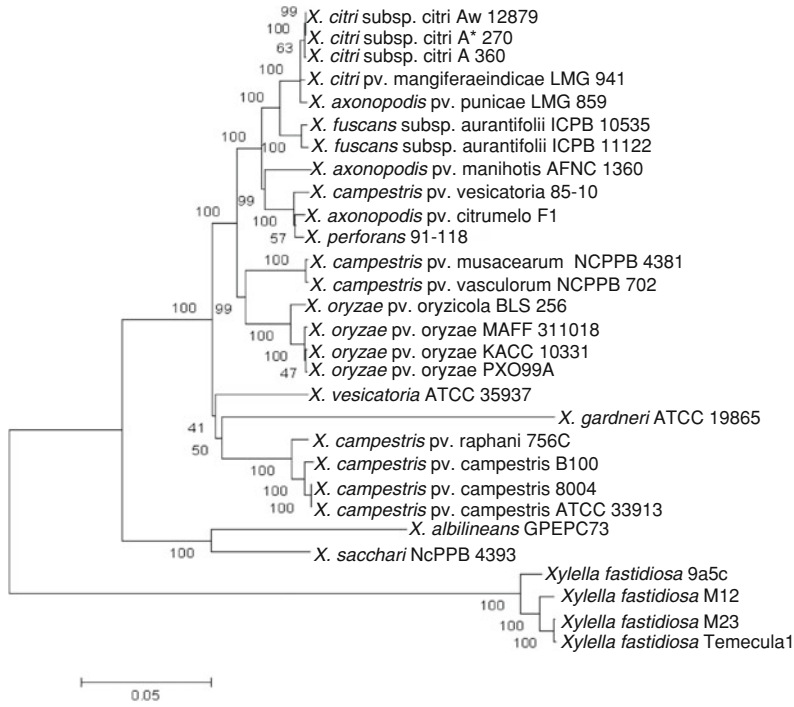

Fig. 7.2 Maximum-likelihood phylogenetic tree of the genomes of *Xanthonomas citri* subsp. *citri* and other sequenced xathomonads. The tree was constructed using concatenated protein sequences of nine housekeeping genes (*uvrD*, *secA*, *carA*, *recA*, *groEL*, *dnaK*, *atpD*, *gyrB,* and *infB*) aligned using Clustal W. A phylogenic tree from concatenated sequences was constructed in MEGA (version 5.0) using the maximum-likelihood method. The strains of *Xylella fastidiosa* were used as out-group species. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (1,000 replicates) is shown next to the branches. Horizontal scale bar (0.1) at the *bottom* represents number of amino acid substitutions per site

Table 7.1 Summary of the complete genome of *Xanthomonas citri* subsp. *citri* A strain 306 and related species

	X. citri subsp. *citri* A strain 306	*X. citri* subsp. *citri* A^w strain 12879	*X. fuscans* subsp. *aurantifolii* strain B	*X. fuscans* subsp. *aurantifolii* strain C	*X. axonopodis* pv. *citrumelo* strain F1
Genome size (bp)	5,274,174	5,398,685	4,877,808	5,012,633	4,967,469
Sequencing status	Finished	Finished	Draft	Draft	Finished
GC content (%)	64.7	64.7	64.9	64.8	64.9
CDSs	4,427	4,760	3,804	3,921	4,202
tRNAs	54	54	51	51	54
rRNA operons	2	2	2	2	2
Plasmids	2	2	0	0	0

shows many inversions, indels, and rearrangements relative to one another (Jalan et al. 2011). Many of the rearranged and deleted blocks are flanked by repetitive mobile genetic elements such as transposons and/or phage-related integrase on at least one border. Since the rest of the genomes are very collinear, it indicates that the rearrangements may be a result of horizontal gene transfer, mediated on occasion by phage vectors. Horizontal gene transfer is recognized as one of the major mechanisms for genome plasticity leading to diversification and speciation of bacteria (Ochman et al. 2000), thus giving rise to variant strains of *X. citri*.

7.4.3 Plasmids

XccA 306 contains two plasmids pXAC33 (33,699 bp) and pXAC64 (64,920 bp). Plasmids are very important for citrus canker-associated *Xanthomonas* spp. as they encode for various virulence factors. The most prominent effector protein is PthA, which belongs to *avrBs3/pthA* family of type III effectors (Al-Saadi et al. 2007). XccA 306 has four copies of *pthA* homologs: *pthA1*, *pthA2*, *pthA3,* and *pthA4* on two plasmids. PthA4 with 17.5 repeats, which is the same as observed for PthA, is known to play an important role in citrus canker as a knockout of *pthA4* abolished citrus canker symptom development (Yan and Wang 2012). PthA is responsible for development of hypertrophic and hyperplasic symptoms and cell death, and its mutation leads to reduction in ability of bacteria to disseminate from infected lesions (Yan and Wang 2012). A type IV secretion system (T4SS) encoded by *vir* genes was identified in the XccA plasmid pXac064. The $XccA^{w}12879$ genome consists of two plasmids pXcaw19 and pXcaw58 that are significantly different from the plasmids found in XccA 306. Plasmid pXcaw19 sequence has no homology with the plasmids of XccA 306, whereas pXcaw58 is only about 35 % similar to pXAC64. Plasmid pXcaw58 contains a *pthAw2* gene, which is a functional homolog of *pthA4*, capable of conferring the ability to cause canker-like symptoms (Swarup et al. 1992). Based on BLAST analysis, 46 % of the plasmid pXAC33 sequence is found in XauB and XauC contigs, respectively, whereas 61 and 55 % of pXAC64 sequence are found in XauB and XauC contigs, respectively (Moreira et al. 2010a). Moreover, Al-Saadi et al. (2007) had established all citrus canker-causing *Xanthomonas* that contain one functional homolog of *pthA* and thus are able to cause canker symptoms (Swarup et al. 1992). The *X. axonopodis* pv. *citrumelo* strain F1 does not harbor any plasmid nor the *pthA4* homolog (Jalan et al. 2011), and thus, it is unable to induce canker-like symptoms on citrus.

7.4.4 IS and CRISPR Elements

A feature shared by all *Xanthomonas* genomes is an abundance of insertion sequence (IS) elements. These IS elements are transposable DNA that can move between bacterial species, thus promoting genome evolution (Monteiro-Vitorello et al. 2005). Other than horizontal gene transfer, IS elements also generate genome rearrangements such as inversions and deletions which result in modification of gene content. XccA 306 contains 108 transposable elements, with the IS*3* family being the most abundant (da Silva et al. 2002). XccA 306 also contains phage-related integrases indicating genome transfer using phage vectors. CRISPRs (Clustered Regularly Interspaced Short Palindromic Repeats) are also found in XccA 306. CRISPRs are short (21–47 bp) direct repeats interspaced with unrelated similarly sized non-repetitive sequences (spacers) (Semenova et al. 2009). CRISPR cassettes function as a prokaryotic immune system and confer resistance to exogenous genetic elements such as plasmids and phages. The spacers incorporated into the genome between CRISPR repeats serve as a 'memory' of past exposures and are used to recognize and silence exogenous genetic elements.

7.4.5 Genomic Islands

Genomic islands are part of the genome and indicate significant evidence of horizontal gene transfer. Distinguishing foreign from indigenous DNA by comparing GC content, GC skew, codon usage bias, and the presence of mobile genetic elements identifies genomic islands. Lima et al. (2005) identified 35 genomic islands in XccA 306 genome by analyzing the GC content bias, the dinucleotide content bias, and the codon usage bias (Lima and Menck 2008). Overall, the total number of genes within genome islands corresponds to almost one quarter of the entire XccA 306 genome. The functional categories that have higher proportions of genes within the genomic islands are those related to pathogenicity and virulence as well as mobile genetic elements (Lima et al. 2008).

The genomic islands also contain genes related to metabolic pathways, including genes involved in the metabolism of the NAD coenzyme, arginine, and cysteine and in energy metabolism (Lima and Menck 2008). Phylogenetic reconstructions for such genes confirm these islands are the result of transfer from distantly related organisms, with some showing homology to the archaea, as well as other distantly related bacteria, such as *Actinobacteria*, *Firmicutes,* and *Bacteroidetes*. This indicates that these genes related to primary metabolic functions were acquired by horizontal gene transfer much earlier in the bacterial evolution cycle (Lima et al. 2008). XccA 306 also contains several recently acquired islands that carry genes related to type II, III, and IV secretion systems, xanthan gum production, and host cell wall degradation (Lima et al. 2005; Van Sluys et al. 2002). Most of the islands are flanked by tRNAs and show the presence of various transposases. Furthermore, the genes in the islands are homologous to distantly related bacteria. Another important characteristic of genomic islands is the presence of orphan genes that have a very limited phylogenetic distribution and have no recognizable homologs. A recent study in *Escherichia coli* demonstrated that most orphan genes encode functional proteins (Daubin and Ochman 2004). Thus, genomic islands in XccA 306 not only encode essential metabolic functions but also facilitate fitness and virulence *in planta*.

7.4.6 Genomic Resources

The complete and draft genome sequences of XccA 306 and related species are available for download along with annotations at the NCBI database. The most up-to-date and comprehensive database is the *Xanthomonas* resource website (www.xanthomonas.org). This database maintains an updated list of all *Xanthomonas* genomes available, the nomenclature of type III effectors as well as new research on *Xanthomonas* available through papers or meetings. The Xanthomonas Genome Browser (http://xgb.fli-leibniz.de/cgi/index.pl) is another resource available to provide genomic and other biological information on a group of *Xanthomonas* spp. with an emphasis on genome comparison. The browser contains only five *Xanthomonas* genomes including XccA 306. Plant Associated and Environmental Microbes Database (PAMDB) (http://genome.ppws.vt.edu/cgi-bin/MLST/home.pl) is a multilocus sequence typing and analysis website and database specifically designed for identification of plant associated and environmental microbes and for the study of their epidemiology, population genetics, and molecular evolution. They support analysis for five different bacterial genera including *Xanthomonas*. The *Xanthomonas* sp. portal is a *Xanthomonas* database intended for comparative genomics, BLAST analysis, and transcriptomics related to the represented strains (https://iant.toulouse.inra.fr//bacteria/annotation/cgi/xansp.cgi). Resources that are not directly associated with *Xanthomonas* but may be useful for annotation, BLAST, and type III effector analyses are the following (1) CPGR: Comprehensive Phytopathogen Genomics Resource (http://cpgr.plantbiology.msu.edu/), (2) DIGAP: Database of Improved Gene Annotation in Phytopathogens (http://ibi.hzau.edu.cn/digap/),

(3) Type III effector database (http://effectors.bic.nus.edu.sg/T3SEdb/index.php), (4) Type III effector prediction (http://www.sysbep.org/sieve/), (5) IMG (http://img.jgi.doe.gov/), (6) KEGG (http://www.genome.jp/kegg/), and others.

7.5 Genome-Enabled Understanding of *X. citri* Virulence and Host Specificity

Comparison of the genome of XccA with other closely related strains revealed that they share 3,183 proteins at 70 % identity (Fig. 7.3). Thus, the gene content is largely conserved within XccA, XccA^{w}, XauB, and XauC. There are more shared genes among XccA and XccA^{w} as compared to XauB and XauC confirming their separation into different clades in phylogenetic comparison (Fig. 7.2). Here, we will mainly discuss T3SS genes and T3SS effector genes shared by XccA and other closely related strains. Readers are encouraged to read the comprehensive genomic analyses published elsewhere (Jalan et al. 2011, 2013a, b).

The *hrp/hrc* genes encoding the T3SS are conserved and found in the same order in XccA, XccA^{w}, XauB, and XauC. Nineteen effectors are present in XccA, XccA^{w}, XauB, and XauC and thus represent the core effector set for xanthomonads that cause citrus canker. The effector genes *avrBs2*, *xopK*, *xopL*, *xopQ*, *xopR*, *xopX,* and *xopZ* are found in all other sequenced *Xanthomonas* genomes, and hence, the seven genes might be a core set of effectors required for phytopathogenicity as suggested by Moreira et al. (2010a).

Among all the effectors, it has been suggested that PthA or its functional homolog(s) is the major pathogenicity determinant of the citrus canker pathogen (Al-Saadi et al. 2007), which linked the strains of *Xanthomonas* with different host range together. The *avrBs3/pthA* family of effectors includes various *pth* genes, but only PthA (Swarup et al. 1992) is known to induce canker. PthA4 in XccA 306 is a functional homolog of PthA located on pXAC64. PthA4 has the same repeat number (17.5) as other

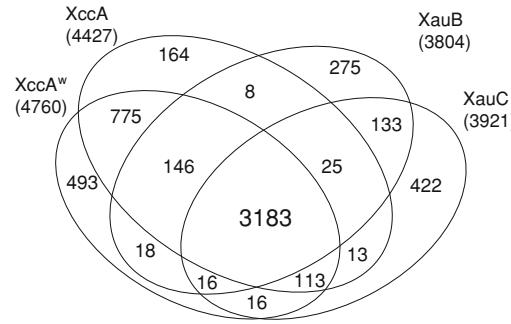

Fig. 7.3 Venn diagram represents the shared genome of *X. citri* subsp. *citri* str. 306 (XccA), *X. citri* subsp. *citri* strain A^{w}12879 (XccA^{w}), *X. fuscans* subsp. *aurantifolii* strain ICPB 11122 (XauB), and *X. fuscans* subsp. *aurantifolii* strain ICPB 10535 (XauC). Numbers in brackets represent the protein-coding genes of each species or pathotype

functional homologs PthAw, PthB, and PthC from the three respective citrus canker-causing strains XccA^{w}, XauB, and XauC (Moreira et al. 2010a). The AvrBs3/PthA family effectors are known as TAL effectors since they reprogram host cells by specifically binding to the promoters of plant genes recognized by the central domain of tandem repeats (Boch et al. 2009). Comparing the DNA-binding TAL effector codes for PthA from XccA as predicted by Boch and Bonas (2010) to other functional homologs indicates that the codes for PthA are quite divergent. This may result in recognition of different target genes in the host plant or differences in strength of induction of plant genes and thus affect virulence of different citrus canker variants.

Comparative analysis has also identified multiple strain-specific effectors that might contribute to the differences in virulence and host range. Among these are two effector genes, *avrGf1* and *xopAF,* identified in XccA^{w}, XauB, and XauC but which were not present in XccA genome (Table 7.2). The presence of these effectors in limited host range strains causing citrus canker and not in the broader host range XccA makes them prime candidates for effectors that could affect host specificity. The *AvrGf1* effector in XccA^{w} belongs to the XopAG family and has been shown to trigger a hypersensitive reaction in grapefruit, thus limiting its host range

Table 7.2 Distribution of type III effectors that vary in *X. citri* subsp. *citri* str. 306 (XccA), *X. citri* subsp. *citri* A^w 12879 (Xcaw), *X. fuscans* subsp. *aurantifolii* str. ICPB 11122 (XauB), *X. fuscans* subsp. *aurantifolii* str. ICPB 10535 (XauC), and *X. axonopodis* pv. *citrumelo* strain F1 (Xacm)

Effector class	XccA	Xcaw	XauB	XauC	Xacm	Pfam domains
PthA (AvrBs3, TAL)	XACa0022 (PthA1) XACa0039 (PthA2) XACb0015 (PthA3) XACb0065 (PthA4)	XCAW_b00018 (PthAw1) XCAW_b00026 (pthAw2)	XAUB_40130 XAUB_28490	XAUC_22430 XAUC_24060 XAUC_09900 XAUC_43080	–	Transcriptional activator, nuclear localization
XopE3 (AvrXacE2)	XCAW_03515	XAC3224	XAUB_14680	XAUC_00040		Putative transglutaminase
XopL	XAC3090	XCAW_03376	XAUB_34130	XAUC_02900/ XAUC_12488Ψ	XACM_3007	LRR protein
XopZ1	XAC2009	XCAW_01815	XAUB_11532/ XAUB_13710Ψ	XAUC_25915	XACM_2036	–
XopAI	XAC3230	XCAW_01099	XAUB_26830	XAUC_23780	–	Putative ADP—ribosyltransferase
XopAP	XAC2990	XCAW_03269	XAUB_13980	XAUC_08760	–	–
HpaA	XAC0400	XCAW_00810	XAUB_19430	XAUC_19990	–	T3S control protein
HrpW (PopW)	XAC2922	XCAW_03200	XAUB_19460	XAUC_20020	–	Pectate Lyase
XopAQ	No annotation between XAC3223 and XAC3224	XCAW_03514	No annotation between XAUB_14670 and XAUB_14680	–	–	–
XopE2 (AvrXacE3, AvrXccE1)	XACb0011	XCAW_03520	XAUB_31660	–	–	Putative transglutaminase
XopN	XAC2786	XCAW_01387	XAUB_07520	–	XACM_2728	ARM/HEAT repeat
XopP	XAC1208	XCAW_01310	XAUB_06720	–	XACM_1178	–
XopAE (HpaF/HpaG)	XAC0393	XCAW_00801	XAUB_19500	–	XACM_0381	LRR protein
XopC2	XAC1209Ψ XAC1210Ψ	XCAW_01311Ψ	–	–	XACM_1180	Haloacid dehalogenase-like hydrolase

(continued)

Table 7.2 (continued)

Effector class	XccA	Xcaw	XauB	XauC	Xacm	Pfam domains
XopAF (AvrXv3)	–	XCAW_b00003	XAUB_02310	XAUC_00300	–	–
XopAG (AvrGf1/ AvrGf2)	–	XCAW_00608	XAUB_03570 Ψ	XAUC_04910	–	–
XopF1 (Hpa4)	–	XCAW_00804/ XCAW_00805Ψ	–	XAUC_31730Ψ	XACM_0384	–
XopB	–	–	XAUB_09070/ XAUB_14842 Ψ	XAUC_00260	–	–
XopE4	–	–	XAUB_23330	XAUC_31730	–	Putative transglutaminase
XopJ1	–	–	XAUB_20830	XAUC_08850	–	C55-family cysteine protease or Ser/Thr acetyltransferase
XopAJ (AvrRxo1)	–	–	–	–	XACM_4204	Zeta toxin
XopW	–	–	–	–	XACM_0435	–

Ψ Inactive/Pseudogene

(Rybak et al. 2009). AvrGf1 from XccA^{w} shows only about 45 % identity to its homolog XAUC_04910 in XauC, whereas the homolog XAUB_03570 in XauB is interrupted by a transposon and might be non-functional. When the mutant strain XccA^{w}Δ*avrGf1* was inoculated in grapefruit, it caused typical canker-like symptoms instead of HR, but the symptoms were visibly reduced (Rybak et al. 2009). Another candidate gene, which might contribute to host specificity, is *xopAF*, which belongs to *avrXv3* family. Homologs of *xopAF* from XccA^{w} are found in XauB and XauC (XAUB_02310 and XAUC_00300) but not in XccA (Table 7.2). A XopAF homolog AvrXv3 from *X. campestris* pv. *vesicatoria* is known to induce HR in tomato line Hawaii 7981 and pepper plants (Astua-Monge et al. 2000). In our recent study, we found that *xopAF* mutant and *xopAF avrGf1* double mutant both have lower growth *in planta* as compared to XccA^{w} and *avrGf1* single mutant, respectively (Jalan et al. 2013b). Mutation of *xopAF* did not make XccA^{w} strain pathogenic in sweet orange Valencia. Instead, it slowed growth of the pathogen in Duncan grapefruit and Mexican lime, which was restored by complementation, indicating that XopAF is important for bacterial growth *in planta* but not host range (Jalan et al. 2013b).

7.6 Transcriptomics of XccA

7.6.1 Transcriptomic Analysis of XccA Using Microarray

The availability of whole genome sequences of XccA and related species has accelerated the transcriptome analyses. Microarray is one of the earliest tools and is still a convenient and affordable method to investigate the bacterial transcriptome. Since 2005, three gene array platforms of XccA have been reported. The first one is a macroarray containing 279 genes of XccA associated with pathogenicity and virulence (Astua-Monge et al. 2005). The second microarray platform consists of 2,365 genes which were selected from shotgun libraries used for genomic sequencing of XccA (Moreira et al. 2010b). This gene array covers 52.7 % of the annotated genes of the XccA genome. The third one is a whole genome DNA Agilent microarray that we developed for XccA, representing all 4,427 annotated protein-coding genes (Guo et al. 2011). This microarray platform has been used extensively in the transcriptomic studies of XccA. Here, we summarize the current knowledge gained from XccA transcriptomic research on regulatory networks controlling the virulence factors of XccA.

DSF-mediated QS has been found in many bacteria, including *X. campestris* pv. *campestris*, *X. oryzae* pv. *oryzae*, *Xylella fastidiosa*, *Stenotrophomonas maltophilia*, and *Burkholderia cenocepacia* as well as XccA (Barber et al. 1997; Boon et al. 2007; Colnaghi Simionato et al. 2007; da Silva et al. 2002; Fouhy et al. 2007; He et al. 2010). This QS pathway contains three major components: RpfF, RpfC, and RpfG, which are DSF signal producer, signal sensor, and response regulator of the two-component system, respectively. A transcriptome study on the *rpfF* mutant of *X. campestris* pv. *campestris* revealed that the DSF regulon consists of genes in 12 functional groups, such as genes involved in flagellar biosynthesis, exopolysaccharide (EPS) production, toxin and stress resistance, and genes encoding extracellular enzymes (He et al. 2006). Comparative genomic analysis showed that the sequences of the *rpf* gene cluster of XccA are highly similar to the homologs of *X. campestris* pv. *campestris*, including *rpfF*, *rpfC,* and *rpfG*. Guo et al. (2012) investigated the transcriptomic profiles of *rpfF*, *rpfC,* and *rpfG* mutants of XccA. The analysis revealed that the three regulons overlap in the majority of genes, 143 genes in the exponential phase and 63 genes in the stationary phase. Those genes are involved in multiple biological activities, such as chemotaxis and flagellar biosynthesis, stress tolerance, adhesion, transport, energy metabolism, and detoxification (Fig. 7.4). It indicates that RpfC-RpfG is the major signal transduction system of DSF-mediated QS in XccA, whose regulatory function is similar to the QS system

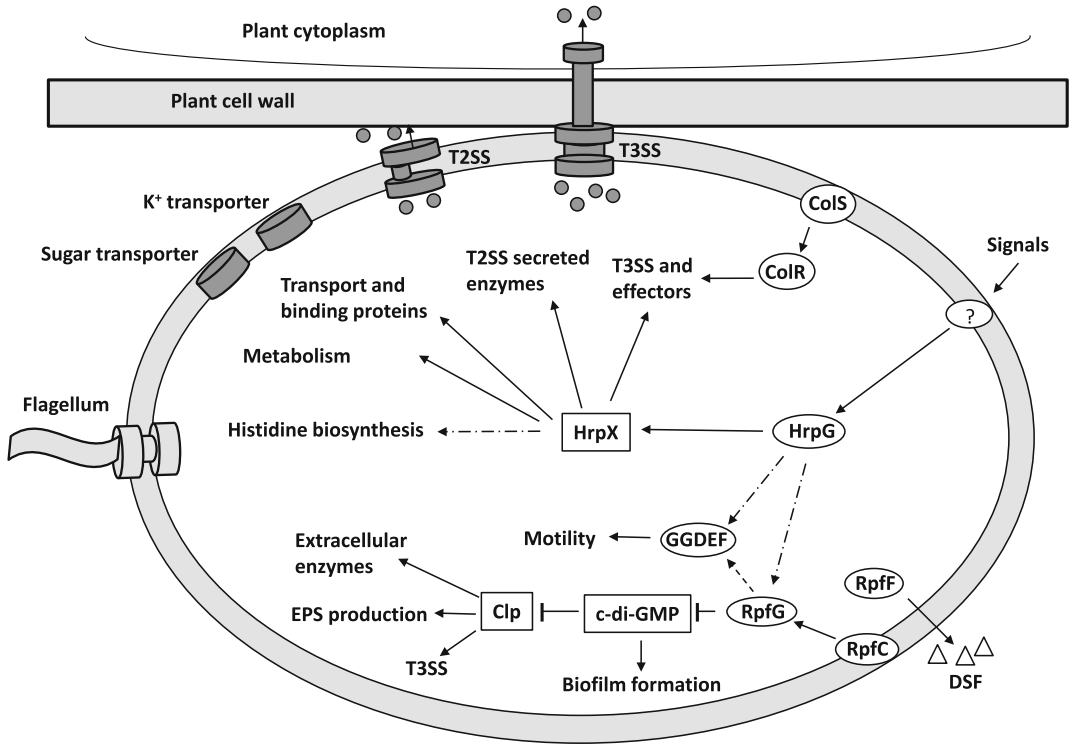

Fig. 7.4 The global regulatory network of virulence factors of *X. citri*. The signals from the environment or plant activate the unknown sensor protein which relays the signal to the regulator HrpG. The activated HrpG induces the transcription of HrpX, which in turn activates the expression of a large set of virulence genes encoding T3SS, T3SS effectors, and T2SS-secreted enzymes. HrpX also regulates the transcription of some genes related to nutrient transport and metabolism. HrpX indirectly represses the transcription of genes involved in histidine biosynthesis. A two-component system ColS/ColR also controls the expression of genes encoding the T3SS translocon. RpfF is involved in the production of DSF signal molecule, which is sensed by a two-component system RpfC/RpfG. The activated RpfG binds to two proteins with a GGDEF domain, resulting in the control of motility. RpfG also decreases the level of the second messenger c-di-GMP by its c-di-GMP phosphodiesterase activity. The low levels of c-di-GMP inhibit biofilm formation and also release the bound transcriptional activator Clp. Clp binds to the promoters of target genes and induces the expression of genes encoding T3SS and genes involved in the synthesis of extracellular enzymes and EPS. HrpG interacts with DSF-mediated quorum-sensing pathway via repressing the transcription of RpfG and two GGDEF proteins. *Arrows* represent positive regulation, the *dash-dot arrow* represents negative regulation, *dash arrows* represent physical binding, and '?' represents unknown proteins

in *X. campestris* pv. *campestris*. However, the comparison of three regulons showed that multiple genes are regulated by only one of the three *rpf* genes: 42 genes in the RpfF regulon, which represents 40 % of the number of genes in this regulon, were not controlled by RpfC and RpfG in stationary phase; 21 genes in the RpfC regulon were not affected by the mutations of *rpfF* and *rpfG* in the exponential phase; and 88 genes in the exponential phase and 156 genes in the stationary phase were controlled by RpfG only. Those unique genes in the RpfG regulon include genes encoding type V secretion system adhesins, T3SS, regulators, iron transporters, and genes involved in energy metabolism. The unique genes in each regulon suggest that the additional signal transduction pathways may be involved in DSF-mediated QS. Additionally, the RpfC-RpfG system may participate in gene regulation other than the DSF pathway. These

findings significantly broaden our perspective of the complexity of DSF-mediated QS.

It is well known that HrpG and HrpX are essential for *hrp* gene expression. HrpG is an OmpR family regulator, which was first reported to activate the expression of *hrcC* and *hrpX* in *X. campestris* pv. *campestris* (Wengelnik et al. 1996). HrpX is an AraC-type transcriptional activator regulating the expression of *hrp* genes (T3SS structural genes) and some T3SS effectors (Wengelnik and Bonas 1996). Due to the regulatory roles of HrpG and HrpX in the expression of virulence genes, a few studies have been done to identify the genes in HrpG and HrpX regulons, resulting in the extension of HrpG and HrpX regulons to genes encoding T2SS-secreted enzymes (Furutani et al. 2004; Wang et al. 2008; Yamazaki et al. 2008) and genes containing a plant-inducible promoter (PIP) box (Fenselau and Bonas 1995; da Silva et al. 2002; Noel et al. 2001). Guo et al. (2011) used an XccA whole genome DNA microarray to explore the transcriptome of XccA and characterized the HrpG and HrpX regulons. The comprehensive study provides us with an overall view of the genes in the HrpG and HrpX regulons, by expanding the HrpG regulon to 232 genes and the HrpX regulon to 181 genes (Guo et al. 2011). All of the 24 *hrp* genes encoding T3SS proteins are upregulated by HrpG and HrpX. Twenty-three putative and known T3SS effectors are also controlled by the two regulators. Besides the previously reported eight genes encoding T2SS-secreted enzymes, 21 additional HrpG- and HrpX-regulated genes were identified encoding putative T2SS enzymes. HrpG and HrpX regulons also include genes involved in amino acid biosynthesis, oxidative phosphorylation, pentose phosphate pathway, transport of sugar, iron and potassium, and phenolic catabolism (Fig. 7.4). Several genes encoding a two-component system and transcriptional regulators are also found to be controlled by HrpG and HrpX. For instance, *pccQ* and XAC1455 encoding transcription regulators were regulated by both HrpG and HrpX. It was suggested that those genes are utilized as downstream regulators in the signaling pathways of HrpG and HrpX. Notably, HrpG alone regulates a number of genes involved in flagellar biosynthesis and chemotaxis and also influences the expression of *rpfG* and two GGDEF family genes (Fig. 7.4), which are involved in DSF-mediated QS. This indicates the possibility of the cross talk between the HrpG regulatory cascade and QS. Overall, the findings of the transcriptomics study revealed that HrpG and HrpX participate in a global signaling network and regulate the expression of multiple virulence genes to modify and adapt to the host environment.

7.6.2 Transcriptomic Sequencing Using RNA-Seq

RNA-seq has been conducted to obtain a comprehensive picture of the HrpX regulon of XccA cultured in XVM2 medium (Kogenaru et al. 2012). RNA-seq is a whole transcriptome profiling method that can measure the RNA levels using a massive parallel deep-sequencing-based approach (Wang et al. 2009). RNA-seq provides qualitative identification as well as quantitative level of expression of genes in a given condition (Wang et al. 2009; Marguerat and Bähler 2010). Various platforms, such as Roche 454 sequencer, Illumina Genome Analyzer, Illumina HiSeq, and Applied Biosystems SOLiD sequencer, are currently available to obtain RNA-seq data. In our study, RNA samples of XccA that passed the quality control were sequenced using the Illumina Genome Analyzer IIx (GAIIx) system (Kogenaru et al. 2012). A total of 4,323 genes were found to have at least one read mapped, constituting >90 % of the reads being mapped onto the reference genome, indicating good sequence coverage. Overall, ~97 % of the annotated genes had more than one read mapped, while merely ~3 % of the annotated genes had no reads mapped, indicating good sequencing depth (Kogenaru et al. 2012). It was also observed that annotated coding genes from the chromosome with a size of 5.18 Mb had 98 % sequence coverage, while it was observed to be

78 % for plasmid pXAC64 with a size of 64 kb and even relatively lower with only 62 % sequence coverage for the second plasmid pXAC33 with a size of 33 kb. The read count of uniquely mapped genes to each annotated gene from both the wild-type as well as *hrpX* mutant for all the biological replicates was used to estimate the genes, which are differentially expressed using DESeq, an open-source package available under the Bioconductor suite under R environment (Anders and Huber 2010). DESeq estimates the variance in RNA-seq data across replicates and conditions and conducts tests for differential expression using negative binomial distribution (Anders and Huber 2010).

Gene expression levels quantified by RNA-seq were compared with microarray (Kogenaru et al. 2012; Guo et al. 2011). The RNA-seq and microarray highly correlated both at absolute as well as relative levels (Spearman correlation-coefficient, $r_s > 0.76$). Further, the expression levels quantified by RNA-seq and microarray for the significantly differentially expressed genes (DEGs) also were highly correlated with qRT-PCR-based quantification ($r_s = 0.58$ to 0.94). In addition to the 55 newly identified DEGs, 72 % of the already known HrpX target genes were detected by both RNA-seq and microarray, while the remaining 28 % could only be detected by either one of the methods.

Eighty-seven statistically significant differentially expressed genes were obtained by applying a cutoff threshold of FDR ≤ 0.05 (5 %) and an absolute $\log_2$ fold-change ≥ 0.6 (Kogenaru et al. 2012). Among them, 21 (25 %) were upregulated while 66 (75 %) were downregulated in the *hrpX* mutant. Among the *hrp* genes, 16 genes, *hrcN*, *hrcJ, hrcU*, *hrcV*, *hrpB1*, *hrpB2*, *hrpB4*, *hrcQ*, *hrcR*, *hrpD5*, *hrpD6*, *hrpE*, *hrpF*, *hpa1*, *hpaB,* and *hpaF,* were detected as differentially expressed with a $\log_2$FC ranging between -0.92 to -2.23, while nine genes from the cluster, *hrpB5*, *hrcS*, *hpaP*, XAC0395, *hrcC*, *hrpB7*, *hpa2*, *hpaA*, and *hrcT,* were not detected as differentially expressed by RNA-seq (Kogenaru et al. 2012). Among the effector genes, 11 genes, *xopE*, *avrXacE3*, *xopI*, *xopK*, *xopN*, *xopP*, *xopR*, *xopX*, *xopAI*, *hrpW*, and *xopAD,* were found to be regulated by HrpX with a $\log_2$FC range varying between -0.8816 to -3.5117, while 11 genes, *xopQ*, *xopV*, *avrBs2*, *xopL*, *xopE*, *xopZ*, *xopAK*, *pthA1*, *pthA2*, *pthA3*, and *pthA4,* were not detected as differentially expressed at the given time point (Kogenaru et al. 2012). Pseudo- or inactive genes XAC2785, XAC1210, and XAC1209 were not included in the RNA-seq analysis. Seventy-five percent of the known T3SS and effectors were detected as differentially expressed at the given time point by RNA-seq. Among the differentially expressed genes newly identified by the transciptome study is gene XAC0755 that is found to be downregulated. This gene encodes KdpF, a component of an integral membrane potassium-transporting system. Four genes, XAC4116, XAC1819, XAC3026, and XAC3363, were found to be involved in signal transduction and gene regulation. Further, the gene products of seven differentially expressed genes, XAC3275, XAC3680, XAC1943, XAC0527, XAC0599, XAC0239, and XAC0755, were predicted to be T2SS substrates. Fifty novel targets have been identified from the RNA-seq transcriptome profiling: they are XAC0239, *sflA*, XAC0335, *metE*, XAC0392, XAC0527, XAC0599, *lacZ*, XAC0755, *kdpA*, XAC0854, *rpoA*, XAC1061, XAC1163, XAC1203, *rpsT*, *cyoD*, *scoF*, *araJ*, *tpmT*, XAC1629, XAC1715, XAC1816, *tspO*, XAC1943, *fliO*, *rpmJ, rpmS*,XAC3026, XAC3275, *blaI,* XAC3446, XAC3680, *putA*, *yhhT,* XAC4085, XAC4116, XAC4131, *appA,* XAC4149, XACb0064, *pcaH*, XAC0315, XAC2787, XAC3646, XAC2654, XAC1172, *btuB*, XAC3445, and *stkXac1*. Sixty-eight percent of these novel targets were found to be hypothetical, which might have potential roles in virulence, while the remaining 32 % novel targets were involved in signaling and transcription regulation (Kogenaru et al. 2012). RNA-seq transcriptome profiling has significantly enhanced the understanding of the HrpX regulome.

7.7 Concluding Remarks

The advent of high-throughput sequencing technologies, for example, Roche-454 Life Sciences, Solexa-Illumina, ABI-SOLiD, and Pacific Bio-Single-molecule real-time sequencing and ever-increasing computing capacity render it possible to sequence multiple strains of the same bacterial species. It has been realized that a significant percentage of each genome sequence was specific to each individual strain, and therefore, each new genome sequenced provided a number of new genes not previously characterized. Thus, pan-genome analysis has become of increasing interest. The pan-genome includes the 'core genome' of genes common to all strains of the species and the 'dispensable or accessory genome', which consists of genes present in at least one, but not all strains of a species (Medini et al. 2005). The core genome is related to its fundamental biological processes and derived traits from a common ancestor, whereas the dispensable genome contributes to the species' diversity and probably provides functions that are not essential to its basic lifestyle but confer selective advantages including niche adaptation, antibiotic resistance, and the ability to colonize new hosts (Tettelin et al. 2008). *X. citri* is a pathogen with wide geographical distribution and contains multiple strains with different host ranges (Ngoc et al. 2010). Thus, a comprehensive sequencing of strains of *X. citri* will help with pan-genome analysis and understand the mechanisms for different strains with distinct traits including differences in host range and resistance to copper. Pathogens always undergo evolution to adapt to ever-changing environments. Genomic studies using high-throughput sequencing will facilitate an ability to understand the mechanism(s) of *X. citri* evolution in the generation of super pathogens that can overcome plant defense of resistant varieties or become antibiotic or copper resistant. Consequently, we could prevent the generation of the super pathogens of *X. citri*. Using both high-throughput DNA sequencing and RNA-seq, we can better understand the bacterial genome biology and novel transcripts. RNA-seq makes it easier to study the regulation of *X. citri* under different environments and the regulons of different regulators for a better understanding of the pathogen. Importantly, RNA-seq makes it easier to simultaneously investigate the transcriptomes of both the host and pathogen during interactions, which in turn will help us understand the virulence mechanism of *X. citri* and the resistance mechanism of the host to design innovative strategies to control citrus canker and other bacterial diseases.

References

Al-Saadi A, Reddy JD, Duan YP, Brunings AM, Yuan Q, Gabriel DW (2007) All five host-range variants of *Xanthomonas citri* carry one *pthA* homolog with 17.5 repeats that determines pathogenicity on citrus, but none determine host-range variation. Mol Plant Microbe Interact 20:934–943

Alegria MC, Docena C, Khater L, Ramos CHI, da Silva ACR, Farah CS (2004) New protein-protein interactions identified for the regulatory and structural components and substrates of the type III secretion system of the phytopathogen *Xanthomonas axonopodis* pathovar *citri*. J Bacteriol 186:6186–6197

Anders S, Huber W (2010) Differential expression analysis for sequence count data. Genome Biol 11:R106

Andrade MO, Alegria MC, Guzzo CR, Docena C, Pareda Rosa MC, Ramos CHI, Farah CS (2006) The HD-GYP domain of RpfG mediates a direct linkage between the Rpf quorum-sensing pathway and a subset of diguanylate cyclase proteins in the phytopathogen *Xanthomonas axonopodis* pv *citri*. Mol Microbiol 62:537–551

Aslam SN, Newman M-A, Erbs G, Morrissey KL, Chinchilla D, Boller T, Jensen TT, De Castro C, Ierano T, Molinaro A (2008) Bacterial polysaccharides suppress induced innate immunity by calcium chelation. Curr Biol 18:1078–1083

Astua-Monge G, Freitas-Astua J, Bacocina G, Roncoletta J, Carvalho SA, Machado MA (2005) Expression profiling of virulence and pathogenicity genes of *Xanthomonas axonopodis* pv. *citri*. J Bacteriol 187:1201–1205

Astua-Monge G, Minsavage GV, Stall RE, Davis MJ, Bonas U, Jones JB (2000) Resistance of tomato and pepper to T3 strains of *Xanthomonas campestris* pv. *vesicatoria* is specified by a plant-inducible avirulence gene. Mol Plant Microbe Interact 13:911–921

Barber C, Tang J, Feng J, Pan M, Wilson T, Slater H, Dow J, Williams P, Daniels M (1997) A novel

regulatory system required for pathogenicity of *Xanthomonas campestris* is mediated by a small diffusible signal molecule. Mol Microbiol 24:555–566

Behlau F, Belasque J, Bergamin Filho A, Graham J, Leite R, Gottwald T (2008) Copper sprays and windbreaks for control of citrus canker on young orange trees in southern Brazil. Crop Protect 27:807–813

Behlau F, Belasque J, Graham J, Leite R (2010) Effect of frequency of copper applications on control of citrus canker and the yield of young bearing sweet orange trees. Crop Protect 29:300–305

Boch J, Bonas U (2010) *Xanthomonas* AvrBs3 family-type III effectors: discovery and function. Annu Rev Phytopathol 48:419–436

Boch J, Scholze H, Schornack S, Landgraf A, Hahn S, Kay S, Lahaye T, Nickstadt A, Bonas U (2009) Breaking the code of DNA binding specificity of TAL-type III effectors. Science 326:1509–1512

Boon C, Deng Y, Wang L-H, He Y, Xu J-L, Fan Y, Pan SQ, Zhang L-H (2007) A novel DSF-like signal from *Burkholderia cenocepacia* interferes with Candida albicans morphological transition. ISME J 2:27–36

Brunings A, Gabriel D (2003) *Xanthomonas citri*: breaking the surface. Mol Plant Pathol 4:141–157

Büttner D, He SY (2009) Type III protein secretion in plant pathogenic bacteria. Plant Physiol 150:1656–1664

Canteros B (2004) Management of citrus canker in Argentina. A review. In: Proceedings of international society citriculture

Casabuono A, Petrocelli S, Ottado J, Orellano EG, Couto AS (2011) Structural analysis and involvement in plant innate immunity of *Xanthomonas axonopodis* pv. *citri* lipopolysaccharide. J Biol Chem 286:25628–25643

Chatterjee S, Wistrom C, Lindow SE (2008) A cell–cell signaling sensor is required for virulence and insect transmission of *Xylella fastidiosa*. Proc Natl Acad Sci USA 105:2670–2675

Civerolo EL (1984) Bacterial canker disease of citrus. J Rio Grande Valley Hort Assoc 37:127–146

Colnaghi Simionato AV, da Silva DS, Lambais MR, Carrilho E (2007) Characterization of a putative *Xylella fastidiosa* diffusible signal factor by HRGC-EI-MS. J Mass Spectrom 42:490–496

Cubero J, Graham JH (2002) Genetic relationship among worldwide strains of *Xanthomonas* causing canker in citrus species and design of new primers for their identification by PCR. Appl Environ Microbiol 68:1257–1264

Cuthbertson L, Kimber M, Whitfield C (2007) Substrate binding by a bacterial ABC transporter involved in polysaccharide export. Proc Natl Acad Sci USA 104:19529–19534

da Silva A, Ferro J, Reinach F, Farah C, Furlan L, Quaggio R, Monteiro-Vitorello C, Van Sluys M, Almeida N, Alves L (2002) Comparison of the genomes of two *Xanthomonas* pathogens with differing host specificities. Nature 417:459–463

Danhorn T, Fuqua C (2007) Biofilm formation by plant-associated bacteria. Annu Rev Microbiol 61:401–422

Das A (2003) Citrus canker-A review. J Appl Hort 5:52–60

Daubin V, Ochman H (2004) Bacterial genomes as new gene homes: the genealogy of ORFans in *E. coli*. Genome Res 14:1036–1042

de Kievit T, Iglewski B (2000) Bacterial quorum sensing in pathogenic relationships. Infect Immun 68:4839

Dekkers LC, Bloemendaal CJP, de Weger LA, Wijffelman CA, Spaink HP, Lugtenberg BJJ (1998) A two-component system plays an important role in the root-colonizing ability of *Pseudomonas fluorescens* strain WCS365. Mol Plant Microbe Interact 11:45–56

Dopson R (1964) The eradication of citrus canker. Plant Dis Reporter 48:30–31

Dow J, Osbourn A, Wilson T, Daniels M (1995) A locus determining pathogenicity of *Xanthomonas campestris* is involved in lipopolysaccharide biosynthesis. Mol Plant Microbe Interact 8:768–777

Dow M (2008) Diversification of the function of cell-to-cell signaling in regulation of virulence within plant pathogenic xanthomonads. Sci Signal 1:pe23. doi:10.1126/stke.121pe23

Duan Y, Castaneda A, Zhao G, Erdos G, Gabriel D (1999) Expression of a single, host-specific, bacterial pathogenicity gene in plant cells elicits division, enlargement, and cell death. Mol Plant Microbe Interact 12:556–560

El Yacoubi B, Brunings A, Yuan Q, Shankar S, Gabriel D (2007) *In planta* horizontal transfer of a major pathogenicity effector gene. Appl Environ Microbiol 73:1612–1621

Fenselau S, Bonas U (1995) Sequence and expression analysis of the *hrpB* pathogenicity operon of *Xanthomonas campestris* pv. *vesicatoria* which encodes eight proteins with similarity to components of the Hrp, Ysc, Spa, and Fli secretion systems. Mol Plant Microbe Interact 8:845–854

Figurski DH, Helinski DR (1979) Replication of an origin-containing derivative of plasmid RK2 dependent on a plasmid function provided *in trans*. Proc Natl Acad Sci USA 76:1648–1652

Fouhy Y, Scanlon K, Schouest K, Spillane C, Crossman L, Avison MB, Ryan RP, Dow JM (2007) Diffusible signal factor-dependent cell-cell signaling and virulence in the nosocomial pathogen *Stenotrophomonas maltophilia*. J Bacteriol 189:4964–4968

Fujikawa T, Ishihara H, Leach JE, Tsuyumu S (2006) Suppression of defense response in plants by the *avrBs3/pthA* gene family of *Xanthomonas* spp. Mol Plant Microbe Interact 19:342–349

Furutani A, Tsuge S, Ohnishi K, Hikichi Y, Oku T, Tsuno K, Inoue Y, Ochiai H, Kaku H, Kubo Y (2004) Evidence for HrpXo-dependent expression of type II secretory proteins in *Xanthomonas oryzae* pv. *oryzae*. J Bacteriol 186:1374–1380

Gabriel D, Kingsley M, Hunter J, Gottwald T (1989) Reinstatement of *Xanthomonas citri* (ex Hasse) and *X. phaseoli* (ex Smith) to species and reclassification

of all *X. campestris* pv. *citri* strains. Int J Syst Bacteriol 39:14–22

González J, Keshavan N (2006) Messing with bacterial quorum sensing. Microbiol Mol Biol Rev 70:859–875

Gottwald T, Graham J, Schubert T (1997a) Citrus canker in urban Miami: an analysis of spread and prognosis for the future. Citrus Ind 78:72–78

Gottwald T, Graham J, Schubert T (1997b) An epidemiological analysis of the spread of citrus canker in urban Miami, Florida, and synergistic interaction with the Asian citrus leafminer. Fruits 52:383–390

Gottwald T, Timmer L (1995) The efficacy of windbreaks in reducing the spread of citrus canker caused by *Xanthomonas campestris* pv. *citri*. Trop. Agric. 72:194–201

Gottwald TR, Graham JH, Schubert TS (2002) Citrus canker: the pathogen and its impact. Plant Health Prog. doi: 10.1094/PHP-2002-0812-1001-RV. http://www.plantmanagementnetwork.org/php/

Gottwald TR, Hughes G, Graham JH, Sun X, Riley T (2001) The citrus canker epidemic in Florida: the scientific basis of regulatory eradication policy for an invasive species. Phytopathology 91:30–34

Gottwald TR, Irey M (2007) Post-hurricane analysis of citrus canker II: Predictive model estimation of disease spread and area potentially impacted by various eradication protocols following catastrophic weather events. Plant, Health Progress doi 10

Graham J, Gottwald T, Civerolo E, McGuire R (1989) Population dynamics and survival of *Xanthomonas campestris* in soil in citrus nurseries in Maryland and Argentina. Plant Dis 43:423–427

Graham J, Gottwald T, Cubero J, Achor D (2004) *Xanthomonas axonopodis* pv. *citri*: factors affecting successful eradication of citrus canker. Mol. Plant Pathol. 5:1–15

Guo Y, Figueiredo JF, Jones J, Wang N (2011) HrpG, together with HrpX, play global roles in coordinating different virulence traits of *Xanthomonas axonopodis* pv. *citri*. Mol Plant Microbe Interact 24:649–661

Guo Y, Sagaram US, Kim J-S, Wang N (2010) Requirement of the *galU* gene for polysaccharide production by and pathogenicity and growth *in planta* of *Xanthomonas citri* subsp. *citri*. Appl Environ Microbiol 76:2234–2242

Guo Y, Zhang Y, Li JL, Wang N (2012) Diffusible signal factor-mediated quorum sensing plays a central role in coordinating gene expression of *Xanthomonas citri* subsp. *citri*. Mol Plant Microbe Interact 25:165–179

Hartung J (1992) Plasmid-based hybridization probes for detection and identification of *Xanthomonas campestris* pv. *citri*. Plant Dis 76:889–893

Hasse CH (1915) *Pseudomonas citri*, the cause of citrus canker. J. Agric. Res. 4:97–100

He Y-W, Wu JE, Cha J-S, Zhang L-H (2010) Rice bacterial blight pathogen *Xanthomonas oryzae* pv. *oryzae* produces multiple DSF-family signals in regulation of virulence factor production. BMC Microbiol 10:187

He Y, Xu M, Lin K, Ng Y, Wen C, Wang L, Liu Z, Zhang H, Dong Y, Dow J (2006) Genome scale analysis of diffusible signal factor regulon in *Xanthomonas campestris* pv. *campestris*: identification of novel cell-cell communication-dependent genes and functions. Mol Microbiol 59:610–622

He YW, Ng AYJ, Xu M, Lin K, Wang LH, Dong YH, Zhang LH (2007) *Xanthomonas campestris* cell–cell communication involves a putative nucleotide receptor protein Clp and a hierarchical signalling network. Mol Microbiol 64:281–292

Hu N, Zhao B (2007) Key genes involved in heavy-metal resistance in *Pseudomonas putida* CD2. FEMS Microbiol Lett 267:17–22

Hu Y, Zhang J, White FF, Wang N, Jones JB (2012) TAL effector PthA4-mediated virulence and host gene induction in citrus canker. Phytopathology 102(S4):55

Huguet E, Hahn K, Wengelnik K, Bonas U (1998) *hpaA* mutants of *Xanthomonas campestris* pv. *vesicatoria* are affected in pathogenicity but retain the ability to induce host-specific hypersensitive reaction. Mol Microbiol 29:1379–1390

Jalan N, Aritua V, Kumar D, Yu F, Jones JB, Graham JH, Setubal JC, Wang N (2011) Comparative genomic analysis of *Xanthomonas axonopodis* pv. *citrumelo* F1, which causes citrus bacterial spot disease, and related strains provides insights into virulence and host specificity. J Bacteriol 193:6342–6357

Jalan N, Kumar D, Yu F, Jones JB, Graham JH, Wang N (2013) Complete genome sequence of *Xanthomonas citri* subsp. *citri* strain AW12879, a restricted-host-range citrus canker causing bacterium. Genome Announcements 1(3):e00235-13

Jalan N, Kumar D, Andrade MO, Yu F, Jones JB, Graham JH, White FF, Setubal JC, Wang N (2013b) Comparative genomic and transcriptomic analyses of pathotypes of *Xanthomonas citri* subsp. *citri* provide insights into mechanisms of bacterial virulence and host range. BMC Genom 14:551

Jansson P, Kenne L, Lindberg B (1975) Structure of the extracellular polysaccharide from *Xanthomonas campestris*. Carbohydr Res 45:275–282

Jenkins AE (1933) Records of Citrus scab mainly from herbarium specimens of the genus *Citrus* in England and the United States. Phytopathology 23:475–482

Kay S, Hahn S, Marois E, Hause G, Bonas U (2007) A bacterial effector acts as a plant transcription factor and induces a cell size regulator. Science 318:648

Keller L, Surette M (2006) Communication in bacteria: an ecological and evolutionary perspective. Nat Rev Microbiol 4:249–258

Kivistik PA (2010) ColR-ColS signalling system and transposition of Tn4652 in the adaptation of *Pseudomonas putida*. http://hdl.handle.net/10062/16165

Kivistik PA, Putrins M, Puvi K, Ilves H, Kivisaar M, Horak R (2006) The ColRS two-component system regulates membrane functions and protects *Pseudomonas putida* against phenol. J Bacteriol 188:8109–8117

Kogenaru S, Yan QY, Guo Y, Wang N (2012) RNA-seq and microarray complement each other in transcriptome profiling. BMC Genom 13:629

Ladaniya MS (2008) Citrus fruit: biology, technology and evaluation. Academic Press, San Diego

Laia ML, Moreira LM, Dezajacomo J, Brigati JB, Ferreira CB, Ferro MI, Silva AC, Ferro JA, Oliveira JC (2009) New genes of *Xanthomonas citri* subsp. *citri* involved in pathogenesis and adaptation revealed by a transposon-based mutant library. BMC Microbiol 9:12

Leduc JL, Roberts GP (2009) Cyclic di-GMP allosterically inhibits the CRP-like protein (Clp) of *Xanthomonas axonopodis* pv. *citri*. J Bacteriol 191:7121–7122

Lee HA (1918) Further data on the susceptibility of rutaceous plants to citrus-canker. J Agric Res 15:661–665

Leigh JA, Coplin DL (1992) Exopolysaccharides in plant-bacterial interactions. Annu Rev Microbiol 46:307–346

Leite R, Mohan S (1990) Integrated management of the citrus bacterial canker disease caused by *Xanthomonas campestris* pv. *citri* in the State of Paraná. Brazil Crop Protect 9:3–7

Lenz DH, Mok KC, Lilley BN, Kulkarni RV, Wingreen NS, Bassler BL (2004) The small RNA chaperone Hfq and multiple small RNAs control quorum sensing in *Vibrio harveyi* and *Vibrio cholerae*. Cell 118:69–82

Li J, Wang N (2011a) Genome-wide mutagenesis of *Xanthomonas axonopodis* pv. *citri* reveals novel genetic determinants and regulation mechanisms of biofilm formation. PLoS ONE 6:e21804

Li J, Wang N (2011b) The *wxacO* gene of *Xanthomonas citri* subsp. *citri* encodes a protein with a role in lipopolysaccharide biosynthesis, biofilm formation, stress tolerance and virulence. Mol Plant Pathol 12:381–396

Lima WC, Menck CF (2008) Replacement of the arginine biosynthesis operon in Xanthomonadales by lateral gene transfer. J Mol Evol 66:266–275

Lima WC, Paquola A, Varani AM, Van Sluys MA, Menck CF (2008) Laterally transferred genomic islands in Xanthomonadales related to pathogenicity and primary metabolism. FEMS Microbiol Lett 281:87–97

Lima WC, Sluys M-AV, Menck CF (2005) Non-gamma-proteobacteria gene islands contribute to the *Xanthomonas* genome. *OMICS*. J Integr Biol 9:160–172

Marguerat S, Bähler J (2010) RNA-seq: from technology to biology. Cell Mol Life Sci 67:569–579

Marois E, Van den Ackerveken G, Bonas U (2002) The *Xanthomonas* type III effector protein AvrBs3 modulates plant gene expression and induces cell hypertrophy in the susceptible host. Mol Plant Microbe Interact 15:637–646

Medini D, Donati C, Tettelin H, Masignani V, Rappuoli R (2005) The microbial pan-genome. Curr Opin Genet Dev 15:589–594

Miller MB, Bassler BL (2001) Quorum sensing in bacteria. Annu Rev Microbiol 55:165–199

Monteiro-Vitorello CB, De Oliveira MC, Zerillo MM, Varani AM, Civerolo E, Sluys M-AV (2005) *Xylella* and *Xanthomonas* mobil'omics. *OMICS*. J Integr Biol 9:146–159

Moreira L, Almeida N, Potnis N, Digiampietri L, Adi S, Bortolossi J, da Silva A, da Silva A, de Moraes F, de Oliveira J (2010a) Novel insights into the genomic basis of citrus canker based on the genome sequences of two strains of *Xanthomonas fuscans* subsp. *aurantifolii*. BMC Genom 11:238

Moreira L, de Laia M, de Souza R, Zaini P, da Silva A, Ferro J (2010b) Development and validation of a *Xanthomonas axonopodis* pv. *citri* DNA microarray platform (XACarray) generated from the shotgun libraries previously used in the sequencing of this bacterial genome. BMC Res Notes 3:150

Muraro RP, Roka FM, Spreen TH (2001) An overview of Argentina's citrus canker control program. University of Florida Cooperative Extension Service, Institute of Food and Agricultural Sciences, EDIS

Nealson K, Hastings J (1979) Bacterial bioluminescence: its control and ecological significance. Microbiol Rev 43:496

Newman MA, Dow JM, Molinaro A, Parrilli M (2007) Invited review: priming, induction and modulation of plant defence responses by bacterial lipopolysaccharides. J Endotoxin Res 13:69

Ngoc LBT, Vernière C, Jouen E, Ah-You N, Lefeuvre P, Chiroleu F, Gagnevin L, Pruvost O (2010) Amplified fragment length polymorphism and multilocus sequence analysis-based genotypic relatedness among pathogenic variants of *Xanthomonas citri* pv. *citri* and *Xanthomonas campestris* pv. *bilvae*. Int J Syst Evol Microbiol 60:515–525

Ninfa AJ, Atkinson MR, Kamberov E, Feng J, Ninfa E (1995) Control of nitrogen assimilation by the NRI-NRII two-component system of enteric bacteria. Twocompon Signal Transduction 67–88 (American Society for Microbiology, Washington, DC)

Noel L, Thieme F, Nennstiel D, Bonas U (2001) cDNA-AFLP analysis unravels a genome-wide hrpG-regulon in the plant pathogen *Xanthomonas campestris* pv. *vesicatoria*. Mol Microbiol 41:1271–1281

Ochman H, Lawrence JG, Groisman EA (2000) Lateral gene transfer and the nature of bacterial innovation. Nature 405:299–304

Papo N, Shai Y (2005) A molecular mechanism for lipopolysaccharide protection of Gram-negative bacteria from antimicrobial peptides. J Biol Chem 280:10378

Patil P, Sonti R (2004) Variation suggestive of horizontal gene transfer at a lipopolysaccharide(lps) biosynthetic locus in *Xanthomonas oryzae* pv. *oryzae*, the bacterial leaf blight pathogen of rice. BMC Microbiol 4:40

Patil PB, Bogdanove AJ, Sonti RV (2007) The role of horizontal transfer in the evolution of a highly variable lipopolysaccharide biosynthesis locus in xanthomonads that infect rice, citrus and crucifers. BMC Evol Biol 7:243

Paul R, Weiser S, Amiot NC, Chan C, Schirmer T, Giese B, Jenal U (2004) Cell cycle-dependent dynamic localization of a bacterial response regulator with a novel di-guanylate cyclase output domain. Genes Dev 18:715–727

Qian W, Han Z-J, He C (2008) Two-component signal transduction systems of *Xanthomonas* spp.: a lesson from genomics. Mol Plant Microbe Interact 21:151–161

Raetz CRH, Whitfield C (2002) Lipopolysaccharide endotoxins. Annu Rev Biochem 71:635–700

Rhoads AS, DeBusk E (1931) Diseases of citrus in Florida. University of Florida Agricultural Experiment Station

Rigano L, Siciliano F, Enrique R, Sendín L, Filippone P, Torres P, Qüesta J, Dow J, Castagnaro A, Vojnov A (2007) Biofilm formation, epiphytic fitness, and canker development in *Xanthomonas axonopodis* pv. *citri*. Mol Plant Microbe Interact 20:1222–1230

Rita Hõrak HI, Pruunsild Priit, Kuljus Martin, Kivisaar Maia (2004) The ColR–ColS two-component signal transduction system is involved in regulation of Tn4652 transposition in *Pseudomonas putida* under starvation conditions. Mol Microbiol 54:795–807

Ryan R, McCarthy Y, Andrade M, Farah C, Armitage J, Dow J (2010) Cell–cell signal-dependent dynamic interactions between HD-GYP and GGDEF domain proteins mediate virulence in *Xanthomonas campestris*. Proc Natl Acad Sci USA 107:5989–5994

Rybak M, Minsavage G, Stall R, Jones J (2009) Identification of *Xanthomonas citri* ssp. *citri* host specificity genes in a heterologous expression host. Mol Plant Pathol 10:249–262

Schoulties CL, Miller JW (1985) A new outbreak of citrus canker in Florida. Plant Dis 69:1

Schubert T, Sun X (1996) Bacterial citrus canker. Plant Pathol Circular 377:110–111 (Florida Department of Agriculture and Consumer Services DoPI, Gainesville, FL)

Schubert TS, Rizvi SA, Sun X, Gottwald TR, Graham JH, Dixon WN (2001) Meeting the challenge of eradicating citrus canker in Florida-Again. Plant Dis 85:340–356

Semenova E, Nagornykh M, Pyatnitskiy M, Artamonova II, Severinov K (2009) Analysis of CRISPR system function in plant pathogen *Xanthomonas oryzae*. FEMS Microbiol Lett 296:110–116

Slater H, Alvarez-Morales A, Barber CE, Daniels MJ, Dow JM (2000) A two-component system involving an HD-GYP domain protein links cell–cell signalling to pathogenicity gene expression in *Xanthomonas campestris*. Mol Microbiol 38:986–1003

Stall R, Civerolo E (1991) Research relating to the recent outbreak of citrus canker in Florida. Annu Rev Phytopathol 29:399–420

Stall RE, Miller J, Marco GM, Echenique BIC (1980) Population dynamics of *Xanthomonas citri* causing cancrosis of citrus in Argentina. P Fl St Hortic Soc 93:10–14

Stall RE, Seymour CP (1983) Canker, a threat to citrus in the Gulf-Coast states. Plant Dis 67:581–585

Stock AM, Robinson VL, Goudreau PN (2000) Two-component signal transduction. Annu Rev Biochem 69:183–215

Subramoni S, Jha G, Sonti R (2006) Virulence functions of xanthomonads. Plant Assoc Bact 535–571

Sun M (1984) The mystery of Florida's citrus canker: scientists haven't seen this bacterial strain before, but it's already led to the destruction of millions of seedlings. Science 226:322–323

Sun X, Stall RE, Jones JB, Cubero J, Gottwald TR, Graham JH, Dixon WN, Schubert TS, Chaloux PH, Stromberg VK, Lacy GH, Sutton BD (2004) Detection and characterization of a new strain of citrus canker bacteria from Key/Mexican lime and alemow in South Florida. Plant Dis 88:1179–1188

Swarup S, De Feyter R, Brlansky R, Gabriel D (1991) A pathogenicity locus from *Xanthomonas citri* enables strains from several pathovars of *X. campestris* to elicit cankerlike lesions on citrus. Phytopathology 81:802–809

Swarup S, Yang Y, Kingsley MT, Gabriel DW (1992) An *Xanthomonas citri* pathogenicity gene, *pthA*, pleiotropically encodes gratuitous avirulence on nonhosts. Mol Plant Microbe Interact 5:204–213

Szurek B, Marois E, Bonas U, Ackerveken GVD (2001) Eukaryotic features of the *Xanthomonas* type III effector AvrBs3: protein domains involved in transcriptional activation and the interaction with nuclear import receptors from pepper. Plant J 26:523–534

Tal R, Wong HC, Calhoon R, Gelfand D, Fear AL, Volman G, Mayer R, Ross P, Amikam D, Weinhouse H (1998) Three *cdg* operons control cellular turnover of cyclic di-GMP in *Acetobacter xylinum*: genetic organization and occurrence of conserved domains in isoenzymes. J Bacteriol 180:4416–4425

Tettelin H, Riley D, Cattuto C, Medini D (2008) Comparative genomics: the bacterial pan-genome. Curr Opin Microbiol 11:472–477

USDA (2006) USDA determines citrus canker eradication not feasible. Department Press Release http://nassau.ifas.ufl.edu/news/citruscanker.html

Van Sluys M, Monteiro-Vitorello C, Camargo L, Menck C, Da Silva A, Ferro J, Oliveira M, Setubal J, Kitajima J, Simpson A (2002) Comparative genomic analysis of plant-associated bacteria. Annu Rev Phytopathol 40:169–189

Vernière C, Hartung JS, Pruvost OP, Civerolo EL, Alvarez AM, Maestri P, Luisetti J (1998) Characterization of phenotypically distinct strains of *Xanthomonas axonopodis* pv. *citri* from Southwest Asia. Eur J Plant Pathol 104:11

Vorholter F, Niehaus K, Pühler A (2001) Lipopolysaccharide biosynthesis in *Xanthomonas campestris* pv. *campestris*: a cluster of 15 genes is involved in the biosynthesis of the LPS O-antigen and the LPS core. Mol Genet Genomics 266:79–95

Wang J, So B, Kim J, Park Y, Lee B, Kang H (2008) Genome-wide identification of pathogenicity genes in *Xanthomonas oryzae* pv. *oryzae* by transposon mutagenesis. Plant Pathol 57:1136–1145

Wang LH, He Y, Gao Y, Wu JE, Dong YH, He C, Wang SX, Weng LX, Xu JL, Tay L (2004) A bacterial cell–cell communication signal with cross-kingdom structural analogues. Mol Microbiol 51:903–912

Wang Z, Gerstein M, Snyder M (2009) RNA-Seq: a revolutionary tool for transcriptomics. Nat Rev Genet 10:57–63

Waters CM, Bassler BL (2005) Quorum sensing: cell-to-cell communication in bacteria. Annu Rev Cell Dev Biol 21:319–346

Waters CM, Lu W, Rabinowitz JD, Bassler BL (2008) Quorum sensing controls biofilm formation in *Vibrio cholerae* through modulation of cyclic di-GMP levels and repression of *vpsT*. J Bacteriol 190:2527–2536

Wengelnik K, Bonas U (1996) HrpXv, an AraC-type regulator, activates expression of five of the six loci in the *hrp* cluster of *Xanthomonas campestris* pv. *vesicatoria*. J Bacteriol 178:3462–3469

Wengelnik K, Van den Ackerveken G, Bonas U (1996) HrpG, a key *hrp* regulatory protein of *Xanthomonas campestris* pv. *vesicatoria* is homologous to two-component response regulators. Mol Plant Microbe Interact 9:704–712

White FF, Potnis N, Jones JB, Koebnik R (2009) The type III effectors of *Xanthomonas*. Mol. Plant Pathol. 10:749–766

Williams P, Winzer K, Chan WC, Camara M, Cámara M (2007) Look who's talking: communication and quorum sensing in the bacterial world. Philos Trans R Soc B: Biol Sci 362:1119–1134

Yamazaki A, Hirata H, Tsuyumu S (2008) HrpG regulates type II secretory proteins in *Xanthomonas axonopodis* pv. *citri*. J Gen Plant Pathol 74:138–150

Yan Q, Hu X, Wang N (2012) The novel virulence-related gene nlxA in the lipopolysaccharide cluster of *Xanthomonas citri* ssp. *citri* is involved in the production of lipopolysaccharide and extracellular polysaccharide, motility, biofilm formation and stress resistance. Mol Plant Pathol. doi: 10.1111/j.1364-3703.2012.00800.x

Yan Q, Wang N (2011) The ColR/ColS two-component system plays multiple roles in the pathogenicity of the citrus canker pathogen *Xanthomonas citri* subsp. *citri*. J Bacteriol 193:1590–1599

Yan Q, Wang N (2012) High-throughput screening and analysis of genes of *Xanthomonas citri* subsp. *citri* involved in citrus canker symptom development. Mol Plant Microbe Interact 25:69–84

Yang B, White FF (2004) Diverse members of the AvrBs3/PthA family of type III effectors are major virulence determinants in bacterial blight disease of rice. Mol Plant Microbe Interact 17:1192–1200

Zansler ML (2004) The economic impacts to an industry associated with an invasive species: the case of ctrus canker in Florida. University of Florida

Zeidler D, Zähringer U, Gerber I, Dubery I, Hartung T, Bors W, Hutzler P, Durner J (2004) Innate immunity in *Arabidopsis thaliana*: lipopolysaccharides activate nitric oxide synthase (NOS) and induce defense genes. Proc Natl Acad Sci USA 101:15811–15816

Zhang S, He Y, Xu L, Chen B, Jiang B, Liao J, Cao J, Liu D, Huang Y, Liang X (2008) A putative *colRXC1049–colSXC1050* two-component signal transduction system in *Xanthomonas campestris* positively regulates *hrpC* and *hrpE* operons and is involved in virulence, the hypersensitive response and tolerance to various stresses. Res Microbiol 159:569–578

Genomic Insights into *Xylella fastidiosa* Interactions with Plant and Insect Hosts

8

Adam C. Retchless, Fabien Labroussaa, Lori Shapiro, Drake C. Stenger, Steven E. Lindow, and Rodrigo P. P. Almeida

8.1 Introduction

The utilization of genomic data and widespread availability of genomics tools is still incipient in plant pathology. The first genome of a plant pathogen, that of the bacterium *Xylella fastidiosa*, was only completed in 2000 (Simpson et al. 2000). Since then, many bacterial plant pathogens have been sequenced, but much of the scientific knowledge extracted from these data is still limited, especially when compared to human pathogens. Nevertheless, significant advances have been made during the last decade, when genomics became widely available to the plant pathology community. Our understanding of the functional role of genes and pathogen taxonomy and evolution has improved significantly. The availability of genome sequences has served as the backbone for much of this work. In this chapter, we take a holistic approach and explore *X. fastidiosa* evolution, biology, and management based on information and insights that would not have been possible, or would have been technically challenging, during the pre-genomics period of plant pathology. Although *X. fastidiosa* is widely thought of as a plant pathogen, its biology is more complex and focus on its pathogenicity to crops clouds our broader understanding of its biology, ecology, and evolution. We predict that genomic data and further research on isolates that do not cause crop diseases will permit a more complete view of *X. fastidiosa*. This is especially important in the context of emerging diseases, as data suggest that novel *X. fastidiosa* isolates are particularly prone to emerge as crop pathogens of economic importance.

In late 2013 *Xylella fastidiosa* was reported in southern Italy (Saponari et al. 2013). This report highlights the importance of *X. fastidiosa* as a quarantine pathogen, and the need to better understand its ecology and evolution.

A. C. Retchless · F. Labroussaa · L. Shapiro · R. P. P. Almeida (✉)
Department of Environmental Science, Policy and Management, University of California, Berkeley, CA, USA
e-mail: rodrigoalmeida@berkeley.edu

D. C. Stenger
United States Department of Agriculture-Agricultural Research Service, Parlier, CA, USA

S. E. Lindow
Department of Plant and Microbial Biology, University of California, Berkeley, CA, USA

8.1.1 History

The first plant disease associated with the bacterium *X. fastidiosa* was described over one century ago (Pierce 1892). However, knowledge about this organism lags significantly behind other well-known plant pathogenic bacteria. Epidemics have been the major factor driving research on this pathogen. The first recorded epidemic caused by *X. fastidiosa* in grapevines in southern California in the late 1800s led to initial characterization of Pierce's disease of

D. C. Gross et al. (eds.), *Genomics of Plant-Associated Bacteria*,
DOI: 10.1007/978-3-642-55378-3_8,

grapevines (Pierce 1892). Later, parallel and inter-related epidemics of Pierce's disease and alfalfa dwarf in the Central Valley of California resulted in the identification of insect vectors, but also in the conclusion that the etiological agent of these diseases was a virus (Hewitt et al. 1949). Only in the mid-1970s, after it was shown that 'yellows' diseases of plants were of bacterial origin (phytoplasmas and spiroplasmas; Doi et al. 1967), were diseases caused by *X. fastidiosa* determined to also have bacterial etiology (Hopkins and Mollenhauer 1973). That breakthrough led to its axenic culture in 1978 (Davis et al. 1978); and the bacterium was named *X. fastidiosa* one decade later (Wells et al. 1987). Despite its importance, research on this pathogen remained limited until a new disease in citrus emerged in Brazil in 1987 and another re-emerged in California in the late 1990s (Hopkins and Purcell 2002). These diseases, principally in Brazil, led to efforts aimed at sequencing the genome of these pathogens, with the expectation that this novel information would assist research aimed at the development of better and new disease management practices. That culminated in the sequencing of the first bacterial plant pathogen, *X. fastidiosa*, in the year 2000 (Simpson et al. 2000). It is fair to say that, other than data available for PCR-based detection of *X. fastidiosa* (Minsavage et al. 1994), very limited information derived from molecular tools was available for this bacterium prior to 2000. We direct readers interested in a more complete perspective of the history of *X. fastidiosa* research to a recently published review by Purcell (2013).

8.1.2 Geographical Distribution

Initially thought to be limited to North America (Hewitt 1958), *X. fastidiosa* is now known to be present throughout the Americas and Taiwan. There is one report of *X. fastidiosa* in Kosovo, Europe (Berisha et al. 1998); however, the bacterium is considered to be absent from the continent. Taiwan represents an interesting case, where two distinct phylotypes of *X. fastidiosa* occur, causing disease in pear trees and grapevines, respectively (Su et al. 2012, 2013). The grapevine disease is caused by isolates that are phylogenetically among those widely distributed in the United States, suggesting that it was recently introduced into the island, which is also evidenced by the lack of genetic diversity among Taiwanese isolates. The pear isolates, however, appear to have evolved in isolation, raising important questions regarding the worldwide distribution of *X. fastidiosa* and *X. fastidiosa*-like bacteria, which may have remained undetected due to significant biological, phenotypic, and genetic differences when compared to known taxa.

Within the Americas, *X. fastidiosa* occurs from the northeast region of the United States and Canada (Goodwin and Zhang 1997) to Argentina and southern Brazil (Hopkins and Purcell 2002). However, as discussed later, the distribution of individual subspecies is more limited, suggesting that geographical isolation has been important in the evolution of this bacterium. Furthermore, as with grapevine isolates in Taiwan, evidence indicates that at least some isolates causing disease in plum in Brazil originated from North America (Nunes et al. 2003) and that grapevine isolates in the United States originated from Central America (Nunney et al. 2010). Therefore, *X. fastidiosa* populations appeared to have remained largely isolated due to geographical barriers, but recent human activity has resulted in dispersal over continental distances.

8.1.3 Impact

X. fastidiosa causes disease in a wide range of host plants, from perennial fruit crops such as grapes and citrus to ornamental trees and shrubs. The economic impact of these diseases is poorly understood; to our knowledge, only Pierce's disease of grapevines in Northern California has been carefully studied in this context (Fuller 2012). The production of sweet orange in Brazil is severely impacted by this pathogen, which now infects up to half of plants in some regions, over one-third of citrus plants in São Paulo State are estimated to be infected with the pathogen. In rare instances, such as oleander leaf scorch in

southern California in the 1990s, most susceptible hosts were eliminated from the landscape due to *X. fastidiosa* infection. Control of *X. fastidiosa* diseases is dependent on host plant species, vector species, geographical location, and management practices, among other factors. We direct readers to other reviews that cover *X. fastidiosa* diseases, their impact, epidemiology, and management alternatives (Hopkins and Purcell 2002; Redak et al. 2004).

8.1.4 Quarantine Importance

This bacterium is of worldwide quarantine importance. Some countries where *X. fastidiosa* is established have enacted policies to avoid the introduction of new genotypes, while countries in which it is considered to be absent tend to prohibit the introduction of the species (i.e., Australia, New Zealand, and the European Union have strict regulations aimed at reducing the likelihood of *X. fastidiosa* introduction). It should be mentioned that these countries also list insect vectors of *X. fastidiosa* as quarantine organisms. The United States Department of Agriculture strictly regulates the introduction of foreign isolates and considered the genotype causing citrus variegated chlorosis (CVC), which is currently limited to South America, a bioterrorism agent until 2012 (Federal Register 2012).

The European Union considers *X. fastidiosa* an EPPO A1 pathogen, which are quarantine pests absent from the region. Strict guidelines exist for positive diagnostic and reporting of *X. fastidiosa* in Europe (European and Mediterranean Plant Protection Organization 2004). European governments routinely survey for this bacterium using ELISA and PCR. In addition to regulations on pathogen introduction and field surveys, some countries have been more proactive, assuming that an eventual introduction will occur. Australia and New Zealand scientists have performed research in California, USA, to determine the host range of *X. fastidiosa* and one of its vectors using native plant species from those countries (Rathé et al. 2012a; Sandanayaka and Backus 2008). Such efforts could have a significant impact on the eradication or management of *X. fastidiosa* introductions, as some biological knowledge would be already available to decision makers upon pathogen detection. Lastly, analysis of the potential threat of an introduction can provide important information to assess risk and develop science-based policies (e.g., Rathé et al. 2012b).

In addition to the above-mentioned risks, the introduction of *X. fastidiosa* into regions where the bacterium has already been detected increases chances of gene flow between endemic and invasive genotypes. The availability of novel loci and alleles can result in the emergence of epidemic isolates that have the potential to exploit host plants previously not susceptible to disease as a result of *X. fastidiosa* infection. Although not conclusively demonstrated, independent studies have suggested that *X. fastidiosa* diseases can emerge via gene exchange between isolates (Almeida et al. 2008; Nunes et al. 2003; Nunney et al. 2012). Therefore, increases in genetic diversity, driven by the introduction of foreign genotypes, carry risks that go beyond the potential host range of the newly arrived isolate.

8.1.5 Taxonomy

Xylella fastidiosa is a gamma-proteobacterium in the order Xanthomonadales, family *Xanthomonadaceae* (Wells et al. 1987). This is the single species in the genus *Xylella*, which is monophyletic, and has *Xanthomonas* ssp. as a sister clade (Fig. 8.1). *X. fastidiosa* is currently subdivided into four subspecies, largely based on DNA–DNA hybridization and multi-locus sequence typing data (Scally et al. 2005; Schaad et al. 2004). Although these groupings are phylogenetically robust, they can be further subdivided into groups with well-supported genetic and biological distinctions (i.e., different host ranges) (e.g., Almeida and Purcell 2003b; Almeida et al. 2008; Nunney et al. 2010, 2013). The taxonomy of *X. fastidiosa* is discussed in detail later within an evolutionary context.

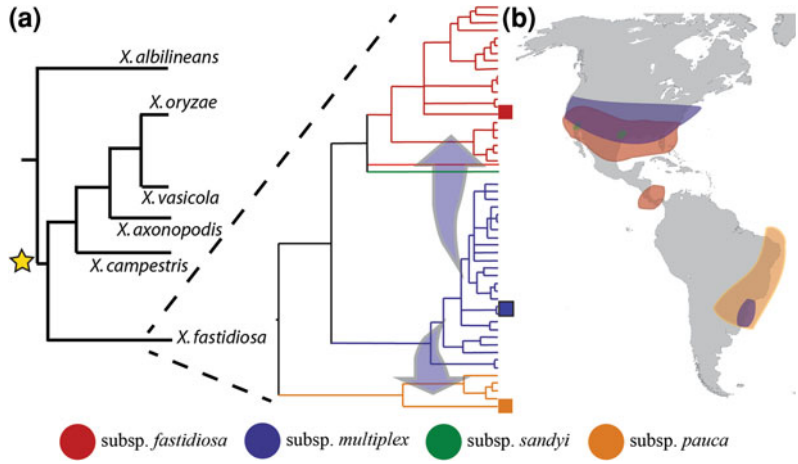

Fig. 8.1 Evolutionary history and geographical distribution of *X. fastidiosa* subspecies in the Americas. **a** Relationship of *X. fastidiosa* to *Xanthomonas* species (after Rodriguez et al. 2012), rooted by *Stenotrophomonas* and *Pseudoxanthomonas*. The position of *X. albilineans* is ambiguous (*star*; see text). *X. fastidiosa* diversity (MLST) is shown in detail, with published genome sequences identified with *squares*; the Taiwanese pear isolate is not included due to limited sequence data. Gene flow has been observed between subsp. *multiplex* and others (*arrows*). **b** Approximate spatial distribution of subspecies; it is assumed that the species extends throughout Central and South America

8.2 General Biology

The biology of *X. fastidiosa* is complex due to its requirement to successfully colonize two very distinct hosts, plants and insects. This bacterium emerged from plant-colonizing *Xanthomonas* ssp., and its divergence is hypothesized to be driven by the novel utilization of insect vectors for dispersal (Killiny et al. 2010). Therefore, a discussion of the biology of *X. fastidiosa* must consider its plant and vector colonization, in addition to mechanisms necessary to switch from one host to another. This section provides an introduction to the general biology of plant and insect colonization by *X. fastidiosa*; genome-derived insights are addressed later in this chapter.

8.2.1 Colonization of Host Plants

Unlike many other bacterial plant pathogens, *X. fastidiosa* has a large host range, both as a species and for each individual subspecies. Hill and Purcell (1995) compiled published data and concluded that plants in twenty-nine families were hosts of this bacterium. This bacterium is capable of multiplying in almost all tested plant species under greenhouse conditions when mechanically inoculated (Purcell and Saunders 1999). It can also be recovered from a wide range of plants in the field (Lopes et al. 2003), although *X. fastidiosa* does not cause disease in most of these species (Purcell and Saunders 1999). Thus, *X. fastidiosa* colonization of plants does not necessarily equal disease development. The list of host plants known to be susceptible to this pathogen has grown during the last decade, as has the even longer list of non-symptomatic host plants.

The details of pathogenicity mechanisms are yet to be fully understood, but disease appears to be a consequence of bacterial multiplication and movement within the xylem vessel network, leading to clogging of water flow through the plant (Chatterjee et al. 2008a; Newman et al. 2003). Although the bacterium is xylem limited, populations are not homogenous in the entire plant (Hopkins 1985). Bacterial accumulation at specific tissues within the plant are also host dependent; as an example, in Pierce's disease-

infected grapevines, *X. fastidiosa* is found at higher populations in symptomatic leaf veins and petioles (Baccari and Lindow 2011; Krivanek and Walker 2005). A positive correlation between symptom severity and pathogen populations within the plant has also been shown (Krivanek and Walker 2005). *X. fastidiosa* genomic analyses revealed a number of genes that encode plant cell wall-degrading enzymes; *X. fastidiosa* is strictly xylem limited, and these enzymes are predicted to solely function in pit membrane degradation. Using a GFP-tagged *X. fastidiosa* strain, symptom development was shown to be highly correlated with the number of vessels clogged (Newman et al. 2003). While structural differences in xylem vessels of susceptible versus nonsusceptible grape varieties did not correlate with probability of disease development (Chatelet et al. 2011), pit membrane degradation was more successful in susceptible versus resistant grapevine varieties (Sun et al. 2011). In addition, defensive responses of susceptible grapevines to *X. fastidiosa* colonization may further exacerbate restrictions on xylem fluid flow and symptom development (Sun et al. 2013).

8.2.2 Colonization of Insect Vectors

Severin (1949) showed that *X. fastidiosa* is persistently transmitted by adult sharpshooter leafhoppers (Hemiptera, Cicadellidae), and Freitag (1951) demonstrated lack of transovarial transmission. Persistence of infection in adults has been confirmed in other studies (Almeida and Purcell 2003a; Hill and Purcell 1995b); although transovarial transmission was not tested further, it is unlikely that vertical transmission of *X. fastidiosa* occurs. The lack of transstadial transmission and absence of a detectable latent period for acquisition or inoculation (Almeida and Purcell 2003a; Purcell and Finlay 1979) are strong indicators that this bacterium does not circulate within vectors. In addition, microscopic studies on the colonization of the foregut of vectors (Almeida and Purcell 2006; Brlansky et al. 1983; Purcell et al. 1979) supported the results obtained from transmission experiments.

After acquisition from source plants, *X. fastidiosa* attaches to and multiplies in the foregut of vectors; multiplication was first shown by culturing (Hill and Purcell 1995b) and more recently by quantitative PCR (Killiny and Almeida 2009a), and inferred by microscopy at different time points after acquisition from plants (Almeida and Purcell 2006). The generation time of *X. fastidiosa* cells within vectors as estimated by quantitative PCR was 7–8 h and remained constant for up to 4 days (Killiny and Almeida 2009a). We estimate that the foregut of vectors may house ~50,000 cells. Two regions of the foregut have been implicated in *X. fastidiosa* transmission based on spatial colonization patterns. Purcell et al. (1979) observed cells in the cibarium, the distal region of the precibarium, and the anterior region of the esophagus; Brlansky et al. (1983) confirmed that the precibarium was colonized by *X. fastidiosa*. However, those studies did not correlate bacterial visualization with vector transmission to plants. A more recent microscopy-based study showed an association between bacterial colonization of the precibarium and transmission to plants (Almeida and Purcell 2006), suggesting that colonization of the esophagus and cibarium was not directly associated with inoculation events. However, the specific site(s) in the foregut as well as the vector probing behavior(s) associated with inoculation is yet to be determined.

The precibarium and cibarium of leafhoppers are highly turbulent environments, and cell attachment is not a trivial process. Although fluid flow dynamics in this system have not been experimentally determined, xylem sap has been estimated to flow through the precibarium at average speeds of 8 cm/s (Purcell et al. 1979). Turbulence is likely present as the muscle connected to the cibarium's diaphragm creates enough tension to pump sap from plants into the midgut by contracting and relaxing approximately every second during ingestion events (Dugravot et al. 2008). Thus, it is possible that very few colonization events of the precibarium

occur given the number of cells or cell aggregates that may be ingested by a vector through feeding on sap from an infected xylem vessel. While few colonization events have been suggested based on microscopy data (Almeida and Purcell 2006), quantitative PCR data indicate that approximately one to five thousand cells can be detected in sharpshooter heads after feeding on infected plant material (Rashed et al. 2011). Many of the ingested cells likely do not adhere to the insect cuticle and pass through the digestive tract, and therefore are considered to not be involved in vector colonization and *X. fastidiosa* transmission to plants. This discrepancy highlights the fact that successful attachment of *X. fastidiosa* to the foregut of vectors is probably a rare event requiring high-affinity ligand–receptor interactions.

8.2.3 Vector Transmission

Transmission experiments identified sharpshooter leafhoppers (Hemiptera, Cicadellidae, subfamily Cicadellinae) and spittlebugs (Hemiptera, Cercopoidea) as insect vectors of this pathogen (Severin 1949, 1950). A third group of xylem sap-sucking insects, cicadas (Hemiptera, Cicadidae), has been reported as a vector of *X. fastidiosa* (Paião et al. 2002; Krell et al. 2007), although more work is necessary to determine the contribution of this group to overall epidemiology and transmission dynamics. An important aspect associated with the transmission of *X. fastidiosa* is the lack of pathogen genotype and vector species specificity (Almeida et al. 2005). Frazier (1965) summarized the current state of knowledge at the time by concluding that all sharpshooter leafhoppers should be considered vectors of the Pierce's disease etiological agent until proven otherwise. The addition of various pathogen genotypes to this statement occurred later, after the advent of molecular tools that eventually allowed the split of the species into subspecies, when various studies showed that all tested vector species were capable of transmitting various *X. fastidiosa* genotypes (Almeida et al. 2005). The best example was the demonstration that a North American vector species transmitted a South American isolate to plants (Damsteegt et al. 2006). Ultimately, more vector–pathogen combinations must be tested, but so far Frazier's 50-year-old statement remains true.

Various factors affect *X. fastidiosa* vector transmission efficiency. The length of time insects are allowed to acquire or inoculate *X. fastidiosa* into plants is proportional to overall transmission efficiency, up to approximately four days (Almeida and Purcell 2003a; Daugherty and Almeida 2009; Purcell and Finlay 1979). The most parsimonious interpretation of these results is that the probability of insects probing into xylem vessels with *X. fastidiosa* increases over time (Almeida and Backus 2004; Backus et al. 2005). In fact, the overall population of *X. fastidiosa* within host plant tissue has been the only parameter consistently correlated with transmission efficiency. Hill and Purcell (1997) were the first to demonstrate such relationship, which has been used to explain why specific vector–pathogen–plant combinations result in higher or lower transmission efficiency [e.g., Lopes et al. (2009)]. Furthermore, within-plant differences in pathogen population also affect transmission efficiency (Daugherty et al. 2010). Thus, host plants harboring larger *X. fastidiosa* populations generally result in higher transmission efficiency, and vice versa. Vector age and sex do not affect overall transmission efficiency (Krugner et al. 2012).

8.3 Genome Structure

As of August 2013, five finished *X. fastidiosa* genome sequences have been published (PubMed Genomes Database; http://www.ncbi.nlm.nih.gov/genome/genomes/173), representing three of the four described subspecies (subspecies *sandyi* being the exception; Fig. 8.2) (Chen et al. 2010; Simpson et al. 2000; van Sluys et al. 2003). These genomes range from 2.5 to 2.7 megabase pairs (Mbp) and include a single chromosome along with zero to two plasmids (ranging from 1.3 to 51 kb). The nucleotide

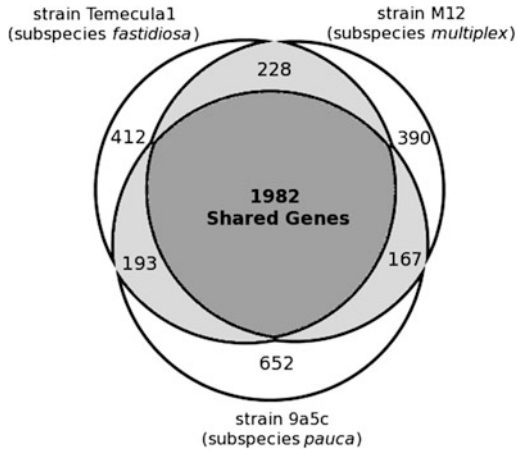

Fig. 8.2 Venn diagram for the number of predicted coding sequences that are shared and unique from representative genomes of subspecies *multiplex*, *fastidiosa*, and *pauca* calculated with EDGAR (Blom et al. 2009), using a protein–protein BLAST e-value cutoff of 10^{-6} and 70 % amino acid identity

sequences are 51.8–52.6 % GC. Published annotations indicate that these genomes each contain from 2,066 to 2,294 protein-coding sequences, 2 ribosomal RNA operons, 49–50 tRNA genes (43 different anti-codons), and 1 tmRNA gene (van Sluys et al. 2003). Draft quality genomes are available for three other isolates, each appearing to be very similar to one of the finished genomes at the coarse scale (Bhattacharyya et al. 2002; Schreiber et al. 2010; Zhang et al. 2011), as are contigs from the shotgun sequencing of a mixed sample of subspecies *multiplex* and *sandyi* (Bhattacharyya et al. 2002; Nunney et al. 2012). The diversity of available sequences will increase rapidly in the near future, as several projects are underway that are sequencing tens of genomes from isolates collected around the world.

8.3.1 Core Genome

The location of the origin of replication can be confidently predicted from three different lines of genomic evidence: the location of the *dnaA* gene, a cluster of DnaA boxes, and inversion of the G/C and A/T skews (Mackiewicz et al. 2004; Simpson et al. 2000). The terminus can be identified by similar methods: Both the predicted *dif* sequence and oligonucleotide asymmetry analysis point to the same region of each sequenced chromosome (Yen et al. 2002; Kono et al. 2011). Based on the predicted origin and terminus of replication, we estimate that 59–60 % of genes are encoded on the leading replication strands (Rocha 2008). Notably, the chromosome arms of strain 9a5c are heavily imbalanced such that one replichore is approximately half the length of the other; in contrast, the origin and terminus of replication for the other complete *X. fastidiosa* chromosomes are approximately opposite from each other. Based on studies in *Escherichia coli*, this degree of replication imbalance is expected to be a substantial impediment to growth (Esnault et al. 2007).

We identified 1982 protein-encoding genes that are shared among subspecies *fastidiosa*, *multiplex*, and *pauca* (Fig. 8.2). This accounts for 71–86 % of the largest *X. fastidiosa* genome, 9a5c, depending on how conservative one is in predicting genes (e.g., Simpson et al. 2000; van Sluys et al. 2003). Only one study addressed this question at the subspecies level (subsp. *pauca*); it showed that the pan-genome of *X. fastidiosa* is potentially large and that differences among isolates were associated with genome island-like fragments (da Silva et al. 2007). However, representative genomes of different subspecies that have been fully sequenced are largely syntenic (Fig. 8.3). Several short sequence repeats (SSR) have been identified in the chromosome (Coletta-Filho et al. 2001). These loci tend to evolve quickly, and several have been confirmed as having variable number of tandem repeats (VNTR), making them useful for high-resolution studies of population structure (Coletta-Filho et al. 2011).

8.3.2 Mobile Elements

Several self-replicating genetic elements are present among *X. fastidiosa* genomes, including plasmids and prophages. These contribute to the gene diversity of *X. fastidiosa* and provide potential avenues for horizontal gene transfer (HGT) both within and between species. Due to the ability of HGT to confer new phenotypes on

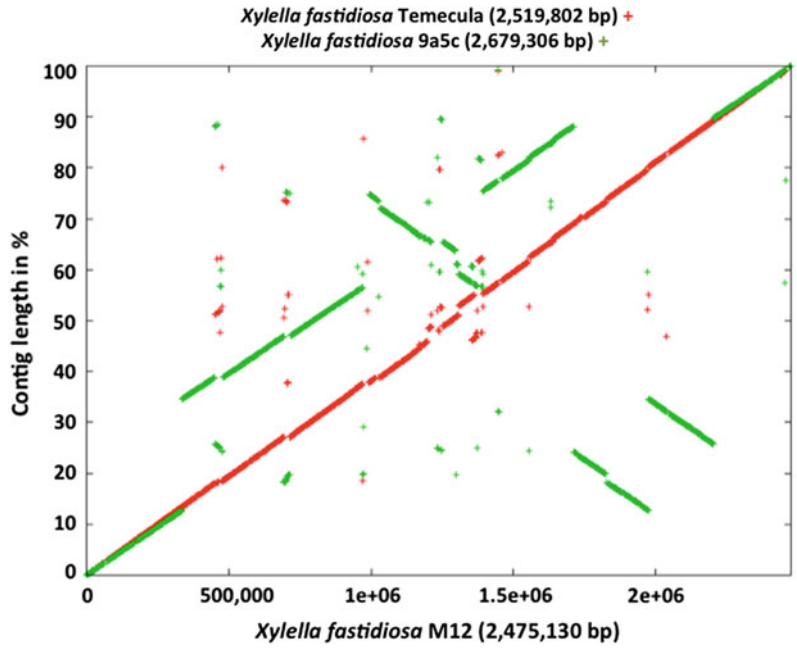

Fig. 8.3 Synteny map of representative genomes from each of the three sequenced subspecies, showing several rearrangements on the *X. fastidiosa* 9a5c (subsp. *pauca*) chromosome compared to *X. fastidiosa* M12 (subsp. *multiplex*) and *X. fastidiosa* Temecula1 (subsp. *fastidiosa*)

bacteria and modify their ecological niche, there is intense interest in exploring the contribution of these elements to the distinct virulence traits of different isolates, including the ability to colonize novel host plants (Moreira et al. 2005; Nunes et al. 2003).

Despite the abundance of putative prophages embedded within the chromosome, the isolation of active phage particles has been difficult, with the first lysogen confirmed only in 2010 (Summer et al. 2010). The prevalence of prophages and phage remnants within *X. fastidiosa* genomes contrasts with the absence of identifiable transposons, which are abundant in *Xanthomonas* genomes (Monteiro-Vitorello et al. 2005). Prophages are also candidates for inducing rearrangements within the *X. fastidiosa* genome, due to the activity of their integrase genes (Moreira et al. 2005; Nunes et al. 2003).

Among subspecies, it is evident that *X. fastidiosa*'s flexible genome is primarily composed of laterally transferred phage-like elements, including prophages (Nunes et al. 2003; van Sluys et al. 2003; Varani et al. 2008). Interestingly, transcription studies have suggested that this laterally transferred gene pool is differentially regulated in *X. fastidiosa* in comparison with genes in the core genome, suggesting that their regulation is yet to be tightly linked to the rest of the genome (Nunes et al. 2003). The finding that most obvious differences among *X. fastidiosa* genomes are connected to phage-like sequences has led several groups to propose that the divergence of lineages leading to host specificity is mediated by laterally transferred elements (e.g., van Sluys et al. 2003; Nunes et al. 2003; Varani et al. 2008). Thus, host adaptation in *X. fastidiosa* could be the result of one or few changes with large fitness benefits, despite the fact that overall pathogenicity mechanisms are similar regardless of host plant. This model is comparable to the effector–host resistance genes evolutionary model for most bacterial pathogens (Lindeberg et al. 2009; Ma et al. 2006). However, a competing hypothesis proposes that quantitative rather than qualitative differences are the major drivers of host specificity (Killiny and Almeida 2011).

The existence of plasmids in *X. fastidiosa* cultures has been demonstrated in two manners: Some plasmids have been purified directly from *X. fastidiosa* cultures, while the existence of others have been inferred based on circular contigs generated during genome sequencing projects. These extra-chromosomal DNA molecules are commonly associated with bacteria, encoding accessory modules conferring selective advantage in specific environments (e.g., resistance to antibiotics or heavy metals, or an ability to degrade toxic organic compounds; Van der Auwera et al. 2009). Conjugative plasmids encode modules for DNA transfer, allowing for their own propagation independent of the host and potentially transferring chromosomal DNA. Complete nucleotide sequences are currently available for several plasmids of *X. fastidiosa*. Below, a brief summary of plasmids associated with *X. fastidiosa* is presented.

Known plasmids of *X. fastidiosa* represent a broad range of diversity. A 51 kilobase pair (kb) circular contig designated pXF51 was generated from genome sequencing of *X. fastidiosa* subspecies *pauca* strain 9a5c from citrus (Simpson et al. 2000; van Sluys et al. 2003). A smaller 6 kb plasmid (pXF5823) has been characterized from another citrus-infecting strain (Qin and Hartung 2001). Four *X. fastidiosa* strains isolated from mulberry in Southern California harbor closely related 25 kb plasmids (pXFRIV11, pXFRIV16, pXFRIV19, and pXFRIV25) that have been assigned to incompatibility group P1 (IncP1) (Stenger et al. 2010). A 38 kb plasmid (pXFRIV5) has been characterized from a subspecies *multiplex* strain (RIV5), isolated from ornamental plum in Southern California (Rogers and Stenger 2012; Chen et al. 2010). A nearly identical plasmid (pXFAS01), varying from pXFRIV5 at only six nucleotide positions, was discovered as a circular contig during genome sequencing of subspecies *fastidiosa* strain M23 (Chen et al. 2010) isolated from almond in California. Multiple strains of *X. fastidiosa* contain a small (1.3 kbp) plasmid that utilizes rolling circle replication (Guilhabert et al. 2006; Pooler et al. 1997). All of the plasmids mentioned above have been completely sequenced. Other, likely distinct, plasmids are present in *X. fastidiosa* strains isolated from a variety of hosts and locations but so far have been characterized only for restriction endonuclease patterns (Hendson et al. 2001).

Beyond inferences from homology, there is limited information regarding function(s) of genes encoded by *X. fastidiosa* plasmids. These plasmids appear to lack modules that would provide selective advantages to their hosts. Interestingly, the only type IV secretion systems in *X. fastidiosa* are found on conjugative plasmids, where they presumably facilitate DNA transfer. Among the sequenced plasmids of *X. fastidiosa*, only those of the 38 kb class (pXFAS01 and pXFRIV5) encode what appear to be a complete type IV secretion system: *tra* and *trb* modules necessary for conjugative transfer and mating pair formation, respectively. Indeed, the occurrence of nearly identical 38 kb plasmids in subspecies *fastidiosa* (pXFAS01) and *multiplex* (pXFRIV5) may be due to recent conjugative transfer among strains representing different subspecies (Rogers and Stenger 2012).

Several genes of the 25 kb (Inc-P1) class of plasmids from mulberry-infecting strains of *X. fastidiosa* (Stenger et al. 2010) have been characterized for function. Sequence comparisons indicate that extensive regions (~75 %) the 25 kb IncP-1 plasmids of *X. fastidiosa* are most closely related to an IncP-1 plasmid (pVEIS01) from the earthworm symbiont *Verminephrobacter eiseniae* (Pinel et al. 2008). IncP-1 plasmids initiate DNA replication through a plasmid-encoded protein (TrfA), which binds to iterative elements of the vegetative origin of replication (*oriV*) (Mei et al. 1995). Phylogenetic analysis revealed that the TrfA homologues of *X. fastidiosa* and *V. eiseniae* plasmids represent a newly discovered and divergent lineage of IncP-1 plasmids (Stenger and Lee 2011). Replication modules (*trfA* and *oriV*), derived from IncP-1 plasmids of *X. fastidiosa* or *V. eiseniae*, placed into a standard *E. coli* cloning vector allowed plasmid replication in *X. fastidiosa* (Lee et al. 2010). Furthermore, hybrid Inc-P1 replication modules (e.g., heterologous combinations of *trfA* and *oriV* from *X. fastidiosa* and *V. eiseniae*) conferred the ability to replicate in *X. fastidiosa* (Stenger and Lee 2011).

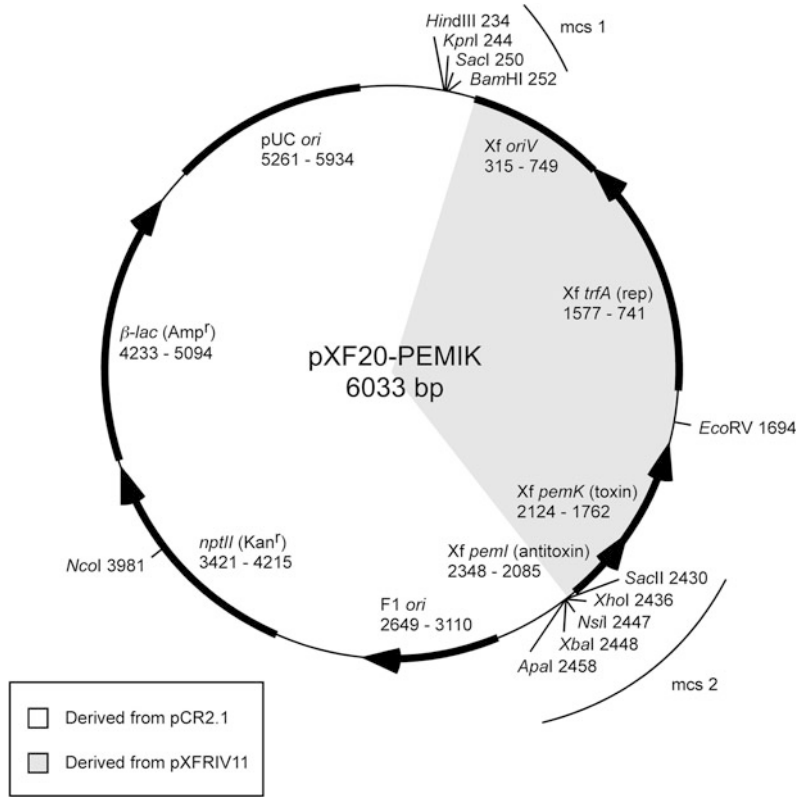

Fig. 8.4 Map of the *X. fastidiosa* shuttle vector pXF20-PEMIK. Colored arcs denote regions derived from *E. coli* plasmid pCR2.1 or *X. fastidiosa* plasmid pXFRIV11. MCS1 and MCS2 denote locations of unique endonuclease restriction sites serving as multiple cloning sites

Plasmids bearing IncP-1 replication modules were stable in *X. fastidiosa* only when cultured under antibiotic selection. Growth of *X. fastidiosa* transformants in the absence of antibiotic selection resulted in extinction of plasmids bearing *trfA* and *oriV*. However, addition of the *pemI/pemK* toxin–antitoxin(TA) system homologue encoded by pXFRIV11 conferred stable inheritance of plasmids bearing the *X. fastidiosa* IncP-1 replication module in the absence of antibiotic selection (Lee et al. 2010). Functional analysis of the *X. fastidiosa pemI/pemK* TA system indicated that PemK toxin is an endoribonuclease and that PemI is the cognate antitoxin that blocks PemK ribonuclease activity via direct and reversible binding (Lee et al. 2012). These results, and the effects of unbound PemK toxin on cell growth, indicate that the *X. fastidiosa pemI/pemK* TA system is a classic plasmid addiction system (stable toxin, labile antitoxin) in which daughter cells not containing plasmid are killed by residual toxin. The construct pXF20-PEMIK (Fig. 8.4) bearing the *X. fastidiosa* Inc-P1 replication module and *pemI/pemK* TA system represents a stable shuttle vector able to replicate in both *E. coli* and *X. fastidiosa*.

Of what benefit are plasmids to *X. fastidiosa* fitness as a plant pathogen? At present, the answer to this question is not known. Currently, no pathogenicity or virulence factors of *X. fastidiosa* are known to be plasmid-encoded. In some cases (e.g., the 1.3 kb rolling circle replicon), plasmids simply may be acting as selfish DNA. In another case (e.g., the 38 kb plasmids), encoding a functional type IV secretion system may allow for (or has facilitated) acquisition of new traits, potentially including pathogenicity islands or other genetic modules of benefit to *X. fastidiosa*.

8.3.3 Comparative Sequence Analysis

Like other xanthomonads, *X. fastidiosa* has a circular chromosome and a variety of plasmids, none of which are present among all isolates. A notable difference between *Xanthomonas* and *Xylella* genomes is size; *X. fastidiosa* shares ~ 74 % of its genome with *Xanthomonas* strains (Moreira et al. 2004), but *X. fastidiosa* genomes range from 2.5 to 2.7 megabase pairs (Mb), the smallest *Xanthomonas* genome (*X. albilineans*) consists of 3.7 Mb and other genomes range from 4.5 to 5.5 Mb. Unlike *X. fastidiosa*, *Xanthomonas* ssp. genomes are highly plastic and there is very high diversity within the genus.

X. fastidiosa appears to be derived from a form of *Xanthomonas* (Fig. 8.1); the relationship between *X. fastidiosa*, *Xanthomonas albilineans*, and the remainder of *Xanthomonas* species is unclear, but the most comprehensive published analysis identifies *X. albilineans* as an outgroup with low confidence (Rodriguez et al. 2012). Our own analysis supports this relationship, but *Xylella* has also been inferred to be the sister clade to *Xanthomonas albilineans* (Pieretti et al. 2009). The family *Xanthomonadaceae* is very diverse, but there is no indication that plant association is common outside of *Xanthomonas* and *Xylella*. The family (and order) itself has a complicated evolutionary history; while it is nominally within the gamma-proteobacteria, its relationship to that group is not straightforward. Phylogenies generally depict the lineage as splitting from the rest of the gamma-proteobacteria, and a gene-by-gene analysis has revealed that many genes are more similar to genes of the alpha or beta proteobacteria, implying that much of the genome was acquired from these other lineages and the placement within gamma-proteobacteria may not be reliable (Comas et al. 2007).

Like *X. fastidiosa*, *X. albilineans* is considered to have a reduced form of the *Xanthomonas* genome. However, gene loss occurred subsequent to their divergence from each other, indicating that the lineage leading to *X. fastidiosa* likely underwent the loss of 2 Mb of genome content independent of any known lineage (Pieretti et al. 2009). This reduction in genome size is consistent with the stereotypical genome degradation process that has been described for assorted host-dependent bacteria (Moran and Plague 2004). The genome erosion syndrome is also revealed in the absence of detectable selection on codon usage and the slow growth rate of *X. fastidiosa* (Sharp et al. 2005; Rocha 2004), relative to *Xanthomonas* species. Two explanations are available for this evolutionary pattern. The ecological explanation posits that bacteria growing in a stable, nutrient-limited environment do not benefit from a diverse gene repertoire or from the ability to rapidly produce proteins (Vieira-Silva and Rocha 2010). Alternatively, the genetic explanation appeals to a reduction in selective efficacy due to reduction in effective population size that results from frequent population bottlenecks associated with transmission between hosts (Sharp et al. 2010). Combined with the deletion-biased mutation processes that are typical of bacterial genomes, this results in a reduction in genome size as genes that provide little fitness advantage are lost.

Only a small number of genes distinguish *X. fastidiosa* from *Xanthomonas* species. Due to evolutionary association of such genes with the defining ecological traits of *X. fastidiosa* (particularly insect transmission), they are candidates for participation in these functions. The pear leaf scorch pathogen from Taiwan appears to represent a lineage of *X. fastidiosa* that branched from the American lineages prior to their diversification. If it is confirmed to exhibit the same basic ecological and physiological traits of *X. fastidiosa*, then its genome sequence will help us to narrow in on the genomic changes responsible for the evolution of those traits.

Although four subspecies are commonly recognized within *X. fastidiosa*, there are reports of variants that may belong to two other subspecies, though these have yet to be well described (Su et al. 2012; Randall et al. 2009). Although various approaches have been used to

study *X. fastidiosa* genetic diversity, multi-locus sequence typing (MLST) has provided the most valuable insights into the evolutionary history of this pathogen (Yuan et al. 2010). MLST is based on the sequencing of seven housekeeping genes showing neutral variation and are distributed throughout the chromosome of *X. fastidiosa* such that they are unlinked during recombination. This approach has been used to characterize the genetic diversity of *X. fastidiosa* subspecies, providing insights into the history of pathogen dispersal and genetic recombination between subspecies, as described below.

Subspecies *fastidiosa* and *sandyi* in the United States are composed of isolates with very little diversity when typed using MLST (Yuan et al. 2010), but subsp. *fastidiosa* isolates from Costa Rica colonizing various plant species have more allelic diversity than those causing disease in grapevines in the United States (Nunney et al. 2010). These results suggest that Pierce's disease in the United States was introduced from Central America. The genetic structure of subsp. *multiplex* is somewhat distinct in that there are multiple clusters of isolates, with indications of host-range differences among clusters (Nunney et al. 2013; Scally et al. 2005; Almeida and Purcell 2003b). Lastly, subsp. *pauca* is limited to South America; while most studied isolates colonize either citrus or coffee sympatrically, there is no overlap in host range and the groups are phylogenetically distinct (Almeida et al. 2008). Interestingly, isolates recovered from plum in Brazil fall within subsp. *multiplex*, which is otherwise limited to North America, or fall between subsp. *multiplex* and *pauca* (Nunes et al. 2003), suggesting it was likely introduced from North America via contaminated plant material.

Horizontal gene transfer between subspecies is a recurring theme in *X. fastidiosa* evolution, as shown by both MLST and whole-genome data, along with distribution of plasmids as described above. Both the coffee and citrus isolates of subspecies *pauca* appear to have acquired alleles from subspecies *multiplex* (Nunney et al. 2012). Likewise, subspecies *multiplex* has participated in recombination with subspecies *fastidiosa* (Nunney et al. 2010; Kung et al. 2013). This process of allele conversion has even gone as far as to produce new lineages with roughly equal contributions from subsp. *multiplex* and *fastidiosa,* as found in the strains isolated from mulberry (Nunney et al. 2014). Overall MLST has proven to be a robust tool for reconstructing the relationship between *X. fastidiosa* isolates despite gene flow between subspecies and has therefore become the paradigm for *X. fastidiosa* taxonomy (Fig. 8.1).

In contrast to effector-mediated host/microbe interactions shown for *Xanthomonas* ssp. and most other bacterial plant pathogens, host specificity in *X. fastidiosa* violates the dominant paradigm of host–pathogen specificity among bacterial plant pathogens, as this bacterium does not have effector-encoding genes or a type III secretion system (Simpson et al. 2000). *X. fastidiosa* is further distinguished from *Xanthomonas* ssp. In that genome comparisons do not reveal gene content differences that could be responsible for the observed host specificity (van Sluys et al. 2003).

All *Xanthomonas* have six secretion systems (type I–VI) to export proteins. Within individual *Xanthomonas* ssp., pathovars show specific host associations, largely driven by secreted effectors (White et al., 2009), of which each strain secretes a unique blend of 20–40 directly into the host cytoplasm. While the exact function of many effectors in host cells is just beginning to be understood, many interfere or suppress induced plant defenses. *X. albilineans* is the only sequenced *Xanthomonas* ssp. without an Hrp-T3SS, but contains an SP1-T3SS. Mutations at the T3SS locus eliminate virulence among *Xanthomonas* isolates; therefore, the absence of a T3SS in *X. fastidiosa* indicates that its pathogenicity mechanisms are distinct and not yet fully described compared to most other bacterial pathogens (Simpson et al. 2000; Chatterjee et al. 2008a).

8.4 Applications from Genomic Data

8.4.1 Genetic Tractability

Determination of the complete genome sequence of *X. fastidiosa* (Simpson et al. 2000) raised questions about the putative function of several genes in plant pathogenicity and insect transmission, among other topics. Genetic analysis of these genes typically requires the ability to transform bacteria with appropriate plasmids carrying complementary wild-type genes. The first description of a system for genetic analysis was the transformation of a construct carrying ORF1 of *X. fastidiosa* along with its origin of replication, originally present in the sequence of one of its plasmids (Qin and Hartung 2001). Several studies described different systems used to transform *X. fastidiosa* (e.g., Silva Neto et al. 2002) or extracted from broad-host-range plasmids (Guilhabert and Kirkpatrick 2003). However, many of these early constructs were not stable once introduced in *X. fastidiosa* chromosome without antibiotic selection, and genetic analysis of knockouts within plants or insects was impossible. Development of tools allowing the stable transformation of *X. fastidiosa* by homologous recombination through the utilization of plasmids carrying its own chromosomal replication origin (*oriC*) circumvented this limitation (Monteiro et al. 2001; Newman et al. 2003, 2004). More recently, the pAX1 plasmid series carrying a *colE1*-like (pMB1) replicon associated with four different antibiotic selection markers allowed the recovery of double recombinants (Matsumoto et al. 2009). The pXF20-PEMIK plasmid with an addiction system, described above, may be successful in allowing for complementation of mutants, which has been a major limiting factor in *X. fastidiosa* research. To our knowledge, only Chatterjee et al. (2010) performed studies with complemented knockouts in both plants and insects using the broad-host-range plasmid pBRR-5 (Kovach et al. 1995), which was previously used by Reddy et al. (2007).

Xylella fastidiosa is naturally competent and able to transform and recombine linear and circular DNA during its growth phase (Kung and Almeida 2011). Transformation and recombination efficiencies increase with increments in the length of homologous flanking regions of inserts, up to 1 kb in size. In addition, efficiency reduces as the size of the non-homologous insert increases, with recombination not detected once the region was 6 kb in length (Kung et al. 2013). Competency, as in other bacteria, appears to be mediated by a type IV pilus-like apparatus, with associated *com* genes that transport DNA fragments through the cell membrane. The ease of transformation with this protocol, in addition to its high efficiency, should facilitate the genetic tractability of *X. fastidiosa*.

8.4.2 Genomics Opens New Research Venues

While there has been much less study of the virulence mechanisms utilized by *X. fastidiosa* compared to its closest relatives in the genus *Xanthomonas*, the availability of genome sequences has enabled several putative virulence factors to be investigated. Both the lack of an apparent type III secretion system, commonly used in other plant pathogens to suppress plant host defense responses, and the fact that this pathogen probably seldom encounters living plant cells while colonizing the xylem vessels suggests that *X. fastidiosa* differs in the factors that contribute to its virulence compared to other plant pathogens. Most attention has been directed toward the study of the role of other secretion systems in its virulence, and it appears that *X. fastidiosa* is capable of type I secretion. Analysis of the genome of *X. fastidiosa* reveals the presence of at least 23 systems comprising 46 proteins that belong to the ABC (ATP-binding cassette) superfamily of proteins (Meidanis et al. 2002). Type I secretion systems are often used in tolerance of toxic compounds and encode efflux pumps for small molecules as well as enabling the secretion of extracellular

proteins such as hemolysins. A central component of many bacterial type I secretion systems is a protein similar to TolC that spans the inner and outer membranes of gram-negative species. Importantly, a *tolC* knockout mutant of *X. fastidiosa* exhibited nearly complete loss of virulence to grape (Reddy et al. 2007). Perhaps more importantly, no viable cells of the *tolC* mutant could be recovered after inoculation into grape. This latter observation suggests strongly that an efflux pump in which TolC participated was required for survival in plants after inoculation, perhaps by enabling export of toxic compounds found in the xylem of the plant. Support for this conjecture was provided by the observation that the *tolC* mutant was more sensitive to toxic phytochemicals such as berberine, certain detergents, as well as the variety of compounds found in crude plant homogenates than the wild-type strain (Reddy et al. 2007). While such efflux systems can also export toxins and proteins that might elicit a host plant response, no studies have appeared to suggest that *X. fastidiosa* uses such a process to interact with living plant cells. Clearly, more work is needed to understand to which extent *X. fastidiosa* interacts with living plant tissues. The fact that tyloses that invade the xylem vessel are induced in plants infected with this pathogen (Hopkins and Purcell 2002) suggests that some communication with living plant tissue does occur. Furthermore, others have posited that many of the water stress symptoms associated with *X. fastidiosa* infection are not due solely to plugging of xylem vessels by cells of the pathogen, but instead are a result of excessive self-induced blockage of vessels by tyloses that are induced by the presence of the pathogen (Fry and Milholland 1990; Hopkins 1989; Hopkins and Purcell 2002; Purcell and Hopkins 1996).

The type II secretion system appears to be particularly important for the virulence of *X. fastidiosa* in the absence of at type III secretion system. The type II secretory system is primarily involved in the export of extracellular enzymes, frequently for the hydrolysis of various plant structural features. The type II secretion system, encoded by a collection of genes commonly known as Xps (Xanthomonas protein secretion), has been widely studied in other *Xanthomonas* species. Not only does *X. fastidiosa* harbor a complete set of Xps homologues, but it also contains genes capable of encoding several different extracellular enzymes such as several β-1,4 endoglucanases, xylanases, xylosidases, and a polygalacturonase (Simpson et al. 2000; van Sluys et al. 2003). Only the role of the polygalacturonase, encoded by *pglA* on the virulence of *X. fastidiosa*, has been addressed experimentally. Importantly, a *pglA* mutant exhibited greatly reduced ability to colonize grape and therefore incited very little symptom development (Roper et al. 2007). While the mutant could be re-isolated from near the point of inoculation, it was severely restricted in its long-distance movement along the grape xylem vessels (Roper et al. 2007). It seems likely that its greatly reduced motility is attributable to its inability to degrade pit membranes that serve to restrict both lateral and longitudinal movement of the pathogen between xylem vessels. As both the growth and movement of the *pglA* mutant was suppressed compared to the parental strain, it seems possible that pectin may serve not only as a barrier to intercellular movement, but might also be a nutrient source for *X. fastidiosa* although pectin does not appear to affect bacterial growth in vitro (Killiny and Almeida 2009b). In fact, the differential polysaccharide composition of pit membranes has been suggested to account for the relative differences in susceptibility of different grape varieties to invasion by *X. fastidiosa*, with homogalacturans and xyloglucans playing an important role in the susceptibility of various varieties to infection (Sun et al. 2011). Further evidence for the role of pectin as a constituent of a physical barrier for intercellular movement of *X. fastidiosa* is provided by the observation that grape expressing a polygalacturonase-inhibiting protein from pear exhibited higher resistance to symptom development after inoculation with this pathogen than the parental line (Agüero et al. 2005). It is also important to note that *rpfF* mutants of *X. fastidiosa* that do not produce diffusible signaling factor (DSF) express *pglA* and genes encoding

certain other extracellular enzymes at a higher level than that of the wild-type strain (Chatterjee et al. 2008b; Wang et al. 2012). Given that *rpfF* mutants are hyper-virulent to grape (Newman, et al. 2004), it is reasonable to assume that the elevated expression of these cell wall-degrading enzymes in this mutant background could have contributed to at least some of the enhanced virulence that they exhibit.

Active motility appears to be an important factor contributing to the virulence of *X. fastidiosa*. While *X. fastidiosa* is a non-flagellating bacterium, it harbors several genes that encode proteins involved in production and function of type IV pili (Simpson et al. 2000; van Sluys et al. 2003). These long, polar located pili have been shown to be involved in twitching motility in *X. fastidiosa* (Meng et al. 2005). Intriguingly, such twitching motility was demonstrated to enable the movement of *X. fastidiosa* not only along abiotic surfaces, but also along xylem vessels. Importantly, *X. fastidiosa* exhibits the ability to actively move against the flow of xylem fluids in vessels by twitching motility (Meng et al. 2005). The apparently high efficiency with which cells of *X. fastidiosa* move through the orifices of pits (Newman et al. 2004), which are apparently enlarged due to the action of extracellular enzymes secreted by this pathogen (Perez-Donoso et al. 2010), may be due at least in part to its ability to move along surfaces by retraction of the type IV pili. *X. fastidiosa* also produces relatively short type I pili, encoded by *fimA*, that apparently act to restrict motility (De La Fuente et al. 2007). *fimA* mutants exhibited a higher rate of twitching motility than that of the wild-type strain, suggesting that it may serve as an "anchor," being involved in attachment but serving to repress motility (De La Fuente et al. 2007). Perhaps not surprisingly, the expression of the genes for type IV pili and that of *fimA* tend to be oppositely regulated by the accumulation of DSF signal molecule (Chatterjee et al. 2008b; Wang et al. 2012). It therefore would seem prudent for *X. fastidiosa* to balance the abundance of these two pilus types depending on those stages of its colonization to which they might primarily contribute. That is, the initial attachment of *X. fastidiosa* to surfaces such as insect vectors may be facilitated by *fimA* (Killiny and Almeida 2009b), but its presence would tend to suppress active movement by twitching within the plant after inoculation. Curiously, a *tonB* homologue was found to be required for twitching motility and appropriate biofilm formation and virulence of *X. fastidiosa*, phenotypes unexpected from its role in transport of non-permeable molecules across the outer membrane and other bacteria (Cursino et al. 2009). Further evidence of the need for complex regulation of type IV pili function in *X. fastidiosa* is the observation that a chemosensory system is operative. An operon named Pil-Chp containing genes homologous to those found in chemotaxis systems of other bacteria were found to be required for proper pilus function but not biogenesis in *X. fastidiosa* (Cursino et al. 2011). Such regulatory mutants were also deficient in colonization of plants, suggesting that complex regulation of pilus function may also occur in plants (Cursino et al. 2011). While such regulators are apparently required for motility, it remains unclear whether *X. fastidiosa* exhibits chemotactic movement toward or away from particular compounds. It is intriguing to consider that the host range or virulence of *X. fastidiosa* might be modulated by the differential presence of plant compounds that might anticipate in regulation of type IV pilus function.

The role of extracellular polysaccharide (EPS) production in the behaviors of *X. fastidiosa* is not completely understood. EPS production is an important virulence factor in many *Xanthomonas* species. *X. fastidiosa* contains homologues of many but not all of the so-called *gum* genes necessary for EPS production in *Xanthomonas* species (Simpson et al. 2000; van Sluys et al. 2003). It therefore has been suggested that *X. fastidiosa* makes an EPS molecule similar to that of xanthan gum made by *Xanthomonas campestris*, but one which is lacking terminal mannosyl residues (Silva et al. 2001). Antibodies that could recognize the EPS produced by *X. fastidiosa* were recently used to illustrate its production both in culture and *in*

planta (Roper et al. 2007). Generation of *gum* mutants (Killiny et al. 2013) recently helped to better characterize the role of EPS in *X. fastidiosa* virulence. In addition to their reduced capacity to form biofilms in culture, *gum* mutants, once mechanically introduced in plants, were avirulent and did not result in the development of any symptoms in grapevines (Killiny et al. 2013). These traits were also associated with an apparent lack of plant colonization and an altered motility compared to that of the wild-type cells. In addition, EPS was also shown having an important role in vector transmission, probably due to its implication in biofilm formation, essential step for a successful insect colonization.

Xylella fastidiosa appears to have a surprisingly large number of adhesin and hemagglutinin-encoding genes compared to other bacteria (Simpson et al. 2000; van Sluys et al. 2003), and such adhesins appear to play a central role in its complex lifestyle. Initial results of studies of virulence factors in *X. fastidiosa* involving the screening of random insertional mutants revealed a surprising role for adhesins. Specifically, *hxfA* and *hxfB* mutants that no longer expressed these two related hemagglutinin-like proteins exhibited a hyper-virulent phenotype in grape (Guilhabert and Kirkpatrick 2005). Furthermore, these mutants exhibited reduced cell–cell aggregation and moved further in grape xylem vessels in the wild-type strain after inoculation (Guilhabert and Kirkpatrick 2005). Because such proteins would be expected to facilitate both cell–cell aggregation as well as cell surface attachment, it would be expected that their presence would tend to reduce the virulence of *X. fastidiosa* by impeding its movement along xylem vessels. Further evidence for the negative effects of a fimbrial adhesins on the virulence of *X. fastidiosa* has come from studies of various mutants altered in DSF-mediated cell–cell signaling. For example, *rpfF* mutants, which are blocked in the accumulation of DSF, express a variety of adhesins such as HxfA, HxfB, FimA, and XadA at a much lower level than that of the wild-type strain (Chatterjee et al. 2008b; Wang et al. 2012). The hyper-virulence of the *rpfF* mutant could be conferred by the lower level of these adhesins, accounting for the ability of such mutant cells to move more extensively within plants (Chatterjee et al. 2008c). Further support for the role of adhesins as anti-virulence factors comes from studies of *rpfC* mutants which accumulate excessive amounts of DSF and *cgsA* mutants of *X. fastidiosa* which are predicted to have relatively low concentrations of the intracellular signaling molecule cyclic di-GMP (Chatterjee et al. 2008b, 2010). In both mutant backgrounds, the expression of various fimbrial and afimbrial adhesins is higher than in a wild-type strain, accounting for their relatively low ability of these mutants to move throughout plants and therefore to cause disease symptoms (Chatterjee et al. 2010). In addition to the hemagglutinin-like proteins and other fimbrial and afimbrial adhesins that have been noted to affect the adhesiveness and cell–cell aggregation capabilities of *X. fastidiosa*, LPS also may play a similar role. Disruption of PD0914, which encodes a Wzy polymerase involved in biosynthesis of a high molecular weight O-antigen in the LPS of *X. fastidiosa*, led to measurable differences in the structure of biofilms formed by such a mutant as well as reducing its virulence (Clifford et al. 2013). It is also fascinating to find that several afimbrial adhesins produced by *X. fastidiosa* can be found in the extracellular milieu of this pathogen. For example, the adhesin XadA1 was found associated not only with the surface of *X. fastidiosa* cells, but was also often found in intercellular spaces of bacterial aggregates (Caserta et al. 2010). In contrast, the related molecule XadA2 was found associated only with intact cells (Caserta et al. 2010). Curiously, a portion of the population of the adhesins HxfA and HxfB produced by *X. fastidiosa* in culture were found in culture supernatants, presumably associated with membranous vesicles (Voegel et al. 2010). The role of such extracellular forms of these adhesins is unknown, as is the identity of any factors that may control the apparent release of these molecules from the producing cell.

It is probably quite significant that several different regulatory systems have been found to coordinate expression of various fimbrial and

afimbrial adhesins in *X. fastidiosa*. Given that adhesins almost certainly play a strong but contextual role in the biology of *X. fastidiosa*, their proper temporal and spatial expression would be essential for this pathogen to maintain optimal fitness in its plant and insect niches. The response regulator XhpT, composed of a receiver domain and a histidine phosphotransferase output domain, was found to control surface attachment, cell–cell aggregation, and EPS production as well as virulence of *X. fastidiosa* (Voegel et al. 2013). While any signal that this response regulator might perceive remains unknown, it is intriguing to consider that it might be involved in habitat-specific expression of those genes such as *hxfA*, *hxfB*, and *tonB* which, as discussed above, play central roles in the behavior of *X. fastidiosa* in different settings. The global regulator GacA was also found to positively regulate several virulence factors in *X. fastidiosa* including the adhesins XadA and Hsf (Shi et al. 2009). GacA has been found to control a variety of physiological processes as well as pathogenicity factors in many other gram-negative bacteria (Heeb and Haas 2001). As such, it is thought to play an important coordinating role in context-dependent expression of virulence factors. The suppression of adhesins in *X. fastidiosa* provides further evidence that the proper temporal and spatial expression of these molecules is important in the context-dependent behavior of this pathogen.

8.4.2.1 Vector Colonization

The chemical composition of the external layer of an insect's exoskeleton (the epicuticle) is not well understood for several insect groups; and to our knowledge, there is no information on its composition for *X. fastidiosa* insect vectors. The epicuticle of insects is composed of several layers: the inner and outer epicuticle covered by a wax layer; and in some insects an additional cement layer also exists above this wax layer. The thickest layer, the inner epicuticle, is 0.5–2 μm thick (Chapman 1998). The cement layer is a thin layer composed of mucopolysaccharides associated with lipids. The wax layer is largely composed of lipids with embedded proteins, and serves as a waterproofing element for the cuticle. In addition, proteinaceous molecules would also be present in the cuticle, along with other potential molecules. Although the chemical composition of the cuticular surface of arthropods is generally not well understood, interactions between bacteria and the exoskeletons have been successfully studied in other systems using chitin as a proxy, as this system mimics bacterial behavior on the surface of actual hosts (Tarsi and Pruzzo 1999).

The first step in determining the nature of *X. fastidiosa*-vector interactions was to learn whether *X. fastidiosa* surface proteins were involved in cell adhesion to vectors. Killiny and Almeida (2009a) demonstrated that *X. fastidosa* cells bind to carbohydrates, and that treating intact cells with proteases reduced adhesion to compounds such as chitin. Thus, surface proteins were involved in cell adhesion to carbohydrates; however, *X. fastidiosa* had variable affinity to different molecules. For example, competition assays showed that *N*-acetylglucosamine (GlcNAc, the monomer of chitin) acted as a strong competitor in binding assays where vector foregut extracts were used as a substrate, reducing cell adhesion. On the other hand, mannose and galactose did not affect binding. In addition, to look for specific *X. fastidiosa* proteins involved in adhesion, several mutants were tested in vitro for binding to foregut extracts (Killiny and Almeida 2009a). Only hemagglutinin and cell–cell signaling mutants were affected in adhesion. Altogether, these biochemical and other biological assays indicated that initial cell adhesion to vectors is mediated by carbohydrate–lectin interactions and that specific *X. fastidiosa* surface proteins can be identified in vitro as potential candidates for more comprehensive studies.

Since it was first cultured in the laboratory, attempts to deliver *X. fastidiosa* cells to vectors from growth media have been unsuccessful (Davis et al. 1978), until it was discovered that plant structural polysaccharides result in phenotypic changes in *X. fastidiosa,* inducing its transmissibility by leafhoppers (Killiny and

Almeida 2009b). Plant polysaccharides induce phenotypic changes leading to higher degrees of adhesiveness and, consequently, attachment to vectors after acquisition from plants (Killiny and Almeida 2009b). The pattern of gene expression of *X. fastidiosa* exposed to pectin-supplemented media is also similar to that of *X. fastidiosa* cells occurring at high cell density (Newman et al. 2004; Chatterjee et al. 2008b).

The gene expression profile is also modified when media is supplemented with chitin, the main component of the insect foregut surface (Killiny et al. 2010). Adhesins involved in *X. fastidiosa* initial adhesion to insect cells are upregulated, and biofilm formation is also enhanced on chitinous surfaces. In addition, a functional chitinase (ChiA; Killiny et al. 2010) was recently discovered, as part of the machinery used by *X. fastidiosa* to degrade and assimilate chitin as its sole carbon source. Molecular mechanisms involved in chitin utilization are not completely understood. Until now, no known chitin-binding domains have been found in that enzyme, and the implication of other *X. fastidiosa* chitin-binding proteins are likely to be involved in the process (Labroussaa and Almeida, unpublished).

8.4.2.2 Host Switching and Cell–Cell Signaling

One intriguing aspect of *X. fastidiosa*'s biology unique among xanthomonads is that it is restricted to colonizing two very different, yet highly specialized environments—plants and insect vectors. Water conducting xylem vessels of plants and the foregut of vectors are extremely different environments. Cells of *X. fastidiosa* are attached to surfaces in the insect foregut and experience very rapid fluid flow (estimated > 5 cm/s, Purcell et al. 1979) caused by a powerful pumping system used by insects to suck sap under negative tension; turbulence is created in the mouthparts once every second when leafhoppers pull sap from plants and push it into the gut (Dugravot et al. 2008; Purcell and Finlay 1979). Sap flow conditions are not nearly as extreme in xylem vessels, where flow was calculated to achieve ca. 1×10^{-2} to 1×10^{-4} cm/s inside grapes growing in the field (Andersen and Brodbeck 1989; Greenspan et al. 1996). The dramatic differences in the flow speed of different environments (100 to 10,000 times faster inside insect foregut compared to xylem in plants) will have an impact on bacterial cell attachment to surfaces and to each other, since they need to overcome the external shear force stress. In plants, *X. fastidiosa* multiplies within individual vessels and moves actively to adjacent vessels in the xylem network by producing enzymes that degrade pit membranes separating individual xylem vessels (Perez-Donoso et al. 2010). In plant xylem vessels, *X. fastidiosa* also has a larger surface area to colonize compared to the foregut of leafhoppers (Newman et al. 2003).

In addition to responding to host or environment specific cues, many bacteria also utilize highly specific quorum-sensing signals that induce gene expression above a certain concentration threshold. By using density-dependent signaling, populations of bacteria can quickly coordinate the expression of metabolically costly traits in response to environmental cues that would be ineffective if expressed by low populations of cells. The role of cell–cell signaling toward host colonization is often not well documented (Bassler and Losick 2006) but is implicated in expression of colonization and virulence traits for an increasing number of plant and animal pathogenic bacteria (Ham 2013; Ryan and Dow 2008).

Like many other bacteria, *X. fastidiosa* has a regulatory system that responds to signal molecules produced by individual cells in a population; these diffusible molecules accumulate in the environment and trigger population-wide phenotypic changes, likely associated with global changes in gene expression when the signal threshold is reached (Wang et al. 2012). Many traits implicated in insect and host plant colonization are also under control of the regulation of pathogenicity factor (*rpf*) operon. In *Xanthomonas* ssp. and *X. fastidiosa*, *rpf* controls synthesis and detection of DSF, a medium-chain fatty acid that functions as a signaling molecule (Colnaghi Simionato et al. 2007; Dow et al.

2003). The *rpf* cluster in *Xanthomonas* ssp. contains 12 genes, 9 of which are conserved in *X. fastidiosa*. In both genera, *rpfF* produces DSF, which diffuses freely through cell membranes and is sensed by a two-component transmembrane receptor RpfC and RpfG (Barber et al. 1997; Newman et al. 2004), but many of the genes in this operon remain functionally uncharacterized.

The *Xanthomonas campestris* pv. *campestris* DSF molecule was described as cis-11-methyl-2-dodecanoic acid (Barber et al. 1997; Wang et al. 2004); while *X. fastidiosa* can respond to *Xanthomonas* ssp. DSF, the weaker response (Newman et al. 2004) suggested that *X. fastidiosa* synthesizes and responds to a distinct DSF molecule. Recently, the structure of an *X. fastidiosa* DSF was identified as 2(Z)-tetradecanoic acid and was shown to control DSF-dependent traits in *X. fastidiosa*, including biofilm formation and attachment to surfaces (Beaulieu et al. 2013).

Experiments with DSF deficient and blind mutants have shown that DSF signals are important in regulating *X. fastidiosa* phenotypes in plant and insect hosts. Disruption of DSF production (in an *rpfF* knockout mutant) results in hyper-virulence within plants, possibly due to up-regulation of plant colonization-related genes and down-regulation of adhesins (Newman et al. 2004; Wang et al. 2012) that decrease rates of both self-aggregation and attachment to xylem vessels. However, the *rpfF* mutant is not capable of colonizing the precibarium of vectors and is very poorly transmitted to plants. It appears that the induction of a vector-transmissible state occurs due to up-regulation of genes under control of the *rpf* system.

In contrast, DSF-blind *rpfC* mutants overproduce *hxfA* and *hxfB*. This 'stickier' phenotype is less virulent than wild type in planta due to reduced movement through and colonization of xylem. While *rpfC* is not impaired in attachment to vector foreguts, transmission is reduced due to increased adhesin expression and lower rates of cell detachment (Chatterjee et al. 2008b). In *X. campestris*, DSF regulates the production of an enzyme that controls cell dispersal from a biofilm (Dow et al. 2003). The inability of *rpfC* to detach from the biofilm in the insect foregut implicates DSF signaling as an important but uncharacterized regulator of expression controlling cell detachment from biofilm bound cells in the insect foregut. Consequently, afimbrial adhesins are over-expressed, while genes associated with plant host colonization are down-regulated (see Chatterjee et al. 2008a for discussion). A complex picture of *X. fastidiosa* gene regulation in relation to vector transmission is emerging based on this research.

Although adhesion is essential for retention in insects, it limits colonization of plants compared to the *rpfF* mutant, highlighting the distinct requirements for life in such different hosts (Newman et al. 2004; Guilhabert and Kirkpatrick 2005; Chatterjee et al. 2008c). Thus, *X. fastidiosa*'s conflicting life history is framed by the contrasting requirement to move within plants to increase its population size and thus its chances of being acquired by insects, but at the same time, it must increase its adhesiveness so it can attach to insects, which consequently reduces within-plant movement.

Recent characterization of *rpfB* in both the rice blight pathogen *Xanthomonas oryzae* pv. *oryzae* (*Xoo*) (He et al. 2010) and *X. fastidiosa* (Almeida et al. 2012) have demonstrated that both species produce at least three distinct DSF signals, that they accumulate at different cell densities, and that *rpfB* has a role in DSF processing (Almeida et al. 2012). In *X. fastidiosa* and *Xoo*, an *rpfB* mutant only produced one of these signals, indicating a regulatory role in DSF processing for RpfB (Almeida et al. 2012). *rpfB* is spatially separated from the *rpf* operon in the *X. fastidiosa* genome compared to *Xanthomonas* ssp., suggesting a more complex role for *rpfB* in *X. fastidiosa*, which is vector transmitted compared to *Xanthomonas* ssp., which are not. The production of multiple DSF signals in wild-type cells and the ability of *rpfB* mutants, which produce only one of these signals, to adhere to but not colonize insect suggests the intriguing but unexplored possibility that host cues influence the production of different ratios of DSF to

modulate complex host colonization behaviors in alternating plant and insect hosts (Almeida et al. 2012).

8.5 Disease Management Strategies

The availability of a collection of approaches to control *X. fastidiosa* diseases is highly desirable, as integration of various strategies will likely be necessary to sustainably control this pathogen. Unfortunately, disease spread in these systems is controlled primarily through the extensive use of pesticides, which often have negative short- and long-term consequences to integrated disease and pest management. In addition, this approach has obvious negative impacts on the environment and communities that rely on agricultural activities. Alternative, efficient, pathogen-specific, environmentally friendly, and safe approaches to control these diseases would lead to long-term sustainability of crop systems. Strategies to control *X. fastidiosa* currently in development, based on genomics-derived knowledge and the production of transgenic plants, focus on either the pathogen (e.g., Chatterjee et al. 2008c; Dandekar et al. 2012) or the vector (such as RNAi to impact insect development, not discussed here; Rosa et al. 2012), or on both partners and their interactions during insect transmission (Killiny et al. 2012). We limit our discussion to these approaches, although it should be mentioned that genomic data has provided great insights into the biology, ecology, evolution, and taxonomy of *X. fastidiosa*, much of which has been useful to devise other disease management strategies and for detection and quarantine purposes.

First, the DSF signaling system of *X. fastidiosa* has been used as the basis to develop a confusion strategy, where presence of DSF in the environment (i.e., xylem stream) at all times should limit bacterial movement and plant colonization. Chatterjee et al. (2008c) showed that the *rpfC* mutant strain is indeed limited within plant movement and multiplication, suggesting that the presence of DSF functions as a suppressor of population growth. If constitutively expressed in transgenic plants, DSF molecules should lead to adherent *X. fastidiosa* populations with limited movement within plants and, as a result, the expression of disease symptoms should be reduced.

Another strategy concerns the utilization of cecropin A or B as antibacterial peptides for limiting the growth of *X. fastidiosa* (Ishida et al. 2004). Recently, the introduction of a construct carrying such a peptide into transgenic grapevines, allowing its specific expression into the xylem of plants, was efficient protecting grapes against the development of Pierce's disease (Dandekar et al. 2012). Other concepts leading to the constitutive expression of exogenous proteins in transgenic plants include the expression of polygalacturonase-inhibiting proteins (PGIPs) that inhibit *X. fastidiosa* polygalacturonase (*pglA*) responsible for the systemic movement of *X. fastidiosa* in plants (Agüero et al. 2005). Strategies may also emerge from research on the identification of xylem compounds produced during plant exposure to low temperatures (Wilhelm et al. 2011; Meyer and Kirkpatrick 2011). Grapevines with *X. fastidiosa* and expressing symptoms of Pierce's disease are cured of infections when subject to cold winters (Purcell 1977, 1980), via a yet to be determined mechanism.

Lastly, a strategy following the hypothesis that blocking interactions between both partners occurring during initial adhesion could lead to the impossibility for the bacteria to successfully colonize its vectors is also being pursued (Killiny et al. 2012). *X. fastidiosa* proteins identified as able to bind to insect receptors will be used as transmission-blocking molecules. Once expressed in grapevines, those molecules will compete for insect receptors with *X. fastidiosa* preventing the attachment of the bacteria on insect cells, essential step for its transmission and consequently, disrupting bacteria transmission from plant to plant.

8.6 Conclusion

The availability of *X. fastidiosa* genome sequences have allowed research on this fastidious organism to move significantly faster than in the past. In addition to improving diagnostic tools and promoting research on the functional role of genes, which has led to the incipient technologies briefly discussed here, it has provided insights into its evolution that would not have been possible otherwise. Because advances in technology will continue to make sequence data available at a larger scale with diminishing costs, we foresee that it will become an integral part of *X. fastidiosa* research. The complete integration of this tool with others now routinely used in plant pathology is yet to be realized, but efforts to use genome sequences have already generated exciting findings and will continue to do so.

Acknowledgments We thank students, researchers, and colleagues that have contributed to the work discussed here. ACR is supported by a postdoctoral fellowship from the Miller Institute for Basic Research in Science. LRS is supported by a postdoctoral fellowship in biology from the NSF.

References

Agüero CB, Uratsu SL, Greve C et al (2005) Evaluation of tolerance to Pierce's disease and *Botrytis* in transgenic plants of *Vitis vinifera* L. expressing the pear PGIP gene. Mol Plant Pathol 6:43–51

Almeida RPP, Backus EA (2004) Stylet penetration behaviors of *Graphocephala atropunctata* (Signoret) (Hemiptera, Cicadellidae): EPG waveform characterization and quantification. Ann Entomol Soc Am 97:838–851

Almeida RPP, Purcell AH (2003a) Transmission of *Xylella fastidiosa* to grapevines by *Homalodisca coagulata* (Hemiptera: Cicadellidae). J Econ Entomol 96:264–271

Almeida RPP, Purcell AH (2003b) Biological traits of *Xylella fastidiosa* strains from grapes and almonds. Appl Environ Microbiol 69:7447–7452

Almeida RPP, Purcell AH (2006) Patterns of *Xylella fastidiosa* colonization on the precibarium of sharpshooter vectors relative to transmission to plants. Ann Entomol Soc Am 99:884–890

Almeida RPP, Blua MJ, Lopes JRS et al (2005) Vector transmission of *Xylella fastidiosa*: applying fundamental knowledge to generate disease management strategies. Ann Entomol Soc Am 98:775–786

Almeida RPP, Nascimento FE, Chau J et al (2008) Genetic structure and biology of *Xylella fastidiosa* strains causing disease in citrus and coffee in Brazil. Appl Environ Microbiol 74:3690–3701

Almeida RPP, Killiny N, Newman KL et al (2012) Contribution of *rpfB* to cell-to-cell signal synthesis, virulence, and vector transmission of *Xylella fastidiosa*. Mol Plant Microbe Interact 25:453–462

Andersen P, Brodbeck B (1989) Temperature and temperature preconditioning on flux and chemical composition of xylem exudate from muscadine grapevines. J Am Soc Hort Sci 114:440–444

Baccari C, Lindow SE (2011) Assessment of the process of movement of *Xylella fastidiosa* within susceptible and resistant grape cultivars. Phytopathology 101:77–84

Backus EA, Habibi J, Yan FM et al (2005) Stylet penetration by adult *Homalodisca coagulata* on grape: Electrical penetration graph waveform characterization, tissue correlation, and possible implications for transmission of *Xylella fastidiosa*. Ann Entomol Soc Am 98:787–813

Barber CE, Tang JL, Feng JX et al (1997) A novel regulatory system required for pathogenicity of *Xanthomonas campestris* is mediated by a small diffusible signal molecule. Mol Microbiol 24:555–566

Bassler BL, Losick R (2006) Bacterially speaking. Cell 125:237–246

Beaulieu ED, Ionescu M, Chatterjee S et al (2013) Characterization of a diffusible signaling factor from *Xylella fastidiosa*. mBio 4:e00539-12

Berisha B, Chen YD, Zhang GY et al (1998) Isolation of Peirce's disease bacteria from grapevines in Europe. Eur J Plant Pathol 104:427–433

Bhattacharyya A, Stilwagen S, Ivanova N et al (2002) Whole-genome comparative analysis of three phytopathogenic *Xylella fastidiosa* strains. Proc Natl Acad Sci USA 99:12403–12408

Blom J, Albaum SP, Doppmeier D et al (2009) EDGAR: a software framework for the comparative analysis of prokaryotic genomes. BMC Bioinf 10:154

Brlansky RH, Timmer LW, French WJ et al (1983) Colonization of the sharpshooter vectors, *Oncometopia nigricans* and *Homalodisca coagulata,* by xylem-limited bacteria. Phytopathology 73:530–535

Caserta R, Takita MA, Targon ML et al (2010) Expression of *Xylella fastidiosa* fimbrial and afimbrial proteins during biofilm formation. Appl Environ Microbiol 76:4250–4259

Chapman RF (1998) The insects, structure and function, 4th edn. Cambridge University Press, Cambridge, p 770

Chatelet DS, Wistrom CM, Purcell AH et al (2011) Xylem structure of four grape varieties and 12 alternative hosts to the xylem-limited bacterium *Xylella fastidious*. Ann Bot 108:73–85

Chatterjee S, Almeida RPP, Lindow SE (2008a) Living in two worlds: the plant and insect lifestyles of *Xylella fastidiosa*. Ann Rev Phytopathol 46:243–271

Chatterjee S, Wistrom C, Lindow SE (2008b) A cell–cell signaling sensor is required for virulence and insect

transmission of *Xylella fastidiosa*. Proc Natl Acad Sci USA 105:2670–2675

Chatterjee S, Newman KL, Lindow SE (2008c) Cell-to-cell signaling in *Xylella fastidiosa* suppresses movement and xylem vessel colonization in grape. Mol Plant Microbe Interact 21:1309–1315

Chatterjee S, Killiny N, Almeida RPP et al (2010) Role of cyclic di-GMP in *Xylella fastidiosa* biofilm formation, plant virulence, and insect transmission. Mol Plant Microbe Interact 23:1356–1363

Chen J, Xie G, Han S et al (2010) Whole genome sequences of two *Xylella fastidiosa* strains (M12 and M23) causing almond leaf scorch disease in California. J Bacteriol 192:4534

Clifford JC, Rapicavoli JN, Roper MC (2013) A rhamnose-rich O-antigen mediates adhesion, virulence, and host colonization for the xylem-limited phytopathogen *Xylella fastidiosa*. Mol Plant Microbe Interact 26:676–685

Coletta-Filho HD, Takita MA, De Souza AA et al (2001) Differentiation of strains of *Xylella fastidiosa* by a variable number of tandem repeat analysis. Appl Environ Microbiol 67:4091–4095

Coletta-Filho HD, Bittleston LS, Almeida RPP (2011) Spatial genetic structure of a vector-borne generalist pathogen. Appl Environ Microbiol 77:2596–2601

Colnaghi Simionato AV, da Silva DS, Lambais M et al (2007) Characterization of a putative *Xylella fastidiosa* diffusible signal factor by HRGC-EI-MS. J Mass Spect 42:1375–1381

Comas I, Moya A, Gonzales-Candelas F (2007) From phylogenetics to phylogenomics: the evolutionary relationships of insect endosymbiotic & #x03B3;-proteobacteria as a test case. Syst Biol 56:1–16

Cursino L, Li Y, Zaini PA et al (2009) Twitching motility and biofilm formation are associated with *tonB1* in *Xylella fastidiosa*. FEMS Microbiol Lett 299:193–199

Cursino L, Galvani CD, Athinuwat D et al (2011) Identification of an operon, Pil-Chp, that controls twitching motility and virulence in *Xylella fastidiosa*. Mol Plant Microbe Interact 24:1198–1206

Damsteegt VD, Brlansky RH, Phillips PA et al (2006) Transmission of *Xylella fastidiosa*, causal agent of citrus variegated chlorosis, by the glassy-winged sharpshooter, *Homalodisca coagulata*. Plant Dis 90:567–570

Dandekar AM, Gouran H, Ibáñez AM et al (2012) An engineered innate immune defense protects grapevines from Pierce disease. Proc Natl Acad Sci USA 109:3721–3725

Daugherty MP, Almeida RPP (2009) Estimating *Xylella fastidiosa* transmission parameters: decoupling sharpshooter number and feeding period. Entomol Exp Appl 132:84–92

Daugherty MP, Lopes JRS, Almeida RPP (2010) Vector within-host feeding preference mediates transmission of a heterogeneously distributed pathogen. Ecol Entomol 35:360–366

Davis MJ, Purcell AH, Thomson SV (1978) Pierce's disease of grapevines: isolation of the causal bacterium. Science 199:75–77

De La Fuente L, Burr TJ, Hoch HC (2007) Mutations in Type I and Type IV pilus biosynthetic genes affect twitching motility rates in *Xylella fastidiosa*. J Bacteriol 189:7507–7510

Doi Y, Teranaka M, Yora K et al (1967) Mycoplasma or PLT group-like microorganisms found in the phloem elements of plants infected with mulberry dwarf, potato witches' broom, aster yellows, or *Paulownia* witches' broom. Ann Phytopath Soc Jpn 33:259–266

Dow JM, Crossman L, Findlay K et al (2003) Biofilm dispersal in *Xanthomonas campestris* is controlled by cell-cell signaling and is required for full virulence to plants. Proc Natl Acad Sci USA 100:10995–11000

Dugravot S, Backus EA, Reardon BJ et al (2008) Correlations of cibarial muscle activities of *Homalodisca* ssp. sharpshooters (Hemiptera: Cicadellidae) with EPG ingestion waveform and excretion. J Insect Physiol 54:1467–1478

Esnault E, Valens M, Espéli O, Boccard F (2007) Chromosome structuring limits genome plasticity in Escherichia coli. PLoS Genet 3:e226

European and Mediterranean Plant Protection Organization (2004) Diagnostic protocols for regulated pests: *Xylella fastidiosa*. OEPP/EPPO Bulletin, pp 155–157

Federal Register (2012) Agricultural bioterrorism protection act of 2002; biennial review and republication of the select agent and toxin list; amendments to the select agent and toxin regulations

Frazier NW (1965) Xylem viruses and their insect vectors. Proceedings international conference on virus and vector on perennial hosts, with special reference to Vitis. University of California, Division of Agricultural Sciences, Davis, California, pp 91–99

Freitag JH (1951) Host range of Pierce's disease virus of grapes as determined by insect transmission. Phytopathology 41:920–934

Fry SM, Milholland RD (1990) Response of resistant, tolerant, and susceptible grapevine tissues to invasion by the Pierce's disease bacterium, *Xylella fastidiosa*. Phytopathology 80:66–69

Fuller KB (2012) The economics of Pierce's Disease in the California winegrape industry. University of California, Davis, PhD, Dissertation

Goodwin PH, Zhang S (1997) Distribution of *Xylella fastidiosa* in southern Ontario as determined by the polymerase chain reaction. Can. J. Plant Pathol. 19:13–18

Greenspan MD, Schultz HR, Matthews MA (1996) Field evaluation of water transport in grape berries during water deficits. Physiol Plant 97:55–62

Guilhabert MR, Kirkpatrick BC (2003) Transformation of *Xylella fastidiosa* with broad host range RSF1010 derivative plasmids. Mol Plant Pathol 4:279–285

Guilhabert MR, Kirkpatrick BC (2005) Identification of *Xylella fastidiosa* antivirulence genes: hemagglutinin adhesins contribute to *X. fastidiosa* biofilm maturation and colonization and attenuate virulence. Mol Plant Microbe Interact 18:856–868

Guilhabert MR, Stewart VJ, Kirkpatrick BC (2006) Characterization of putative rolling-circle plasmids

from the Gram-negative bacterium *Xylella fastidiosa* and their use as shuttle vectors. Plasmid 55:70–80

Ham JJ (2013) Intercellular and intracellular signalling systems that globally control the expression of virulence genes in plant pathogenic bacteria. Mol Plant Pathol 14:308–322

He YW, Wu J, Cha JS, Zhang LH (2010) Rice bacterial blight pathogen *Xanthomonas oryzae pv. oryzae* produces multiple DSF-family signals in regulation of virulence factor production. BMC Microbiol 10:187

Heeb S, Haas D (2001) Regulatory roles of the GacS/GacA two-component system in plant-associated and other gram-negative bacteria. Mol Plant Microbe Interact 14:1351–1363

Hendson M, Purcell AH, Chen D et al (2001) Genetic diversity of Pierce's disease strains and other pathotypes of *Xylella fastidiosa*. Appl Environ Microbiol 67:895–903

Hewitt WB (1958) The probable home of Pierce's disease virus. Am J Enol Vitic 9:94–98

Hewitt WB, Frazier NW, Freitag JH (1949) Pierce's disease investigations. Hilgardia 19:207–264

Hill BL, Purcell AH (1995a) Multiplication and movement of *Xylella fastidiosa* within grapevine and four other plants. Phytopathology 85:1368–1372

Hill BL, Purcell AH (1995b) Acquisition and retention of *Xylella fastidiosa* by an efficient vector, *Graphocephala atropunctata*. Phytopathology 85:209–212

Hill BL, Purcell AH (1997) Populations of *Xylella fastidiosa* in plants required for transmission by an efficient vector. Phytopathology 87:1197–1201

Hopkins DL (1985) Physiological and pathological characteristics of virulent and avirulent strains of the bacterium that causes Pierce's disease of grapevine. Phytopathology 75:713–717

Hopkins DL (1989) *Xylella fastidiosa*: Xylem-limited bacterial pathogen of plants. Annu Rev Phytopathol 27:271–290

Hopkins DL, Mollenhauer HH (1973) Rickettsia-like bacterium associated with Pierce's disease of grapes. Science 179:298–300

Hopkins DL, Purcell AH (2002) *Xylella fastidiosa*: cause of Pierce's disease of grapevine and other emergent diseases. Plant Dis 86:1056–1066

Ishida ML, Andersen PC, Leite B (2004) Effect of *Vitis vinifera* L. cv. Chardonnay xylem fluid on cecropin B activity against *Xylella fastidiosa*. Mol Plant Pathol 64:73–81

Killiny N, Almeida RPP (2009a) *Xylella fastidiosa* afimbrial adhesins mediate cell transmission to plants by leafhopper vectors. Appl Environ Microbiol 75:521–528

Killiny N, Almeida RPP (2009b) Host structural carbohydrate induces vector transmission of a bacterial plant pathogen. Proc Natl Acad Sci USA 106:22416–22420

Killiny N, Prado SS, Almeida RPP (2010) Chitin utilization by the insect-transmitted bacterium *Xylella fastidiosa*. Appl Environ Microbiol 76:6134–6140

Killiny N, Almeida RPP (2011) Gene regulation mediates host specificity of a bacterial pathogen. Environ Microbiol Rep 3:791–797

Killiny N, Rashed A, Almeida RPP (2012) Disrupting the transmission of a vector-borne plant pathogen. Appl Environ Microbiol 78:638–643

Killiny N, Hernandez Martinez R, Dumenyo CK et al (2013) The exopolysaccharide of *Xylella fastidiosa* is essential for biofilm formation, plant virulence and vector transmission. Mol Plant Microbe Interact 26:1044–1053

Kono N, Arakawa K, Tomita M (2011) Comprehensive prediction of chromosome dimer resolution sites in bacterial genomes. BMC Genom 12:19

Kovach M, Elzer P, Hill D et al (1995) Four new derivatives of the broad-host-range cloning vector pBBR1MCS, carrying different antibiotic-resistance cassettes. Gene 166:175–176

Krell RK, Boyd EA, Nay JE et al (2007) Mechanical and insect transmission of *Xylella fastidiosa* to *Vitis vinifera*. Am J Enol Vitic 58:211–216

Krivanek AF, Walker MA (2005) *Vitis* resistance to Pierce's disease is characterized by differential *Xylella fastidiosa* populations in stems and leaves. Phytopathology 95:44–52

Krugner R, Sisterson MS, Lin H (2012) Effects of gender, origin, and age on transmission of *Xylella fastidiosa* to grapevines by *Homalodisca vitripennis* (Hemiptera: Cicadellidae). Ann Entomol Soc Am 105:280–286

Kung SH, Almeida RPP (2011) Natural competence and recombination in the plant pathogen *Xylella fastidiosa*. Appl Environ Microbiol 77:5278–5284

Kung SH, Retchless AC, Kwan JY et al (2013) Effects of DNA size on transformation and recombination efficiencies in *Xylella fastidiosa*. Appl Environ Microbiol 79:1712–1717

Lee MW, Rogers EE, Stenger DC (2010) Functional characterization of replication and stability factors of an incompatibility group P-1 plasmid from *Xylella fastidiosa*. Appl Environ Microbiol 76:7734–7740

Lee MW, Rogers EE, Stenger DC (2012) *Xylella fastidiosa* plasmid-encoded PemK toxin is an endoribonuclease. Phytopathology 102:32–40

Lindeberg M, Cunnac S, Collmer A (2009) The evolution of *Pseudomonas syringae* host specificity and type III effector repertoires. Mol Plant Pathol 10:767–775

Lopes JRS, Daugherty MP, Almeida RPP (2009) Context-dependent transmission of a generalist plant pathogen: host species and pathogen strain mediate insect vector competence. Entomol Exp Appl 131:216–224

Lopes SA, Marcussi S, Torres SCZ et al (2003) Weeds as alternative hosts of the citrus, coffee, and plum strains of *Xylella fastidiosa* in Brazil. Plant Dis 87:544–549

Ma W, Dong FFT, Stavrinides J et al (2006) Type III effector diversification via both pathoadaptation and horizontal transfer in response to a coevolutionary arms race. PLoS Genet 2:e209

Mackiewicz P, Zakrzewska-Czerwinska J, Zawilak A et al (2004) Where does bacterial replication start?

Rules for predicting the *oriC* region. Nucleic Acids Res 32:3781–3791

Matsumoto A, Young GM, Igo MM (2009) Chromosome-based genetic complementation system for *Xylella fastidiosa*. Appl Environ Microbiol 75:1679–1687

Mei J, Benashski S, Firshein W (1995) Interactions of the origin of replication (*oriV*) and initiation proteins (TrfA) of plasmid RK2 with submembrane domains of *Escherichia coli*. J Bacteriol 177:6766–6772

Meidanis J, Braga MDV, Verjovski-Almeida S (2002) Whole-genome analysis of transporters in the plant pathogen *Xylella fastidiosa*. Microbiol Mol Biol Rev 66:272–299

Meng Y, Li Y, Galvani CD et al (2005) Upstream migration of *Xylella fastidiosa* via pilus-driven twitching motility. J Bacteriol 187:5560–5567

Meyer MM, Kirkpatrick BC (2011) Exogenous applications of abscisic acid increase curing of Pierce's disease-affected grapevines growing in pots. Plant Dis 95:173–177

Minsavage GV, Thompson CM, Hopkins DL et al (1994) Development of a polymerase chain reaction protocol for detection of *Xylella fastidiosa* in plant tissue. Phytopathology 84:456–461

Monteiro PB, Teixeira DC, Palma RR et al (2001) Stable transformation of the *Xylella fastidiosa* citrus variegated chlorosis strain with *oriC* plasmids. Appl Environ Microbiol 67:2263–2269

Monteiro-Vitorello CB, De Oliveira MC, Zerillo MM et al (2005) *Xylella* and *Xanthomonas* mobil'omics. OMICS 9:146–159

Moran NA, Plague GR (2004) Genomic changes following host restriction in bacteria. Curr Opin Genet Dev 14:627–633

Moreira LM, De Souza RF, Almeida NF Jr, Setubal JC, Oliveira JC, Furlan LR, Ferro JA, da Silva AC (2004) Comparative genomics analyses of citrus-associated bacteria. Annu Rev Phytopathol 42:163–184

Moreira LM, De Souza RF, Digiampietri LA, Da Silva AC, Setubal JC (2005) Comparative analyses of *Xanthomonas* and *Xylella* complete genomes. OMICS 9:43–76

Newman KL, Almeida RPP, Purcell AH et al (2003) Use of a green fluorescent strain for analysis of *Xylella fastidiosa* colonization of *Vitis vinifera*. Appl Environ Microbiol 69:7319–7327

Newman KL, Almeida RPP, Purcell AH et al (2004) Cell-cell signaling controls *Xylella fastidiosa* interactions with both insects and plants. Proc Natl Acad Sci USA 101:1737–1742

Nunes LR, Rosato YB, Muto NH et al (2003) Microarray analyses of *Xylella fastidiosa* provide evidence of coordinated transcription control of laterally transferred elements. Genome Res 13:570–578

Nunney L, Yuan X, Bromley R et al (2010) Population genomic analysis of a bacterial plant pathogen: novel insight into the origin of Pierce's disease of grapevine in the US. PLos One 5:e15488

Nunney L, Yuan X, Bromley RE et al (2012) Detecting genetic introgression: high levels of intersubspecific recombination found in *Xylella fastidiosa* in Brazil. Appl Environ Microbiol 78:4702–4714

Nunney L, Vickerman DB, Bromley RE et al (2013) Recent evolutionary radiation and host plant specialization in the *Xylella fastidiosa* subspecies native to the United States. Appl Environ Microbiol 79:2189–2200

Nunney L, Schuenzel EL, Scally M, Bromley RE, Stouthamer R (2014) Large-scale intersubspecific recombination in the plant-pathogenic bacterium *Xylella fastidiosa* is associated with the host shift to mulberry. Appl Environ Microbiol 80:3025–3033

Paião FG, Meneguim AM, Casagrande EC et al (2002) Envolvimento de cigarras (Homoptera, Cicadidae) na transmissão de *Xylella fastidiosa* em cafeeiro. Fitopatol Brasil 27:S67

Perez-Donoso AG, Sun Q, Roper MC et al (2010) Cell wall-degrading enzymes enlarge the pore size of intervessel pit membranes in healthy and *Xylella fastidiosa*-infected grapevines. Plant Physiol 152:1748–1759

Pierce NB (1892) The California vine disease. US Dept Agric Div Veg Pathol Bull 2:22

Pieretti I, Royer M, Barbe V et al (2009) The complete genome sequence of *Xanthomonas albilineans* provides new insights into the reductive genome evolution of the xylem-limited *Xanthomonadaceae*. BMC Genom 10:616

Pinel N, Davidson SK, Stahl DA (2008) *Verminephrobacter eiseniae* gen. nov, sp nov, a nephridial symbiont of the earthworm *Eisenia foetida* (Savigny). Int J Syst Evol Microbiol 58:2147–2157

Pooler MR, Hartung JS, Fenton RG (1997) Sequence analysis of a 1296-nucleotide plasmid from *Xylella fastidiosa*. FEMS Microbiol Lett 155:217–222

Purcell AH (2013) Paradigms: examples from the bacterium *Xylella fastidiosa*. Annu Rev Phytopathol 51:339–356

Purcell AH (1977) Cold therapy of Pierce's disease of grapevines. Plant Dis Rep 61:514–518

Purcell AH (1980) Environmental therapy for Pierce's disease of grapevines. Plant Dis 64:388–390

Purcell AH, Finlay AH (1979) Evidence for noncirculative transmission of Pierce's disease bacterium by sharpshooter leafhoppers. Phytopathology 69:393–395

Purcell AH, Hopkins DL (1996) Fastidious xylem-limited bacterial plant pathogens. Ann Rev Phytopathol 34:131–151

Purcell AH, Saunders SR (1999) Fate of Pierce's disease strains of *Xylella fastidiosa* in common riparian plants in California. Plant Dis 83:825–830

Purcell AH, Finlay AH, McLean DL (1979) Pierce's disease bacterium: mechanism of transmission by leafhopper vectors. Science 206:839–841

Qin X, Hartung JS (2001) Construction of a shuttle vector and transformation of *Xylella fastidiosa* with plasmid DNA. Curr Microbiol 43:158–162

Randall JJ, Goldberg NP, Kemp JD et al (2009) Genetic analysis of a novel *Xylella fastidiosa* subspecies found in the southwestern United States. Appl Environ Microbiol 75:5631–5638

Rashed A, Killiny N, Kwan J et al (2011) Background matching behaviour and pathogen acquisition: plant site preference does not predict the bacterial acquisition efficiency of vectors. Arthropod Plant Interact 5:97–106

Rathé AA, Pilkington LJ, Gurr GM et al (2012a) Potential for persistence and within-plant movement of *Xylella fastidiosa* in Australian native plants. Aus Plant Pathol 41:405–412

Rathé AA, Pilkington LJ, Gurr GM et al (2012b) Incursion preparedness: anticipating the arrival of an economically important plant pathogen *Xylella fastidiosa* Wells (Proteobacteria: *Xanthomonadaceae*) and the insect vector *Homalodisca vitripennis* (Germar) (Hemiptera: Cicadellidae) in Australia. Aus J Entomol 51:209–220

Redak RA, Purcell AH, Lopes JRS et al (2004) The biology of xylem fluid-feeding insect vectors of *Xylella fastidiosa* and their relation to disease epidemiology. Ann Rev Entomol 49:243–270

Reddy JD, Reddy SL, Hopkins DL et al (2007) TolC is required for pathogenicity of *Xylella fastidiosa* in *Vitis vinifera* grapevines. Mol Plant Microbe Interact 20:403–410

Rocha EPC (2004) Codon usage bias from tRNA's point of view: redundancy, specialization, and efficient decoding for translation optimization. Genome Res 14:2279–2286

Rocha EPC (2008) The organization of the bacterial genome. Ann Rev Genet 42:211–233

Rodriguez LM, Grajales A, Arrieta-Ortiz ML et al (2012) Genomes-based phylogeny of the genus *Xanthomonas*. BMC Microbiol 12:43

Rogers EE, Stenger DC (2012) A conjugative 38 kB plasmid is present in multiple subspecies of *Xylella fastidiosa*. PloS One 7:e52131

Roper MC, Greve LC, Warren JG et al (2007) *Xylella fastidiosa* requires polygalacturonase for colonization and pathogenicity in *Vitis vinifera* grapevines. Mol Plant Microbe Interact 20:411–419

Rosa C, Kamita SG, Falk BW (2012) RNA interference is induced in the glassy winged sharpshooter *Homalodisca vitripennis* by actin dsRNA. Pest Managem Sci 68:995–1002

Ryan RP, Dow JM (2008) Diffusible signals and interspecies communication in bacteria. Microbiology 154:1845–1858

Sandanayaka WRM, Backus EA (2008) Quantitative comparison of stylet penetration behaviors of glassy-winged sharpshooter on selected hosts. J Econ Entomol 101:1183–1197

Saponari M, Boscia D, Nigro F, Martelli GP (2013) Identification of DNA sequences related to *Xylella fastidiosa* in oleander, almond and olive trees exhibiting leaf scorch symptoms in Apulia (Southern Italy). J Plant Pathol 95:668

Scally M, Schuenzel EL, Stouthamer R et al (2005) Multilocus sequence type system for the plant pathogen *Xylella fastidiosa* and relative contributions of recombination and point mutation to clonal diversity. Appl Environ Microbiol 71:8491–8499

Schaad NW, Postnikova E, Lacy G et al (2004) *Xylella fastidiosa* subspecies: *X. fastidiosa* subsp *piercei*, subsp. nov, *X. fastidiosa* subsp. *multiplex* subsp. nov, and *X. fastidiosa* subsp. *pauca* subsp. nov. Syst Appl Microbiol 27:290–300

Schreiber HL, Koirala M, Lara A et al (2010) Unraveling the first *Xylella fastidiosa* subsp *fastidiosa* genome from Texas. Southwest Entomol 35:479–483

Severin HHP (1949) Transmission of the virus of Pierce's diseasae of grapevines by leafhoppers. Hilgardia 19:190–206

Severin HHP (1950) Spittle-insect vectors of Pierce's disease virus II. Life history and virus transmission. Hilgardia 19:357–382

Sharp PM, Bailes E, Grocock RJ et al (2005) Variation in the strength of selected codon usage bias among bacteria. Nucl Acids Res 33:1141–1153

Sharp PM, Emery LR, Zeng K (2010) Forces that influence the evolution of codon bias. Phil Trans Royal Soc London B 365:1203–1212

Shi XY, Dumenyo CK, Hernandez-Martinez R et al (2009) Characterization of regulatory pathways in *Xylella fastidiosa*: genes and phenotypes controlled by *gacA*. Appl Environ Microbiol 75:2275–2283

Silva FR, Vettore AL, Kemper EL et al (2001) Fastidian gum: the *Xylella fastidiosa* exopolysaccharide possibly involved in bacterial pathogenicity. FEMS Microbiol Lett 203:165–171

Silva VS, Shida CS, Rodrigues FB et al (2007) Comparative genomic characterization of citrus-associated *Xylella fastidiosa* strains. BMC Genom 8:474

Silva Neto JF, Koide T, Gomes SL et al (2002) Site-directed gene disruption in *Xylella fastidiosa*. FEMS Microbiol Lett 210:105–110

Simpson AJ, Reinach FC, Arruda P et al (2000) The genome sequence of the plant pathogen *Xylella fastidiosa*. Nature 406:151–159

Stenger DC, Lee MW (2011) Phylogeny of replication initiator protein TrfA reveals a highly divergent clade of incompatibility group P1 plasmids. Appl Environ Microbiol 77:2522–2526

Stenger DC, Lee MW, Rogers EE et al (2010) Plasmids of *Xylella fastidiosa* mulberry-infecting strains share extensive sequence identity and gene complement with pVEIS01 from the earthworm symbiont *Verminephrobacter eiseniae*. Physiol Mol Plant Pathol 74:238–245

Su CC, Chang CJ, Yang WJ et al (2012) Specific characters of 16 rRNA gene and 16S-23S rRNA internal transcribed spacer sequences of *Xylella fastidiosa* pear leaf scorch strains. Eur J Plant Pathol 132:203–216

Su CC, Chang CJ, Chang CM et al (2013) Pierce's disease of grapevines in Taiwan: isolation, cultivation

and pathogenicity of *Xylella fastidiosa*. J Phytopathol 161:389–396

Summer EJ, Enderle CJ, Ahern SJ et al (2010) Genomic and biological analysis of phage XFas53 and related prophages of *Xylella fastidiosa*. J Bacteriol 192:179–190

Sun Q, Greve LC, Labavitch JM (2011) Polysaccharide compositions of intervessel pit membranes contribute to Pierce's disease resistance of grapevines. Plant Physiol 155:1976–1987

Sun Q, Sun Y, Walker MA et al (2013) Vascular occlusions in grapevines with pierce's disease make disease symptom development worse. Plant Physiol 161:1529–1541

Tarsi R, Pruzzo C (1999) Role of surface proteins in *Vibrio cholerae* attachment to chitin. Appl Environ Microbiol 65:1348–1351

Van der Auwera GA, Król JE, Suzuki H et al (2009) Plasmids captured in *C. metallidurans* CH34: defining the PromA family of broad-host-range plasmids. Antonie Van Leeuw 96:193–204

van Sluys MA, de Oliveira MC, Monteiro-Vitorello CB (2003) Comparative analyses of the complete genome sequences of Pierce's disease and citrus variegated chlorosis strains of *Xylella fastidiosa*. J Bacteriol 185:1018–1026

Varani AM, Souza RC, Nakaya HI (2008) Origins of the *Xylella fastidiosa* prophage-like regions and their impact in genome differentiation. PloS One 3:e4059

Vieira-Silva S, Rocha EPC (2010) The systemic imprint of growth and its uses in ecological (meta)genomics. PloS Genet 6:e1000808

Voegel TM, Warren JG, Matsumoto A et al (2010) Localization and characterization of *Xylella fastidiosa* haemagglutinin adhesins. Microbiology 156: 2172–2179

Voegel TM, Doddapaneni H, Cheng DW et al (2013) Identification of a response regulator involved in surface attachment, cell–cell aggregation, exopolysaccharide production and virulence in the plant pathogen *Xylella fastidiosa*. Mol Plant Pathol 14:256–264

Wang LH, He YW, Gao YF et al (2004) A bacterial cell-cell communication signal with cross-kingdom structural analogues. Mol Microbiol 51:903–912

Wang N, Li JL, Lindow SE (2012) RpfF-dependent regulon of *Xylella fastidiosa*. Phytopathology 102:1045–1053

Wells JM, Raju BC, Hung HY et al (1987) *Xylella fastidiosa* gen. nov, sp. nov: Gram-negative, xylem-limited, fastidious plant bacteria related to *Xanthomonas* ssp. Int J Syst Bact 37:136–143

Wilhelm M, Brodbeck BV, Andersen PC et al (2011) Analysis of xylem fluid components in almond cultivars differing in resistance to almond leaf scorch disease. Plant Dis 95:166–172

White FF, Potnis N, Jones JB et al (2009) The type III effectors of *Xanthomonas*. Mol Plant Pathol 10:749–766

Yen MR, Lin NT, Hung CH et al (2002) *oriC* region and replication termination site, *dif*, of the *Xanthomonas campestris* pv. *campestris* 17 chromosome. Appl Environ Microbiol 68:2924–2933

Yuan X, Morano L, Bromley R et al (2010) Multilocus sequence typing of *Xylella fastidiosa* causing Pierce's disease and oleander leaf scorch in the United States. Phytopathology 100:601–611

Zhang S, Flores-Cruz Z, Kumar D et al (2011) The *Xylella fastidiosa* biocontrol strain EB92-1 genome is very similar and syntenic to Pierce's disease strains. J Bacteriol 193:5576–5577

Comparative Genomics of the Liberibacteral Plant Pathogens

9

Hong Lin and Edwin L. Civerolo

9.1 Introduction

9.1.1 Liberibacteral Species

'*Candidatus* Liberibacter' species are bacteria that are associated with apparently healthy and diseased plants and vectored by several psyllid species. Liberibacters are Gram-negative, phloem-inhabiting bacteria which are members of the *Rhizobiaceae* α-Proteobacteria (Jagoueix et al. 1997). Most of the plant- and disease-associated Liberibacters are designated '*Candidatus*' species because they have not yet been successfully cultured in vitro. Thus, Koch's postulates have not been fulfilled *sensu stricto* for any of the plant disease-associated Liberibacters. Currently, the taxonomy of '*Ca.* Liberibacter' species is based on sequence analyses of the 16S rRNA gene, the intergenic 16S/23S rDNA spacer region, and ribosomal genes of the *β* operon (Bastianel et al. 2005; Garnier et al. 1991, 1996; Jagoueix et al. 1997; Magomere et al. 2009; Villechanoux et al. 1993).

Plant disease-associated Liberibacters include '*Ca.* Liberibacter asiaticus' (Bové 2006), '*Ca.* Liberibacter africanus' (Alfaro-Fernández et al. 2012a, b; Bové 2006; Garnier et al. 2000b), '*Ca.* Liberibacter americanus' (Bové 2006; Texeira et al. 2005a), and '*Ca.* Liberibacter solanacearum' (syn. '*Ca.* Liberibacter psyllaurous') (Hansen et al. 2008; Liefting et al. 2008b; Lin et al. 2009; Wen et al. 2009).

In addition to citrus, '*Ca.* Liberibacter asiaticus' also occurs in several non-citrus rutaceous plants, including *Murraya paniculata*/*exotica*, *Bergera* (syn. *Murraya*) *koenigii*, *Clausena lansium*, *Atalantia* (*Severinia*) *buxifolia*, *Toddalia lanceolat*, and *Limonia acidissma* (da Graça et al. 2009). '*Ca.* Liberibacter solanacearum' is an apparent endophyte in other solanaceous weeds, such as silverleaf nightshade (*Solanum elaegnifolium*), wolfberry (*Lycium barbarum*), and black nightshade (*S. ptychanthum*) (Wen et al. 2009).

'*Ca.* Liberibacter europaeus' occurs in healthy-appearing pear trees in Italy (Raddadi et al. 2011). *Liberibacter crescens* was isolated from mountain papaya without any disease or disease-like symptoms (Leonard et al. 2012).

The Liberibacters detected in, or isolated from, plants are transmitted or associated with psyllids, including *Bactericera cockerelli*, ten species of the genus *Cacopsylla* (*C. affinis*, *C. ambigua*, *C. breviantennata*, *C. crataegi*, *C. melanonerva*, *C. nigrita*, *C. peregrina*, *C pyri*, *C. pyricola*, and *C. pyrisuga*), *Diaphorina citri*, and *Trioza erytreae*. Plant-associated Liberibacters are part of the complex microbiota and may be naturally occurring endosymbionts in psyllids (Fagen et al. 2012; Hansen et al. 2008).

H. Lin (✉) · E. L. Civerolo
USDA-ARS San Joaquin Valley Agricultural Sciences Center, 9611 South Riverbend Avenue, Parlier, CA 93648, USA
e-mail: hong.lin@ars.usda.gov

D. C. Gross et al. (eds.), *Genomics of Plant-Associated Bacteria*,
DOI: 10.1007/978-3-642-55378-3_9, © Springer-Verlag Berlin Heidelberg (outside the USA) 2014

9.1.2 Diseases Associated with 'Ca. Liberibacter' Species

9.1.2.1 Citrus Huanglongbing (Greening)

Huanglongbing (HLB), also known as greening, is a destructive disease of citrus worldwide that severely reduces productivity, resulting in catastrophic economic losses (Bové 2006; Gottwald 2010; Gottwald et al. 2007). Generally, most commercial citrus species, cultivars, and hybrids are susceptible to HLB. However, some current cultivars and hybrid seedlings exhibit tolerance (Stover et al. 2013). HLB-affected trees begin to decline within a few years of infection, produce reduced yields of poor-quality fruit, and may die or become otherwise unproductive. There are three forms of HLB, namely the Asian form, African form, and American form (Bové 2006; da Graça 1991; Gottwald 2010; Gottwald et al. 2007). The Asian form is the most extensive and severe form of HLB. '*Ca.* Liberibacter asiaticus' is associated with the Asian form that is heat tolerant and can develop at temperatures above about 30 °C. '*Ca.* Liberibacter africanus' is associated with the African form that is suppressed at elevated temperatures above about 30 °C. The American form of HLB is similar to the Asian form with respect to symptom expression and disease severity. However, the American form of HLB, with which '*Ca.* Liberibacter americanus' is associated, is heat intolerant, similar to that of the African form (Bové 2006; Coletta-Filho et al. 2004; Gottwald 2010; Lopes et al. 2009; Shokrollah et al. 2010). '*Ca.* Liberibacter asiaticus' and '*Ca.* Liberibacter americanus' are transmitted by the citrus psyllid *D. citri* in Asia and the Americas (Bové 2006), while '*Ca.* Liberibacter africanus' is transmitted by psyllid *T. erytreae* only found in South Africa. Although psyllids collected from *Calodendrum capense* carry high levels of '*Ca.* Liberibacter africanus subspecies capensis,' there is no information regarding insect transmission of '*Ca.* Liberibacter africanus subsp. capensis' (Gerhard Pietersen "personal communication"). HLB-associated Liberibacters can also be transmitted by grafting. There are no commercially effective, economically feasible HLB control measures. HLB management is currently based on preventive and phytosanitary practices, including removal of HLB-affected trees to reduce availability of inoculum, citrus psyllid vector control via application of insecticides and area-wide psyllid management programs, protection of budwood or propagating material source trees, production and protection of nursery trees free from HLB-associated Liberibacters, and quarantines to prevent dissemination of HLB-associated Liberibacters.

9.1.2.2 Cape Chestnut Leaf Mottle

In South Africa, leaf mottle symptoms of ornamental rutaceous trees, Cape chestnut (*Calodendrum capense*), were associated with a Liberibacter genetically close to '*Ca.* Liberibacter africanus' (Garnier et al. 2000a, b). This Liberibacter was not detected in HLB-affected citrus trees in South Africa, was phylogenetically and serologically more closely related to '*Ca.* Liberibacter africanus' than to '*Ca.* Liberibacter asiaticus,' and was genotypically distinguishable from the HLB-associated '*Ca.* Liberibacter africanus.' Accordingly, it was designated as '*Ca.* africanus subsp. capensis' (Garnier et al. 2000a, b).

9.1.2.3 Potato Zebra Chip and Diseases of Other Solanaceous Crops

Zebra chip (ZC) disease (or zebra chip complex) is an economically important disease (or disease complex) of fresh table and processed potatoes. ZC was first reported in commercial potato production fields in Mexico in 1994 and first identified in 2000 in commercial potato production fields in the United States in Texas (Gudmestad and Secor 2007; Lin and Gudmestad 2013; Munyaneza et al. 2007; Secor et al. 2009; Wen et al. 2009). Since that time, ZC disease has occurred in most potato production regions in the southern plains, northern plains,

and Pacific Northwest states in the United States (Rondon et al. 2012), in Belize, in Guatemala and Honduras in Central America (EPPO 2009), and in Oceania New Zealand (Liefting et al. 2009b). ZC reduces the marketability of potatoes due to distinct internal brown discoloration or necrotic flecking of the medullary rays in raw tubers and intensely dark, striped blackish discoloration of chips processed from ZC-affected tubers (Crosslin et al. 2010; Gudmestad and Secor 2007; Munyaneza et al. 2007; Secor et al. 2009). Aboveground symptoms include reddening of leaves, leaf scorching, swollen nodes, axillary bud development, and plant death (Gudmestad and Secor 2007; Wen et al. 2009). ZC disease is associated with '*Ca.* Liberibacter solanacearum' (Liefting et al. 2008a, b) that is transmitted by the potato–tomato psyllid *B. cockerelli*. Subsequently, '*Ca.* Liberibacter solanacearum' was detected in other symptomatic solanaceous crops, including tamarillo (*Solanum betaceum*), cape gooseberry (*Physalis peruviana*), and greenhouse-grown tomato (*Solanum lycopersicum*) and pepper (*Capsicum annuum*) plants in New Zealand (Liefting et al. 2008a, 2009a), and in greenhouse-grown tomatoes in Arizona (Brown et al. 2010) and in Jocotitlan, Mexico (Ling et al. 2011).

An α-proteobacterium with 99.95 % similarity in the 16S rRNA gene sequence to '*Ca.* Liberibacter solanacearum' was associated with psyllid yellows disease of tomato in California and was named '*Ca.* Liberibacter psyllaurous' (Hansen et al. 2008). Accordingly, '*Ca.* Liberibacter solanacearum' and '*Ca.* Liberibacter psyllaurous' are currently considered to be synonymous (Lin and Gudmestad 2013; Secor et al. 2009; Wen et al. 2009). '*Ca.* Liberibacter solanacearum' is the preferred designation of this bacterium since this name and a description of this bacterium were published in the *International Journal of Systematic and Evolutionary Microbiology*, following the rules of the International Code of Nomenclature of Bacteria, 1990 Revision (Lapage et al. 1992). In addition, a reference isolate from a tomato plant in New Zealand was designated (NZ082226), and DNA and freeze-dried plant material of this isolate are available for reference (Liefting et al. 2009a).

9.1.2.4 Pysllid Yellows Disorders of Carrot and Celery

'*Ca.* Liberibacter solanacearum' haplotypes C, D, and E (see below) have been detected in several European countries in carrot and celery in plants with 'psyllid yellows' or 'psyllid yellows-like' symptoms (Alfaro-Fernández et al. 2012a, b; EPPO 2012; Munyaneza et al. 2010a, b, 2011; Teresani et al. 2014). In Finland, Norway, and Sweden, '*Ca.* Liberibacter solanacearum' is also associated with carrots and the psyllid vector *T. apicalis* (Munyaneza et al. 2010a, b, 2011). '*Ca.* Liberibacter solanacearum' is associated with celery and the psyllid vector *B. trigonica* in Spain and the Canary Islands (Alfaro-Fernández et al. 2012a, b; EPPO 2012; Teresani et al. 2014).

9.1.3 '*Candidatus* Liberibacter europaeus'

Raddadi et al. (2011) described a novel Liberibacter species, '*Ca.* Liberibacter europaeus' associated with, and transmitted by, the pear psyllid, *Cacopsylla pyri*, in Italy. Subsequently, '*Ca.* Liberibacter europaeus' was found in *C. affinis*, *C. ambigua*, *C. breviantennata*, *C. crataegi*, *C. melanonerva*, *C. nigrita*, *C. peregrina*, *C. pyricola*, and *C. pyrisuga* in northwestern Italy and in *C. pyri* and *C. pyricola* in Hungary (Camerota et al. 2012). *C. pyri* is widely prevalent in different areas in the Piedmont and Valle d'Aosta regions in Italy. No specific disease symptoms were associated with the presence of '*Ca.* Liberibacter europaeus' in pears in Italy. However, '*Ca.* Liberibacter europaeus' was transmitted by *C. pyri* to healthy pear plants in experimental transmission tests. In addition, '*Ca.* Liberibacter europaeus' was also detected in several *C. pyri* host plants, including apple, blackthorn, hawthorne, and pear, in northwestern Italy (Camerota et al. 2012). Most recently,

'*Ca.* Liberibacter europaeus' was associated with broom psyllid (*Arytainilla spartiophila*)-infested Scotch broom (*Cytisus scoparius*), an invasive leguminous exotic shrub, with symptoms that included stunted shoot growth, shortened internodes, and leaf dwarfing and leaf tip necrosis in the South Island of New Zealand where the broom psyllid is common (Thompson et al. 2013). This novel Liberibacter apparently is an endophyte rather than a pathogen in pear in Italy (Camerota et al. 2012; Raddadi et al. 2011).

9.1.3.1 Liberibacter crescens

Isolation and in vitro culture of a bacterium associated with mountain papaya (babaco; mountain pawpaw; *Vasconcellea x heilbornii; Carica stiplata* x *C. pubescens* hybrid; also known as *C. pubescens* and *C. candamarcensis*) and designated *Liberibacter crescens* (strain BT-1) were recently reported (Leonard et al. 2012). This bacterium is apparently non-pathogenic as it has not been associated with any disease or disease symptoms. Based on comparison of the complete genome sequences of '*Ca.* Liberibacter asiaticus' and *L. crescens*, these species are 75.5 % similar (Leonard et al. 2012). However, the prophage regions in these genomes are not similar.

9.2 Evolution and Speciation of '*Ca.* Liberibacter Species'

9.2.1 Discovery and Identification of Liberibacter

HLB has been present in Asian countries for more than a century and has now been reported to occur in many citrus-growing regions around the world. The origins of Liberibacters associated with HLB are not fully understood. The prevailing assumption is that HLB originated in Southeast Asian countries, most likely in China or India or in nearby regions (Bové 2006; Zhao 1981). Historically, various names were given based on the characteristics of HLB. Reinking reported it as a 'yellow shoot disease' of citrus in southern China in 1919 (Reinking 1919). In India, the disease was referred to as 'dieback' based on its characteristic symptom and was believed to be a physiological disorder rather than a disease (Capoor 1963). Likewise, the name 'Likubin' was used in Taiwan (Chen et al. 1971) and was also known as 'mottle leaf' in Philippines and 'vein phloem degeneration' in Indonesia (Tirtawidjaja et al. 1965). In South Africa, the disease was characterized based on fruit symptoms and was given the name 'greening' based on the abnormal green coloring of the fruit. In the first description of 'greening' in 1937, the disease was attributed to mineral toxicity (Van der Merwe and Andersen 1937). It was not until 1965 when the infectious nature of 'greening' was demonstrated by transmission of the causal agent(s) by phloem-feeding insects, *T. erytreae*, and by grafting from symptomatic plant materials (McClean and Oberholzer 1965a, b). In spite of its perplexity, it is now clear that HLB is a disease associated with bacteria that are transmitted in nature by the Asian citrus psyllid (*D. citri*) found in most citrus-growing regions worldwide or by *T. erytreae* only found in South Africa (Bové 2006).

Dr. Kung Hsiang Lin, a Chinese plant pathologist, in the 1940s carried out a series of classic transmission experiments in southern China. It was his research that clarified HLB is associated with a graft-transmissible, infectious agent(s) and not due to nutritional deficiency or to other abiotic stresses (Bové 2006; Lin 1956) and correctly characterized the symptom as 'huanglongbing' (literally meaning 'yellow shoot disease' in a local dialect in southern China) as early as the 1950s (Bové 2006). As recognition for his contributions, huanglongbing or 'HLB' was proposed as the name for this disease at the 13th Conference of the International Organization of Citrus Virologists (IOCV) in 1995 in Fuzhou, China. While the nature of disease was determined, the causal agent was not identified largely because of the fastidious nature of the pathogen and inability to isolate and grow it on artificial cultural media. Using transmission

electron microscopy (TEM), Garnier and Bové (1977), Garnier et al. (1984) provided the first visual evidence of putative bacterial cells with Gram-negative cell walls present in sieve cells of leaf midribs with yellow mottle, a typical symptom of leaves from HLB-affected trees. The TEM evidence clarified two properties of the HLB-associated bacteria—the exclusive location of HLB bacterium in sieve tube cells—and the bacterial cells are surrounded by a triple-layered cell envelope consisting of an outer cell wall, peptidoglycan layer, and inner cytoplasmic membrane distinguishing it from wall-less bacteria (i.e., mycoplasmas) which are associated with 'yellows'-type diseases that often resemble HLB symptoms (Bové 2006; Garnier et al. 1984; Shokrollah et al. 2010). Information derived from these TEM studies agreed with characteristics of HLB observed in Asia and Africa. However, TEM techniques could not distinguish morphological and genetic differences between the HLB-associated bacteria from Asia and Africa (Bové 2006).

9.2.2 Phylogenetic Relationships Among Liberibacter Species

Due to the fastidious nature of plant disease-associated Liberibacters, standard microbiological methods cannot be applied at that time for determining the biological characteristics of these bacteria and their roles in plant disease development. Thus, details of the etiology of the disease are lacking. The taxonomic position of HLB-associated bacteria could not be determined until DNA technology became available. Using universal PCR primer pairs, f-D1, and r-P1 for prokaryotes (Weisburg et al. 1991), researchers were able to amplify ~1,500 bp of the 16S ribosomal gene DNA sequences in HLB-affected samples (Jagoeix et al. 1994). Comparison of these DNA sequences with those in the GenBank database indicated that the Asian and African HLB-associated bacteria belonged to the class of α-Proteobacteria. DNA sequence divergence data in the 16S rRNA genes are widely used for defining bacterial species (Cohan 2002). Bacteria with >3 % divergence in 16S rRNA gene sequences are nearly always members of different species. Therefore, a cutoff of 3 % divergence was recommended as a criterion for demarcating species (Stackebrandt and Goebel 1994). This is particularly useful for classification of unculturable bacterial taxa, such as '*Candidatus* Liberibacter,' for which phenotypic and ecological characteristics are generally not available (Lin et al. 2009). Thus, sequence analyses determined that the 16S rRNA gene sequences of Asian and Africa HLB-associated Liberibacter had 87.5 % identity with the members of α-2 subgroup. However, both have about ~3 % dissimilarity in the 16S rRNA gene region. Therefore, they were defined as two species and named '*Ca.* Liberibacter asiaticus' and '*Ca.* Liberibacter africanus,' respectively, based on their presumptive origins (Jagoeix et al. 1997).

In March 2004, HLB was reported in sweet orange trees near the city of Araraquara, in São Paulo State (SPS), Brazil (Coletta-Filho et al. 2004), and was the first occurrence of HLB in the Americas. Leaf symptoms of the disease resembled classic blotchy mottle symptoms described previously (Bové 2006). '*Ca.* Liberibacter asiaticus' was soon confirmed by PCR in mottled leaf samples. However, many blotchy mottle leaves (~80 %) repeatedly tested negative by PCR designed for '*Ca.* Liberibacter asiaticus' and '*Ca.* Liberibacter africanus.' These results led to the discovery of a new type of HLB-associated bacterium in SPS (Texeira et al. 2005a). Indeed, sequencing analysis of the new SPS HLB-associated bacterium indicated that it shared only 96 % similarity with the 16S rRNA gene sequences of '*Ca.* Liberibacter asiaticus' and '*Ca.* Liberibacter africanus,' but shared 66.0 and 79.5 % sequence homology with 16S–23S intergenic ribosomal region, respectively. Thus, the designation '*Ca.* Liberibacter americanus' was given to this new Liberibacter species, reflecting its occurrence in the American continent (Teixeira et al. 2005b). So far, '*Ca.* Liberibacter americanus' is only found in Brazil.

In 2008, researchers at the Plant Health and Environment Laboratory, MAF Biosecurity New Zealand, reported that a new Liberibacter

species, based on comparative 16S rRNA gene sequence analyses, was associated with potato ZC (Liefting et al. 2008a, 2009b). Along with transmission electron microscopy evidence (De Boer et al. 2007; Liefting et al. 2009a; Secor and Rivera-Varas 2004) confirmed the presence of phloem-inhabiting bacterial-like organisms in diseased plants, but not in healthy plants. Sequence analysis of the 16S rRNA gene PCR amplicon indicated that it shared a high degree of similarity with, but distinct from, the Liberibacteral species associated with citrus HLB. Thus, it was given the name '*Ca* L. solanacearum' (Liefting et al. 2009b). This was the first time that Liberibacter bacteria were found in bacteria were found in *solanaceous* hosts. Subsequently, '*Ca* L. solanacearum' was detected in other solanaceous crops (Liefting et al. 2008b; Wen et al. 2009).

In August, 2008, a research group at University of California, Riverside, also reported a new citrus 'huanglongbing species' bacterium named '*Ca* Liberibacter psyllaurous' responsible for a tomato yellowing disease and that was associated with bacteria transmitted by the tomato/potato psyllid, *B. cockerelli* (Hansen et al. 2008). Since the sequences of '*Ca* Liberibacter solanacearum' and '*Ca* Liberibacter psyllaurous' share >99 % sequence similarity in the 16S rRNA gene, it is likely that these two bacterial names are synonymous.

Shortly after the report of '*Ca* Liberibacter solanacearum' in New Zealand, a new strain of '*Ca* Liberibacter solanacearum' associated with carrot (*Daucus carota*), a non-solanaceous crop, and vectored by carrot psyllid (*T. apicalis*) was identified in Finland, Canary Islands, and mainland Spain (Munyaneza et al. 2010a, 2011; Alfaro-Fernández et al. 2012a; Teresani et al. 2014).

A non-pathogenic Liberibacter species, named '*Ca.* Liberibacter europaeus,' was isolated from pear in Italy and is naturally vectored by pear psyllid *Cacopsylla pyri* in Italy (Raddadi et al. 2011). More recently, another new Liberibacter species named 'Liberibacter *crescens*' (strain BT-1) was reported (Raddadi et al. 2011). This Liberibacter species was isolated from mountain papaya and appears to be non-pathogenic.

Phylogenetic analyses based on the sequences of the 16S rRNA genes of Liberibacter species, along with the members of *Rhizobiaceae* and other α-Proteobacteria, show that '*Ca.* Liberibacter' species belong to the α-Proteobacteria group, and their closest relatives are members of the group of bacteria known as the α-Proteobacteria subgroup 2 (Jagoueix et al. 1994). All six '*Ca*. Liberibacter' species and one subspecies are clustered into a monophyletic group (Fig. 9.1). Among them, '*Liberibacter crescens*' is closely positioned proximally to the basal node, followed by '*Ca*. L. europaeus,' '*Ca*. L. americanus,' and '*Ca*. L. solanacearum,' indicating early divergence of '*L. crescens.*' This result agrees with the orthologous gene-based phylogenetic analysis that the lineage of Liberibacters is an early-branching member of the *Rhizobiaceae* (Doddapaneni et al. 2008; Duan et al. 2009). It appears that the genome of '*Ca.* Liberibacter asiaticus' has evolved rapidly (Duan et al. 2009). Such rapid genome evolution is typical of host-restricted symbionts or pathogens (Moran et al. 2008). Presumably, the adaptation of Liberibacters to crops is a relatively recent occurrence. Consequently, since there has not been sufficient coevolutionary time between the crop cultivation and the pathogen, this may explain the lack of disease-resistant hosts in both citrus and citrus relatives and in solanaceous crops.

9.2.3 Evolutionary Potential and Genetic Diversity

The genetic variation and evolutionary potential of bacteria have stimulated researchers to investigate the origins and evolutionary processes that have driven the emergence of highly adapted and specialized mechanisms for plant pathogens. These studies will provide an in-depth insight into the genetic and epidemiological factors that have contributed to various disease outbreaks. Based on current research findings, it is expected that '*Ca.* Liberibacter' species have undergone rapid evolutionary

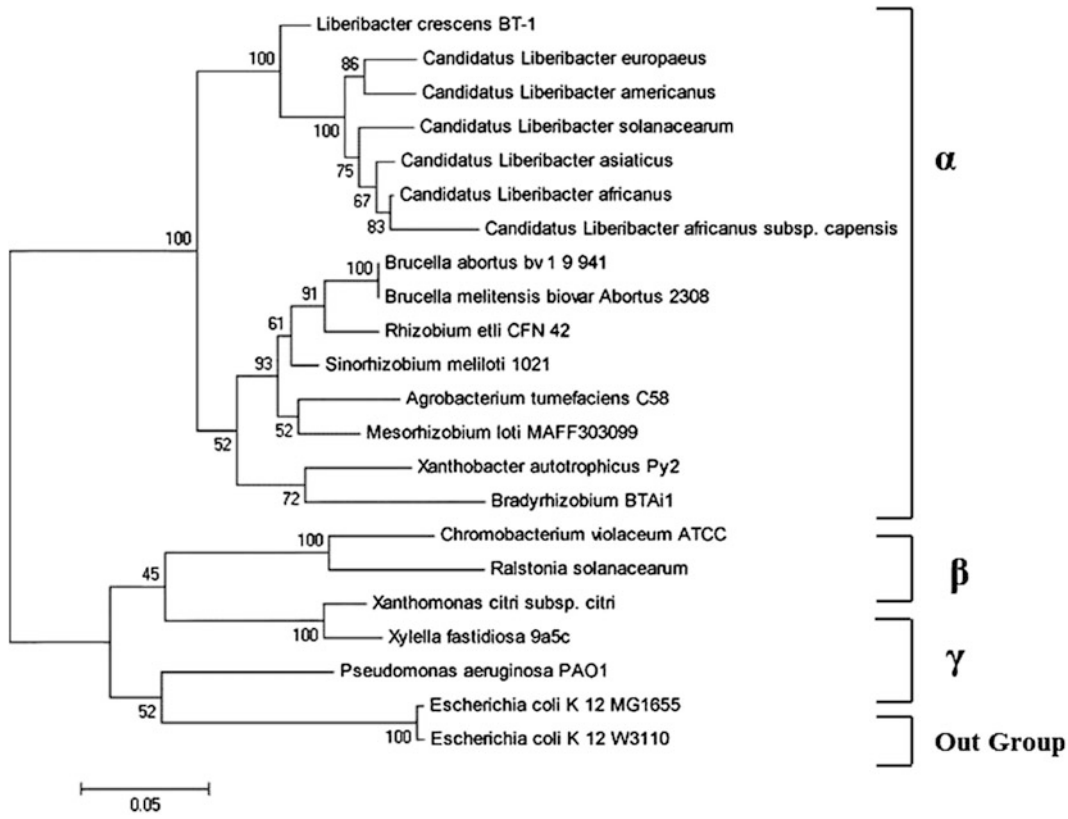

Fig. 9.1 Neighbor-joining phylogenetic tree constructed using 16S rDNA sequences. Bootstrap values are indicated at the nodes

divergence, speciation, and adaptation to a wider range of hosts since the divergence from their ancestors. Genetic diversity of Liberibacters is perhaps much larger than previously expected. Beattie et al. (2011) proposed that all three citrus HLB-associated Liberibacter species could be derived from a single origin and evolved into three forms after the separation of Gondwana. However, this is not supported by the situation for '*Ca*. Liberibacter americanus' as citrus hosts are not native to the Americas (Beattie et al. 2011). An alternative assumption could be that Liberibacters are of insect or animal origin.

It was not until recently (~500 years ago or earlier) that Liberibacters adapted to new hosts and environments. Discovery of new species of '*Ca.* L. solanacearum,' '*Ca.* L. europaeus,' and '*Liberibacter crescens*' associated with plants outside the *Rutaceae* sparked research interest in searching for Liberibacter origins and evolutionary relationships of these bacteria with other Liberibacter species. Current evidence supports the hypothesis that Liberibacter bacteria may have the capacity to adapt rapidly to new hosts (Lin and Gudmestad 2013). Allopatric speciation or geographical speciation is a process that occurs when organisms of the same species become isolated due to geographical changes. The separated populations then independently undergo evolutionary selection and genotypic and/or phenotypic divergence imposed by selective pressures. Using 16S rRNA gene nucleotide sequences as molecular clocks, Eveillard et al. (2008) predicted that the 'speciation' of '*Ca.* Liberibacter asiaticus' and '*Ca.* Liberibacter africanus' occurred some 150 million years ago, while speciation of '*Ca.* Liberibacter americanus' might have started some

300 million years ago. Doddapaneni et al. (2008) estimated that speciation of the Liberibacters may have occurred 110–120 million years ago, prior to the breakup of Gondwana.

To determine the genetic diversity of Liberibacter, various molecular markers have been developed. Glynn et al. (2012) developed multilocus sequence typing markers (MLST) for Lso. This marker system is based on 7–10 housekeeping gene loci located throughout the genome of the organism of interest (Ibarz Pavón and Maiden 2009). Sequence variations due to insertion/deletion (INDEL) or SNPs within these loci can be used to determine genetic relationships in different geographical sources. Using a panel of MLST markers, two sequence types (ST-1 and ST-2) were identified, representing two major lineages in US '*Ca.* Liberibacter solanacearum' populations (Glynn et al. 2012). A multilocus simple sequence repeat (SSR) marker system (also known as variable-number tandem repeat or microsatellite) was also developed for the genetic analysis of '*Ca.* Liberibacter solanacearum' and '*Ca.* Liberibacter asiaticus' populations (Islam et al. 2012; Katoh et al. 2011; Lin et al. 2012). In contrast to conserved gene-based marker systems, this marker system possesses high discriminating power for differentiating isolates within populations that have evolved over a relative short timescale. SSR markers have been used for the genetic analyses of '*Ca.* Liberibacter asiaticus' from regional and global populations (Chen et al. 2010; Islam et al. 2012; Katoh et al. 2011). Genetic characterization of '*Ca.* Liberibacter asiaticus' populations from various geographical and epidemiological sources can facilitate the identification of introduced patterns and predict possible relationships and distribution of HLB-associated Liberibacter in different epidemiological situations. In Florida, HLB was first discovered in the United States in Miami-Dade County in August 2005. Since then, HLB has been reported in most citrus-producing counties in Florida (Gottwald 2010). The rapid, widespread distribution of this disease and the level of genetic diversity of '*Ca.* Liberibacter asiaticus' among citrus-growing counties in Florida suggested that multiple introductions of HLB-associated '*Ca.* Liberibacter asiaticus' likely occurred since 2005. The exact routes of the dissemination of HLB-associated Liberibacters are not known. However, SSR genetic analysis indicated that there are at least more than two '*Ca.* Liberibacter asiaticus' genetic lineages in Florida, indicating that such genetic diversity in Florida populations likely resulted in multiple introduction events (Islam et al. 2012). The characterization of strains by nucleotide-based molecular markers has provided useful tools for insights into genetic diversity and evolutionary potential among Liberibacters in agricultural populations. This information, in combination with the advances in understanding the spatial and temporal dynamics of the epidemiology of disease and disease tracking, will facilitate the development of effective integrated pest and disease management systems.

9.3 Genome Features

9.3.1 Sequencing Strategy

The emergence of high-throughput next-generation sequencing technologies has enabled researchers to obtain a comprehensive view of the metagenome of complex microbial communities (Hongoh and Toyoda 2011). Development of a method to sequence genomes of unculturable bacteria provides new strategies to gain insights into the metabolic and physiological capabilities and potential biological functions of bacteria including pathogenic bacteria. Traditional microbiological approaches rely on in vitro culturing and analysis of pure-cultured microbial strains from environmental samples. This approach is largely limited to only a small fraction of culturable microbes, as it is estimated that more than 99 % of microorganisms on the earth are yet unculturable (Hongoh and Toyoda 2011). Because the majority of bacteria and archaea remain unculturable, the diversity of complex bacterial communities is inevitably underestimated using standard in vitro cultivation methods (Vartoukian et al. 2010). The metagenomic strategy, a culture-independent cloning

approach, provides unprecedented opportunities for researchers to capture information of microbial communities, including culturable and unculturable species. However, the mixture of numerous genomic fragments of unidentifiable genera and species impedes further functional identification of individual species. This limitation is overcome by combining the next-generation sequencing approach with amplification of whole-genome sequences obtained from only a few bacterial cells or even a single bacterial cell (Rodrigue et al. 2009; Spits et al. 2006; Zhang et al. 2006). These new sequencing strategies, combined with increasing bioinformatic computational capability, have advanced the field of metagenomics to provide glimpses into the biology and ecology of uncultured prokaryotes.

9.3.2 DNA Enrichment

Since disease-associated '*Ca.* Liberibacter' species are usually present in very low titers and unevenly distributed in their plant hosts, attempts to obtain the complete Liberibacter genome directly using HLB-affected plant tissue have generally failed. Alternatively, '*Ca.* Liberibacter asiaticus' genomic DNA has been obtained from Asian citrus psyllid vectors. Li et al. (2008) reported that while most psyllids collected from HLB-affected citrus trees carry low titers of '*Ca.* Liberibacter asiaticus,' some adult psyllids can carry up to 10^{10} '*Ca.* Liberibacter asiaticus' cells per head. With the development of a multiple displacement amplification method (MDA) (Paez et al. 2004), researchers were able to amplify genomic DNA from a single high-titer Liberibacter-infected psyllid to obtain a sufficient quantity of DNA for whole-genome sequencing (Duan et al. 2009). Using this technique, the sequence of the whole '*Ca.* Liberibacter asiaticus' strain psy66 genome was obtained. Similarly, the '*Ca.* Liberibacter solanacearum' strain ZC-1 genome sequence was also obtained via a metagenomic approach. In this case, the target bacterium was enriched from '*Ca.* Liberibacter solanacearum'-affected potato psyllids using immune capture (Lin et al. 2011). This approach resulted in the enrichment of '*Ca.* Liberibacter solanacearum' bacterial cells from the homogenized Lso-infected potato psyllid tissues prior to DNA extraction and MDA. About 1,000-fold increase in '*Ca.* Liberibacter solanacearum' genomic DNA was obtained compared with unenriched psyllid DNA based on quantitative PCR estimation. Sequences of both Liberibacter genomes were successfully obtained using 454 pyrosequencing technologies. *De novo* assembly resulted in sequence contigs that covered more than 98 % of the target genomes. For example, for '*Ca.* Liberibacter asiaticus' (NCBI GenBank accession # CP001677), 38 contigs were identified, covering 1.2 Mbp. These '*Ca.* Liberibacter asiaticus' contigs were assembled from 91,875 reads, providing about 16-fold coverage for the '*Ca.* Liberibacter asiaticus' genome. For the '*Ca.* Liberibacter solanacearum' genome (NCBI GenBank accession # CP002371), 27 contigs ranging from 1,000 to 279,292 bp were identified with sequences homologous to '*Ca.* Liberibacter asiaticus' genomic DNA sequences and were subsequently confirmed by PCR. Together, these '*Ca.* Liberibacter solanacearum' DNA sequences represent at least 30-fold coverage for the '*Ca.* Liberibacter solanacearum' genome (Lin et al. 2011).

9.3.2.1 Genome Size and Structure

The sequences of the '*Ca.* L. asiaticus' and '*Ca.* L. solanacearum' genomes using metagenomic approaches have been published (Duan et al. 2009; Lin et al. 2011). Both genomes are assembled in a circular chromosome, composed of 1,226,704 and 1,258,278 bp, respectively. No plasmids were detected in either of these Liberibacter genomes. Genome analyses identified 1,136 putative proteins (CDS) in the '*Ca.* Liberibacter asiaticus' genome and 1,192 putative proteins in the Lso genome. Among them, 362 in '*Ca.* Liberibacter asiaticus' and 405 in '*Ca.* Liberibacter solanacearum', were designated as hypothetical proteins with unknown functions representing 31.8 and 31.9 % in each genome, respectively (Table 9.1). Interestingly,

Table 9.1 General genome features of '*Candidatus* Liberibacter solanacearum' and '*Candidatus*. Liberibacter asiaticus'

Genome features	'*Ca.* Liberibacter solanacearum'	'*Ca.* Liberibacter asiaticus'
Size (bp)	1,258,278	1,227,204
GC %	35.2	36.5
Protein-coding genes	1,192	1,136
Hypothetical proteins	405	362
rRNA operons	3	3
tRNA genes	45	44
Pseudogenes	35	32

both '*Ca.* Liberibacter asiaticus' and '*Ca.* Liberibacter solanacearum' genomes have a GC content of 35.2 and 36.5 %, considerably lower than the ~60 % GC content observed for most bacterial species in the *Rhizobiaceae* (Capela et al. 2001; Okubo et al. 2013; Wood et al. 2001). Both genomes contain three copies of complete rRNA operons (16S, 23S, and 5S) and 44 genes encoding for tRNAs in '*Ca.* Liberibacter asiaticus' and 45 genes encoding for tRNAs in '*Ca.* Liberibacter solanacearum.' A two-way comparative analysis of the '*Ca.* Liberibacter solanacearum' and '*Ca.* Liberibacter asiaticus' genomes revealed that both genomes share 828 protein-coding sequences, while 236 coding proteins were unique in '*Ca.* Liberibacter asiaticus' and 186 coding proteins unique in '*Ca.* Liberibacter solanacearum.'

Comparative genomic analysis exploits genome structure and putative function in '*Ca.* Liberibacter asiaticus' and '*Ca.* Liberibacter solanacearum' and reveals similarities and differences in the RNA, protein biosynthesis, and regulatory regions of different species to understand how evolutionary selection has operated upon these elements. Those elements responsible for similarities among species reflect stabilizing selection, while those responsible for differences are considered to be the result of divergent selection. Whole-genome alignment shows that Liberibacters have undergone genome evolutionary processes. Recombination events cause frequent genome rearrangements. Horizontal gene transfer introduces new sequences into bacterial chromosomes, and deletions remove segments of the genome. Consequently, each genome is a mosaic of conserved sequences shared with a subset of other genomes. Figure 9.2 depicts the alignment of genome rearrangements in '*Ca.* Liberibacter asiaticus' and '*Ca.* Liberibacter solanacearum.'

9.3.2.2 Prophage, Genome Rearrangements, and Repetitive Sequences

Prophages are viral sequences integrated in bacterial genomes and contribute to interstrain genetic variability. Virulence-associated genes of many phytopathogenic bacteria are prophage-encoded (de Mello Varani et al. 2008; Srividhya et al. 2007). Computational analyses identified prophage sequences in both '*Ca.* Liberibacter asiaticus' and '*Ca.* Liberibacter solanacearum' genomes. Alignment of both of these Liberibacter genomes indicates that several recombination events have occurred since the divergence of these two species (Lin and Gudmestad 2013). In addition to the prophage genome sequences, there are a number of prophage-like elements and phage remnants distributed throughout the '*Ca.* Liberibacter asiaticus' and '*Ca.* Liberibacter solanacearum' genomes presumably derived from multiple ancestral bacteriophage integration events (Lin et al. 2011). Genome rearrangement is a widespread process in prokaryotes (Eisen et al. 2000; Suyama and Bork 2001). Such structural variations have profound impacts on bacterial phenotype, yet little is known about how natural selection acts on genome arrangement (Darling et al. 2008). Two homologous segments of prophage sequences were identified in the '*Ca.* Liberibacter solanacearum' genome (Lin et al. 2011) as well as in the Chinese '*Ca.* Liberibacter asiaticus' strain GX-1 (NCBI GenBank accession # CP004005). Both genomes possess integrated prophage sequences, which is in agreement with the report of SC1 and SC2 phage sequences in the genome of '*Ca.* Liberibacter asiaticus' strain UF506 (Zhang et al. 2011).

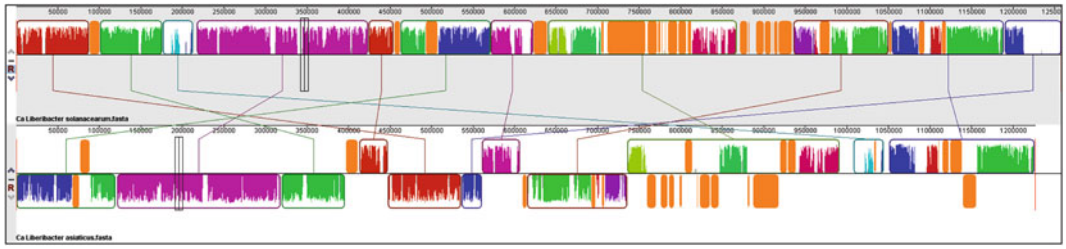

Fig. 9.2 Whole-genome alignment between '*Candidatus* Liberibacter solanacearum' (*Lso*) (*top*) and '*Candidatus* Liberibacter asiaticus' (*Las*) (*bottom*) using MAUVE (http://gel.ahabs.wisc.edu/mauve/). Locally collinear blocks (*LCBs*), regions without rearrangement of homologous backbone sequences, are presented as contiguously colored regions for Lso and Las genomes. LCBs below a genome's centerline are in the reverse complement orientation relative to the reference genome (Lso). Lines between genomes trace each orthologous LCB through each genome. Comparative genome alignment shows considerable genome rearrangements between Lso and Las genomes

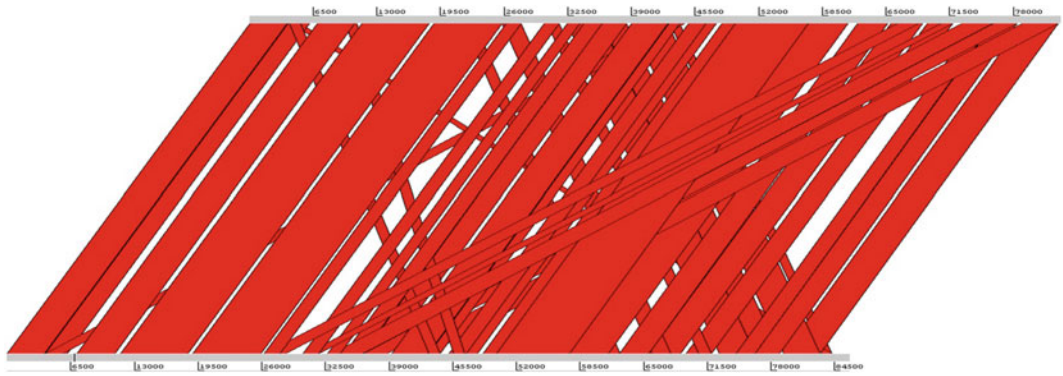

Fig. 9.3 Comparison of the prophage sequences of '*Candidatus* Liberibacter asiaticus' in Chinese strain GX-1 and the SC1 region in a Florida strain. Artemis Comparison Tool (*ACT*) comparison (http://www.sanger.ac.uk/Software/ACT) of amino acid matches between the complete six-frame translations (computed using TBLASTX) of prophage regions. Syntenic conservations of prophage sequences are shown in *red bars* spanning between the genomes represent individual TBLASTX matches

Comparative genome analyses showed that the overall distribution of predicted genes based on functional categories for the Chinese '*Ca.* Liberibacter asiaticus' strain GX-1 was nearly identical to that of previously sequenced Florida strain Psy66 '*Ca.* Liberibacter asiaticus.' However, a tandem alignment of two prophage segments was found in the genome of '*Ca.* Liberibacter asiaticus' strain GX-1. The average GC content in this prophage region is ~40 %, significantly different from that for the core bacterial genome. The prophage sequence of strain GX-1 has 96–99 % similarity to that of '*Ca.* Liberibacter asiaticus' strain UF506 (Fig. 9.3). In the genome of '*Ca.* Liberibacter asiaticus' strain UF506, two circular phage genomes (SC1 and SC2) were identified in '*Ca.* Liberibacter asiaticus'-infected periwinkle (*Catharanthus roseus*) and dodder (*Cuscuta campestris*) (Zhang et al. 2011). When sequences were obtained from psyllid hosts, however, both sequences were integrated in tandem in the UF506 genome (Zhang et al. 2011). It is not clear whether a lack of a circular phage genome is because the genome sequences were derived from infected psyllids or due to other factors. Alignment of the prophage sequences of UF506 and the GX-1 strain shows shared synteny with some rearrangements (Fig. 9.3), indicating the divergence between American and Asian strains.

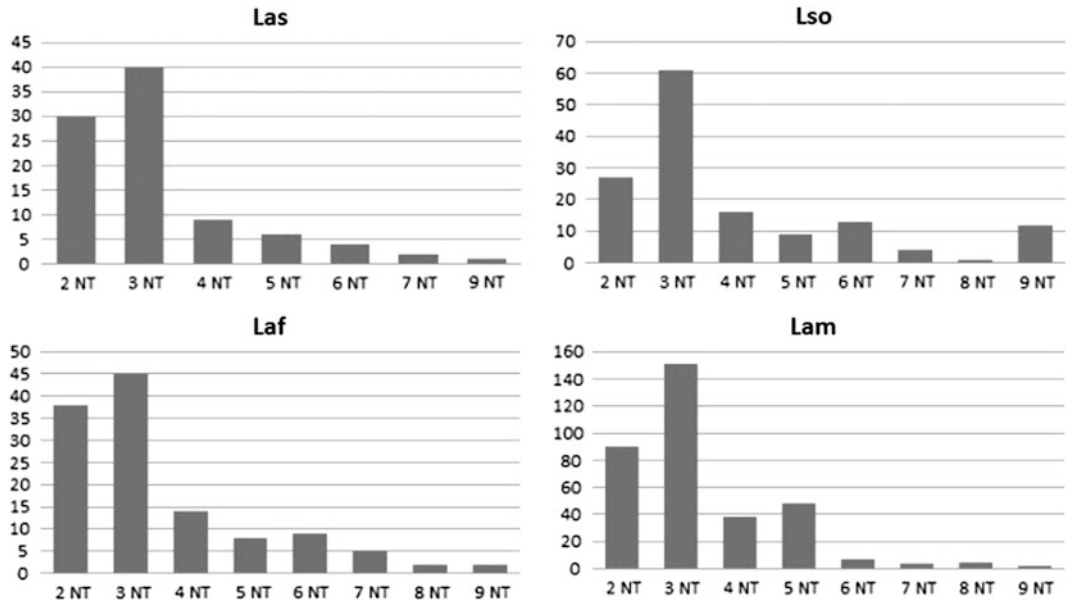

Fig. 9.4 Numbers of simple sequence repeat (*SSR*) motifs identified in '*Candidatus* Liberibacter asiaticus' (*Las*), '*Candidatus* Liberibacter africanus' (*Laf*), '*Candidatus* Liberibacter americanus' (*Lam*), and '*Candidatus* Liberibacter solanacearum' (*Lso*) genomes. SSR motifs include 2–9 nucleotide (*NT*) repeats with perfect, imperfect, and compound repeats

Prophage integration seems to play a key factor in rearrangements in Liberibacter genomes.

Repetitive DNA sequences are ubiquitous in bacterial genomes (van Belkum 1999). Various classes of simple sequence repeat (SSR) motifs have been identified in non-coding and coding regions. Comparative sequence analyses of four Liberibacter genomes (i.e., '*Ca.* Liberibacter asiaticus,' '*Ca.* Liberibacter africanus,' '*Ca.* Liberibacter americanus,' and '*Ca.* Liberibacter solanacearum') revealed that tri-nucleotide repeat motifs are the most abundant in Liberibacter genomes followed by di-, tetra-, and penta-nucleotide repeat motifs (Fig. 9.4). A majority of these motifs is located in hypothetical protein regions, non-coding, or upstream/downstream of coding regions. Those repeat motifs with five or more repeat units are usually useful for the development of multilocus SSR (or microsatellite or VNTR) molecular markers for strain genotyping and identification and population genetic analysis (Islam et al. 2012; Katoh et al. 2011; Lin et al. 2012). SSR DNA markers associated with hypervariable sequence regions can provide sufficient discriminating power for differentiating closely related isolates (Islam et al. 2012). Zhou et al. (2011) identified tandem repetitive DNA sequences in *hyv*I and *hyv*II genes in prophage regions. This type of tandem repeat contains longer repeat motifs (123 base pairs) than SSR repeat motifs. Due to the compact nature of Liberibacter genomes, some SSR loci reside in or near coding regions and, therefore, may have been subjected to selection. Sequence variation in these loci could be associated with the biological, physiological, and biochemical properties of HLB-associated Liberibacters.

9.4 Comparative Genomics and Virulence Genes

9.4.1 Common and Unique Core Genes in Liberibacter Genomes

Comparative pathogen genome analysis identifies both common and divergent pathogenic strategies and reveals the evolutionary lineages

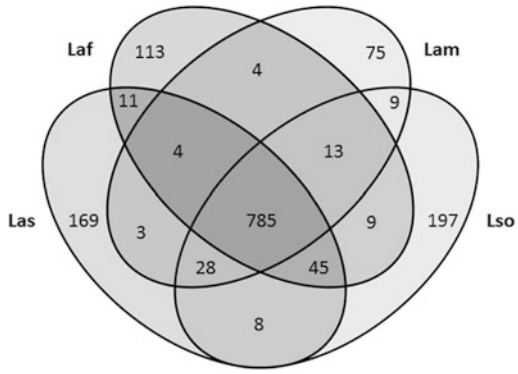

Fig. 9.5 Summary of the four-way analysis of the annotated genes among four. '*Candidatus* Liberibacter *species*': '*Candidatus* Liberibacter asiaticus' (*Las*), '*Candidatus* Liberibacter africanus' (*Laf*), '*Candidatus* Liberibacter americanus' (*Lam*), and '*Candidatus* Liberibacter solanacearum'(*Lso*). There are 785 genes that are conserved and shared in all four Liberibacter species; the remaining genes show different degrees of overlap

of pathogens. Shared virulence strategies can be retained in different bacterial genomes through vertical inheritance from common ancestors due to selective advantage or as a result of the exposure to common host-defensive selection forces (Gardiner et al. 2012). Recently, two additional Liberibacter genomes, '*Ca.* Liberibacter africanus' genome (strain PTSAPSY, NCBI accession # CP004021) and '*Ca.* Liberibacter americanus' genome (strain PW_SP, NCBI accession # AOFG00000000), have been completed (Lin and Gudmestad 2013). Using the available genome data, a four-way comparison of proteins encoding in pathogenic Liberibacter genomes, '*Ca.* Liberibacter asiaticus,' '*Ca.* Liberibacter americanus,' '*Ca.* Liberibacter africanus,' and '*Ca.* Liberibacter solanacearum,' was made (Fig. 9.5). A total of 785 proteins were shared between all four Liberibacter species. Except for hypothetical proteins, most of these proteins were involved in core metabolic pathways. Additionally, 169 unique genes were found in '*Ca.* Liberibacter asiaticus,' 197 in '*Ca.* Liberibacter solanacearum,' 113 in '*Ca.* Liberibacter africanus,' and 75 in '*Ca.* Liberibacter americanus.' There are only four common proteins shared between the three citrus-associated Liberibacters ('*Ca.* Liberibacter asiaticus,' '*Ca.* Liberibacter americanus,' and '*Ca.* Liberibacter africanus'). In contrast, 45 proteins were shared among '*Ca.* Liberibacter asiaticus,' '*Ca.* Liberibacter africanus,' and '*Ca.* Liberibacter solanacearum,' 28 proteins were shared among '*Ca.* Liberibacter asiaticus,' '*Ca.* Liberibacter americanus,' and '*Ca.* Liberibacter solanacearum,' and 13 proteins were shared among '*Ca.* Liberibacter americanus,' '*Ca.* Liberibacter africanus,' and '*Ca.* Liberibacter solanacearum.' Thus, '*Ca.* Liberibacter americanus' has the fewest proteins in common among these four Liberibacters (Fig. 9.5). '*Ca.* Liberibacter solanacearum,' which is associated with potato zebra chip disease, is genetically closer to, and shares more common genes with, citrus HLB-associated '*Ca.* Liberibacter asiaticus' and '*Ca.* Liberibacter africanus' than with '*Ca.* Liberibacter americanus.' These relationships are congruent with phylogenetic analysis based on orthologous coding genes and on 16S rRNA gene sequences.

9.4.2 Virulence Genes and Pathogenicity Determinants

The annotation of Liberibacter genomes has given insight into the genomic inventory that enables prediction of putative biochemical information regarding these bacteria. While the precise substrate specificity for those significant pathways cannot be determined without functional confirmation, genomic information in general offers a good overview of the biochemical, metabolic capabilities and potential pathogenicity of these pathogens. This is particularly important for unculturable bacteria whereby conventional characterizations by standard bacteriological approaches are often limited or not possible.

Many Gram-negative plant pathogenic bacteria possess a type III secretion system (TTSS) to inject effector proteins directly into host cells (Alfano and Collmer 2004). TTSSs are encoded by hypersensitive response and pathogenicity genes, which cause disease in susceptible plants, but elicit the hypersensitive response in resistant

plants (Büttner and He 2009). Unexpectedly, no TTSS was identified in either '*Ca.* Liberibacter asiaticus' or '*Ca.* Liberibacter solanacearum' genome. Interestingly, the absence of a TTSS system was observed in some insect-mediated, bacteria-associated diseases such as citrus variegated chlorosis (CVC) and grape Pierce's disease (PD), caused by xylem-limited *Xylella fastidiosa* (Simpson et al. 2000; Van Sluys et al. 2003). Since the route of entry for these pathogens into the host phloem or xylem cells is via injection by the insect vectors, a secretion apparatus may not be necessary to initially establish infection and colonization in hosts (Lin and Gudmestad 2013).

While a TTSS is absent in Liberibacter genomes, Liberibacters might encode other uncharacterized effectors in their genomes. Flagellum-encoding gene clusters have been identified in both '*Ca.* Liberibacter asiaticus' and '*Ca.* Liberibacter solanacearum' (Duan et al. 2013; Lin and Gudmestad 2013). A high degree of sequence similarity between some conserved TTSS proteins and flagellar paralogs has suggested that TTSS and flagella evolved from a common ancestor (Aizawa 2001). However, phylogenetic relationships based on 16S rRNA and protein sequences between TTSS and flagella have led to the identification of several major lateral transfer events involving clusters of TTSS genes (Gophna et al. 2003; Nguyen et al. 2000). It is, therefore, hypothesized that horizontal gene transfer is a major force driving the evolution of species that harbor TTSSs (Aizawa 2001; Blocker et al. 2003; Gophna et al. 2003). Many bacterial species are motile by means of flagella. Motility is an important virulence factor in pathogenic species. Several lines of evidence indicate that the activity of the flagellum may have an impact on virulence gene regulation (Ottemann and Miller 1997). Recent studies have shown that flagella are involved in multiple roles including adhesion, biofilm formation, virulence factor secretion, and modulation of the immune system of eukaryotic cells besides motility and chemotaxis (Duan et al. 2013). Deakin (2000) demonstrated that the flagellum export apparatus was capable of transporting other proteins into the extracellular milieu. One of these exported proteins was identified as a virulence-associated phospholipase. Vascular disease pathogens require motility in order to establish systemic infection and spread beyond the initially infected tissue. Although '*Ca.* Liberibacter asiaticus' and '*Ca.* Liberibacter solanacearum' have significantly reduced genome sizes, both retain the entire flagellum-encoding gene clusters, resembling those found in other members of the *Rhizobiaceae* (Duan et al. 2009; Lin et al. 2011). Using a heterologous expression system, Zou et al. (2012) demonstrated that flagellin genes encoded by '*Ca.* Liberibacter asiaticus' Psy62 could be involved in the filament production. It has been shown that the *fla* gene possesses a conserved N-terminal part with a 22-amino acid sequence (flg22) which shares an 86 % amino acid sequence identity with '*Ca.* Liberibacter solanacearum' strain ZC-1. The flg22 protein is known to activate plant's defense mechanisms. A transient expression assay demonstrated that Fla_{Las} induced cell death and callose deposition in tobacco leaves (Zou et al. 2012). Therefore, it has been suggested that flagellin may trigger host defense response and act as a pathogen-associated molecular pattern (PAMP). Although flagellar genes were identified in '*Ca.* Liberibacter asiaticus' and '*Ca.* Liberibacter solanacearum' genomes, flagellar structures have not been observed microscopically in either Liberibacter species.

Obligate intracellular prokaryotes depend on constantly exchanging chemical compounds with their hosts. An ATP-binding cassette (ABC) transporter is believed to be associated with virulence in some Gram-negative bacteria. The activity of ABC transporters results in the imbalance of host metabolites (Li et al. 2012). Genome annotation indicates that a greater proportion of genes in Liberibacters involve a transporter than in other *Rhizobiaceae* bacterial genomes. Duan et al. (2009) identified 40 ATP-binding cassette (ABC) transporters in the '*Ca.* Liberibacter asiaticus' (strain Psy62) genome. More recently, 15 more transporter genes were predicted using in-depth sequence and structure analysis (Li et al. 2012). The large number of transport proteins allows Liberibacter bacteria to acquire metabolites and nutrients from their

hosts, resulting in a metabolic imbalance in their host. For example, virulence genes associated with *ZnuABC* were found in Liberibacter genomes. This zinc transport system is composed of a cluster of three genes (*ZnuABC*) that encode a prokaryotic-type ABC family protein complex (Gabbianelli et al. 2011; Vahling-Armstrong et al. 2012).

It is postulated that the putative zinc transport system is a high-affinity cation carrier that shuttles zinc ions across the inner membrane (Ammendola et al. 2007). Both '*Ca.* Liberibacter asiaticus' and '*Ca.* Liberibacter solanacearum' genomes harbor zinc transport systems. The yellowing symptoms in Liberibacter-associated HLB and ZC diseases in citrus and potato, respectively, are presumably attributed to the presence of this putative high-affinity zinc carrier. A recent study in '*Ca.* Liberibacter asiaticus' suggested that disruption of this system by making more zinc available to the plant may decrease the virulence of the bacterium and help mitigate HLB symptom severity (Vahling-Armstrong et al. 2012).

All bacterial ATP/ADP translocases characterized thus far have been found in few obligate intracellular bacteria including *Rickettsia prowazekii*, *Chlamydia psittaci*, and *Lawsonia intracellularis* (Hatch et al. 1982; Schmitz-Esser et al. 2008). Sequence annotation identified an ATP/ADP translocase in the NttA family for the '*Ca.* Liberibacter asiaticus' and '*Ca.* Liberibacter solanacearum' genomes. Functional analysis confirmed that this transport protein was able to import ATP directly from the cytosol of eukaryotic host cells into the bacterial cells. Unexpectedly, orthologs are absent from all other *Rhizobiaceae* bacteria. Given the reduced size of the '*Ca.* Liberibacter asiaticus' genome (1.2 Mb), it appears that this transport protein was obtained through horizontal gene transfer early in the evolution, or prior to the speciation, of '*Ca*. Liberibacter.' Though the biochemical function of the translocase was confirmed via heterologous expression in *Escherichia coli* (Vahling et al. 2010), its role in pathogenicity has yet to be determined. The capability of acquiring ATP from the host could play an important role in thriving as energy parasites in obligate intracellular bacteria with reduced genomes. Further study is needed to evaluate the significance of this feature.

Sequence analysis also indicated that the '*Ca.* Liberibacter asiaticus' genome encoded a NahG-like salicylate hydroxylase. Interestingly, the orthologous gene was also identified in the '*Ca.* Liberibacter solanacearum' genome (Lin et al. 2011). It has become clear that salicylic acid (SA) accumulation is essential to a number of processes related to disease resistance (Brodersen et al. 2005). The plant defense response to microbial infection, known as systemic acquired resistance (SAR), results in broad-spectrum disease resistance. SA has been implicated as a component of the signaling pathway that leads to basal defense, the hypersensitive response, and SAR. Transgenic tobacco and tomato plants that express the bacterial gene encoding salicylate hydroxylase have been shown to accumulate very little SA and to be defective in their ability to induce SAR (Block et al. 2005; Friedrich et al. 1995). Similar results were observed in transgenic NahG *Arabidopsis* plants (van Wees and Glazebrook 2003). This defect, caused by the breakdown of SA to catechol as a result of *nahG* gene enzymatic activity (Bartsch et al. 2006; Yalpani et al. 1991), not only makes the plants unable to induce SAR, but also leads to increased susceptibility to viral, fungal, and bacterial pathogens. Thus, functional confirmation and suppression of *nahG* genes in '*Ca.* Liberibacter asiaticus' and '*Ca.* Liberibacter solanacearum' could enhance host SAR to Liberibacter-associated diseases.

9.4.3 Bacterium–Host Interactions

Plant defense systems constitute a complex signal transduction network and metabolic pathways in response to diverse pathogens. These responses are mediated by signal transduction intermediaries including SA and jasmonic acid and other phytohormones, leading to metabolic reprogramming (Ham and Bent 2002). Phytopathogenic bacteria attack plant hosts

using various virulence strategies and employ a suite of pathogenic factors to cause diseases. In response to pathogen invasion, host plants have evolved various defense mechanisms as protection from pathogen attack, including inducible and constitutive defense responses. A key to enhancing host resistance to Liberibacter-associated diseases is to identify the virulence strategies employed by Liberibacters and understand how these bacteria modify or suppress their hosts' pathogen defense mechanisms. Furthermore, it is critical to induce host defense responses in early infection because host defense responses that occur at a relatively late stage of infection would be ineffective in suppressing disease development.

All commercially grown citrus cultivars are susceptible to HLB. However, disease severity and symptom expression characteristics vary among different species and cultivars of HLB-affected plants. For example, lemon (*Citrus limon*) and pummelo (*C. maxima* syn. *C. grandis*) seem to be more tolerant than sweet orange cultivars (Bové 2006; Fan et al. 2010). Similarly, some potato cultivars and germplasm selections appear to have more tolerance to ZC disease than 'Atlantic,' a highly susceptible potato cultivar (Miller et al. 2011). Apparently, the degree of tolerance varies depending on experimental conditions, including the time of year, plant age, nutritional status, and variable virulence of '*Ca.* Liberibacter solanacearum' strains used for screening experiments. Since no genetically resistant crops have been found against Liberibacter-associated diseases, research efforts have been focused on understanding the molecular basis of host responses to HLB and ZC diseases.

Transcriptomic studies using microarray technology have been conducted to elucidate the effect of '*Ca.* Liberibacter asiaticus' infection on the total mRNA expression profiles in tissues of sweet orange (*C. sinensis*) plants (Albrecht and Bowman 2008; Kim et al. 2009). These studies concluded that genes commonly involved in the defense responses to bacterial infection were incapable of induction at an early stage of infection and, thus, were not able to suppress disease development. Meanwhile, a large number of genes possibly associated with pathogen defense in HLB tolerance citrus (US-897, a hybrid of trifoliate orange) were expressed at much higher levels but were independent of infection. It appears that HLB causes alteration or imbalance of host metabolism including the accumulation of starch in HLB-affected leaves coupled with the decrease in photosynthesis activity (Albrecht and Bowman 2008, 2012; Kim et al. 2009). To exploit the molecular mechanisms involved in the response of citrus plants to '*Ca.* Liberibacter asiaticus' infection at a cellular level, a proteomic study of global gene expression was performed using grapefruit (*C. paradisi*) and lemon (*C. limon*) plants (Nwugo et al. 2013a, b). These studies identified several proteins including Cu/Zn superoxide dismutase, chitinases, lectin-related proteins, miraculin-like proteins, peroxiredoxins, and a CAP 160 protein that were upregulated in '*Ca*. Liberibacter asiaticus'-infected but presymptomatic grapefruit plants.

The use of molecular markers that respond specifically to '*Ca.* Liberibacter asiaticus' infection at an asymptomatic stage could be developed as host-based molecular diagnostic methods for the early detection of '*Ca.* Liberibacter asiaticus'-infected plants. The increase in starch accumulation in leaves is concomitant with an increase in expression of starch synthase in HLB-diseased lemon plants. Interestingly, upregulation of starch synthase was coordinated with the accumulation of potassium even in early stages of infection in asymptomatic grapefruit plants (Nwugo et al. 2013b). Potassium is required for the activation of starch synthase. However, it is not clear whether an increase in the accumulation of starch is also partly due to a slowdown of the starch's breakdown metabolic pathways since photosynthetic activity is decreased as a result of HLB development (Nwugo et al. 2013a, b). Disease symptoms are often associated with the altered nutritional status of plants, and nutrient–disease interactions are well documented in plant pathosystems (Huber and Haneklaus 2007). The reduction in root nutrient uptake or a blocked vascular system, such as that implicated in HLB disease development, can induce a systemic or localized

nutrient sufficiency or deficiency. Physiological symptoms of HLB resemble those of Zn deficiency (Cevallos-Cevallos et al. 2011), and the productive life of HLB-affected plants has been shown to be compensated by fertilizer application (Huber and Haneklaus 2007; Shokrollah et al. 2011). Most macro- and micronutrients (Ca, Mg, Fe, Mn, Zn, and Cu) except K are decreased during HLB development (Nwugo et al. 2013a, b). However, an accumulation of zinc was observed in lemon plants, which is in contrast to observations in grapefruit and other previously studied citrus plants (Cevallos-Cevallos et al. 2011). More recently, a microRNA, miR399, was identified in sweet orange and was reported to be specifically induced by '*Ca.* Liberibacter asiaticus' infection (Zhao et al. 2013). The up-regulation of miR399 is induced by phosphorus deficiency which in turn enhances P acquisition. The study showed that the miR399-mediated regulatory mechanism for phosphorus homeostasis is conserved in citrus (Zhao et al. 2013). While additional P or other macro- and micronutrient applications can mitigate the effects of '*Ca.* Liberibacter asiaticus' infection and temporarily improve tree growth and fruit yield, these measures cannot stop HLB development once the tree is infected. Nutrient homeostasis forms part of a delicately balanced interdependent system with plant gene regulation; however, there is limited information on the relationships between the nutritional status and protein expression profiles of citrus plants during HLB development (Nwugo et al. 2013a, b).

Potato plants have a shorter host regeneration time than that of citrus plants and visibly respond much faster to '*Ca.* Liberibacter solanacearum' infection compared with the time for visual citrus response to '*Ca.* Liberibacter asiaticus' infection (Alvarado et al. 2012). Thus, potato ZC disease could serve as a model system for understanding the mechanisms in Liberibacter-associated diseases. An increase in total protein production has been observed in ZC-diseased potatoes (Alvarado et al. 2012). '*Ca.* Liberibacter solanacearum' infection results in reprogramming cellular metabolism in ZC-diseased potato plants. For example, increase in starch and elevated accumulation of patatin proteins, abundantly present in healthy tubers, were present in stems of ZC-affected potato plants (Alvarado et al. 2012). In addition, ZC-affected plants contained more phenolics than the healthy plants as well as higher levels of polyphenol oxidase (Alvarado et al. 2012; Navarre et al. 2009; Wallis et al. 2012).

More recently, a study of transcriptional and translational profiles of host responses to '*Ca.* Liberibacter solanacearum' infection revealed differential expression patterns between aboveground (AG) and belowground (BG) parts of potato plants (Nwugo and Lin 2012). Studies showed that in spite of a marked down-regulation of photosynthesis-related genes/proteins, a consistent '*Ca.* Liberibacter solanacearum'-mediated up-regulation of pathogen response-related genes/proteins was observed in AG and BG tissues. However, while proteinase inhibitors, patatin, and polyphenol oxidase were down-regulated in BG tissues, there was evidence of alternative splicing in a proteinase inhibitor II precursor gene in AG tissues upon '*Ca.* Liberibacter solanacearum' infection, which generally agrees with the results reported earlier. Furthermore, several antioxidative proteins, including superoxide dismutase, were down-regulated in AG tissues but not in BG tissues, which might play a role in potato susceptibility to ZC (Nwugo, unpublished data). These studies shed light on biochemical mechanisms involved in potato-'*Ca.* Liberibacter solanacearum' interactions and suggest the reprogramming of AG tissues to behave like BG tissues during ZC development.

9.5 Evolutionof Pathogenic Bacteria—Gene Acquisition and Gene Loss

9.5.1 Obligate Intracellular Liberibacter Bacteria

Plant and insect hosts provide luxurious nutrient and physical protection niches for intracellular bacteria. In order to survive and multiply

successfully in such habitats, intracellular bacteria must be able to enter the hosts, multiply inside the host cells, move within the host, and eventually reinvade a new host. Throughout these processes, the bacterium must avoid being recognized and killed by the innate host immune system (Moulder 1985). Obligate intracellular bacteria survive and multiply strictly in their hosts in contrast to free-living or facultative intracellular parasites that can be cultured in vitro. The molecular mechanisms associated with these different lifestyles of plant- and insect-associated microbes are not clear. It is believed that the fastidious nature of obligate host-associated bacteria is largely due to the lack of genes encoding essential metabolic pathways and reliance on compounds available in the intracellular environment. Unculturable Liberibacters exemplify such obligate intracellular bacteria in plants. Despite numerous attempts made to culture these disease-associated Liberibacters on artificial growth media, it has not yet been possible to maintain their growth in vitro (Davis et al. 2008).

Bacterial genomes differ greatly in their sizes and gene repertoires. Genome size and gene number in bacteria often reflect their lifestyles. Obligate intracellular bacteria generally have smaller genome sizes and fewer genes than their closely related free-living bacteria. As more complete bacterial genome sequences become available, comparative genomic analysis becomes an essential approach to identify orthologous genes that have been inherited from their ancestors and non-orthologous genes acquired through horizontal gene transfer. Such comparative analyses are usually conducted by comparing uncharacterized bacterial genomes with those of well-characterized organisms. For example, genomic information from *A. tumefaciens* and *S. meliloti* facilitates identification and annotation of phylogenetically related and orthologous genes in '*Ca.* Liberibacter' species. Most phloem-limited bacteria have resisted in vitro cultivation, even though they multiply actively in their insect vectors or in phloem sap within the sieve tubes (Bové and Garnier 2003). This implies that the phloem sap in plants or microbial communities in insect guts may provide indispensable nutrients or growth factors for multiplication that have not yet been identified for incorporation into artificial culture media. Like obligate intracellular bacteria, the phloem-restricted Liberibacter bacteria have relatively small genomes of ~1.2 Mbp. Likewise, the reduction in genome size through gene loss is associated with the evolutionary adaption to a lifestyle as an obligate intracellular pathogen that largely depends on intermediate metabolites acquired from its hosts (Lin and Gudmestad 2013).

9.5.2 Reduced Genome Size

The size and gene content of Liberibacter genomes are greatly reduced compared with other bacterial species in the *Rhizobiaceae* which are reflective of adaptation of these bacteria to an intracellular lifestyle in their insect and plant hosts (Hartung et al. 2011; Lin and Gudmestad 2013). Comparative orthologous genes in the '*Ca.* Liberibacter asiaticus' and '*Ca.* Liberibacter solanacearum' and phylogenetically related *A. tumefaciens* and *S. meliloti* genomes are presented in Table 9.2. In addition, these are compared with the smallest genomes of free-living bacteria, namely *Mycoplasma genitalium* (Fraser et al. 1995), and the strictly endosymbiont, *Carsonella ruddii*, a bacteriocyte-associated γ-Proteobacterial symbiont present in most species of phloem sap-feeding insects, such as psyllids (Nakabachi et al. 2006).

Interestingly, genes encoding for sulfur metabolism are completely absent in both '*Ca.* Liberibacter asiaticus' and '*Ca.* Liberibacter solanacearum' genomes. Sulfur is a ubiquitous element and present as inorganic and organic forms in cells. Oxidation of sulfur can generate ATP. In some archaea, sulfur is the sole energy source. Despite the reduced size of Liberibacter genomes, the genes necessary for replication, transcription, and ribosomal protein synthesis are maintained. For example, both '*Ca.* Liberibacter asiaticus' and '*Ca,* L. solanacearum' genomes have 74 genes (9 %) and 72 genes (8.7 %)

involved in nucleotide metabolism, respectively. In contrast, 4.9 and 5.0 % of the genes are involved in the same categories in *A. tumefaciens* and *S. meliloti* genomes. Likewise, in both Liberibacter genomes, 15.9 % of the genes are associated with translation processes, while 6.0 and 4.9 % of the genes are associated with translation in *A. tumefaciens* and *S. meliloti*, respectively. On the other hand, significant gene content reduction was found in categories involved in carbohydrate and amino acid metabolism. For instance, about 102 genes (12.4 %) were identified to function in carbohydrate metabolism, of which 14, 19, and 9 genes are identified with the function in glycolysis, citrate and pentose phosphate pathways in both Liberibacter genomes. There are 408 genes (16.3 %) in *A. tumefaciens* and 486 genes (16.7 %) in *S. meliloti* that participate in carbohydrate metabolism, respectively. Analysis of amino acid metabolism revealed only 59 genes (7.2 %) in '*Ca.* L. solanacearum' and 69 genes (8.4 %) in '*Ca.* Liberibacter asiaticus' are identified as compared to 358 genes (14.5 %) in *A. tumefaciens* and 452 genes (15.5 %) in *S. meliloti.* Genes involved in valine, leucine, and isoleucine biosynthesis, and genes required for histidine, phosphonate, and phosphinate metabolism, are missing in both Liberibacter genomes. However, in '*Ca.* Liberibacter solanacearum,' no gene was found for cysteine and methionine metabolism pathways (Lin et al. 2011). It is believed that unicellular organisms do not normally take up proteins from the outside environment. Despite genome size reduction, genes associated with pathways in anabolism to construct essentially living macromolecules are indispensable for forming macromolecules essential for life. Bacteria retained the genes for nucleic acid replication and protein synthesis and also enzymes for metabolic pathways to form basic building blocks, such as amino acids, and sugar, that are available in the immediate environment (i.e., the plant or insect host). In fact, a large number of genes (12 %) encoding transporter proteins were identified in Liberibacter genomes (Duan et al. 2009; Lin et al. 2011), suggesting the direct acquisition of nutrients and intermediate metabolites from a metabolically static phloem environment, compensate for missing metabolic pathways caused by a massive genome reduction. Indeed, extensive genome reduction also has been noted for obligate intracellular bacteria and endosymbionts that have developed intimate relationships with their eukaryotic hosts (McCutcheon and Moran 2010; Moya et al. 2008; Thao et al. 2000). Among them is the extreme case of *C. ruddii*, a strictly endosymbiont that has the smallest prokaryote genome known to date. Its genome consists of a circular chromosome of 159,662 bp and has a high coding density (97 %) with many overlapping genes and reduced gene length. The number of predicted genes is 182, the lowest on record thus far in the NCBI genome databases. Analysis shows that genes involved in lipid and amino acid metabolism are completely absent (Table 9.2).

Reduction in genome size is also correlated with decreased GC content (Hartung et al. 2011; Nakabachi et al. 2006). The relationship between GC content and genome size based on the characterization of the genomes of more than 300 bacteria in the α-, β-, and γ-Proteobacterial groups ranged from 16.5 % of GC content for *C. ruddii* to 71.4 % of GC content of *Sorangium cellulosum* (with genome size 13,033,779 base pairs) (Schneiker et al. 2007) (Figure 9.6). A positive correlation between bacterial genome size and GC content was observed with $R^2 = 0.376$. While the basis for this correlation is not clear, a plausible explanation is that the synthesis of GTP and CTP is energetically more costly than that of ATP and TTP (Rocha and Danchin 2002). Under resource-limiting conditions, it could be a favorable or selective advantage for microbial genomes to drift toward AT richness. Additionally, mutations from guanine–cytosine (GC) to adenosine–thymine (AT) are more common than mutations from AT to GC (Hershberg and Petrov 2010; Lind and Andersson 2008) largely due to the deamination of cytosine to form uracil, which is then replicated as thymidine (Glass et al. 2006). Consequently, genomes develop a bias toward less GC-rich chromosomes through the course of evolution. The extensive reduction in many bacterial genomes has inspired researchers to

Table 9.2 Summary of comparative gene contents *Candidatus* Liberibacter solanacearum' (*Lso*), '*Candidatus* Liberibacter asiaticus' (*Las*), *Carsonella ruddii* (*C. rud*), *Mycoplasma genitalium* (*M. gen*), *Agrobacterium tumefaciens* (*A. tum*), and *Sinorhizobium meliloti* (*S. mel*) sorted by functional category

Metabolism	Lso	Las	*C. rud*	*M. gen*	*A. tum*	*S. mel*
Carbohydrate metabolism	**102**	**102**	**8**	**58**	**403**	**486**
Glycolysis /gluconeogenesis	14	14		15	34	47
Citrate cycle (TCA cycle)	19	19	2	4	24	24
Pentose phosphate pathway	9	9		8	28	30
Pentose and glucuronate interconversions	2	2	1	2	26	30
Fructose and mannose metabolism	2	3		5	17	26
Galactose metabolism	4	4		3	18	28
Ascorbate and aldarate metabolism	1				14	16
Starch and sucrose metabolism	3	3		3	27	40
Amino sugar and nucleotide sugar metabolism	9	9		4	32	41
Pyruvate metabolism	14	14	2	8	43	50
Glyoxylate and dicarboxylate metabolism	7	7	2	1	52	46
Propanoate metabolism	6	6		3	30	40
Butanoate metabolism	7	7	1	2	35	40
C5-branched dibasic acid metabolism	2	2			9	9
Inositol phosphate metabolism	3	3		1	14	19
Energy metabolism	**61**	**63**	**8**	**16**	**140**	**213**
Oxidative phosphorylation	33	34	8	9	52	79
Methane metabolism	7	8		7	27	33
Nitrogen metabolism	21	21			50	89
Sulfur metabolism					11	12
Lipid metabolism	**28**	**28**	**0**	**12**	**121**	**147**
Fatty acid biosynthesis	14	14			28	24
Fatty acid metabolism					25	40
Synthesis and degradation of ketone bodies		1			7	5
Secondary bile acid biosynthesis					1	21
Glycerolipid metabolism	3	3		4	17	29
Glycerophospholipid metabolism	9	8		8	21	1
Sphingolipid metabolism					3	3
Arachidonic acid metabolism					2	3
Linoleic acid metabolism					3	8
Alpha-linolenic acid metabolism					1	1
Biosynthesis of unsaturated fatty acids	2	2			13	12
Nucleotide metabolism	**74**	**72**	**12**	**45**	**133**	**147**
Purine metabolism	40	39	5	21	78	93
Pyrimidine metabolism	34	33	7	24	55	54
Amino acid metabolism	**59**	**69**	**39**	**7**	**358**	**452**
Alanine, aspartate, and glutamate metabolism	14	14	4		33	36
Glycine, serine, and threonine metabolism	12	13	5	3	53	63
Cysteine and methionine metabolism		7	4	2	20	25
Valine, leucine, and isoleucine degradation	3	3	1	1	32	40
Valine, leucine, and isoleucine biosynthesis			8		17	20

(continued)

Table 9.2 (continued)

Metabolism	Lso	Las	*C. rud*	*M. gen*	*A. tum*	*S. mel*
Lysine biosynthesis	11	12	9		22	21
Lysine degradation	2	3			19	21
Arginine and proline metabolism	12	12	3	1	51	63
Histidine metabolism					25	37
Tyrosine metabolism	1	1			21	45
Phenylalanine metabolism	1	1			22	29
Tryptophan metabolism	1	1			19	23
Phenylalanine, tyrosine, and tryptophan biosynthesis	2	2	5		24	29
Metabolism of other amino acids	**19**	**19**	**5**	**6**	**72**	**92**
beta-alanine metabolism					16	23
Taurine and hypotaurine metabolism	1	1		2	3	6
Phosphonate and phosphinate metabolism				3	4	6
Selenocompound metabolism	3	3	3		10	11
Cyanoamino acid metabolism	1	1	1	1	6	6
D-Glutamine and D-glutamate metabolism	4	4			4	4
D-Arginine and D-ornithine metabolism					2	1
D-Alanine metabolism	2	2			5	5
Glutathione metabolism	8	8	1	1	22	30
Glycan biosynthesis and metabolism	**21**	**21**	**1**	**2**	**31**	**37**
Lipopolysaccharide biosynthesis	8	8			9	9
Peptidoglycan biosynthesis	13	13			20	20
Other glycan degradation					2	8
Metabolism of cofactors and vitamins	**68**	**60**	**1**	**2**	**160**	**171**
Thiamine metabolism	3				9	12
Riboflavin metabolism	5	5		1	8	9
Vitamin B6 metabolism	5	5	1		7	6
Nicotinate and nicotinamide metabolism	5	5		4	16	20
Pantothenate and CoA biosynthesis	5	5	4	2	20	23
Biotin metabolism	11	11			22	16
Lipoic acid metabolism	2	2		1	2	2
Folate biosynthesis	7	4		1	13	15
One carbon pool by folate	9	9	2	6	16	17
Porphyrin and chlorophyll metabolism	11	10	1	1	38	40
Ubiquinone and other terpenoid-quinone biosynthesis	5	4			9	11
Metabolism of terpenoids and polyketides	**20**	**20**	**1**	**1**	**58**	**62**
Terpenoid backbone biosynthesis	9	9		1	13	15
Carotenoid biosynthesis	1	1			1	1
Limonene and pinene degradation					16	18
Geraniol degradation					11	7
Biosynthesis of ansamycins	1	1	1	1	3	3
Tetracycline biosynthesis	4	4			6	6
Polyketide sugar unit biosynthesis	4	4			6	7
Non-ribosomal peptide structures					0	1
Biosynthesis of vancomycin group antibiotics	1	1			2	4

(continued)

Table 9.2 (continued)

Metabolism	Lso	Las	*C. rud*	*M. gen*	*A. tum*	*S. mel*
Biosynthesis of other secondary metabolites	**10**	**10**	**0**	**0**	**24**	**30**
Streptomycin biosynthesis	8	8			2	0
Butirosin and neomycin biosynthesis	1	1			2	2
Novobiocin biosynthesis	1	1			14	19
Butirosin and neomycin biosynthesis					1	1
Novobiocin biosynthesis					5	8
Xenobiotics biodegradation and metabolism	**7**	**7**	**0**	**0**	**122**	**216**
Benzoate degradation					23	30
Aminobenzoate degradation					11	20
Fluorobenzoate degradation					1	4
Chloroalkane and chloroalkene degradation					22	44
Chlorocyclohexane and chlorobenzene degradation					9	14
Toluene degradation	4	4			9	15
Xylene degradation					3	7
Nitrotoluene degradation					2	1
Ethylbenzene degradation					3	8
Styrene degradation					0	6
Atrazine degradation					3	7
Caprolactam degradation					7	8
Bisphenol degradation					7	14
Dioxin degradation	1	1			3	8
Naphthalene degradation	1	1			11	19
Polycyclic aromatic hydrocarbon degradation	1	1			8	11
Genetic information processing						
Transcription	**4**	**4**	**3**	**4**	**4**	**4**
RNA polymerase	4	4	3	4	4	4
Translation	**131**	**131**	**81**	**77**	**148**	**143**
Ribosome	62	62	37	52	67	65
Aminoacyl-tRNA biosynthesis	69	69	44	25	81	78
Folding, sorting and degradation	**22**	**22**	**3**	**15**	**38**	**42**
Protein export	11	12		9	15	16
Sulfur relay system	2	2	1	1	9	9
RNA degradation	9	8	2	5	14	17
Replication and repair	**65**	**58**	**9**	**44**	**91**	**99**
DNA replication	14	12	4	13	16	17
Base excision repair	9	8		5	16	18
Nucleotide excision repair	7	7		6	11	12
Mismatch repair	16	14	2	9	21	23
Homologous recombination	19	17	3	11	21	22
Non-homologous end joining					6	7
RNA family	**44**	**53**	**29**	**2**	**66**	**59**
Non-coding RNAs	44	53	29	2	66	59
Environmental information processing						
Membrane transport	**38**	**38**	**0**	**28**	**335**	**286**

(continued)

Table 9.2 (continued)

Metabolism	Lso	Las	*C. rud*	*M. gen*	*A. tum*	*S. mel*
ABC transporters	28	28		17	282	228
Phosphotransferase system (PTS)	1			3	3	3
Bacterial secretion system	9	10		8	50	55
Signal transduction	**16**	**15**	**0**	**2**	**100**	**131**
Two-component system	16	15		2	100	131
Cellular processes						
Cell motility	**33**	**32**	**0**	**0**	**71**	**95**
Bacterial chemotaxis	7	6			40	58
Flagellar assembly	26	26			31	37

Data were obtained from Kyoto Encyclopedia of Genes and Genomes (*KEGG* http://www.genome.jp/kegg/)

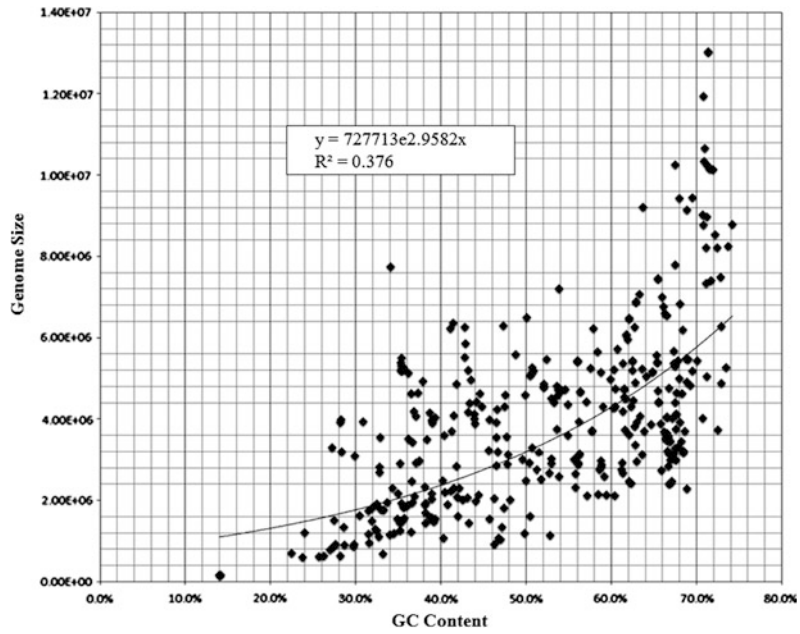

Fig. 9.6 Correlation between GC content and genome size of 325 complete bacterial genomes

search for 'a minimal gene set for a cellular life.' The concept of a minimal gene set refers to the smallest group of genes that would be, in principle, sufficient to sustain a functioning life form in the presence of a full complement of essential nutrients and in the absence of environmental stress (Hutchison et al. 1999; Koonin 2000). In spite of vast differences in bacterial genome sizes, a minimal gene set needed for a free-living organism should meet the following criteria: (1) ability to maintain metabolic homeostasis (essential housekeeping functions) and (2) multiplication relying on its own metabolic machineries (Gil et al. 2004). *M. genitalium* has the smallest genome of free-living prokaryote with 521 predicted genes (470 protein-encoding genes) in one circular chromosome (Fraser et al. 1995). Its small genome size and cultivable nature made *M. genitalum* the organism of choice in research on minimal genome. Using global transposon knockout mutagenesis, researchers determined the minimal gene set to range from 206 to 387 (Gil et al. 2004; Glass et al. 2006). Unexpectedly, the essential gene set also includes not yet annotated and hypothetical proteins.

Understanding of the principles of minimal gene sets for cell functions and deciphering the roles of essential, but functionally uncharacterized hypothetical proteins in prokaryote genomes, will advance knowledge of fundamental cellular biology. Information derived from this research will also facilitate development or improvement of in vitro culture media for Liberibacter bacteria.

9.6 Conclusion

Liberibacters comprise a recently recognized group of Gram-negative bacteria in the class α-Proteobacteria associated with a diverse and complex range of associations with plants and psyllid insects. The currently recognized species are associated with asymptomatic cultivated and non-cultivated plants, plant diseases, and psyllid insects. These include four species ('*Ca.* Liberibacter asiaticus,' '*Ca.* Liberibacter africanus,' '*Ca*. Liberibacter americanus,' and '*Ca*. Liberibacter solanacearum (syn. psyllaurous)') that are associated with diseases of cultivated agricultural crops (citrus huanglongbing or greening; potato zebra chip; yellows-like disorders of carrots, celery, tomato, and disorders of other solanaceous hosts) around the world; one subspecies ('*Ca.* Liberibacter africanus subspecies capensis') associated with a disease of Cape chestnut in South Africa); and two species isolated from asymptomatic pear in Italy ('*Ca.* Liberibacter europaeus') and papaya (*L. crescens*) in Puerto Rico. In plants, these Liberibacters are restricted to the phloem. These Liberibacters are associated with, and may also be endosymbionts in, a number of psyllids (including *Arytainilla spartiophila, Bacticera cockerelli*, *B. trigonica, Diaphorini citri, T. erytreae*, *T. apicalis,* and several *Cacopsylla* species). In addition, some psyllids, such as *B. cockerelli*, *D. citri*, *T. erytreae*, and *C. pyri*, have also been shown to transmit '*Ca.* Liberibacter solanacearum,' '*Ca.* Liberibacter asiaticus and *Ca.* Liberibacter americanus,' '*Ca.* Liberibacter africanus,' and '*Ca.* Liberibacter europaeus,' respectively.

L. crescens is the only Liberibacter species that has been cultured in vitro so far. Accordingly, information regarding the biology, ecology, epidemiology, and intraspecific relationships between plant- and insect-associated Liberibacter species is not fully understood. However, significant insights into the evolution, genetic diversity, and phylogenetic relationships have been gained through comparative analyses of multiple Liberibacter genomes. In addition, genomic analyses of plant disease-associated Liberibacters have led to the identification of potential or putative pathogenicity and virulence factors, as well as biomarkers associated with Liberibacter–host plant interactions. The availability of specific biomarkers associated with Liberibacter–host plant interactions will aid in the development of improved, reliable diagnostic protocols for early (i.e., presymptomatic), rapid Liberibacter detection as part of disease management strategies. Knowledge of specific Liberibacter genes and their products associated with pathogenicity and virulence, as well as interactions with psyllid vectors, should lead to the identification of potential targets for mitigating Liberibacter acquisition and transmission by psyllids, host plant infection, and disease development. Finally, genomics-based research will facilitate improved understanding of the complex and diverse mechanisms of Liberibacter–plant interactions involving different species with the same hosts (e.g., HLB-associated '*Ca.* Liberibacter asiaticus,' '*Ca.* Liberibacter africanus,' and '*Ca.* Liberibacter americanus' in citrus), the same species in different hosts (e.g., zebra chip disease-associated '*Ca.* Liberibacter solanacearum' in solanaceous and other hosts), the same species in asymptomatic and symptomatic associations with plants (e.g., '*Ca.* europaeus' in pear and Scotch broom), and those that do not apparently result in disease (e.g., '*Ca*. Liberibacter europaeus' in pear and *L. crescens* in papaya).

References

Aizawa SI (2001) Bacterial flagella and type III secretion systems. FEMS Microbiol Lett 202:157–164

Albrecht U, Bowman KD (2008) Gene expression in *Citrus sinensis* (*L.*) *Osbeck* following infection with the bacterial pathogen '*Candidatus* Liberibacter asiaticus' causing Huanglongbing in Florida. Plant Sci 175:291–306

Albrecht U, Bowman KD (2012) Transcriptional response of susceptible and tolerant citrus to infection with '*Candidatus* Liberibacter asiaticus'. Plant Sci 185–186:118–130

Alfano JR, Collmer A (2004) Type III secretion system effector proteins: double agents in bacterial disease and plant defense. Annu Rev Phytopathol 42:385–414

Alfaro-Fernández A, Cebrián MC, Villaescusa FJ, de Mendoza AH, Ferrándiz JC, Sanjuán S, Font MI (2012a) First report of '*Candidatus* Liberibacter solanacearum' in carrot in mainland Spain. Plant Dis 96:582

Alfaro-Fernández A, Siverio F, Cebrián MC, Villaescusa FJ, Font MI (2012b) '*Candidatus* Liberibacter solanacearum' associated with *Bactericera trigonica*-affected carrots in the Canary Islands. Plant Dis 96:581

Alvarado VY, Odokonyero D, Duncan O, Mirkov TE, Scholthof HB (2012) Molecular and physiological properties associated with Zebra complex disease in potatoes and its relation with '*Candidatus* Liberibacter' contents in psyllid vectors. PLoS ONE 7:e37345

Ammendola S, Pasquali P, Pistoia C, Petrucci P, Petrarca P, Rotilio G, Battistoni A (2007) High-affinity $Zn2^{+}$ uptake system ZnuABC is required for bacterial zinc homeostasis in intracellular environments and contributes to the virulence of *Salmonella enterica*. Infect Immun 75:5867–5876

Bartsch M, Gobbato E, Bednarek P, Debey S, Schultze JL, Bautor J, Parker JE (2006) Salicylic acid-independent ENHANCED DISEASE SUSCEPTIBILITY 1 signaling in *Arabidopsis* immunity and cell death is regulated by the monooxygenase FMO1 and the nudix hydrolase NUDT7. Plant Cell 18:1038–1051

Bastianel C, Garnier-Semancik M, Renaudin J, Bové JM, Eveillard S (2005) Diversity of '*Candidatus* Liberibacter asiaticus' based on the *omp* gene sequence. Appl Environ Microbiol 71:6473–6478

Beattie GAC, Holford P, Mabberley DJ, Haigh AM, Bayer R, Broadbent P (2011) On the origins of citrus, huanglongbing, diaphorina citri and *Trioza erytreae*. In: The 2nd international research conference on Huanglongbing. Orlando, USA

Block A, Schmelz E, O'Donnell PJ, Jones JB, Klee HJ (2005) Systemic acquired tolerance to virulent bacterial pathogens in tomato. Plant Physiol 138:1481–1490

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

Bové JM (2006) Huanglongbing: a destructive, newly-emerging, century-old disease of citrus. J Plant Pathol 88:7–37

Bové JM, Garnier M (2003) Phloem-and xylem-restricted plant pathogenic bacteria. Plant Sci 164:423–438

Brodersen P, Malinovsky FG, Hématy K, Newman M-A, Mundy J (2005) The role of salicylic acid in the induction of cell death in Arabidopsis *acd11*. Plant Physiol 138:1037–1045

Brown JK, Rehman M, Rogan D, Martin RR, Idris AM (2010) First report of '*Candidatus* Liberibacter psyllaurous' (synonym "*Ca*. L. solanacearum") associated with 'Tomato Vein-Greening' and 'Tomato Psyllid Yellows' diseases in commercial greenhouses in Arizona. Plant Dis 94:376

Büttner D, He SY (2009) Type III protein secretion in plant pathogenic bacteria. Plant Physiol 150: 1656–1664

Camerota C, Raddadi N, Pizzinat A, Gonella E, Crotti E, Tedeschi R, Mozes-Daube N, Ember I, Acs Z, Kolber M, Zchori-Fein E, Daffonchio D, Alma A (2012) Incidence of '*Candidatus* Liberibacter europaeus' and *phytoplasmas* in *Cacopsylla* species (*Hemiptera*: *Psyllidae*) and their host/shelter plants. Phytoparasitica 40:213–221

Capela D, Barloy-Hubler F, Gouzy J, Bothe G, Ampe F, Batut J, Boistard P, Becker A, Boutry M, Cadieu E, Dreano S, Gloux S, Godrie T, Goffeau A, Kahn D, Kiss E, Lelaure V, Masuy D, Pohl T, Portetelle D, Puhler A, Purnelle B, Ramsperger U, Renard C, Thebault P, Vandenbol M, Weidner S, Galibert F (2001) Analysis of the chromosome sequence of the legume symbiont *Sinorhizobium meliloti* strain 1021. Proc Natl Acad Sci USA 98:9877–9882

Capoor SP (1963) Decline of citrus trees in India. Bull Natl Inst Sci India 24:48–64

Cevallos-Cevallos JM, Garcia-Torres R, Etxeberria E, Reyes-De-Corcuera JI (2011) GC-MS analysis of headspace and liquid extracts for metabolomic differentiation of citrus Huanglongbing and zinc deficiency in leaves of 'Valencia' sweet orange from commercial groves. Phytochem Anal 22:236–246

Chen J, Deng X, Sun X, Jones D, Irey M, Civerolo E (2010) Guangdong and Florida populations of '*Candidatus* Liberibacter asiaticus' distinguished by a genomic locus with short tandem repeats. Phytopathol 100:567–572

Chen MH, Miyakawa T, Matsui C (1971) Mycoplasma-alike bodies associated with likubin-diseased Ponkan citrus. Phytopathol 61:598

Cohan FM (2002) What are bacterial species? Ann Rev Microbiol 56:457–487

Coletta-Filho HD, Targon MLPN, Takita MA, De Negri JD, Pompeu J, Machado MA, do Amaral AM, Muller GW (2004) First report of the causal agent of

Huanglongbing (‘*Candidatus* Liberibacter asiaticus’) in Brazil. Plant Dis 88:1382

Crosslin JM, Munyaneza JE, Brown JK, Liefting LW (2010) A history in the making: potato zebra chip disease associated with a new psyllid-borne bacterium—a tale of striped potatoes. Online. APSnet features. doi:10.1094/APSnetFeature-2010-0110. doi:10.1094/PHP-2010-0317-01-RV

da Graça J, Pietersen G, van Vuuren SP (2009) Research on HLB in South Africa. Phytopathol 99:S157

da Graça JV (1991) Citrus greening disease. Ann Rev Phytopathol 29:109–136

Darling AE, Miklos I, Ragan MA (2008) Dynamics of genome rearrangement in bacterial populations. PLoS Genet 4:e1000128

Davis MJ, Mondal SN, Chen H, Rogers ME, Brlansky RH (2008) Co-cultivation of ‘*Candidatus* Liberibacter asiaticus’ with *Actinobacteria* from citrus with Huanglongbing. Plant Dis 92:1547–1550

De Boer SH, Secor G, Li X, Gourley J, Ross P, Rivera V (2007) Preliminary characterization of the etiologic agent causing zebra chip symptoms in potato. In: Proceedings of EAPR Pathol Section Seminar 2–6, July

de Mello Varani A, Souza RC, Nakaya HI, de Lima WC, Paula de Almeida LG, Kitajima EW, Chen J, Civerolo E, Vasconcelos AT, Van Sluys MA (2008) Origins of the *Xylella fastidiosa* prophage-like regions and their impact in genome differentiation. PLoS ONE 3:e4059

Deakin W (2000) Secretion by bacterial flagella. Genome Biol 1:reports003

Doddapaneni H, Liao H, Lin H, Bai X, Zhao X, Civerolo EL, Irey M, Coletta-Filho H, Pietersen G (2008) Comparative phylogenomics and multi-gene cluster analyses of the Citrus Huanglongbing (HLB)-associated bacterium ‘*Candidatus* Liberibacter’. BMC Res Notes 1:72. doi:10.1186/1756-0500-1-72

Duan Q, Zhou M, Zhu L, Zhu G (2013) Flagella and bacterial pathogenicity. J Basic Microbiol 53:1–8. doi:10.1002/jobm.201100335

Duan Y, Zhou L, Hall DG, Li W, Doddapaneni H, Lin H, Liu L, Vahling CM, Gabriel DW, Williams KP, Dickerman A, Sun Y, Gottwald T (2009) Complete genome sequence of citrus Huanglongbing bacterium, ‘*Candidatus* Liberibacter asiaticus’ obtained through metagenomics. Mol Plant-Microbe Interact 22:1011–1020

Eisen JA, Heidelberg JF, White O, Salzberg SL (2000) Evidence for symmetric chromosomal inversions around the replication origin in bacteria. Genome Biol 1: RESEARCH0011. doi:10.1186/gb-2000-1-6-research0011

EPPO (2009) Quarantine alert-‘*Candidatus* Liberibacter solanacearum’. Online: http://www.eppo.int/QUARANTINE/Alert_List/bacteria/Liberibacter_psyllaurous.htm

EPPO (2012) First report of ‘Liberibacter solanacearum’ on carrots and celery in Spain, in association with *Bactericera trigonica*. EPPO reporting service, Paris 2012-06-01

Eveillard S, Teixeira DC, Bové JM (2008) Huanglongbing disease of citrus and the genus ‘*Candidatus* Liberibacter’. In: Porta-Puglia A, Gonthier (eds) Abstracts of the ninth international congress of plant pathology, Torino, Italy, 24–29 Aug 2008. J Plant Pathol 90:S2:43

Fagen JR, Giongo A, Brown CT, Davis-Richardson AG, Gano KA, Triplett EW (2012) Characterization of the relative abundance of the citrus pathogen ‘*Ca.* Liberibacter asiaticus’ in the microbiome of its insect vector, *Diaphorina citri*, using high throughput 16S rRNA sequencing. Open Microbiol J 6:29–33

Fan J, Chen C, Brlansky RH, Gmitter FG Jr, Li ZG (2010) Changes in carbohydrate metabolism in citrus sinensis infected with ‘*Candidatus* Liberibacter asiaticus’. Plant Pathol (Oxford) 59:1037–1043

Fraser CM, Gocayne JD, White O, Adams MD, Clayton RA, Fleischmann RD, Bult CJ, Kerlavage AR, Sutton G, Kelley JM, Fritchman RD, Weidman JF, Small KV, Sandusky M, Fuhrmann J, Nguyen D, Utterback TR, Saudek DM, Phillips CA, Merrick JM, Tomb JF, Dougherty BA, Bott KF, Hu PC, Lucier TS, Peterson SN, Smith HO, Hutchison CA 3rd, Venter JC (1995) The minimal gene complement of *Mycoplasma genitalium*. Science 270:397–403

Friedrich L, Vernooij B, Gaffney T, Morse A, Ryals J (1995) Characterization of tobacco plants expressing a bacterial salicylate hydroxylase gene. Plant Mol Biol 29:959–968

Gabbianelli R, Scotti R, Ammendola S, Petrarca P, Nicolini L, Andrea B (2011) Role of ZnuABC and ZinT in *Escherichia coli* O157:H7 zinc acquisition and interaction with epithelial cells. BMC Microbiol. doi:10.1186/1471-2180-11-36

Gardiner DM, McDonald MC, Covarelli L, Solomon PS, Rusu AG, Marshall M, Kazan K, Chakraborty S, McDonald BA, Manners JM (2012) Comparative pathogenomics reveals horizontally acquired novel virulence genes in fungi infecting cereal hosts. PLoS Pathog 8:e1002952. doi:10.1371/journal.ppat.1002952

Garnier M, Bové JM (1977) Structure trilamellaire des deux membranes qui entourent les organismes procaryotes associés à la maladie du “greening” des agrumes. Fruits 32:749–752

Garnier M, Bové JM, Cronje CPR, Sanders GM, Korsten L, Le Roux HF (2000a) Presence of ‘*Candidatus* Liberibacter africanus’ in the Western Cape province of South Africa. In: Proceedings of 14th conference IOCV. IOCV, Riverside, pp 369–372

Garnier M, Danel N, Bové JM (1984) Aetiology of citrus greening disease. Ann Microbiol (Paris) 135A:169–179

Garnier M, Gao SJ, He Y, Villechanoux S, Gandar J, Bové JM (1991) Study of the greening organism (GO) with monoclonal antibodies: serological identification, morphology serotypes and purification of the GO. In: Proceedings of 11th conference IOCV. IOCV, Riverside, pp 428–435

Garnier M, Jagoueix-Eveillard S, Cronje PR, Le Roux HF, Bové JM (2000b) Genomic characterization of a Liberibacter present in an ornamental rutaceous tree, *Calodendrum capense*, in the Western Cape Province of South Africa. Proposal of '*Candidatus* Liberibacter africanus subsp. capensis'. Int J Syst Evol Microbiol 50 Pt 6:2119–2125

Garnier M, Jagoueix S, Toorawa P, Grisoni M, Mallessard R, Dookun A, Saumtally S, Autrey JC, Bové JM (1996) Both huanglongbing (greening) Liberibacter species are present in Mauritius and Reunion. In: Proceedings of 13th conference IOCV. IOCV, Riverside, pp 392–394

Gil R, Silva FJ, Pereto J, Moya A (2004) Determination of the core of a minimal bacterial gene set. Microbiol Mol Biol Rev 68:518–537

Glass JI, Assad-Garcia N, Alperovich N, Yooseph S, Lewis MR, Maruf M, Hutchison CA 3rd, Smith HO, Venter JC (2006) Essential genes of a minimal bacterium. Proc Natl Acad Sci USA 103:425–430

Glynn JM, Islam MS, Bai Y, Lan S, Wen A, Gudmestad NC, Civerolo EL, Lin H (2012) Mutilocus sequence typing of '*Candidatus* Liberibacter solanacearum' isolates from North America and New Zealand. J Plant Pathol 94:223–228

Gophna U, Ron EZ, Graur D (2003) Bacterial type III secretion systems are ancient and evolved by multiple horizontal-transfer events. Gene (Amsterdam) 312:151–163

Gottwald TR (2010) Current epidemiological understanding of citrus Huanglongbing. Annu Rev Phytopathol 48:119–139

Gottwald TR, da Graça JV, Bassanezi RB (2007) Citrus huanglongbing: the pathogen and its impact. Plant Health Prog. doi:10.1094/PHP-2007-0906-1001-RV

Gudmestad NC, Secor GA (2007) Zebra chip: a new disease of potato. Nebr Potato Eyes 19:1–4

Ham JH, Bent A (2002) Recognition and defence signalling in plant/bacterial and fungal interactions. In: Scheel E, Westernack C (eds) Plant signal transduction. Oxford University Press, New York, pp 198–225

Hansen AK, Trumble JT, Stouthamer R, Paine TD (2008) A new huanglongbing species, '*Candidatus* Liberibacter psyllaurous', found to infect tomato and potato, is vectored by the psyllid *Bactericera cockerelli* (Sulc). Appl Environ Microbiol 74:5862–5865

Hartung JS, Shao J, Kuykendall LD (2011) Comparison of the '*Ca*. Liberibacter asiaticus' genome adapted for an intracellular lifestyle with other members of the *Rhizobiales*. PLoS ONE 6:e23289. doi:10.1371/journal.pone.0023289

Hatch TP, Al-Hossainy E, Silverman JA (1982) Adenine nucleotide and lysine transport in *Chlamydia psittaci*. J Bacteriol 150:662–670

Hershberg R, Petrov DA (2010) Evidence that mutation is universally biased towards AT in bacteria. PLoS Genet 6:e1001115. doi:10.1371/journal.pgen.1001115

Hongoh Y, Toyoda A (2011) Whole-genome sequencing of unculturable bacterium using whole-genome amplification. Methods Mol Biol 733:25–33

Huber DM, Haneklaus S (2007) Managing nutrition to control plant disease. Landbauforschung Völkenrode 57:313–322

Hutchison CA, Peterson SN, Gill SR, Cline RT, White O, Fraser CM, Smith HO, Venter JC (1999) Global transposon mutagenesis and a minimal *Mycoplasma* genome. Science 286(5447):2165

Ibarz Pavón AB, Maiden MCJ (2009) Multilocus sequence typing. In: Caugant DA (ed) Epidemiology of microorganisms: methods and protocols, vol 551. Springer, Heidelberg, pp 129–140

Islam MS, Glynn JM, Bai Y, Duan YP, Coletta-Filho HD, Kuruba G, Civerolo EL, Lin H (2012) Multilocus microsatellite analysis of '*Candidatus* Liberibacter asiaticus' associated with citrus Huanglongbing worldwide. BMC Microbiol 12:39. doi:10.1186/1471-2180-12-39

Jagoueix S, Bove J-M, Garnier M (1994) The phloem-limited bacterium of greening disease of citrus is a member of the alpha subdivision of the Proteobacteria. Int J Syst Bacteriol 44:379–386

Jagoueix S, Bove JM, Garnier M (1997) Comparison of the 16S/23S ribosomal intergenic regions of '*Candidatus* Liberobacter asiaticum' and '*Candidatus* Liberobacter africanus', the two species associated with citrus huanglongbing (Greening) disease. Int J Syst Bacteriol 47:224–227

Katoh H, Subandiyah S, Tomimura K, Okuda M, Su HJ, Iwanami T (2011) Differentiation of '*Candidatus* Liberibacter asiaticus' isolates by variable-number tandem-repeat analysis. Appl Environ Microbiol 77:1910–1917

Kim J-S, Sagaram US, Burns JK, Li J-L, Wang N (2009) Response of sweet orange (*Citrus sinensis*) to '*Candidatus* Liberibacter asiaticus' infection: microscopy and microarray analyses. Phytopathol 99:50–57

Koonin EV (2000) How many genes can make a cell: the minimal-gene-set concept. Annu Rev Genomics Hum Genet 1:99–116. doi:10.1146/annurev.genom.1.1.99

Lapage SP, Sneath PHA, Lessel EF, Skerman VBD, Seeliger HPR, Clark WA (1992) International code of nomenclature of bacteria (1990 Revision). ASM Press, Am Society for Microbiol Press, Washington, D.C

Leonard MT, Fagen JR, Davis-Richardson AG, Davis MJ, Triplett EW (2012) Complete genome sequence of Liberibacter crescens BT-1. Stand Genomic Sci 7:271–283

Li W, Cong Q, Pei J, Kinch LN, Grishin NV (2012) The ABC transporters in '*Candidatus* Liberibacter asiaticus'. Proteins: Struct, Funct, Bioinf 80:2614–2628

Li W, Duan Y, Brlansky R, Twieg E, Levy L (2008) Incidence and population of '*Candidatus* Liberibacter asiaticus' in Asian citrus psyllids (*Diaphorina citri*) on citrus plants affected by huanglongbing in Florida.

In: Proceedings of the 2008 international research conference on Huanglongbing, Orlando, FL, USA

Liefting LW, Perez-Egusquiza ZC, Clover GRG, Anderson JAD (2008a) A new '*Candidatus* Liberibacter' species in *Solanum tuberosum* in New Zealand. Plant Dis 92:1474

Liefting LW, Sutherland PW, Ward LI, Paice KL, Weir BS, Clover GRG (2009a) A new '*Candidatus* Liberibacter' species associated with diseases of *Solanaceous* crops. Plant Dis 93:208–214

Liefting LW, Ward LI, Shiller JB, Clover GRG (2008b) A new '*Candidatus* Liberibacter' species in *Solanum betaceum* (*Tamarillo*) and *Physalis peruviana* (*Cape Gooseberry*) in New Zealand. Plant Dis 92:1588

Liefting LW, Weir BS, Pennycook SR, Clover GRG (2009b) '*Candidatus* Liberibacter solanacearum', associated with plants in the family *Solanaceae*. Int J Syst Evol Microbiol 59:2274–2276

Lin H, Doddapaneni H, Munyaneza JE, Civerolo EL, Sengoda VG, Buchman JL, Stenger DC (2009) Molecular characterization and phylogenetic analysis of 16S rRNA from a new '*Candidatus* Liberibacter' strain associated with zebra chip disease of potato (*Solanum tuberosum* L.) and the potato psyllid (*Bactericera cockerelli* sulc). J Plant Pathology 91:215–219

Lin H, Gudmestad NC (2013) Aspects of pathogen genomics, diversity, epidemiology, vector dynamics, and disease management for a newly emerged disease of potato: zebra chip. Phytopathol 103:524–537

Lin H, Islam MS, Bai Y, Wen A, Lan S, Gudmestad NC, Civerolo EL (2012) Genetic diversity of '*Candidatus* Liberibacter solanacearum' strains in the United States and Mexico revealed by simple sequence repeat markers. Eur J Plant Pathol 132:297–308

Lin H, Lou B, Glynn JM, Doddapaneni H, Civerolo EL, Chen C, Duan Y, Zhou L, Vahling CM (2011) The complete genome sequence of '*Candidatus* Liberibacter solanacearum', the bacterium associated with potato zebra chip disease. PLoS ONE 6:e19135. doi:10.1371/journal.pone.0019135

Lin KH (1956) Yellow shoot of citrus. Symptomatology. Investigations in the cause of huanglongbing. Natural transmission and spread. General conclusions. Acta Phytopathologica Sin 2:1–42

Lind PA, Andersson DI (2008) Whole-genome mutational biases in bacteria. Proc Natl Acad Sci USA 105:17878–17883

Ling KS, Lin H, Lewis Ivey ML, Zhang W, Miller SA (2011) First report of '*Candidatus* Liberibacter solanacearum' naturally infecting tomatoes in the state of Mexico, Mexico. Plant Dis 95:1026

Lopes SA, Frare GF, Bertolini E, Cambra M, Fernandes NG, Ayres AJ, Marin DR, Bové JM (2009) Liberibacters associated with citrus huanglongbing in Brazil: '*Candidatus* Liberibacter asiaticus' is heat tolerant, '*Ca*. L. americanus' is heat sensitive. Plant Dis 93:257–262

Magomere TO, Obukosia SD, Mutitu E, Ngichabe C, Olubayo F, Shibairo S (2009) Molecular characterization of '*Candidatus* Liberibacter' species/strains causing huanglongbing disease of citrus in Kenya. Electron J Biotechnol 12(2):5–6. doi:10.2225/vol12-issue2-fulltext-2

McClean APD, Oberholzer PCJ (1965a) Citrus psylla, a vector of the greening disease of sweet orange. S Afr J Agr Sci 8:297–298

McClean APD, Oberholzer PCJ (1965b) Greening disease of sweet orange: evidence that it is caused by a transmissible virus. S Afr J Agr Sci 8:253–276

McCutcheon JP, Moran NA (2010) Functional convergence in reduced genomes of bacterial symbionts spanning 200 My of evolution. Genome Biol Evol 2:708–718

Miller C, Scheuring D, Koym J, Henne D, Jifon J, Levy J, Tamborindeguy C, Pierson B (2011) Further studies on host-plant tolerance to ZC. In: Proceedings of the 11th annual SCRI Zebra chip reporting section, pp 134–136

Moran NA, McCutcheon JP, Nakabachi A (2008) Genomics and evolution of heritable bacterial symbionts. Ann Rev Genet 42:165–190

Moulder JW (1985) Comparative biology of intracellular parasitism. Microbiol Rev 49:298–337

Moya A, Pereto J, Gil R, Latorre A (2008) Learning how to live together: genomic insights into prokaryote-animal symbioses. Nat Rev Genet 9:218–229

Munyaneza JE, Crosslin JM, Upton JE (2007) Association of *Bactericera cockerelli* (Homoptera: Psyllidae) with "zebra chip," a new potato disease in southwestern United States and Mexico. J Econ Entomol 100:656–663

Munyaneza JE, Fisher TW, Sengoda VG, Garczynski SF, Nissinen A, Lemmetty A (2010a) Association of '*Candidatus* Liberibacter solanacearum' with the Psyllid, *Trioza apicalis* (Hemiptera: Triozidae) in Europe. J Econ Entomol 103:1060–1070

Munyaneza JE, Fisher TW, Sengoda VG, Garczynski SF, Nissinen A, Lemmetty A (2010b) First report of '*Candidatus* Liberibacter solanacearum' associated with psyllid-affected carrots in Europe. Plant Dis 94:639

Munyaneza JE, Lemmetty A, Nissinen AI, Sengoda VG, Fisher TW (2011) Molecular detection of Aster yellows phytoplasma and '*Candidatus* Liberibacter solanacearum' in carrots affected by the psyllid, *Trioza apicalis* (Hemiptera: Triozidae) in Finland. J Plant Pathology 93:697–700

Nakabachi A, Yamashita A, Toh H, Ishikawa H, Dunbar HE, Moran NA, Hattori M (2006) The 160-kilobase genome of the bacterial endosymbiont *Carsonella*. Science 314:267

Navarre DA, Shakya R, Holden J, Crosslin JM (2009) LC-MS analysis of phenolic compounds in tubers showing zebra chip symptoms. Am J Potato Res 86:88–95

Nguyen L, Paulsen IT, Tchieu J, Hueck CJ, Saier MH Jr (2000) Phylogenetic analyses of the constituents of Type III protein secretion systems. J Mol Microbiol Biotechnol 2:125–144

Nwugo CC, Duan Y, Lin H (2013a) Study on citrus response to huanglongbing highlights a down-regulation of defense-related proteins in lemon plants upon '*Ca.* Liberibacter asiaticus' infection. PLoS One 8:e67442. doi:10.1371/journal.pone.0067442

Nwugo CC, Lin H (2012) Proteomic and physiological characterization of potato plant in response to Zebra Chip disease. In: Proceedings of the 12th annual SCRI Zebra Chip reporting section, San Antonio, Texas, 30 Oct–2 Nov 2012, pp 94–98

Nwugo CC, Lin H, Duan Y, Civerolo EL (2013b) The effect of '*Candidatus* Liberibacter asiaticus' infection on the proteomic profiles and nutritional status of pre-symptomatic and symptomatic grapefruit (*Citrus paradisi*) plants. BMC Plant Biol 13:59. doi:10.1186/1471-2229-13-59

Okubo T, Fukushima S, Itakura M, Oshima K, Longtonglang A, Teaumroong N, Mitsui H, Hattori M, Hattori R, Hattori T, Minamisawa K (2013) Genome analysis suggests that the soil oligotrophic bacterium *Agromonas oligotrophica* (*Bradyrhizobium oligotrophicum*) is a nitrogen-fixing symbiont of *Aeschynomene indica*. Appl Environ Microbiol 79:2542–2551

Ottemann KM, Miller JF (1997) Roles for motility in bacterial-host interactions. Mol Microbiol 24:1109–1117

Paez JG, Lin M, Beroukhim R, Lee JC, Zhao X, Richter DJ, Gabriel S, Herman P, Sasaki H, Altshuler D, Li C, Meyerson M, Sellers WR (2004) Genome coverage and sequence fidelity of ϕ29 polymerase-based multiple strand displacement whole genome amplification. Nucleic Acids Res 32:e71. doi:10.1093/nar/gnh069

Raddadi N, Gonella E, Camerota C, Pizzinat A, Tedeschi R, Crotti E, Mandrioli M, Attilio Bianco P, Daffonchio D, Alma A (2011) '*Candidatus* Liberibacter europaeus' sp. nov. that is associated with and transmitted by the psyllid *Cacopsylla pyri* apparently behaves as an endophyte rather than a pathogen. Environ Microbiol 13:414–426

Reinking OA (1919) Diseases of economic plants in southern China. Philippine Agricul 8:109–135

Rocha EP, Danchin A (2002) Base composition bias might result from competition for metabolic resources. Trends Genet 18:291–294

Rodrigue S, Malmstrom RR, Berlin AM, Birren BW, Henn MR, Chisholm SW (2009) Whole genome amplification and *de novo* assembly of single bacterial cells. PLoS ONE 4:e6864. doi:10.1371/journal.pone.0006864

Rondon S, Schreiber A, Jensen A, Hamm P, Munyaneza J, Nolte P, Olsen N, Wenninger E, Henne D, Wohleb C, Waters T (2012) Potato psyllid vector of zebra chip disease in the Pacific Northwest. Pacific Northwest Extension Publication. Oregon State University, June 2012. 2018

Schmitz-Esser S, Haferkamp I, Knab S, Penz T, Ast M, Kohl C, Wagner M, Horn M (2008) *Lawsonia intracellularis* contains a gene encoding a functional rickettsia-like ATP/ADP translocase for host exploitation. J Bacteriol 190:5746–5752

Schneiker S, Perlova O, Kaiser O, Gerth K, Alici A, Altmeyer MO, Bartels D, Bekel T, Beyer S, Bode E, Bode HB, Bolten CJ, Choudhuri JV, Doss S, Elnakady YA, Frank B, Gaigalat L, Goesmann A, Groeger C, Gross F, Jelsbak L, Jelsbak L, Kalinowski J, Kegler C, Knauber T, Konietzny S, Kopp M, Krause L, Krug D, Linke B, Mahmud T, Martinez- Arias R, McHardy AC, Merai M, Meyer F, Mormann S, Munoz-Dorado J, Perez J, Pradella S, Rachid S, Raddatz G, Rosenau F, Rueckert C, Sasse F, Scharfe M, Schuster SC, Suen G, Treuner-Lange A, Velicer GJ, Vorhoelter F-J, Weissman KJ, Dwelch R, Wenzel SC, Whitworth DE, Wilhelm S, Wittmann C, Bloecker H, Puehler A & Mueller R (2007) Complete genome sequence of the myxobacterium *Sorangium cellulosum*. Nat Biotechnol 25:1281–1289

Secor GA, Rivera-Varas VV (2004) Emerging diseases of cultivated potato and their impact on Latin America. Revista Latinoamericana de la Papa (Suppl.) 1:1–8

Secor GA, Rivera VV, Abad JA, Lee IM, Clover GRG, Liefting LW, Li X, De Boer SH (2009) Association of '*Candidatus* Liberibacter solanacearum' with zebra chip disease of potato established by graft and psyllid transmission, electron microscopy, and PCR. Plant Dis 93:574–583

Shokrollah H, Abdullah TL, Sijam K, Abdullah SNA (2010) Ultrastructures of '*Candidatus* Liberibacter asiaticus' and its damage in huanglongbing (HLB) infected citrus. African J Biotechnol 9:5897–5901

Shokrollah H, Abdullah TL, Sijam K, Abdullah SNA (2011) Identification of physical and biochemical characteristic of mandarin (*Citrus reticulata*) fruit infected by huanglongbing (HLB). Aust J of Crop Sci 5:181–186

Simpson AJ, Reinach FC, Arruda P, Abreu FA, Acencio M, Alvarenga R, Alves LM, Araya JE, Baia GS, Baptista CS, Barros MH, Bonaccorsi ED, Bordin S, Bové JM, Briones MR, Bueno MR, Camargo AA, Camargo LE, Carraro DM, Carrer H, Colauto NB, Colombo C, Costa FF, Costa MC, Costa-Neto CM, Coutinho LL, Cristofani M, Dias-Neto E, Docena C, El-Dorry H, Facincani AP, Ferreira AJ, Ferreira VC, Ferro JA, Fraga JS, Franca SC, Franco MC, Frohme M, Furlan LR, Garnier M, Goldman GH, Goldman MH, Gomes SL, Gruber A, Ho PL, Hoheisel JD, Junqueira ML, Kemper EL, Kitajima JP, Krieger JE, Kuramae EE, Laigret F, Lambais MR, Leite LC, Lemos EG, Lemos MV, Lopes SA, Lopes CR, Machado JA, Machado MA, Madeira AM, Madeira HM, Marino CL, Marques MV, Martins EA, Martins EM, Matsukuma AY, Menck CF, Miracca EC, Miyaki CY, Monteriro-Vitorello CB, Moon DH, Nagai MA, Nascimento AL, Netto LE, Nhani A, Jr Nobrega FG, Nunes LR, Oliveira MA, de Oliveira MC, de Oliveira RC, Palmieri DA, Paris A, Peixoto BR, Pereira GA, Pereira HA, Jr Pesquero JB, Quaggio RB, Roberto PG, Rodrigues V, de MRAJ, de Rosa

VE, Jr de Sa RG, Santelli RV, Sawasaki HE, da Silva AC, da Silva AM, da Silva FR, da Silva WA, Jr da Silveira JF, Silvestri ML, Siqueira WJ, de Souza AA, de Souza AP, Terenzi MF, Truffi D, Tsai SM, Tsuhako MH, Vallada H, Van Sluys MA, Verjovski-Almeida S, Vettore AL, Zago MA, Zatz M, Meidanis J, Setubal JC (2000) The genome sequence of the plant pathogen *Xylella fastidiosa*. The *Xylella fastidiosa* consortium of the organization for nucleotide sequencing and analysis. Nature 406:151–159

Spits C, Le Caignec C, De Rycke M, Van Haute L, Van Steirteghem A, Liebaers I, Sermon K (2006) Whole-genome multiple displacement amplification from single cells. Nat Protoc 1:1965–1970

Srividhya KV, Alaguraj V, Poornima G, Kumar D, Singh GP, Raghavenderan L, Katta AV, Mehta P, Krishnaswamy S (2007) Identification of prophages in bacterial genomes by dinucleotide relative abundance difference. PLoS ONE 2:e1193. doi:10.1371/journal.pone.0001193

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

Stover E, McCollum T, Driggers R, Duan P, Shatters R (2013) Huanglongbing resistance and tolerance in citrus. International Research Conference on Huanglongbing, 4–8 Feb 2013, Orlando, Florida

Suyama M, Bork P (2001) Evolution of prokaryotic gene order: genome rearrangements in closely related species. Trends Genet 17:10–13

Teixeira DC, Saillard C, Eveillard S, Danet JL, Ayres AJ, Bové JM (2005a) '*Candidatus* Liberibacter americanus', associated with citrus huanglongbing (greening disease) in São Paulo State, Brazil. Int J Syst Evol Microbiol 55:1857–1862

Teresani GR, Bertolini E, Alfaro-Fernandez A, Martínez C, Tanaka FAO, Kitajima E, Rosello M, Sanjuan S, Ferrandiz JC, López MM, Cambra M, Font-San-Ambrosio MI (2014) Association of '*Candidatus* Liberibacter solanacearum' with a vegetative disorder of celery in Spain and development of a real-time PCR method for its detection Phytopathol. doi:org/10.1094/PHYTO-07-13-0182-R

Texeira DC, Ayres J, Kitajima EW, Tanaka FAO, Danet JL, Jagoueix-Eveillard S, Saillard C, Bové JM (2005b) First report of a huanglongbing-like disease of citrus in São Paulo State, Brazil, and association of a new Liberibacter Species, '*Candidatus* Liberibacter americanus', with the disease. Plant Dis 89:107

Thao ML, Moran NA, Abbot P, Brennan EB, Burckhardt DH, Baumann P (2000) Cospeciation of psyllids and their primary prokaryotic endosymbionts. Appl Environ Microbiol 66:2898–2905

Thompson S, Fletcher JD, Ziebell H, Beard S, Panda P, Jorgensen N, Fowler SV, Liefting LW, Berry N, Pitman AR (2013) First report of '*Candidatus* Liberibacter europaeus' associated with psyllid infested Scotch broom. New Dis Reports 27. doi:10.5197/j.2044-0588.2013.5027

Tirtawidjaja S, Hadewidjaja T, Lasheen AM (1965) Citrus vein phloem degeneration virus, a possible cause of citrus chlorosis in Java. Proc Am Soc Hortic Sci 86:235–243

Vahling-Armstrong CM, Zhou H, Benyon L, Morgan JK, Duan Y (2012) Two plant bacteria, *S. meliloti* and '*Ca.* Liberibacter asiaticus', share functional znuABC homologues that encode for a high affinity zinc uptake system. PLoS ONE 7:e37340. doi:10.1371/journal.pone.0037340

Vahling CM, Duan Y, Lin H (2010) Characterization of an ATP translocase identified in the destructive plant pathogen '*Candidatus* Liberibacter asiaticus'. J Bacteriol 192:834–840

van Belkum A (1999) Short sequence repeats in microbial pathogenesis and evolution. Cell Mol Life Sci 56:729–734

Van der Merwe AJ, Andersen FG (1937) Chromium and manganese toxicity. Is it important in Transvaal citrus greening? Farming S Afr 12:439–440

Van Sluys MA, de Oliveira MC, Monteiro-Vitorello CB, Miyaki CY, Furlan LR, Camargo LE, da Silva AC, Moon DH, Takita MA, Lemos EG, Machado MA, Ferro MI, da Silva FR, Goldman MH, Goldman GH, Lemos MV, El-Dorry H, Tsai SM, Carrer H, Carraro DM, de Oliveira RC, Nunes LR, Siqueira WJ, Coutinho LL, Kimura ET, Ferro ES, Harakava R, Kuramae EE, Marino CL, Giglioti E, Abreu IL, Alves LM, do Amaral AM, Baia GS, Blanco SR, Brito MS, Cannavan FS, Celestino AV, da Cunha AF, Fenille RC, Ferro JA, Formighieri EF, Kishi LT, Leoni SG, Oliveira AR, Rosa VE, Jr. Sassaki FT, Sena JA, de Souza AA, Truffi D, Tsukumo F, Yanai GM, Zaros LG, Civerolo EL, Simpson AJ, Almeida NF, Jr Setubal JC, Kitajima JP (2003) Comparative analyses of the complete genome sequences of Pierce's disease and citrus variegated chlorosis strains of *Xylella fastidiosa*. J Bacteriol 185:1018–1026

van Wees SCM, Glazebrook J (2003) Loss of non-host resistance of Arabidopsis NahG to *Pseudomonas syringae* pv. *phaseolicola* is due to degradation products of salicylic acid. Plant J 33:733–742

Vartoukian SR, Palmer RM, Wade WG (2010) Strategies for culture of 'unculturable' bacteria. FEMS Microbiol Lett 309:1–7

Villechanoux S, Garnier M, Laigret F, Renaudin J, Bové J-M (1993) The genome of the non-cultured, bacterial-like organism associated with citrus greening disease contains the *nusG-rplKAJL-rpoBC* gene cluster and the gene for a bacteriophage type DNA polymerase. Curr Microbiol 26:161–166

Wallis CM, Chen J, Civerolo EL (2012) Zebra chip-diseased potato tubers are characterized by increased levels of host phenolics, amino acids, and defense-related proteins. Physiol Mol Plant Pathol 78:66–72. doi:10.1016/j.pmpp.2012.02.001

Weisburg WG, Barns SM, Pelletier DA, Lane DJ (1991) 16S ribosomal DNA amplification for phylogenetic study. J Bacteriol 173:697–703

Wen A, Mallik I, Alvarado VY, Pasche JS, Wang X, Li W, Levy L, Lin H, Scholthof HB, Mirkov TE, Rush CM, Gudmestad NC (2009) Detection, distribution, and genetic variability of '*Candidatus* Liberibacter' species associated with zebra complex disease of potato in North America. Plant Dis 93:1102–1115. doi:10.1094/pdis-93-11-1102

Wood DW, Setubal JC, Kaul R, Monks DE, Kitajima JP, Okura VK, Zhou Y, Chen L, Wood GE, Almeida NF Jr, Woo L, Chen Y, Paulsen IT, Eisen JA, Karp PD, Bovee D Sr, Chapman P, Clendenning J, Deatherage G, Gillet W, Grant C, Kutyavin T, Levy R, Li MJ, McClelland E, Palmieri A, Raymond C, Rouse G, Saenphimmachak C, Wu Z, Romero P, Gordon D, Zhang S, Yoo H, Tao Y, Biddle P, Jung M, Krespan W, Perry M, Gordon-Kamm B, Liao L, Kim S, Hendrick C, Zhao ZY, Dolan M, Chumley F, Tingey SV, Tomb JF, Gordon MP, Olson MV, Nester EW (2001) The genome of the natural genetic engineer *Agrobacterium tumefaciens* C58. Science 294: 2317–2323

Yalpani N, Silverman P, Wilson TM, Kleier DA, Raskin I (1991) Salicylic acid is a systemic signal and an inducer of pathogenesis-related proteins in virus-infected tobacco. Plant Cell 3:809–818

Zhang K, Martiny AC, Reppas NB, Barry KW, Malek J, Chisholm SW, Church GM (2006) Sequencing genomes from single cells by polymerase cloning. Nat Biotechnol 24:680–686

Zhang S, Flores-Cruz Z, Zhou L, Kang BH, Fleites LA, Gooch MD, Wulff NA, Davis MJ, Duan YP, Gabriel DW (2011) '*Ca*. Liberibacter asiaticus' carries an excision plasmid prophage and a chromosomally integrated prophage that becomes lytic in plant infections. Mol Plant Microbe Interact 24:458–468

Zhao H, Sun R, Albrecht U, Padmanabhan C, Wang A, Coffey MD, Girke T, Wang Z, Close TJ, Roose M, Yokomi RK, Folimonova S, Vidalakis G, Rouse R, Bowman KD, Jin H (2013) Small RNA profiling reveals phosphorus deficiency as a contributing factor in symptom expression for citrus huanglongbing disease. Mol Plant 6:301–310

Zhao XY (1981) Citrus yellow shoot disease (huanglongbing) in China: a review. Proc Int Soc Citriculture 1:466–469

Zhou L, Powell CA, Hoffman MT, Li W, Fan G, Liu B, Lin H, Duan Y (2011) Diversity and plasticity of the intracellular plant pathogen and insect symbiont '*Candidatus* Liberibacter asiaticus' as revealed by hypervariable prophage genes with intragenic tandem repeats. Appl Environ Microbiol 77:6663–6673

Zou H, Gowda S, Zhou L, Hajeri S, Chen G, Duan Y (2012) The destructive citrus pathogen, '*Candidatus* Liberibacter asiaticus' encodes a functional flagellin characteristic of a pathogen-associated molecular pattern. PLoS One 7:e46447. doi:10.1371/journal.pone.0046447

10 Phytoplasma Genomes: Evolution Through Mutually Complementary Mechanisms, Gene Loss and Horizontal Acquisition

Yan Zhao, Robert E. Davis, Wei Wei, Jonathan Shao, and Rasa Jomantiene

10.1 Introduction

Plants are affected by a large number of diseases that are collectively referred to as 'yellows diseases.' For decades, the causal agents of plant yellows diseases were mistakenly presumed to be viruses. In 1967, plant pathologists were startled by the news that yellows-diseased plants were infected by microbes resembling mycoplasmas, cell wall-less bacteria already known in humans and animals (Doi et al. 1967; Ishiie et al. 1967). Following this discovery, scientists around the world reexamined many plant diseases of unsolved cause, giving rise to the finding of mycoplasma-like organisms (MLOs, now known as phytoplasmas) in numerous and diverse diseased plants. It is rather remarkable that, within a few years following the discovery of phytoplasmas, several other new types of plant pathogens, including spiroplasmas and viroids, were also discovered (Davis et al. 1972; Davis and Worley 1973; Diener 1971).

Y. Zhao · R. E. Davis (✉) · W. Wei · J. Shao
Molecular Plant Pathology Laboratory, Agricultural Research Service, United States Department of Agriculture, Beltsville, MD 20705, USA
e-mail: robert.davis@ars.usda.gov

R. Jomantiene
Phytovirus Laboratory, Nature Research Centre, 08412, Vilnius, Lithuania

Phytoplasmas are cell wall-less bacteria that inhabit sieve cells in the phloem tissue of infected plants and are transmitted from plant-to-plant by phloem-feeding insect vectors, principally leafhoppers. Together with mycoplasmas, spiroplasmas, and acholeplasmas, phytoplasmas are classified in class *Mollicutes*. Having descended from a Gram-positive, low G+C walled bacterium (Gundersen et al. 1994; Zhao et al. 2005), and more recently from an *Acholeplasma*-like ancestor (Wei et al. 2008a), the phytoplasma clade underwent substantial evolutionary genome shrinkage (Marcone et al. 1999; Oshima et al. 2004). Yet, phytoplasma genomes contain numerous repeated sequences that appear in genomic islands. These islands, first termed sequence-variable mosaics (SVMs) (Jomantiene and Davis 2006), were formed by recurrent and targeted attacks by mobile genetic elements found to be ancient phages (Wei et al. 2008a). Apparently, the evolutionary shrinkage of phytoplasmal genomes was partially countered by horizontal acquisition of repeated mobile genetic elements. While the loss of genes encoding diverse biosynthetic pathways must have led to increased host dependence, new capabilities were acquired that enabled and enhanced phytoplasma–host interactions. Thus, two mutually complementary, genome-sculpting mechanisms acting in concert played key roles in phytoplasma evolution.

To date, the genomes of four phytoplasmas have been completely sequenced. The four phytoplasmas are responsible for diseases that

D. C. Gross et al. (eds.), *Genomics of Plant-Associated Bacteria*,
DOI: 10.1007/978-3-642-55378-3_10,

seriously impact affected host plant species. '*Candidatus* Phytoplasma australiense' is the cause of Australian grapevine yellows (Davis et al. 1997); this phytoplasma is also the apparent cause of *Phormium* yellow leaf, strawberry lethal yellows, *Cordyline* sudden decline, pumpkin yellow leaf curl, and diseases of potato and other plants (Liefting et al. 2011; Tran-Nguyen et al. 2008). '*Ca*. Phytoplasma asteris', including strains OY-M and AY-WB, has a very wide plant host range (Bai et al. 2006; Lee et al. 2004; Oshima et al. 2004). '*Ca*. Phytoplasma mali' is the cause of apple proliferation disease, one of the most economically important diseases of fruit trees in Europe (Kube et al. 2008; Carraro et al. 2004).

Three of the four completely sequenced phytoplasmas have circular chromosomes, and the fourth has a linear chromosome. Significant differences in gene content, gene arrangement, and genome size are observed among the four genomes. Despite such differences, the presence of SVMs (Jomantiene and Davis 2006) is common to all four genomes. Recent studies have indicated that formation of these phage-derived SVMs was among key factors that launched evolutionary emergence of the phytoplasma clade (Wei et al. 2008a). Based on this theory, while mobile genetic elements are widely acknowledged as major drivers of bacterial evolution, perhaps nowhere is this role more apparent than in the genomes of phytoplasmas.

10.2 Evolving Concepts in Classification and Taxonomy Driven by Genome-Based Information

10.2.1 Gene-Based Classification

Due to difficulties in establishing phytoplasma cultures in cell-free medium and the consequent inaccessibly of measurable phenotypic characters, diverse phytoplasmas are presently classified into groups based on the genetic information coded in their 16S rRNA genes. Two major classification schemes have been developed. One is based on phylogenetic analysis of 16S rRNA gene sequences (Seemüller et al. 1998), and the other is based on RFLP analysis of PCR-amplified 16S rRNA gene fragments by a defined set of 17 restriction enzymes (Lee et al. 1993, 1998, 2000). Since the RFLP analysis exploits an adequate subset of sequence characters (restriction enzyme recognition sites) present in the 16S rRNA genes, phytoplasma groups delineated under this scheme are generally consistent with the phylogeny of the respective 16S rRNA gene sequences. Advantageously, the RFLP analysis-based scheme also distinguishes subtle pattern differences, making it possible to differentiate distinct subgroups within a phytoplasma group. Following inception of the scheme, 19 groups and more than 40 subgroups were delineated through conventional RFLP analysis (Lee et al. 2006; Al-Saady et al. 2008). Over the past six years, a significant expansion of the scheme has been achieved through the use of computer-simulated or virtual RFLP analysis (Wei et al. 2007b, 2008b); as a result, 32 groups (Zhao et al. 2009a; Lee et al. 2011; Nejat et al. 2013) and over 120 subgroups have been delineated as of this writing. This virtual RFLP analysis approach also facilitated discovery of new phytoplasmas and unveiled extensive genetic diversity among phytoplasma strains present in the same host species and/or same geographic location (Wei et al. 2007a, c; Cai et al. 2008; Quaglino et al. 2009; Jomantiene et al. 2011; Bagadia et al. 2013). To streamline virtual RFLP analysis, a suite of computer programs have been developed and integrated into a web-based interactive research tool named *i*PhyClassifier (Zhao et al. 2009b; URL).

10.2.2 Gene- and Genome-Facilitated Taxonomy

Bacterial species are presently circumscribed through the use of a consensus polyphasic approach that involves morphological, physiological, chemotaxonomic, serologic, and genotypic characterizations (Stackebrandt 2007).

Unfortunately, many of the phenotypic characters contributing to bacterial species delineation remain inaccessible for phytoplasmas. To resolve this encumbrance on phytoplasma taxonomy, the International Committee of Systematic Bacteriology Subcommittee for the Taxonomy of Mollicutes, the International Research Program for Comparative Mycoplasmology (IRPCM) turned to guidelines that had been proposed earlier for recording properties of uncultured organisms (Murray and Schleifer 1994; Murray and Stackebrandt 1995) and proposed to erect a provisional genus-level taxon '*Candidatus* Phytoplasma' to accommodate phytoplasmas (IRPCM 2004). Taxa affiliated with this genus are characterized by the presence of a signature sequence in their 16S rRNA genes (5′-CAAGAYBATKAT GTKTAGCYGGDCT-3′). The IRPCM established guidelines for naming new taxa within the genus: A novel '*Ca.* Phytoplasma' species description should refer to a single, unique 16S rRNA gene sequence that shares less than 97.5 % sequence similarity with that of any previously described '*Ca.* Phytoplasma' species, unless the phytoplasma under consideration clearly represents an ecologically separated population. The 16S rRNA gene of a novel '*Ca.* Phytoplasma' species should possess at least one unique sequence region in addition to the above-defined genus-level signature sequence (IRPCM 2004).

According to the IRPCM (2004) guidelines, the phytoplasma strain whose 16S rRNA gene is used to describe a '*Ca.* Phytoplasma' species is called the 'reference strain'; strains whose 16S rRNA genes have even minimal difference from the corresponding reference strain are referred as 'related strains.' This strict rule of reserving the name of a given '*Ca.* Phytoplasma' species to only its reference strain apparently conflicts with the general understanding that a bacterial species consists of a population of cells with similar characteristics. To better define a '*Ca.* Phytoplasma' species, Davis et al. (2013) proposed that a '*Candidatus* Phytoplasma' species should be attributed to 'phytoplasma strains whose 16S rRNA gene sequences contain the oligonucleotide sequences of unique regions that are designated in the formally published description of the species.' In other words, all strains that possess the same species-specific signature sequences in their 16S rRNA genes belong to the same '*Ca.* Phytoplasma' species. Quaglino et al. (2013) further suggested that, in addition to species-specific signature sequences, combinations of distinguishing sequence blocks (DSBs) in 16S rRNA genes can also be used to identify strains belonging to a given '*Ca.* Phytoplasma' species. Further, it has been proposed that, wherever possible, signature sequences of other genetic loci from the reference strain of a '*Ca.* Phytoplasma' species should be incorporated in emended descriptions and as part of future descriptions of '*Ca.* Phytoplasma' species (Davis et al. 2013). Such new ideas embrace the concept that a '*Ca.* Phytoplasma' species constitutes a well-defined set of genotypically similar phytoplasma strains, rather than a single strain.

10.2.3 Toward Erection of a New Family,Phytoplasmataceae

Thus far, 36 '*Ca.* Phytoplasma' species have been formally described, and an additional 13 potentially new species have been suggested (Davis et al. 2013; Nejat et al. 2013; Quaglino et al. 2013; Win et al. 2013). Based on the similarity of 16S rRNA gene sequences, some phytoplasmas are more distantly related to each other than they are to *Acholeplasma palmae*, member of a separate genus, raising a question as to what taxonomic rank should be assigned to the taxon phytoplasma (Zhao et al. 2010). The topology of phylogenetic tree constructed from 16S rRNA gene sequences divides the phytoplasma clade into three distinct subclades (Fig. 10.1). Furthermore, DNA physical signatures such as dinucleotide relative abundance (DRA) are notably different among the three phylogenetic subclades (Zhao et al. unpublished). Such results encourage us to propose ranking each of the three phytoplasma subclades at genus level and increasing the '*Candidatus* Phytoplasma' clade to a family-level taxon.

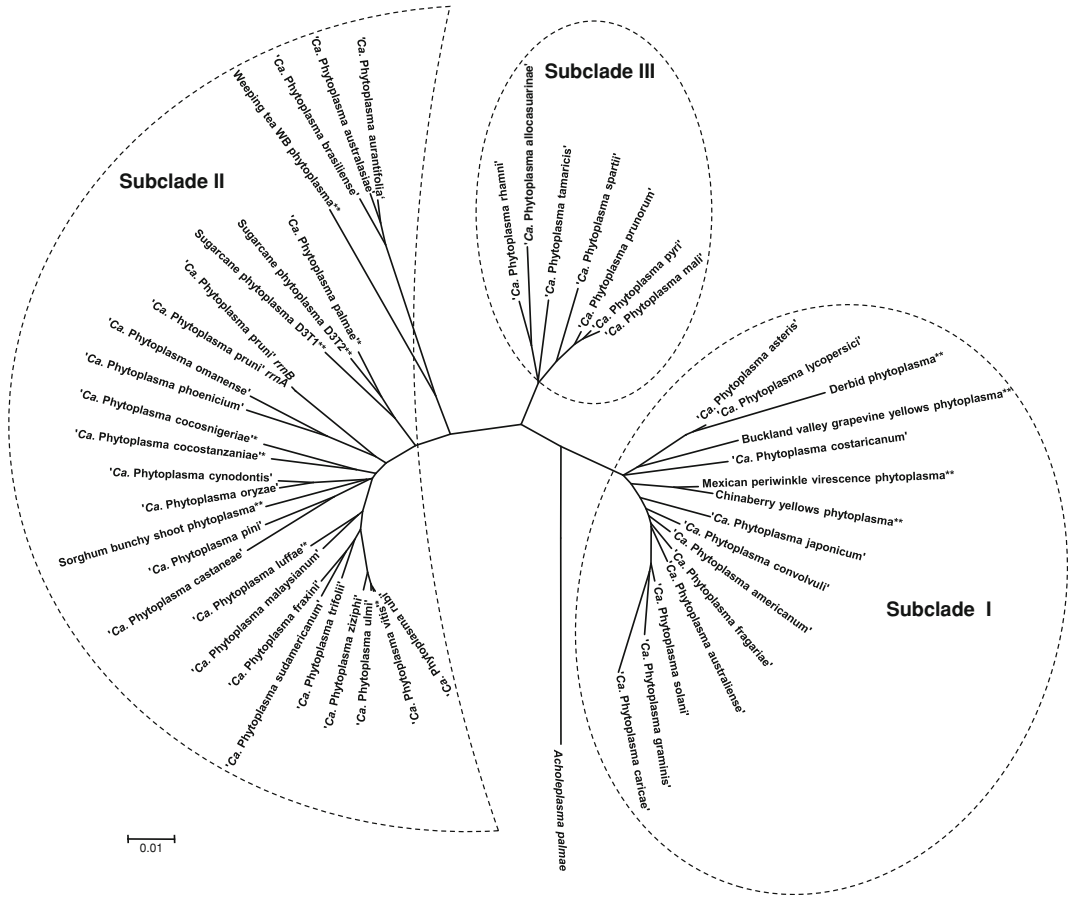

Fig. 10.1 Phylogenetic tree inferred from analysis of 16S rRNA gene sequences. Minimum evolution analysis was conducted using the close-neighbor-interchange (CNI) algorithm implemented in the Molecular Evolutionary Genetics Analysis program (MEGA5, Tamura et al. 2011). The initial tree for the CNI search was obtained by the neighbor-joining method. The reliability of the analysis was subjected to a bootstrap test with 1,000 replicates. The taxa used in the phylogenetic tree reconstruction are reference strains of 36 formally described '*Candidatus* Phytoplasma' species, five incidentally cited '*Ca.* Phytoplasma' species (marked with an asterisk '*'), and eight suggested potentially new '*Ca.* Phytoplasma' species (marked with two asterisks '**'). *Acholeplasma palmae* served as an out-group during the phylogenetic tree reconstruction. The scale bar represents the number of nucleotide substitutions per site

10.3 Genetic Co-linearity Between and Within Species

10.3.1 Global Synteny

After their evolutionary divergence from an *Acholeplasma*-like ancestor and emergence as a discrete clade, phytoplasmas have evolved to give rise to widely divergent lineages. Assuming that phytoplasmas have a monophyletic origin, the pace of phytoplasma evolution may be inferred from the degrees of genetic co-linearity (synteny) and the levels of genomic rearrangements between pairs of extant phytoplasmas. Currently, only limited phytoplasma genome sequence data are available for such analyses at the whole genome level. The four phytoplasma strains whose genomes are completely sequenced belong to three distinct species: '*Ca*. Phytoplasma asteris,'

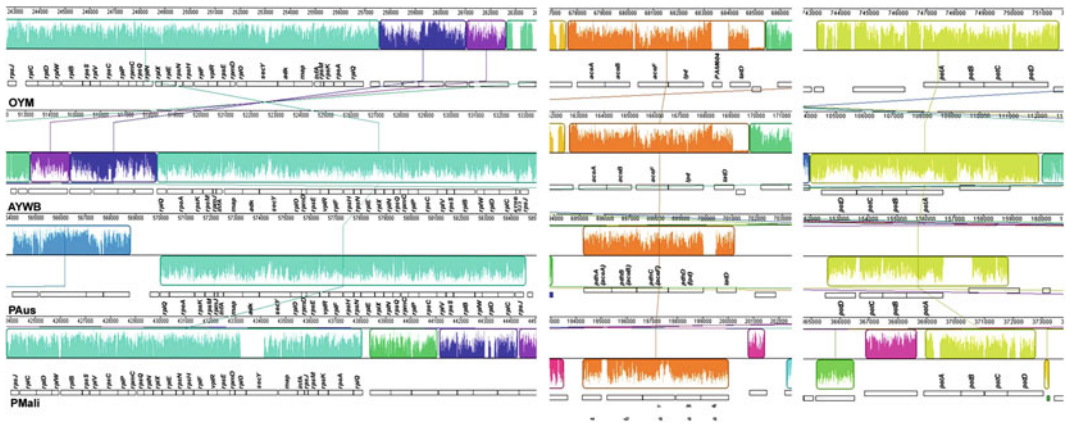

Fig. 10.2 Microsynteny of conserved gene blocks among different phytoplasmas. Multi-genome alignment and identification of locally collinear gene blocks were carried out using MAUVE aligner (Darling et al. 2004). The genome of '*Ca.* Phytoplsma asteris'-related strain OYM was used as a reference genome. Identically colored *blocks* represent a collinear set of matching regions. One connecting line is drawn per collinear block

'*Ca*. Phytoplasma australiense,' and '*Ca*. Phytoplasma mali.' To assess conservation of gene order among the strains, a pairwise comparison of gene arrangement was performed using a sequence alignment program implemented in MolliGen 3 (http://cbib1.cbib.u-bordeaux2.fr/molligen/, Barré et al. 2004). The results revealed that global synteny barely exists between any pair of phytoplasma strains belonging to two different species, a finding that indicates significant evolutionary divergence between the species. The lack of global synteny between strains representing '*Ca*. Phytoplasma asteris' and '*Ca*. Phytoplasma australiense' is surprising, as both species belong to the same subclade (Fig. 10.1). On the other hand, the genomes of OYM and AYWB, both '*Ca*. Phytoplasma asteris'-related strains, exhibit extensive synteny (Bai et al. 2006), and the non-syntenic regions consist mainly of prophage sequences (Wei et al. 2008a).

10.3.2 Microsynteny

Despite extensive genomic rearrangements, the conserved nature, functionality, and evolutionary importance of various coordinately regulated genes are reflected in microsyntenies observed across '*Candidatus* Phytoplasma' species. Examples include genomic segments encoding ribosomal proteins (Fig. 10.2, left panel), the operon encoding key enzymes involved in glycolysis (*acoA*, *acoB*, *aceF*, and *lpd*) (Fig. 10.2, middle panel), and the operon encoding subunits of ABC-type spermidine/putrescine transport system (*potA*, *potB*, *potC*, and *potD*) (Fig. 10.2, right panel). High-level conservation of synteny within a super operon encoding key components of translational apparatus (S10-spc-alpha cluster) was also found in genomes of phytoplasmas affiliated with '*Ca*. Phytoplasma pruni,' '*Ca*. Phytoplasma solani,' '*Ca*. Phytoplasma vitis' (an incidental citation, not formally described '*Ca.* Phytoplasma' species), and '*Ca*. Phytoplasma trifolii' (Durante et al. 2012; Saccardo et al. 2012; ongoing phytoplasma genome project at the Molecular Plant Pathology Laboratory, ARS-USDA). Nevertheless, even in the case of such conserved arrangements of genes, minor subclade-specific differences in gene order are present. Such gene-order differences provide molecular markers for distinguishing varied phytoplasma species/groups/lineages (Fig. 10.2, left panel).

10.4 Distinctive Genome Architecture

10.4.1 Sequence-Variable Mosaics

Despite their unusually small sizes, each of the four completely sequenced phytoplasma genomes (Oshima et al. 2004; Bai et al. 2006; Kube et al. 2008; Tran-Nguyen et al. 2008) contains large amounts of repeated DNA sequences. Dozens of multiple-copy genes with uncharacterized functions are clustered repetitively in non-randomly distributed segments (Jomantiene and Davis 2006). Targeted genome sequencing of multiple strains representing four distinct groups revealed the same phenomenon (Davis et al. 2007, Jomantiene et al. 2007). Such genome organization is unique to phytoplasmas, and this distinct genome architecture is termed sequence-variable mosaics (SVMs) (Jomantiene and Davis 2006; Jomantiene et al. 2007). It was suggested that the SVMs were assembled through multiple events of targeted mobile element attack, duplication, recombination, and rearrangement and that differences in genome size between related phytoplasmas could be largely explained by differences in size and number of SVMs in the chromosomes (Jomantiene and Davis 2006). Terms such as putative mobile units (PMUs) (Bai et al. 2006) and mobile unit genes (MUGs) (Arashida et al. 2008a) were also proposed for the mobile elements comprising these repetitive gene clusters (SVMs).

10.4.2 Phage-Derived Genomic Islands

Extensive computational analyses revealed that the mobile elements comprising SVMs were cryptic prophages or prophage remnants (Wei et al. 2008a), as the chromosomal locations of these phage-derived sequences coincided with those of SVMs. Three types of hypothetical prophage genomes were deduced from the highly mosaic prophage genome remnants; each type of presumed ancestral phage possessing three functional modules: a regulation and DNA packaging module, a morphogenesis and cell lysis module, and a replication and recombination module (Wei et al. 2008a). These phage-derived sequences possess distinct physical properties such as a mol% G+C content lower than that of the host phytoplasma genes (therefore, forming low G+C isochors), and a DRA value drastically different from that of their respective host DNA (thus a remarkable DRA distance or DRAD). The physical properties and the clustered distributions of the phage-derived sequences suggest that the prophage-derived SVMs are genomic islands in phytoplasma genomes. In fact, the DRAD values of phytoplasmal prophage islands (Wei et al. 2008a) are much higher than those of bacterial pathogenicity islands identified in other pathogens (Karlin 2001; Hsiao et al. 2003). New research data have implicated prophage island genes in phytoplasma virulence and pathogen–host interactions (see Sect. 10.8 of this chapter).

While the prophage-derived genomic islands occupy 31.0 and 22.7 % of the circular chromosomes of OYM and AYWB, respectively, the SVM architecture and prophage islands in the genomes of '*Ca*. Phytoplasma australiense' (PAus) and '*Ca*. Phytoplasma mali' (PMali) were not previously delineated. Results from an analysis for this writing revealed that, of the 840 annotated protein-encoding genes (CDSs) in the PAus genome, 374 are clustered, multiple-copy genes of prophage origin. The prophage-derived sequences account for 36.84 % of the PAus genome, with the largest prophage island spanning 77.5 Kbp. PMali has a linear chromosome that bears many fewer genes than the genomes of the other three completely sequenced phytoplasmas (Kube et al. 2008). Results from our analysis indicated that the PMali genome also contains multiple copies of genes that are homologous to prophage sequences that are present in the OYM, AYWB, and PAus genomes. Such genes include those annotated as *hflB*, *gepA*, *fliA*, *yqaJ,* and *tmk*. Without exception, these genes are clustered in discrete regions rather than being randomly distributed across the chromosome, a characteristic of phytoplasmal SVMs (Fig. 10.3).

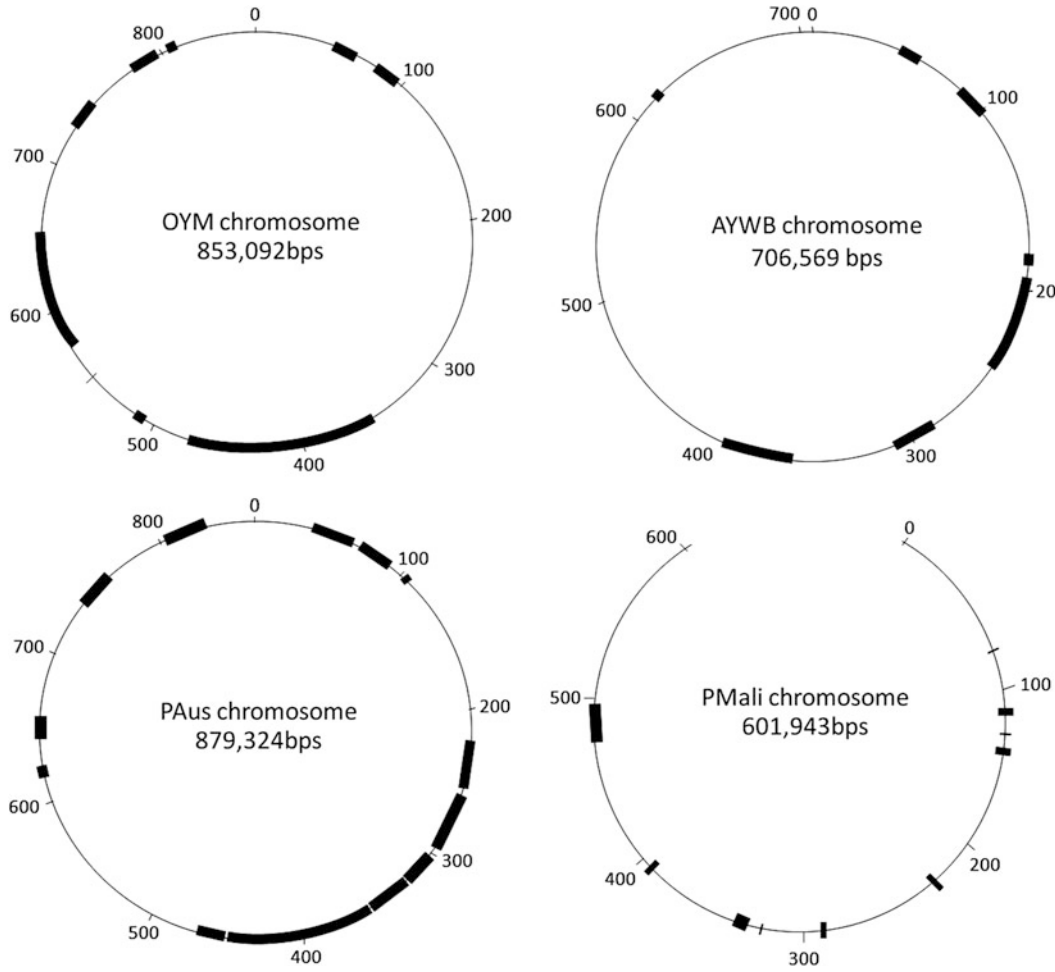

Fig. 10.3 Distribution of phage-derived sequences in the four phytoplasma genomes. The *circular* (OYM, AYWB, and PAus) and *linear* (PMali) chromosomes are represented by *thin lines*. Regions of prophage integration are shown by *thick lines* and regions of sequence-variable mosaic (*SVM*) in the OYM genome are shown by *boxes*. Positions of prophage integration sites in the OYM and AYWB chromosomes are from Wei et al. (2008a). Positions of OYM SVMs are from Jomantiene and Davis (2006)

10.4.3 Hypothesis on the Emergence of the Phytoplasma Clade

Since no SVM-like structures could be identified in genomes of ancestral relatives including low G+C walled bacteria in the *Bacillus/Clostridium* group and in genomes of cell wall-less bacteria *Acholeplasma* spp., we hypothesized that ancient phage attacks leading to SVM formation occurred at or shortly after divergence of phytoplasmas from acholeplasmas, triggering evolution of the phytoplasma clade (Fig. 10.4; Jomantiene and Davis 2006; Jomantiene et al. 2007; Wei et al. 2008a).

10.4.4 Hyper-variable Regions and Foreign Genes Introduced by Phages

Within SVM segments of phytoplasma genomes, there are hyper-variable regions (HVRs), each of which is flanked by a conserved ATP-dependent Zn protease gene and a conserved

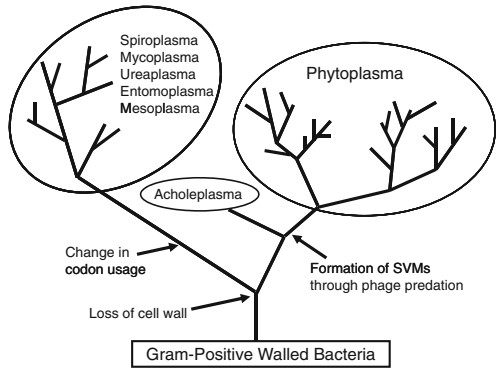

Fig. 10.4 Diagrammatic tree of cell wall-less bacteria. The formation of sequence-variable mosaics following phage predation is depicted as the key evolutionary event at the root of emergence of the phytoplasma clade. The tree was drawn by Katherine T. Davis

palindromic DNA sequence (Jomantiene et al. 2007). Such HVR structure bears some similarities with the integron/mobile gene cassette system found in other bacteria (Recchia and Hall 1997). With phytoplasmal prophage-derived genomic islandsbeing brought to light, and based on the phage-derived nature of the sequences surrounding the HVRs, we hypothesized that HVRs may represent regions of mobile gene cassettes inside a larger mobile genetic element, phage (Wei et al. 2008a). Phages are known to have the ability to carry genes, termed morons, to their host bacterial genomes. Morons are not essential for the phage life cycle but may alter the phenotype or fitness of the lysogen (Brüssow et al. 2004; Cumby et al. 2012). In some pathogenic bacteria, morons introduced by phages are proven or suspected virulence factors (Cenens et al. 2013). At least 23 and 17 morons were identified in the prophage loci of the OYM and AYWB genomes, respectively, and most of them appeared to be strain specific (Wei et al. 2008a). One moron in an OYM prophage was found to be a group II intron that contained a reverse-transcriptase gene. The recognition of a group II intron as a moron draws increased attention to prophages as platforms for acquisition of foreign genes and to their role in phytoplasma evolution and strain diversity. Group II introns were also found in prophages integrated in the PAus chromosome (Tran-Nguyen et al. 2008; this study).

Interestingly, a significant number of morons are targeted for insertion into specific sites in the HVRs, possibly by mechanisms similar to those of integron/mobile gene cassette systems (Jomantiene et al. 2007). One example of morons targeted into HVRs is a phospholipase (PL) gene, first identified in the genome of '*Ca*. Phytoplasma asteris'-related strain Malaysian periwinkle yellows phytoplasma (GenBank no. EF200534). This gene encodes a protein possessing a transmembrane region at its N-terminus. By contrast, a putative PL protein in the western X-disease phytoplasma (GenBank no. AF533231) lacks a transmembrane region and is encoded by a gene located outside of HVRs and SVMs. Similarly, a PL encoded by the *A. laidlawii* genome (GenBank no. NC_010163), which is devoid of SVMs, also lacks a transmembrane region. These observations lead us to suggest that a phytoplasma genome may contain a vertically inherited PL gene and/or a horizontally acquired PL gene. Interestingly, the genome of PMali (GenBank no. NC_011047) contains both a gene (located in a region of decayed SVM) encoding a membrane-targeted PL and a gene (located outside of SVM regions) encoding a non-membrane targeted, presumably vertically inherited PL.

10.5 Genes Unique to Phytoplasmas

Besides the presence of a large amount of repetitive prophage sequences, morons, and transduced genes, phytoplasma genomes contain additional genes that are absent in all other mollicutes. Apparently, along with massive gene loss and decay, gene acquisition and duplication events occurred during, as well as following, the emergence of the phytoplasma clade. In this section, we examine genes that are unique to phytoplasmas, with the optimistic expectation that such analysis will eventually provide an answer to the question of what makes a phytoplasma a phytoplasma.

In this communication, we describe genes that are present in the genomes of at least two phytoplasmas but are absent in the genomes of all other sequenced mollicutes (including mycoplasmas, mesoplasmas, spiroplasmas, ureaplasmas, and acholeplasmas) as 'phytoplasma-unique.' In order to identify genes unique to phytoplasmas, we conducted differential analyses using both bidirectional best hit (BDBH, with an E-value threshold 10^{-4}) and best blast hits (BBH, with an E-value threshold 10^{-8}) approaches. An online database dedicated to the mollicute genomes, MolliGen 3.0 (http://cbib1.cbib.u-bordeaux2.fr/molligen3b/TOOLBOX/DQ/select.php, Barré et al. 2004), was used to formulate and execute differential queries.

10.5.1 Phytoplasma-Unique Genes Within SVMs

Using the OYM genome as a reference, we identified 260 phytoplasma-unique protein-coding genes, of which 187 were prophage-derived genes clustered mainly in four SVM regions. Similarly, in the AYWB genome, we found 261 phytoplasma-unique genes, of which 191 were prophage-derived genes located in SVM regions. Likewise, in the PAus genome, among the 232 phytoplasma-unique genes identified in the present study, 180 were multiple-copy genes sharing significant homology with previously identified OYM and AYWB phytoplasmal prophage remnant sequences. Fewer phytoplasma-unique genes (63) were identified from the PMali genome; yet, 28 shared significant sequence similarities with previously identified phytoplasmal prophage genes or had been annotated by others (Kube et al. 2008) as encoding phage-related or phage-associated proteins. Such findings echo our hypothesis, put forward previously, that targeted integration of prophage sequences and subsequent formation of SVMs characterize the phytoplasma genomes and distinguish phytoplasmas from all other mollicutes including their closest relatives (Jomantiene and Davis 2006; Jomantiene et al. 2007; Wei et al. 2008a).

10.5.2 Phytoplasma-Unique Genes Outside of SVMs

In each of the four completely sequenced phytoplasma genomes, 35–73 phytoplasma-unique genes reside outside of the prophage islands/SVM regions. While many of these phytoplasma-unique, non-phage genes encode hypothetical proteins of unknown functions, a set of about 20 genes can be functionally annotated. It is worth noting that a majority of these genes is common to all four phytoplasmas, regardless of genome size and phylogenetic position of a given phytoplasma (Table 10.1). Significantly, nearly half of these 20 genes encode components of diverse transporters (Fig. 10.5). Since phytoplasmas have limited metabolic capacities, they must be sustained by steady import of nutrients and timely efflux of toxins, in a constant exchange of metabolites with host cells. In addition, cross-membrane transportation may be also required for maintaining intracellular electrode and redox potentials and for mediating secretion of potential virulence factors. The observed abundance of phytoplasma-unique transport systems draws into sharp focus the intimate relationship between phytoplasma and host cell. Improved understanding of the functions of these transporters and their substrate specificities could aid efforts to identify molecular targets for devising practical measures to curb phytoplasma growth and mitigate damage from phytoplasmal diseases of plants.

10.5.2.1 ABC-Type Mn/Zn Transport System

Zinc and manganese are essential nutrients for bacteria since they play important catalytic and structural roles in a variety of enzymes and metallo-regulatory proteins (Moore and Helmann 2005). Manganese is crucial in detoxification of superoxide free radicals (Abreu and Cabelli 2010). Zinc is found in at least 100 enzymes that cover all enzyme classes (Trumbo et al. 2001). In *Staphylococcal* bacteria, zinc is involved in biofilm formation by facilitating

Table 10.1 A subset of phytoplasma-unique genes outside of phage-derived genomic islands/sequence-variable mosaic (SVM) regions[a]

Strain	CDS	Begin	End	Strand	Gene	Product
OYM	PAM005	5,467	6,222	–	*pgpB*	Membrane-associated phospholipid phosphatase
OYM	PAM092	115,631	116,776	+	*znuA*	ABC-type Mn/Zn transport system, periplasmic Mn/Zn-binding protein
OYM	PAM097	123,489	124,316	+	*nlpA*	ABC-type uncharacterized transport system, periplasmic component
OYM	PAM122	144,248	144,949	+	*amp*	Antigenic membrane protein
OYM	PAM133	159,694	160,737	+	*nlpA*	ABC-type uncharacterized transport system, periplasmic component
OYM	PAM188	225,774	226,019	–		Zn-dependent carboxypeptidase
OYM	PAM191	231,479	233,050	–	*oppA*	ABC-type dipeptide/oligopeptide transport system, periplasmic component
OYM	PAM254	290,447	290,857	+	*secE*	Preprotein translocase subunit SecE
OYM	PAM280[c]	322,195	323,790	–	*norM*	Na^+-driven multidrug efflux pump
OYM	PAM282[d]	325,982	327,466	+	*norM*	Na^+-driven multidrug efflux pump
OYM	PAM286	332,415	333,560	+	*citS*	Malate/citrate symporter
OYM	PAM306	350,438	351,805	+	*ksgA*	Dimethyladenosine transferase
OYM	PAM429	476,118	476,294	+	*malK*	ABC-type sugar transport systems, ATPase components
OYM	PAM476	540,857	541,366	–	*mutT*	MutT/Nudix family protein
OYM	PAM494[b]	566,709	566,987	–	*artM*	ABC-type amino acid transport system, permease component
OYM	PAM495	567,076	568,083	–	*artI*	ABC-type amino acid transport system, periplasmic component
OYM	PAM580	657,662	658,177	–	*folK*	7,8-Dihydro-6-hydroxymethylpterin-pyrophosphokinase
OYM	PAM581	658,387	659,268	–	*folP*	Dihydropteroate synthase
OYM	PAM594	677,731	679,407	+	*asnB*	Asparagine synthase
OYM	PAM612	698,172	699,170	–	*psd*	Phosphatidylserine decarboxylase
OYM	PAM613	699,183	69,9917	–	*pssA*	Phosphatidylserine synthase
OYM	PAM667	765,736	766,866	+	*clpX*	ATP-dependent protease Clp, ATPase subunit
OYM	PAM720	818,800	819,915	–	*citS*	Malate/citrate symporter
OYM	PAM721	820,334	821,506	–	*sfcA*	Malic enzyme
OYM	PAM734[e]	836,678	838,177	–	*norM*	Na^+-driven multidrug efflux pump (see foot note of the Table)
AYWB	AYWB_051	58,057	59,229	+	*sfcA*	NAD-dependent malic enzyme
AYWB	AYWB_052	59,226	60,761	+	*citS*	Malate-sodium symport
AYWB	AYWB_120	141,714	142,448	+	*pssA*	CDP-diacylglycerol-serine O-phosphatidyltransferase
AYWB	AYWB_121	142,461	143,462	+	*psd*	Phosphatidylserine decarboxylase
AYWB	AYWB_302	316,503	317,012	+	*mutT*	MutT/Nudix family phosphohydrolase
AYWB	AYWB_415	427,223	428,575	–		Dimethyladenosine transferase
AYWB	AYWB_435	444,283	445,650	–	*citS*	Citrate-sodium symport
AYWB	AYWB_439[d]	450,303	451,811	–	*norM*	Na^+-driven multidrug efflux pump

(continued)

Table 10.1 (continued)

Strain	CDS	Begin	End	Strand	Gene	Product
AYWB	AYWB_441[c]	453,835	455,430	+	*norM*	Na^+-driven multidrug efflux pump
AYWB	AYWB_467	486,386	486,940	–	*secE*	protein translocase subunit SecE
AYWB	AYWB_529	541,595	543,136	+	*dppA*	ABC-type dipeptide-binding protein, solute-binding protein
AYWB	AYWB_599	623,951	624,445	–		Immune-dominant membrane protein precursor
AYWB	AYWB_609	633,238	633,399	–		IscU protein (NifU homolog involved in Fe-S cluster formation)
AYWB	AYWB_610	633,608	633,772	–		Uncharacterized ABC-type transport system, periplasmic component
AYWB	AYWB_611	633,789	634,196	–		Uncharacterized ABC-type transport system, periplasmic component
AYWB	AYWB_621[b]	645,598	646,692	–	*mntB*	ABC-type Mn/Zn transport system, membrane protein
AYWB	AYWB_624	648,572	649,714	–	*znuA*	ABC-type Mn/Zn transport system, periplasmic Mn/Zn-binding protein
AYWB	AYWB_651[e]	681,922	683,382	–	*norM*	Na^+-driven multidrug efflux pump (see foot note of the Table)
PAus	PA0110[e]	118,738	120,204	+	*norM*	Na^+-driven multidrug efflux pump (see foot note of the Table)
PAus	PA0139	164,912	165,760	–	*cbiQ*	ABC-type cobalt transport system, permease protein
PAus	PA0165[c]	197,185	198,630	+	*norM*	Na^+-driven multidrug efflux pump
PAus	PA0171[d]	205,945	207,435	–	*norM*	Na^+-driven multidrug efflux pump
PAus	PA0221	260,515	261,975	+	*gtfA*	Sucrose phosphorylase (Glucosyltransferase-A)
PAus	PA0454	455,152	455,985	+	*nlpA*	ABC-type uncharacterized transport system, periplasmic component
PAus	PA0455	456,017	456,907	+		ABC-type metal ion transport system, periplasmic component
PAus	PA0456	457,305	457,604	+	*iscU*	NifU homolog involved in Fe-S cluster formation
PAus	PA0481	476,283	477,404	+	*znuA*	ABC-type Zn/Mn transport system periplasmic component
PAus	PA0483[b]	478,522	479,427	+	*znuB*	ABC-type Zn/Mn transport system, permease component
PAus	PA0484[b]	479,443	480,531	+	*znuB*	ABC-type Zn/Mn transport system, permease component
PAus	PA0626	621,568	622,233	+	*pssA*	Phosphatidylserine synthase
PAus	PA0627	622,233	623,105	+	*psd*	Phosphatidylserine decarboxylase
PAus	PA0633[c]	629,135	630,658	+	*norM*	Na^+-driven multidrug efflux pump
PAus	PA0671	672,042	672,452	–	*secE*	Protein translocase subunit SecE
PAus	PA0755	768,393	769,952	+		ABC-type peptide/nickel transport system substrate-binding protein
PAus	PA0756	770,273	771,865	+		ABC-type peptide/nickel transport system substrate-binding protein
PAus	PA0768	785,343	786,809	–	*citS*	Putative malate/citrate symporter
PAus	PA0769	786,806	787,978	–	*sfcA*	Malate dehydrogenase

(continued)

Table 10.1 (continued)

Strain	CDS	Begin	End	Strand	Gene	Product
PAus	PA0776	796,496	7,97053	+	*mutT*	Nudix hydrolase
PMali	ATP_00003[b]	2,412	3,506	–	*znuB*	ABC-type Mn/Zn transport system, permease component
Pmali	ATP_00006	538,9	6,462	–	*znuA*	ABC-type Mn/Zn transport system, periplasmic Mn/Zn-binding protein
PMali	ATP_00019	253,81	2,6754	–	*mleP*	Malate/Na^+ symporter
PMali	ATP_00047	601,76	60,919	+	*pssA*	CDP-diacylglycerol-serine-phosphatidyltransferase
PMali	ATP_00048	609,32	618,07	+	*psd*	Phosphatidylserine decarboxylase
PMali	ATP_00068	817,17	83,342	–	*dppA*	ABC-type dipeptide-binding protein, solute-binding protein
PMali	ATP_00085	107,015	107,788	+	*pgpB*	Probable membrane-associated phospholipid phosphatase
PMali	ATP_00104	139,660	141,141	–	*mleP*	Malate/Na^+ symporter
PMali	ATP_00107[c]	143,168	144,679	–	*norM*	Na^+-driven multidrug efflux pump
PMali	ATP_00108	144,746	145,216	–	*mutT*	MutT/Nudix family protein
PMali	ATP_00151	192,248	193,657	–	*mleP*	Malate/Na^+ symporter
PMali	ATP_00192	248,530	249,600	–		ABC-type methionine transport system, periplasmic component
PMali	ATP_00279	358,370	359,794	+		Probable dimethyladenosine transferase
PMali	ATP_00384	460,267	460,659	+	*secE*	Protein translocase subunit SecE
PMali	ATP_00450	535,373	536,545	+	*sfcA*	NADP-dependent malic enzyme
PMali	ATP_00479	575,190	576,563	+	*mleP*	Malate/Na^+ symporter
PMali	ATP_00492	595,482	596,555	+	*znuA*	ABC-type Mn/Zn transport system, periplasmic Mn/Zn-binding protein
PMali	ATP_00495[b]	598,438	599,532	+	*znuB*	ABC-type Mn/Zn transport system, permease component

[a] Gene names and their products are mainly based on previously published genome annotations (Oshima et al. 2004; Bai et al. 2006; Tran-Nguyen et al. 2008; Kube et al. 2008)

[b] Permease components of some phytoplasma-unique ABC transporters exhibit weak sequence homology with those of certain uncharacterized transporters in other mollicutes. Since the E-values are beyond the cutoff value adopted in our analysis, these components are still considered as phytoplasma-unique. Importantly, the periplasmic or substrate-binding domains of all ABC transporters listed are exclusively phytoplasma-unique

[c, d, and e] Genes encoding three subtypes of NorM paralogs are identified from the four completely sequenced phytoplasma genomes. The first subtype (denoted with c) is phytoplasma-unique; the second subtype (denoted with d) bears weak sequence similarity with *Acholeplasma* NorMs but the E-values are higher than the cutoff value adopted in our differential analysis; therefore, the second subtype is also considered as phytoplasma-unique; the third type (denoted with e) shares moderate sequence similarity with *Acholeplasma* NorMs. The third subtype of phytoplasmal NorMs is not considered as phytoplasma-unique in this chapter, and the inclusion of the third subtype *norM* genes in this table is for comparative analysis

adhesive contacts between cells (Conrady et al. 2013). In *Legionella pneumophila*, one of the effector proteins, PlcC/CegC1, is a member of the zinc metallophospholipase C family (Aurass et al. 2013). While many walled bacteria use an ABC transporter for high-affinity uptake of zinc and manganese (Hantke 2005), genomes of cell wall-less bacteria studied thus far, including that of *Spiroplasma kunkelii*, a phloem-inhabiting bacterium that is closely related to phytoplasmas, do not appear to encode an ABC-type Mn/Zn transport (Zhao et al. 2004). Phytoplasmas are the only cell wall-less bacteria known to have genes encoding ABC-type Mn/Zn transport systems. The demand for zinc in phytoplasma cells must be high, as the genomes of all

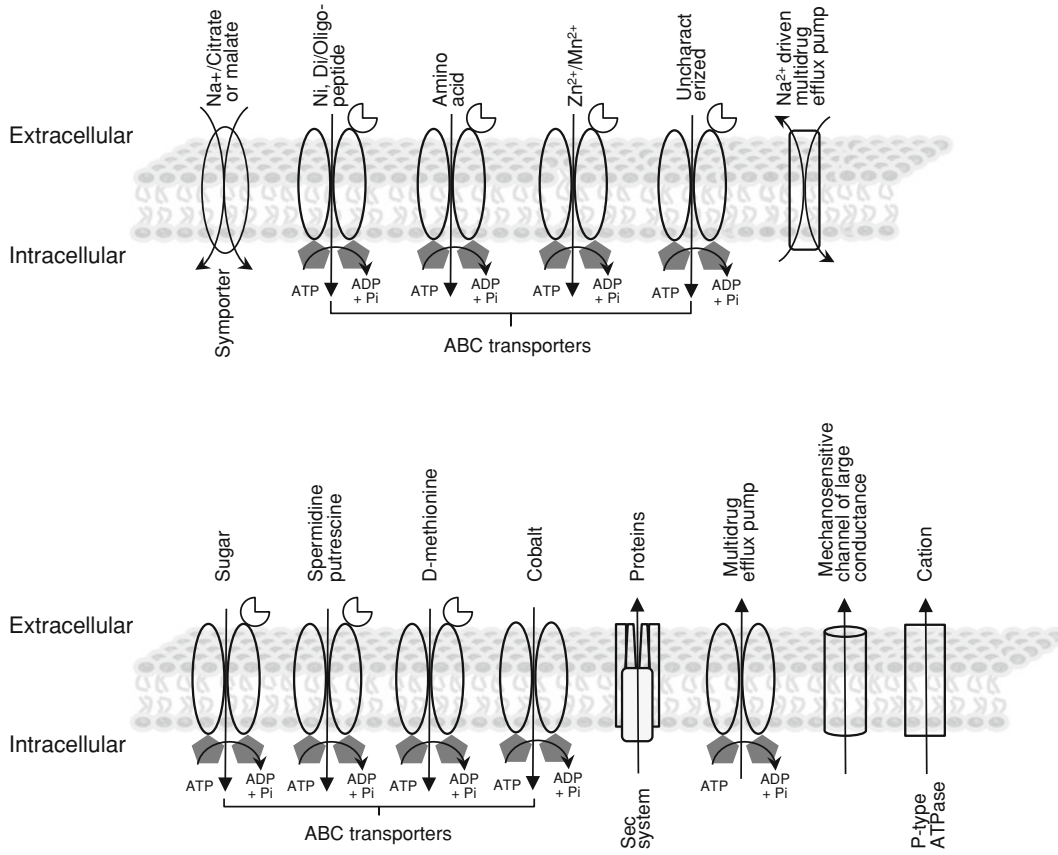

Fig. 10.5 Phytoplasmal transporters. The *upper panel* shows transporters encoded by genomes of phytoplasmas but not other cell wall-less bacteria studied thus far. The *lower panel* shows transporters encoded by genomes of other branches of cell wall-less bacteria in addition to phytoplasmas. Transmembrane components of some phytoplasma-unique ABC transporters illustrated in the upper panel exhibit weak sequence homology with the permease component of certain uncharacterized ABC transporters in mollicutes other than phytoplasmas. These components are still considered as phytoplasma-unique since the E-values are higher than the cutoff value adopted in our differential analysis. Importantly, the sequences of the periplasmic or substrate-binding domains of all ABC transporters shown in the *upper panel* are exclusively phytoplasma-unique. Three types of Na^+-driven multidrug efflux pump paralogs have been identified in phytoplasmas. Refer Table 10.1 for phytoplasma-unique Na^+-driven multidrug efflux pumps

sequenced phytoplasmas possess numerous copies of genes encoding several different classes of Zn protease-like proteins; perhaps significantly, a majority of those Zn protease genes reside in SVM/phage-derived genomic islands.

10.5.2.2 ABC-Type Amino Acid and Dipeptide/Oligopeptide Transport Systems

The four completely sequenced phytoplasma genomes lack many genes encoding enzymes responsible for metabolism of amino acids, especially cysteine, methionine, histidine, lysine, arginine, proline, and tyrosine. Presence of phytoplasma-unique ABC-type amino acid and dipeptide/oligopeptide transport systems may reflect a necessity for amino acid importation from host cells. Imported dipeptides and oligopeptides may be digested into individual amino acids by peptidases, as several peptidase genes are present in phytoplasmas. The genome of PMali also encodes a distinct ABC-type D-methionine transport system whose

periplasmic substrate-binding subunit is phytoplasma-unique (ATP_00192).

10.5.2.3 Malate/Citrate Symporter (Citrate-Sodium Symport, Malate-Sodium Symporter)

Phytoplasmas have limited capacity for carbohydrate metabolism and energy production. Presence of a malate/citrate symporter gene in all studied phytoplasma genomes indicates that malate is imported and possibly utilized as an alternative carbon source for energy generation. Since malate is abundant in phytoplasma plant host cells (Ziegler 1975), and since the genes encoding most of the enzymes required for converting malate to acetate and generation of ATP are present in phytoplasma genomes (Kube et al. 2012), it is quite possible that phytoplasmas use this alternative carbon source.

10.5.2.4 Na^+-Driven Multidrug Efflux Pump

Besides having limited biosynthetic capacity, phytoplasmas also have limited catabolic capacity. Conceivably, instead of degrading toxic chemicals and other cellular wastes metabolically, phytoplasmas may rely on membrane transport systems to discharge such harmful substances. The genomes of all four sequenced phytoplasma strains encode an ABC-type multidrug/protein/lipid transport system. In addition, all four genomes, except for PMali, have multiple genes encoding NorM-type Na^+-driven multidrug efflux pumps. In the present study, genes encoding three subtypes of Na^+-driven multidrug efflux pump paralogs were identified in the OYM, AYWB, and PAus genomes. The first subtype is phytoplasma-unique (PAM280, AYWB_441, PA0165, and PA0633). The second subtype bears weak sequence similarity with *Acholeplasma* multidrug efflux pumps, but the E-values are higher than the cutoff value adopted in our differential analysis; therefore, they are also considered as phytoplasma-unique (PAM282, AYWB_439, and PA0171). The third subtype shares moderate sequence similarity with *Acholeplasma* multidrug efflux pumps (PAM734, AYWB_651, and PA0110). We hypothesize that genes of the three subtypes of phytoplasmal Na^+-driven multidrug efflux pumps had a common ancestor that was also shared by genes of *Acholeplasma* multidrug efflux pumps. In the course of phytoplasma evolution, following gene duplication events, phytoplasmal *norM* paralogs apparently have evolved to maximize phytoplasmas' adaptation to their complex niche, enhancing host switching between plant and insect, and have gradually become phytoplasma-unique. Alternatively, the genes encoding phytoplasma-unique Na^+-driven multidrug efflux pumps may have been acquired horizontally, facilitating phytoplasmas' transkingdom parasitic lifestyle. PMali, which has a much smaller genome, possesses only one copy of the *norM* gene (ATP_00107), and its nucleotide sequence shares highest similarity with the PAus phytoplasma-unique paralog PA0633.

10.5.2.5 Translocon Subunit Composition and Interaction with Ribosomal Proteins

In many bacteria, a majority of secreted as well as membrane-associated proteins are translocated through a heterotrimeric protein-conducting channel that is formed by preprotein translocase subunits SecY, SecE, and SecG. The SecYEG complex, often referred to as the translocon, is evolutionarily conserved across prokaryotes and even across all three domains of life (Rapoport 2007). It has been noted that the phytoplasmas lack the gene encoding SecG (Kube et al. 2012). Phytoplasma genomes also lack genes encoding SecB. SecB is a molecular chaperon that maintains preproteins in an unfolded state after translation and targets these preproteins to the membrane for secretion (Muller 1999). The lack of SecB implies that the Sec system in phytoplasmas accomplishes preprotein translocation in a co-translational mode, in which the translocation

system is in direct contact with the exit tunnel of the ribosome when a nascent polypeptide chain emerges from the ribosome, as indicated in other bacteria (Ménétret et al. 2007).

In the present study, we found that the SecE-encoding genes from phytoplasmas are distinctly divergent from those of all other mollicutes. It would be interesting to learn whether the loss of the *secG* and *secB* genes and the evolution of the phytoplasma-unique *secE* gene are co-related events and to learn how those changes affect the assembly and conformation, thus function, of the phytoplasmal translocon. In *Escherichia coli*, when the SecYEG translocation complex is docked with the ribosome, SecE plays a role in contacting ribosomal proteins (Yahr and Wickner 2000). The ribosomal protein composition of the phytoplasmal translation machinery is quite unique among mollicutes. We found that phytoplasma genomes encode ribosomal protein L33a, whose amino acid sequences are significantly different from those of their counterparts in other mollicutes; on the other hand, OYM and AYWB phytoplasmas lack ribosomal protein L21, and PMali and PAus have a ribosomal protein L21 that is very divergent from those of other mollicutes. Do these alterations in ribosomal protein composition/amino acid sequence affect the ribosome's three-dimensional structure and, in turn, its ability to interface with the translocon, thus requiring corresponding structural changes in the phytoplasma *secE* gene, and SecE amino acid sequence?

10.5.2.6 Antigenic, Membrane-embedded Proteins

The genes that are unique to phytoplasmas also include those that encode immunodominant or antigenic membrane proteins (IMPs and AMPs). It has been reported that, in diverse phytoplasmas, IMP/AMP genes are highly expressed, and therefore, the antigenic membrane proteins are abundant at the surface of the phytoplasma cells (Morton et al. 2003; Kakizawa et al. 2004, Arashida et al. 2008b). The *amp* and *imp* genes from different phytoplasma lineages share limited sequence similarity (Barbara et al. 2002; Morton et al. 2003). Nevertheless, our analysis revealed that their gene products, AMPs and IMPs, are generally rich in positively charged amino acids and therefore have an isoelectric point (pI) greater than 8.0. At physiological pH, such membrane proteins would tend to present, at the cell surface, positively charged sites or pockets that are critical for ligand binding, signal perception, and other biochemical functions during pathogen–host interactions. AMPs and IMPs also hold promise as useful targets for strain- and cluster-specific serological detection and identification of diverse phytoplasmas and for devising novel immunoaffinity procedures aimed at purification of phytoplasmas for genomic and proteomic studies.

The other phytoplasma-unique genes encode enzymes involved in metabolic pathways of carbohydrate, lipid, and energy metabolisms.

10.5.2.7 Sucrose Phosphorylase or Glucosyltransferase-A

Phytoplasma genomes contain no genes encoding enzymes for the citrate (TCA) cycle or oxidative phosphorylation. Based on metabolic pathway genes present in phytoplasma genomes, it has been suggested that anaerobic hydrolysis of glucose to lactic acid, i.e., glycolysis, may be the main energy-yielding process in phytoplasma cells (Oshima et al. 2004). In other bacterial cells, the entry point of glycolysis is glucose-6-phosphate imported through the phosphoenol-pyruvate-dependent sugar phosphotransferase system (PTS). As phytoplasmas lack the PTS for sugar phosphate uptake and lack a hexokinase for glucose phosphorylation, presence of a *gtfA* gene in OYM and PAus raises the possibility that some phytoplasmas may obtain sugar phosphate for glycolysis through the catalytic activity of GtfA, which allows formation of α-D-glucose-1-phosphate from sucrose and phosphate (Fig. 10.6) (Kube et al. 2012).

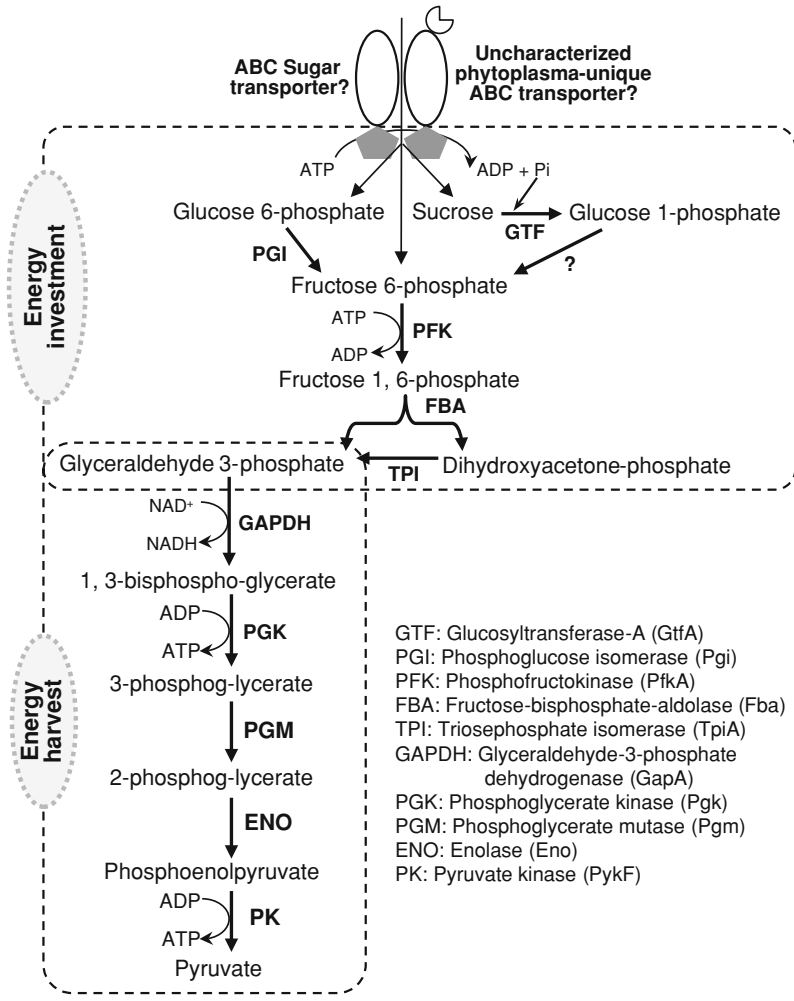

Fig. 10.6 Diagrammatic representation of the Embden–Meyerhof–Parnas pathway (glycolysis) in phytoplasmas, showing a hypothesized ABC sugar transporter and roles of energy investment and energy harvest enzymes. The existence of an as yet unknown enzyme for conversion of glucose-1-phosphate to glucose-6-phosphate is proposed. Genes encoding enzymes involved in the energy investment portion of the pathway (*pgi*, *pfkA*, *fba*, and t*ipA*) are present in the genomes of all four completely sequenced phytoplasmas. Genes encoding enzymes in the energy harvest portion of the pathway (*gapA*, *pgk*, *pgm*, *eno*, and *pykF*) are missing in the genome of PMali

10.5.2.8 NAD-Dependent Malic Enzyme or Malate Dehydrogenase (SfcA)

By catalyzing oxidative decarboxylation, malate dehydrogenase converts malate to pyruvate, and the latter may be further oxidized to produce acetyl CoA, an energy-charged molecule that can be used in diverse biochemical reactions as both a source of metabolic energy and an intermediate. As mentioned above, phytoplasmas have limited energy-yielding capacity; an alternative carbon source and energy-yielding pathway would certainly provide advantages for survival and growth.

10.5.2.9 Dimethyladenosine Transferase

Three of the four sequenced phytoplasma chromosomes each harbor multiple genes that have been annotated by others as dimethyladenosine transferase genes. In the present study, a

comparison of the multiple *ksgA* gene sequences revealed that these genes belong to two mutually distinct groups. One group is represented by ORFs PAM731, AYWB_648, PA0272, and ATP00226. Genes in this group share significant sequence identity with dimethyladenosine transferase genes that are present in the genomes of other mollicutes including *M. genitalium* and *Acholeplasma* spp. In other bacteria as well as in ribosome-containing eukaryotic organelles, Dimethyladenosine Transferase (KsgA) methylates two adjacent adenosines in the loop of a conserved hairpin near the 3′-end of 16S rRNA, playing a critical role in biogenesis of ribosomal 30S subunits (O'Farrell and Rife 2012). The second group of the annotated phytoplasmal *ksgA* genes is represented by PAM306, AYWB_415, and ATP_00279. The genes in this group appear to be phytoplasma-unique among mollicutes. A conserved domain search revealed that an *S*-adenosylmethionine-dependent methyltransferase (AdoMet-MTase) domain is present near the C-terminus of the amino acid sequence deduced from each of these genes; this observation may explain why these genes were previously annotated as encoding dimethyladenosine transferase, which is a member of the AdoMet-MTase family. AdoMet-MTases use *S*-adenosyl-L-methionine as a substrate, producing *S*-adenosyl-L-homocysteine and, at the same time, transferring a methyl group to a variety of biomolecules including nucleotides, amino acids, and lipids. *S*-adenosyl-L-homocysteine can be broken down to adenosine and L-homocysteine, and the latter can subsequently be either remethylated to methionine or converted to cysteine (Clarke and Banfield 2001). Therefore, in addition to methyl group transfer, AdoMet-MTases also can play a role in the conversion of methionine to cysteine. Although our analysis of the gene sequence failed to indicate what specific methyl group transfer reaction(s) the putative phytoplasmal AdoMet-MTases may catalyze, it would be worthwhile to investigate possible involvement of these enzymes in conversion of methionine to cysteine, since phytoplasmas lack the enzymes for cysteine biosynthesis, and the conversion could serve as a source of needed cysteine.

10.5.2.10 Nudix Family Protein

Nudix family proteins (MutT) are diverse hydrolases with a Nudix motif, GXXXXXEXX XXXXXREUXEEXGU, where U is isoleucine, leucine, or valine, and X is any amino acid (McLennan 2006). Nudix proteins hydrolyze potentially hazardous materials, eliminate excess metabolites, and modulate the accumulation of intermediates in biochemical pathways; they are therefore considered as versatile cellular housecleaning enzymes (Bessman et al. 1996). Conceivably, presence of such a housekeeper in the phytoplasma cell is indispensable, because phytoplasmas lack many metabolic enzymes that are capable of pathway-specific degradation of diverse, potentially deleterious, surplus metabolites. The genome of *A. laidlawii* also possesses a gene that encodes a Nudix family protein, but the amino acid sequence of that protein is quite different from those found in phytoplasmas.

10.5.2.11 Phospholipid Phosphatase

Genes encoding putative membrane-associated phospholipid phosphatases are present in the genomes of OYM (PAM005) and PMali (ATP_00085); both genes were previously annotated as *pgpB* (Oshima et al. 2004; Kube et al. 2008). The amino acid sequences deduced from these phytoplasma genes each possess a PAP2-like domain. PAP2 is a super-family of histidine phosphatases that includes type 2 phosphatidic acid phosphatase, phosphatidylglycerophosphatase B, glucose-6-phosphatase, and bacterial acid phosphatases (PFAM01569). Results from a Conserved Domain Database (CDD) search indicated that the two putative phytoplasmal phospholipid phosphatases may belong to the Aur1-like subfamily (cd03386) within the PAP2 super-family (This study). The yeast inositol phosphorylceramide (IPC)

synthase (Aur1p) is the type member of the Aur1-like subfamily; Aur1p is required for the addition of inositol phosphate to ceramide, an essential step in sphingolipid biosynthesis (Nagiec et al. 1997). Recently, a functional ortholog of Aur1p was found in *Toxoplasma gondii* and was reportedly responsible for de novo synthesis of sphingolipids in this intracellular apicomplexan parasite (Pratt et al. 2013). Since IPC is a component of eukaryotic cell membranes, IPC synthase genes are mostly found in yeast, plants, and protozoa including insect vector-borne pathogenic protozoa in the order of *Kinetoplastida* (Mina et al. 2009). It remains to be determined whether the putative membrane-associated phytoplasmal phospholipid phosphatase genes are functional orthologs of Aur1p or are, as the original authors annotated (Oshima et al. 2004; Kube et al. 2008), orthologs of PgpB. Nevertheless, the fact that all other mollicutes examined lack nucleotide sequences homologous to the two noted phytoplasma genes raises the possibility that some phytoplasmas acquired these phospholipid phosphatase genes horizontally during evolutionary adaptation in a lineage-specific manner. Could they have been horizontally acquired from an insect-dwelling eukaryote?

10.5.2.12 Phosphatidylserine Decarboxylase and Phosphatidylserine Synthase

As reviewed by Kube et al. (2012), a common set of genes encoding enzymes for biosynthesis of essential phospholipids are present in all four completely sequenced phytoplasma genomes. These enzymes include PlsX, GpsA, PlsY, PlsC, and CdsA. The presence of such a gene set indicates that phytoplasmas may be able to synthesize CDP-diacylglycerol from acyl phosphate and dihydroxyacetone phosphate. Phytoplasmas also have genes encoding CDP-diacylglycerol-glycerol-3-phosphatidyltransferase (PgsA), enabling the formation of L-1-phosphatidyl-glycerol. In addition, phytoplasmas have genes encoding phosphatidylserine decarboxylase (Psd) and phosphatidylserine synthase (PssA), leading to biosynthesis of L-1-phosphatidylserine (PS) and L-1-phosphatidylethanolamine (PE), respectively. Notably, among mollicutes, phytoplasmas are the only group of organisms whose genomes encode Psd and PssA. In each of the four sequenced phytoplasma chromosomes, *psd* and *pssA* genes are clustered together in a single operon (PAM612, PAM613; AYWB_120, AYWB_121; PA0626, PA0627; ATP_00047, ATP_00048) and are flanked by a gene (*pyrG*) encoding CTP synthase (PAM611, AYWB_122, PA0628, and ATP_00049). CTP is required for the immediate upstream reaction that leads to formation of CDP-diacylglycerol, the very substrate that feeds into the PssA reaction. Such a gene arrangement indicates the conserved nature, and therefore functional importance, of the gene cluster in phytoplasmas. Besides membrane biogenesis, phospholipid biosynthetic pathways play important roles in the virulence of diverse pathogens including fungi (Chen et al. 2010) and bacteria (Conde-Alvarez et al. 2006; Bukata et al. 2008). Since phosphatidylcholine is a typical eukaryotic phospholipid and is absent from most prokaryotes, its presence in the intracellular *Brucella* may constitute host mimicry (Conde-Alvarez et al. 2006). These observations raise the intriguing possibility that the unusually comprehensive phospholipid biosynthesis pathway in phytoplasmas plays a role in phytoplasma pathogenesis and bring to mind that OYM and PMali possibly possess an unusual Aur1-like phospholipid phosphatase that is similar to eukaryotic inositol phosphorylceramide synthase.

10.6 Lineage-Specific Acquisition and Loss of Metabolic Genes in the Four Completely Sequenced Genomes

It seems reasonable to envision that, during their evolutionary lineage radiation, phytoplasma species experienced differing patterns of lineage-specific genome evolution in their adaptations to

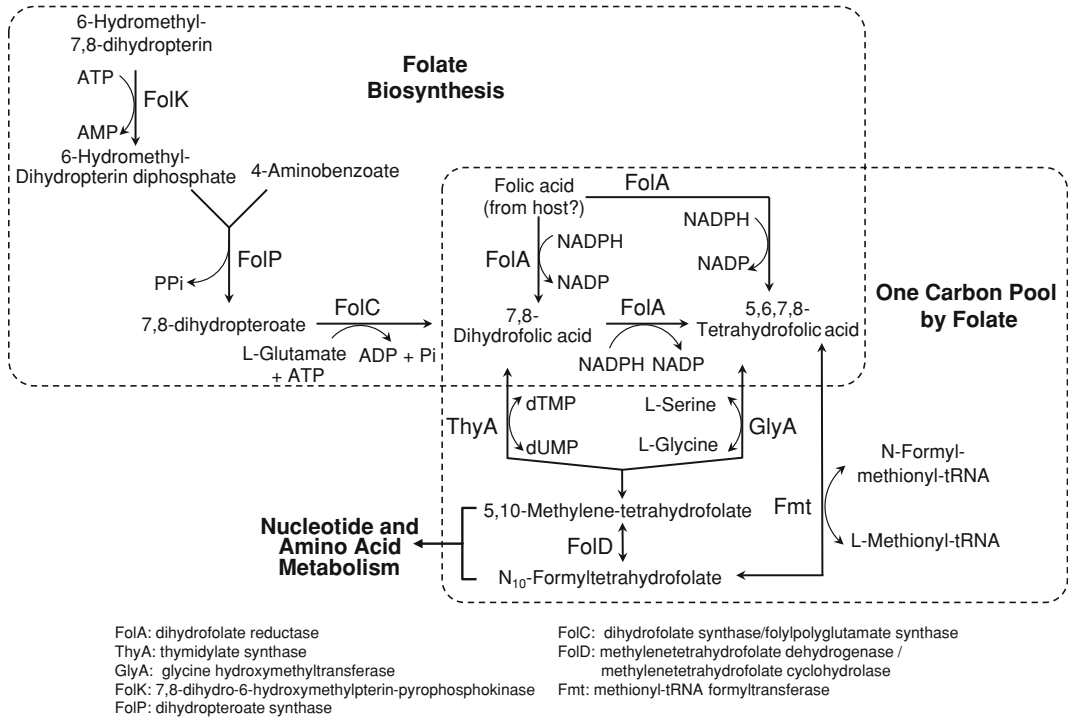

Fig. 10.7 Folate biosynthesis and one-carbon pool by folate pathways in phytoplasmas. Genes encoding enzymes shown in *gray* are either absent or are undergoing lineage-specific decay in the genomes of some phytoplasmas. Genes encoding enzymes shown in *black* are present and intact in the genomes of all four completely sequenced phytoplasmas

diverse plant and insect hosts. Evolving phytoplasma lineages likely were exposed to differing sources of potentially horizontally acquired genes, and availabilities of essential nutrients required by the evolving microbes may also have differed with their hosts. These presumptions could explain both lineage-specific acquisition and lineage-specific loss of metabolic, as well as other, genes.

10.6.1 Folate Biosynthesis and One-Carbon Transfer

Among the four completely sequenced phytoplasma genomes, only the genome of OYM encodes a set of four intact enzymes (FolK, FolP, FolC, and FolA) required for synthesis of tetrahydrofolate from 6-hydroxymethyl-7,8-dihydropterin and 4-aminobenzoate (Fig. 10.7). The gene encoding FolK is missing in the genomes of PAus and PMali, and the genes encoding FolP and FolC are absent in the sequenced genomes other than that of strain OYM. Even before multiple whole genome sequence data became available for comparisons, our previous work revealed that genes encoding FolP and FolK were under ongoing lineage-specific decay, and we predicted that lineage-related loss of recognizable *folP* and *folK* homologs in phytoplasma genomes would eventually occur (Davis et al. 2003, 2005). The completely sequenced phytoplasma genomes now provide further evidence that phytoplasmas are experiencing an ongoing evolutionary process of losing the ability to synthesize folate and consequently must rely on their hosts for folate repletion.

In view of this observation, the consistent presence of genes encoding apparently intact FolA proteins must be explained. The available genome data suggest that the *folA* gene is

universally present in all phytoplasma strains studied thus far. To envision its role in phytoplasmas, we note that folate is a precursor of tetrahydrofolate and that the latter is an important carrier of one-carbon groups vital for metabolism of purines and pyrimidines (therefore of nucleic acids) as well as amino acids (particularly serine, glutamic acid, histidine, and methionine). Through NADPH reduction of folate or dihydrofolate, FolA could conceivably transform imported folate and dihydrofolate, respectively, to tetrahydrofolate, the acceptor of one-carbon groups. Tetrahydrofolate could then accept a methyl group from the breakdown of glycine through the catalytic activity of glycine hydroxymethyltransferase (GlyA) and accept a methyl group from the conversion of dTMP to dUMP through the catalytic activity of thymidylate synthase (ThyA). In turn, the methylated tetrahydrofolate can serve as one-carbon donor, participating in nucleotide and amino acid metabolisms (Fig. 10.7). The genes that encode GlyA and ThyA are present and intact in all four sequenced phytoplasma genomes. Therefore, we hypothesize that, rather than playing a role in folate biosynthesis, FolA is dedicated to one-carbon transfer in phytoplasmas. As in the case of phytoplasmas, the genomes of *Ureaplasma parvum*, *Mycoplasma crocodyli*, *M. gallisepticum*, and *Mycoplasma synoviae* also lack the folate biosynthesis genes *folK*, *folC*, or *folP*, but each of these mycoplasmas possesses a *folA* gene; thus, FolA may perform the same function in some human- and animal-infecting mycoplasmas as we propose in phytoplasmas.

10.6.2 Glycolysis

Glycolysis consists of two main phases: The first phase is the energy investment phase that converts glucose to fructose-1,6-bisphosphate and then to two 3-carbon units (glyceraldehyde-3-phosphate); the second phase is the energy harvest phase that converts glyceraldehyde-3-phosphate to pyruvate (Fig. 10.6). The first phase consumes two moles of ATP for each mol of glucose, and the second phase gains four mols of ATP and two mols of NADH from each initial mol of glucose. While the genomes of OYM, AYWB, and PAus possess the genes responsible for both phases of glycolysis, PMali lacks five genes encoding the enzymes required for the second phase, i.e., glyceraldehyde-3-phosphate dehydrogenase (GapA), phosphoglycerate kinase (Pgk), phosphoglycerate mutase (GpmI), enolase (Eno), and pyruvate kinase (pykF). The absence of these genes implies that PMali is unable to use glycolysis for ATP production (Kube et al. 2008).

10.6.3 Pentose Phosphate Pathway

OYM, AWYB, and PAus genomes lack genes encoding enzymes of the Entner–Doudoroff pathway (or pentose phosphate pathway), except for those genes that are shared with glycolysis (*pgi*, *pfkA*, and *fba*). However, the PMali genome possesses, in addition to *pgi*, *pfkA*, and *fba*, one of the key pentose phosphate pathway genes (*eda*) encoding 2-dehydro-3-deoxyphosphogluconatealdolase/4-hydroxy-2-oxoglutarate aldolase (Eda). Eda could possibly serve two functions, i.e., to catalyze the conversion of 4-hydroxy-2-oxoglutarate to pyruvate and glyoxylate and to catalyze the conversion of 6-phospho-2-dehydro-3-deoxy-D-gluconate to pyruvate and glyceraldehyde-3-phosphate (Kube et al. 2012). While both functions could produce pyruvate independent of glycolysis, it seems doubtful that such conversions would play a significant role in an alternative energy-yielding process in PMali. On the other hand, Eda may play a role in the phytoplasma's amino acid metabolism, since absence of *eda* impacts arginine and proline metabolism. A recent study revealed that, in *Vibrio cholera*, Eda activity may contribute to pathogenicity: Selective activation of Entner–Doudoroff pathway genes including *eda* leads to concurrent increase in transcripts of prime virulence genes (*ctxA* and *tcpA*) and their regulator (*toxT*) (Patra et al. 2012).

10.6.4 Pyrimidine Metabolism

Cytidine/uridine kinase (Udk): Udk participates in pyrimidine metabolism and phosphorylates both uridine and cytidine, using ATP, GTP, or dGTP as phosphate donor. The gene encoding Udk is present in the genomes of OYM and AYWB but is missing in PAus and PMali. As in the case of other mollicutes (Bizarro and Schuck 2007), phytoplasmas lack a gene set essential for de novo synthesis of pyrimidine bases. As outlined by Kube et al. (2012), phytoplasmas may use imported uridine and cytidine as the entry point for the synthesis of UTP/dTTP and CTP/dCTP, respectively, building blocks for DNA and RNA. Since Udk is required for converting uridine to UMP, and cytidine to CMP, the lack of Udk in PAus and PMali indicates that PAus and PMali may have to import UMP and CMP from host cells. While CMP may also be available as a by-product from phospholipid metabolism (Kube et al. 2012), the limited source seems unlikely to be sufficient to support the needs of DNA and RNA biosynthesis.

10.6.5 Amino Acid Metabolism

Phytoplasmas lack most of the genes encoding enzymes required for amino acid biosynthesis and metabolism, while a few of the genes are present in a lineage-specific manner.

Asparagine synthase (AsnB): AsnB catalyzes the ATP-dependent conversion of aspartic acid to asparagine, using either glutamine or ammonia as the nitrogen source. Among the four completely sequenced phytoplasma genomes, only OYM has a full-length *asnB* gene (PAM594). In AYWB, the *asnB* gene (AYWB_130) has decayed substantially, and only a partial C-terminal AsnB domain remains recognizable. PAus and PMali lack the gene entirely. It is worth noting that mycoplasmas also lack the *asnB* gene and, consequently, need to import both aspartic acid and asparagine among other amino acids from their environment or host cells (Razin et al. 1998). Presence of a full-length *asnB* gene in the genome of OYM indicates that this phytoplasma may be able to synthesize asparagine from aspartic acid and an available nitrogen source.

S-adenosylmethionine synthetase (MetK) and C-5 cytosine-specific DNA methylase (C5 Mtase): While MetK catalyzes the formation of S-adenosylmethionine (AdoMet) from methionine and ATP, C5 Mtase methylates the C-5 carbon of cytosines in DNA, using AdoMet as the methyl donor. Although both enzymes are considered key components of cysteine and methionine metabolism, they impact multiple metabolic pathways. The product of the MetK reaction, AdoMet, is the primary alkylating agent in various organisms (Lu and Markham 2002). With a myriad of AdoMet-dependent methyltransferases, the methyl group of AdoMet can be transferred to a multitude of biological targets in the cell (Klimasauskas and Weinhold 2007). DNA methylation mediated by C5 Mtase is one of the key epigenetic mechanisms controlling gene expression, as DNA methylation can switch off genes. Among the four sequenced phytoplasma genomes, the *metK* gene is present only in the genome of PMali and the gene encoding C5 Mtase is present only in the genome of PAus.

Acetylornithine deacetylase (ArgE): ArgE is a zinc-dependent hydrolase that catalyzes the deacylation of N^2-acetyl-L-ornithine, yielding ornithine and acetate. In *E. coli*, ornithine is an obligatory intermediate in the arginine biosynthetic pathway and a branch point in the synthesis of polyamines (Javid-Majd and Blanchard 2000). The *argE* gene is present in OYM and AYWB, but absent in PAus and PMali. It is doubtful that *argE* actually participates in arginine biosynthesis in OYM and AYWB, since none of the three genes (*argF/I*, *argG*, and *argH*), encoding enzymes necessary for the downstream conversion of ornithine to arginine (Caldara et al. 2008), is present in any of the four phytoplasma genomes.

10.6.6 Protein Turnover

Leucyl aminopeptidase (PepA): The gene encoding a putative leucyl aminopeptidase is present in the genomes of PAus and PMali, but absent in the genomes of OYM and AYWB. By preferentially catalyzing the hydrolysis of leucine residues at the N-terminus of substrate proteins, leucyl aminopeptidases play an important role in protein turnover and cellular stress responses in diverse organisms (Boulila-Zoghlami et al. 2011; Cappiello et al. 2004). It has long been recognized that peptidase activities of various pathogens are correlated with their virulence; peptidases thus are potential molecular targets for therapeutic drugs (Pinna et al. 1991; Carroll et al. 2012). It would be interesting to learn what functional role PepA may have in PAus and PMali, and whether the gene or gene product can be targeted for mitigating grapevine yellows and apple proliferation diseases.

10.6.7 Sugar Metabolism

Sucrose phosphorylase (glucosyltransferase-A, GtfA): As discussed in a previous section (10.5.2), sucrose phosphorylase converts sucrose to D-fructose and α-D-glucose-1-phosphate and thus may provide a means for some phytoplasmas to obtain sugar phosphate. The *gtfA* gene is present in OYM and PAus, but absent in AYWB and PMali.

10.6.8 Glycerolipid and Glycerophospholipid Metabolism

1,2-Diacylglycerol 3-glucosyltransferase (RfaG): RfaG is a cell surface membrane-associated enzyme that participates in glycerolipid synthesis. In *Bacillus subtilis*, RfaG is able to successively transfer up to four glucose residues to 1,2-diacylglycerol (1,2-DAG, Jorasch et al. 1998). In *A. laidlawii*, RfaG catalyzes the synthesis of the major non-bilayer-prone lipid alpha-monoglucosyl DAG from 1,2-DAG (Li et al. 2003). Since RfaG plays a major role in determining the core structure of lipopolysaccharides and therefore cell surface properties, it would not be surprising if RfaG plays an important role in phytoplasma–host interactions. The *rfaG* gene is present in OYM and PMali, but not in the AYWB and PAus genomes.

Dimethyladenosine transferase (KsgA): Genes encoding phytoplasma-unique KsgA proteins are present in the genomes of OYM, AYWB, and PMali, but absent in the PAus genome. The possible role of this gene in methyl group transfer as well as in methionine/cysteine interconversion was discussed in a previous section (10.5.2) of this chapter.

10.6.9 Other Lineage-Specific Genes

A gene encoding riboflavin kinase (RibF) is present in PAus and is absent from the other three sequenced phytoplasma genomes. As in other prokaryotes, the protein encoded by the PAus *ribF* gene has two functional domains, an N-terminus FMN adenylyltransferase domain and a C-terminus riboflavin kinase domain. The riboflavin kinase domain catalyzes the phosphorylation of riboflavin, giving rise to FM; and the FMN adenylyltransferase domain catalyzes the adenylylation of FMN, generating FAD (Frago et al. 2008). FMN is a strong oxidizing agent; it can take part in both one- and two-electron transfers, while itself being reduced to FMNH and $FMNH_2$, respectively. Similarly, FAD can exist in two different redox states, FAD and $FADH_2$, as a result of accepting or donating electrons. Both FMN and FAD are important cofactors in numerous redox reactions. As phytoplasmas lack a canonical NH_3/glutamine-dependent NAD synthetase and other enzymes required for NAD^+ biosynthesis (see Sect. 10.7.5), we wonder whether the presence of the bi-functional riboflavin kinase gene in the PAus genome may provide alternative electron carriers for at least some of the redox reactions in '*Ca*. Phytoplasma australiense.'

A gene encoding pyridoxal kinase (PdxK), an enzyme involved in pyridoxal phosphate synthesis, is present in the OYM genome, but not in the other three completely sequenced genomes. Commonly known as vitamin B6, the product of the pyridoxal kinase reaction, pyridoxal 5′-phosphate, is a cofactor for various enzymes involved in transamination, decarboxylation, deamination, and racemization reactions of amino acids. Since phytoplasmas lack most of the genes encoding enzymes responsible for amino acid biosynthesis and metabolism, the demand for pyridoxal cofactor might be low; consequently, the pyridoxal kinase gene may have become dispensable in phytoplasmas. If it is true, gradual decay and eventual loss of the *pdxK* gene in OYM may occur as predicated for other lineage-specific gene loss in phytoplasmas (Davis et al. 2005).

A gene annotated as encoding putative endo-1,4-beta-glucanase (YsdC or FrvX) is present in the genomes of OYM and AYWB, but absent in the genomes of PAus and PMali. Endo-1,4-beta-glucanase, also termed cellulase, catalyzes endohydrolysis of 1-4-β-D-glucosidic linkages in cellulose, lichenin, and cereal β-D-glucans, yielding oligosaccharides. In many plant pathogenic fungi and bacteria, endo-1,4-beta-glucanase is a secreted protein that acts on host components and contributes to virulence of the respective pathogens (Pérez-Donoso et al. 2010; Van Vu et al. 2012). It is worth noting that, unlike the above-mentioned fungal and bacterial endo-1,4-beta-glucanases, the putative OYM and AYWB endo-1,4-beta-glucanases do not have an N-terminal signal peptide nor a transmembrane domain, indicating that the enzymes are most likely to stay in the cytosol of phytoplasmal cells. Such an intracellular location of the phytoplasmal endo-1,4-beta-glucanases casts doubt on their direct involvement in pathogen–host interactions. Based on the structural motifs in their amino acid sequences, the putative OYM and AYWB endo-1,4-beta-glucanases are members of the M42 family of glucanase/peptidase; therefore, they could also have peptidase activities. Functional analyses will be required to determine the actual catalytic activities and biological functions of the phytoplasmal glucanases.

Cobalt, a transition metal, is an essential nutrient for bacteria; it plays crucial roles in cobalt-dependent enzymes such as methionine aminopeptidase and certain metallo-beta-lactamases (Kobayashi and Shimizu 1999). Since in natural environments, soluble cobalt exists only in trace amounts, the synthesis of metalloenzymes requires high-affinity uptake of the metal ions. Structural and functional genomic analyses have revealed that there exist different groups of ABC-type cobalt transporters (Rodionov et al. 2006). While all four completely sequenced phytoplasma genomes harbor genes encoding an ABC-type cobalt transport system (including a permease subunit and an ATP-binding subunit), the PAus and AYWB genomes encode an additional permease component (PA0139 and AYWB_015, respectively) that is absent in the OYM and PMali genomes and is unmatched by permease components of all other mollicutes studied thus far. The genome of *Spiroplasma kunkelii*, another phloem-inhabiting cell wall-less bacterium, also possesses genes encoding an ABC-type cobalt transporter system (Zhao et al. 2004). Interestingly, neither the *S. kunkelii* nor the phytoplasmal cobalt transporters contain an extracytoplasmic solute-binding protein.

10.7 Missing Genes Essential to a Minimal Genome of a Free-Living Bacterium

Over more than four decades, numerous efforts have been devoted to isolation of phytoplasmas in axenic culture. Several papers have claimed success, but attempts at independent confirmation have not substantiated the claims, and in some cases, the data provided were unconvincing and/or critical information was withheld (Contaldo et al. 2012; Ghosh et al. 1971; Hampton et al. 1969; Lin et al. 1970; Lombardo and Pignattelli 1970). Without doubt, cultivation of phytoplasmas in cell-free medium remains a formidable challenge. The availability of complete genome sequence data from four phytoplasma

strains promises new insights and raises renewed hope that it will become possible to formulate phytoplasma culture media through a genomics-guided approach. We envision that the fastidious growth requirements of phytoplasmas demand not only proper nutrient composition but also a well-defined physical environment, including a favorably poised redox potential, and congenial pH and osmotic pressure conditions. We believe that the conditions required for phytoplasma survival/viability and growth/multiplication are governed by metabolic capability, cross-membrane trafficking activity, and other biochemical capacities, all of which may be inferred from information accessible in the phytoplasma genome sequence data.

To decipher such information, we formulated and executed queries to identify essential genes that are missing in phytoplasmas compared with those present in *M. genitalium*, a model wall-less bacterium thought to have the minimum gene complement among known cellular organism (Fraser et al. 1995). *M. genitalium* is culturable in vitro. Our differential analysis identified 100 protein-encoding genes that are present in *M. genitalium,* but are absent in phytoplasmas. Since previous global transposon mutagenesis studies demonstrated that, of the total 482 protein-encoding genes in the in *M. genitalium* genome, 100 were non-essential (Glass et al. 2006), we eliminated 37 non-essential genes from the phytoplasma 'missing gene' list that were identified from our initial differential analysis. The resulting 63 genes are considered as essential genes that are missing in phytoplasma genomes (Table 10.2) and, therefore, may be implicated in phytoplasmas' losing capabilities that would otherwise make possible their axenic cultivation.

10.7.1 Carbohydrate Intake and Metabolism, and Energy Generation

Phytoplasma genome data revealed that phytoplasmas do not possess any genes encoding enzymes involved in aerobic metabolism of carbohydrates, indicating that glycolysis is the main route for sugar metabolism and energy generation. Phytoplasmas lack genes encoding subunits of the F_1F_0-ATP synthase complex (*atpA*, *atpB*, *atpC*, *atpD*,*atpE*, *atpF*, *atpG*, and *atpH*) (Oshima et al. 2004). In other bacteria, the F_1F_0-ATP synthase complex catalyzes ATP synthesis from ADP and phosphate driven by a cross-membrane proton gradient. The F_1F_0-ATP synthase complex can also hydrolyze ATP and generate a proton electrochemical gradient in the opposite direction for locomotion, nutrient uptake, and other functions. Prior to the completion of the first phytoplasma genome sequencing, the F_1F_0-ATP synthase complex had been found in the genomes of all previously sequenced bacteria, including mollicutes; therefore, it was believed indispensable to cellular life. Lack, in phytoplasmas, of the genes encoding components of ATP synthase complex implied that alternative ATP synthesis and energy-coupling mechanisms must exist in phytoplasmas (Oshima et al. 2004). In addition, our analysis revealed that the gene (*nox*) encoding NADH oxidase (NADH dehydrogenase), which transfers electrons along a chain of acceptors and releases energy for ATP formation is also absent in phytoplasmas.

Furthermore, genes encoding components of PTS, the phosphoenolpyruvate-dependent sugar phosphotransferase system (PtsG, PtsH, and PtsI) are absent in the genomes of all four phytoplasmas. PTS is a distinct sugar uptake system essential to free-living bacteria. It utilizes energy stored in phosphoenolpyruvate (PEP), powering cross-membrane transport of glucose, mannose, fructose, cellobiose, and other sugar molecules. The lack of PTS in phytoplasmas indicates the presence of other, as yet unidentified, transmembrane system(s) responsible for transport of sugars into the metabolizing phytoplasma cells. Given that genes encoding ABC-type sugar transporter systems, for example *malEGFK* and *ugpBEAC*, are present in all four sequenced phytoplasmas (Kube et al. 2012), conceivably, some sugars could be imported through ABC-type sugar transporters. In addition, sugars could possibly be

transported into phytoplasma cells through cation symporter(s) or by other membrane channels in the form of complexes with other classes of compounds. Deciphering possible substrate specificities of the presumed phytoplasmal sugar transporters from their respective gene sequences would be an important step toward determining proper type and form of sugar(s) to be included in prospective culture media.

Selection of appropriate carbon source(s) is particularly crucial, since phytoplasmas also lack genes encoding various enzymes required for the pentose phosphate pathway (deoxyribose-phosphate aldolase and ribose-5-phosphate isomerase), for pentose and glucuronate interconversions (UTP–glucose-1-phosphate uridylyltransferase), for galactose metabolism (UDP-glucose 4-epimerase), for fructose and mannose metabolism (phosphomannomutase), and for pyruvate metabolism (phosphotransacetylase). Lack of these genes further indicates that phytoplasmas have a relatively limited resource for energy generation.

Since all four completely sequenced phytoplasma genomes contain genes that encode a malate-sodium symporter, malate could be an alternative carbon source for phytoplasmas. The genes that encode enzymes necessary for converting malate to pyruvate (phosphate acetyltransferase), and then pyruvate to acetyl CoA (pyruvate dehydrogenase multienzyme complex), are present in phytoplasmas. However, a gene encoding phosphotransacetylase, the enzyme that catalyzes the conversion of acetyl CoA to acetyl-phosphate, is absent in all known phytoplasmas (Table 10.2). Unless their genomes encode an alternative enzyme (for example phosphotransacetylase, as suggested by Kube et al. 2012), phytoplasmas may not be capable of utilizing the energy carried in acetyl CoA to produce ATP.

10.7.2 Nucleotide Metabolism

Among genes essential in the smallest known free-living bacterium *M. genitalium* and close relatives *Acholeplasma* spp., but absent in the completely sequenced phytoplasma genomes, are those encoding seven key enzymes involved in the metabolism of nucleotides (adenine phosphoribosyltransferase, hypoxanthine-guanine phosphoribosyltransferase, ribose-phosphate pyrophosphokinase, uracil phosphoribosyltransferase, purine-nucleoside phosphorylase, thioredoxin reductase, and ribonucleotide reductase stimulatory protein). The lack of genes encoding these enzymes negatively impacts pathways of both de novo and salvage biosynthesis of purine and pyrimidine nucleotides. Apparently, phytoplasmas must import certain nucleotides among other nutrients and preformed metabolic substrates, an observation encouraging inclusion of various nucleotides in media formulated for attempts at axenic culture of phytoplasmas.

10.7.3 Lipid Metabolism and Cell Membrane Genesis

As noted previously, each of the four completely sequenced phytoplasma genomes possesses a gene annotated as encoding CDP-diglyceride synthetase or phosphatidate cytidylyltransferase (CdsA) (PAM169, AYWB_550, PA0515, and ATP_00236) (Kube et al. 2012). CdsA is crucial to lipid biosynthesis: by catalyzing the activation of phosphatidic acid, CdsA reaction yields CDP-diglyceride, the phosphatidyl moiety donor for biosynthesis of various phospholipids (Sparrow and Raetz 1985). It is interesting to note that, in the present study, a CDD search using the amino acid sequences of the four presumed phytoplasmal CdsA as queries revealed that the phytoplasmal CdsA proteins contain only a partial CdsA domain at the C-terminus. Furthermore, none of the four presumed phytoplasmal CdsA proteins bears significant sequence homology to the CdsA protein of *M. genitalium* (MG437). Therefore, with the search criteria adopted in the current study, our differential analysis indicated that phytoplasmas 'lack' a *cdsA* gene that is essential to a free-living bacterium. Thus far, whether or not phytoplasmas have a functional

Table 10.2 Genes essential to the minimal gene complement of a free-living bacterium, *Mycoplasma genitalium*, but absent in the genomes of phytoplasmas[a]

CDS	Begin	End	Strand	Gene	Product
MG013	14,396	15,316	–	*folD*	Methylenetetrahydrofolate dehydrogenase/ cyclohydrolase (FolD)
MG028	30,547	31,152	+		Conserved hypothetical protein
MG030	31,704	32,324	–	*upp*	Uracil phosphoribosyltransferase (UPP)
MG038	44,751	46,277	–	*glpK*	Glycerol kinase (GlpK)
MG041	49,377	49,643	+	*ptsH*	Phosphocarrier protein HPr (PtsH), a component of the PTS
MG049	58,117	59,079	+	*deoD*	Purine-nucleoside phosphorylase (DeoD)
MG050	59,083	59,754	+		Deoxyribose-phosphate aldolase
MG053	61,407	63,059	+	*manB*	Phosphomannomutase (ManB)
MG057	65,713	66,249	–		Conserved hypothetical protein
MG058	66,228	67,121	–	*prs*	Ribose-phosphate pyrophosphokinase (PRS)
MG069	88,228	90,954	+	*ptsG*	PTS system, glucose-specific IIABC component (PtsG)
MG086	112,718	113,866	+	*lgt*	Prolipoprotein diacylglyceryl transferase (LGT)
MG099	125,852	127,285	+	*gatA*	Aspartyl/glutamyl-tRNA amidotransferase subunit A
MG100	127,278	128,711	+	*gatB*	Aspartyl/glutamyl-tRNA amidotransferase subunit B
MG102	129,347	130,294	+	*trxB*	NADPH-dependent thioredoxin reductase (TrxB)
MG106	134,146	134,826	–	*def*	Polypeptide deformylase (DEF)
MG118	143,935	144,957	+	*galE*	UDP-glucose 4-epimerase (GalE)
MG120	146,673	148,235	+		Conserved hypothetical protein
MG128	155,443	156,222	+	*ppnK*	Inorganic polyphosphate/ATP-NAD kinase (conserved hypothetical protein)
MG135	160,071	160,913	–		Conserved hypothetical protein (putative transmembrane protein)
MG221	266,626	267,090	+		Conserved hypothetical protein (MraZ cell division protein)
MG224	269,249	270,358	+	*ftsZ*	Cell division protein (FtsZ)
MG230	276,166	27,627	+	*nrdI*	Ribonucleotide reductase stimulatory protein, NADPH-dependent FMN reductase)
MG241	286,884	288,746	+		Conserved hypothetical protein (with multiple transmemberance domians)
MG242	288,752	290,644	+		Conserved hypothetical protein (with multiple transmemberance domians)
MG259	309,008	310,378	+		Conserved hypothetical protein
MG265	321,576	322,412	–		Conserved hypothetical protein
MG267	324,810	325,157	–		Conserved hypothetical protein (with multiple transmemberance domians)
MG276	334,855	335,397	–		Adenine phosphoribosyltransferase
MG299	368,732	369,694	–	*pta*	Phosphotransacetylase (PTA)
MG306	376,793	377,974	–		Conserved hypothetical protein
MG319	398,297	398,833	–		Conserved hypothetical protein (Multi-pass membrane protein)
MG322	403,722	405,398	–		Cation transporter, putative
MG323	405,455	406,138	+		Conserved hypothetical protein
MG332	415,610	416,329	–		Conserved hypothetical protein
MG342	438,730	439,236	–		Conserved hypothetical protein

(continued)

Table 10.2 (continued)

CDS	Begin	End	Strand	Gene	Product
MG347	444,413	445,045	+		Conserved hypothetical protein
MG356	453,911	454,753	–		Conserved hypothetical protein
MG364	461,015	461,689	+		Conserved hypothetical protein
MG365	461,682	462,617	+	*fmt*	Methionyl-tRNA formyltransferase (FMT)
MG374	472,067	472,891	–		Conserved hypothetical protein
MG383	481,329	482,075	–		NH(3)-dependent NAD + synthetase, putative
MG384.1	483,369	483,818	–		Conserved hypothetical protein
MG389	491,147	491,530	–		Conserved hypothetical protein (with single transmembrane domain)
MG396	500,261	500,719	–		Ribose-5-phosphate isomerase, putative
MG398	502,422	502,823	–	*atpC*	ATP synthase F1, subunit epsilon (AtpC)
MG399	502,828	504,258	–	*atpD*	ATP synthase F1, subunit beta (AtpD)
MG400	504,260	505,099	–	*atpG*	ATP synthase F1, subunit gamma (AtpG)
MG401	505,099	506,655	–	*atpA*	ATP synthase F1, subunit alpha (AtpA)
MG402	506,671	507,201	–	*atpH*	ATP synthase F1, subunit delta (AtpH)
MG403	507,194	507,820	–	*atpF*	ATP synthase subunit B (AtpF)
MG404	507,823	508,131	–	*atpE*	ATP synthase subunit C (AtpE)
MG405	508,134	509,012	–	*atpB*	ATP synthase subunit A (AtpB)
MG408	510,903	511,376	+	*pmsR*	Peptide methionine sulfoxide reductase (PmsR)
MG427	533,270	533,695	+		Conserved hypothetical protein (Predicted redox protein)
MG429	534,318	536,036	–	*ptsI*	Phosphoenolpyruvate-protein phosphotransferase (PtsI)
MG432	538,350	539,546	–		Conserved hypothetical protein
MG437	541,860	542,984	+	*cdsA*	Putative phosphatidate cytidylyltransferase, CDP-diglyceride synthetase)
MG448	552,445	552,897	–	*msrB*	Conserved hypothetical protein
MG453	556,435	557,313	+	gtaB	UTP–glucose-1-phosphate uridylyltransferase (GtaB)
MG456	558,937	559,941	–		Conserved hypothetical protein (with multiple transmembrane domains)
MG458	562,780	563,307	–	hpt	Hypoxanthine-guanine phosphoribosyltransferase (Hpt)
MG464	568,397	569,554	–		Conserved hypothetical protein (Multi-pass membrane protein)

[a] Gene names and their products are based on previously published genome annotations (Fraser et al. 1995)

cdsA gene remains an open question. If they do not, phytoplasmas will need to import from host cells activated phosphatidic acid, i.e., CDP-diglyceride, for phospholipid biosynthesis.

Phytoplasmas also lack a gene encoding glycerol kinase (GlpK) that provides glycerol-3-phosphate for phospholipid synthesis. Without GlpK, phytoplasma cells will need either to draw glycerol-3-phoshphate from glycolysis or to import it from host cells.

Also missing in phytoplasma genomes is the gene encoding prolipoprotein diacylglyceryl transferase (LGT). LGTs are integral membrane proteins that catalyze transfer of the n-acyl diglyceride group onto proteins that then become anchored in the phytoplasma membrane. In diverse bacteria, LGT is required for membrane genesis, for membrane transport activity, and for normal growth (Pailler et al. 2012; Chimalapati et al. 2012). Lack of LGT in

phytoplasma genomes will conceivably impact membrane genesis and homeostasis of phytoplasma cells.

10.7.4 One-Carbon Transfer Reactions

As noted above, one-carbon transfer reactions are important for nucleotide and amino acid metabolism, and therefore DNA synthesis, cell division, growth, and survival. In light of the universal presence of *folA*, *glyA*, and *thyA* genes and the absence or lineage-specific loss of other folate biosynthesis genes in phytoplasma genomes, we hypothesized that folate is likely to be imported into phytoplasma cells (Davis et al. 2003, 2005) and fed into one-carbon pools (Sect. 10.6.1 of this chapter), providing one-carbon units for nucleotide and amino acid metabolism. It appears that, compared with *M. genitalium* and *Acholeplasma* spp., phytoplasmas lack a full set of enzymes involved in one-carbon transfer reactions. Among the missing genes are those encoding methylenetetrahydrofolate dehydrogenase/methylenetetrahydrofolate cyclohydrolase (FolD) and methionyl-tRNA formyltransferase (Fmt). FolD is a bifunctional enzyme that catalyzes sequentially (and reversibly) the oxidation of N_5,N_{10}-methylenetetrahydrofolate to N_5,N_{10}-methenyltetrahydrofolate and the hydrolysis of N_5,N_{10}-methenyltetrahydrofolate to N_{10}-formyltetrahydrofolate (Shen et al. 1999; Eadsforth et al. 2012). These two functions permit N_5,N_{10}-methylenetetrahydrofolate to supply one-carbon units for the synthesis of carbons 2 and 8 of purines and the formyl group of *N*-formylmethionyl-tRNA$^{(f)}$. Lack of FolD would significantly limit the utilization of N_5,N_{10}-methylene-tetrahydrofolate, other than in synthesis of thymidylate and methionine. Fmt catalyzes the reversible addition of a formyl group to the free amino group of the aminoacyl moiety of methionyl-tRNA(fMet). Therefore, the absence of Fmt also would be expected to impact the formation of N_{10}-formyltetrahydrofolate, as well as impacting modification of methionyl-tRNA and in consequence affecting efficiency of protein synthesis, among other functions, if not compensated in some other way. These observations lead us to suggest that phytoplasmas likely need to import N_{10}-formyltetrahydrofolate from host cells and that this metabolite probably should be a component of culture media if axenic culture is attempted.

10.7.5 Phytoplasma Genomes and Redox Homeostasis

Most interestingly, among the genes missing from phytoplasma genomes are those encoding NH_3/glutamine-dependent NAD^+ synthetase (NADS), inorganic polyphosphate /ATP-NAD kinase (PPNK), and thioredoxin reductase (TrxB). These enzymes are essential for biosynthesis of nicotinamide adenine dinucleotide and nicotinamide adenine dinucleotide phosphate, and for redox regulation.

Nicotinamide adenine dinucleotide is a coenzyme that is found in all living cells heretofore studied for this feature. The coenzyme exists in two forms: the oxidized form NAD^+ and the reduced form NADH. NAD^+ can accept electrons from other molecules and become reduced; on the other hand, NADH can donate electrons to other molecules and become oxidized. Transferring electrons from one metabolite to another, the coenzyme participates in numerous metabolic redox reactions (Belenky et al. 2007; Pollak et al. 2007). Additionally, some NAD^+ is converted into the coenzyme nicotinamide adenine dinucleotide phosphate ($NADP^+$); the electron-transfer function of this latter coenzyme is similar to that of NAD^+. The ratios NAD^+/NADH and $NADP^+$/NADPH are vital factors in maintaining cellular redox homeostasis in diverse organisms, and the apparent absence of genes encoding enzymes required for their synthesis has serious implications.

NAD synthetase, catalyzing a two-step reaction that transforms deamido-NAD into NAD, is an essential enzyme involved in both the de novo biosynthesis and salvage of NAD^+ (Bi et al. 2011; Magni et al. 2009). Phytoplasma

genomes lack a gene encoding a canonical NH_3/glutamine-dependent NAD synthetase such as is present in all known bacteria with walls and in cell wall-less bacteria including mycoplasmas and spiroplasmas. Instead, each of the four completely sequenced phytoplasma genomes contains an ORF that has been annotated by others as NAD synthetase, because the C-terminus portion of the amino acid sequence deduced from the ORFs shared weak homology with that of canonical NH_3/glutamine-dependent NAD synthetase. Whether such annotated phytoplasmal NAD^+ synthetase gene actually encodes a *bona fide* NAD synthetase remains unknown, but if it does, phytoplasmas may nevertheless need to import deamido-NAD, substrate of the presumed NAD synthetase, since other enzymes required for NAD^+ biosynthesis, such as nicotinate nucleotide adenylyltransferase (NadD) and nicotinate phosphoribosyltransferase protein (PncB), are also absent in phytoplasma genomes.

In both *E. coli* and *B. subtilis*, ATP-NAD kinase is an allosteric enzyme (Garavaglia et al. 2003). The enzyme catalyzes the phosphorylation of NAD to NADP, using ATP and other nucleoside triphosphates as well as inorganic polyphosphate as source of phosphorus. ATP-NAD kinase activity appears tightly coupled to NADPH/$NADP^+$ and NADH/NAD^+ ratios in the cells and therefore plays an important role in the regulation of NADP turnover and the size of the NADP pool (Kawai and Murata 2008). Since phytoplasmas lack ATP-NAD kinase, in addition to affecting the overall redox homeostasis, all redox reactions that require NADPH or $NADP^+$ as the cofactor could be impacted by the lack of this kinase.

Thioredoxin is a ubiquitous redox protein supposedly present in all eukaryotic and prokaryotic cells (Buchanan et al. 2012). It also exists in two forms: the oxidized form (Trx-S2) and the reduced form [Trx(SH)2]. Trx-S2 contains a disulfide bridge that can be reduced by NADPH with thioredoxin reductase; Trx(SH)2 is a powerful hydrogen/electron donor that reduces the disulfide in various substrate proteins. In addition to its function in nucleotide metabolism and DNA synthesis, thioredoxin plays an important role in redox signaling, cell defense against oxidative stress, and maintenance of redox homeostasis (Meyer et al. 2009; Hanschmann et al. 2013). Each of the four completely sequenced phytoplasmas has a gene encoding thioredoxin (*trxA*); however, none of the phytoplasmas has a gene encoding thioredoxin reductase. Since both thioredoxin reductase and NADPH are required to maintain the reducing potential of thioredoxin, the lack of both a thioredoxin reductase gene and an ATP-NAD kinase gene (and therefore NADPH) in the genomes of phytoplasmas raises a question as to how oxidized thioredoxin is being reduced in phytoplasma cells, and ultimately, how redox homeostasis is achieved in phytoplasma cells. This question is particularly intriguing, because phytoplasma genomes also lack genes encoding glutaredoxin and glutathione reductase. Conceivably, oxidized thioredoxin molecules in phytoplasma cells would have to be exported to host cells, where they become reduced by host thioredoxin reductase and NADPH; the 'recharged' thioredoxin molecules would then be shuttled back into phytoplasma cells. As yet undetermined, specific ABC-type transporters or redox-sensitive channels of large conductance may provide phytoplasma cells with such a conduit for exchange of oxidized and reduced thioredoxins. These hypotheses suggest that a condition required for in vitro culture of phytoplasmas would be one that will allow extracellular reduction of phytoplasmal thioredoxin.

10.8 Potential Virulence Factors

Plants infected by phytoplasmas often exhibit symptoms such as general stunting, excessive shoot proliferation, witches' broom growth, rapid senescence, and abnormal floral development (Davis and Lee 1982, Lee et al. 2000). Such growth and developmental abnormalities are suggestive of profound perturbations in plant hormone balance. Phytoplasmal infection also causes impaired amino acid and carbohydrate translocation, and inhibited photosynthesis

(Lepka et al. 1999; Bertamini et al. 2002a, b; Curković-Perica et al. 2007). Studies have revealed that phytoplasma infection suppresses auxin signaling and biosynthesis (Hoshi et al. 2009). A recent study also revealed that potato purple top phytoplasma infection disrupts gibberellin homeostasis in host plants, possibly caused by desensitizing GA biosynthesis negative feedback regulation (Ding et al. 2013). Studies also showed that alterations in expression patterns of floral organ identity genes are associated with phytoplasma-induced floral deformations (Pracros et al. 2006; Himeno et al. 2011; Su et al. 2011; Cettul and Firrao 2011). Since mutually distinct phytoplasmas often induce similar symptoms in their host plants, distinctly different phytoplasmas may trigger indistinguishable host responses, a phenomenon consistent with the hypothesis of 'a common ancestral origin of phytoplasma pathogenicity genes and a limited repertoire of plant responses to phytoplasma pathogen signals' (Davis and Sinclair 1998).

As with other bacterial pathogens, to cause disease in plants, a phytoplasma must produce an array of effector molecules or virulence factors to evade host immune response, establish itself within the host, and enhance its potential to survival by altering the physiology of the host. Among bacterial effector molecules are surface-localized proteins that mediate bacterial attachment, and secreted small proteins that are translocated into host cells. Several research groups have devoted efforts to identification of phytoplasma genes encoding cell surface proteins and secreted proteins, and to functional characterization of such proteins, as noted below.

10.8.1 Tengu

Through transgenic technology, an OYM phytoplasma secretory protein was found to be able to move from phloem to other plant tissues and to induce phenotypes resembling disease symptoms (Hoshi et al. 2009). Termed as Tengu, this small OYM secretory protein has a molecular weight of 4.5 kDa and is encoded by a gene located outside of the SVM regions. When expressed in *Arabidopsis thaliana* plants via *Agrobacterium*-mediated transformation, Tengu induced witches' broom and dwarf phenotypes on transgenic lines. Immunohistochemical assays revealed that, in OYM phytoplasma-infected plant, while phytoplasma cells were restricted within the phloem sieve elements, Tengu was able to move to parenchyma cells and undifferentiated meristem cells. Tengu was also detected in the tip region of the stem and the branching region of axillary buds, as well as in the apical meristem. The expression level of the OYM phytoplasma *tengu* gene was higher in the plant host than in the insect host (Hoshi et al. 2009).

10.8.2 SAP11

A secreted protein of AYWB phytoplasma was reported to modulate plant developmental processes and induce morphological alterations in phenotypes of host plants (Bai et al. 2009). Designated as SAP11, this protein contains an eukaryotic bipartite nuclear localization signal and accumulates in nuclei of host cells (Bai et al. 2009). In the AYWB chromosome, the gene encoding SAP11 lies within a potential mobile unit (PMU)-like region (Bai et al. 2009), one of the chromosomal SVM regions with clustered prophage-derived sequences (Wei et al. 2008a). SAP11 interacts and destabilizes class II CINCINNATS (CIN)-related TCP transcription factors, leading to down-regulation of the host lipoxygenase 2 (LOX2) gene and reduced jasmonic acid production, therefore impairing plant defense responses (Sugio et al. 2011). Transgenic expression of SAP11 in *A. thaliana* induced a bushy phenotype with severely crinkled leaves. Insect vectors (*Macrosteles quadriline*) allowed to feed upon SAP11 transgenic *A. thaliana* lines tended to produce more progeny (Sugio et al. 2011), indicating that SAP11 may be beneficial to the vector and, conceivably, transmission of the phytoplasma, in addition to weakening the host plant defense system.

10.8.3 SAP54

Among secreted proteins encoded by the AYWB phytoplasma chromosome, one induced severe alteration in flower morphology (Maclean et al. 2011). This AYWB protein, designated as SAP54, is also encoded by a gene located in one of the SVM regions. Transgenic *Arabidopsis* lines expressing the *SAP54* gene exhibited leaf-like sepals and virescent petals, and the number of stamens was increased. Due to loss of floral meristem determinacy, in place of carpel, new flowers were produced in the central flower whorl. Such a phenotype suggests that SAP54 might have multiple molecular targets in the plant host (Maclean et al. 2011).

10.8.4 OY-AMP

An abundant, cell-surface-localized membrane protein, termed antigenic membrane protein (AMP), from the OYM phytoplasma was found to be a key player in phytoplasma–insect host interactions (Suzuki et al. 2006). AMP forms a complex in vitro with three insect proteins: actin, myosin heavy chain, and myosin light chain. The observation that AMP-microfilament complexes exist in all OYphytoplasma-transmitting leafhopper species, but not in the OY phytoplasma non-transmitting leafhoppers, prompted a hypothesis that the AMP-microfilament complex is correlated with the phytoplasma-transmitting capability of leafhoppers (Suzuki et al. 2006). It seems reasonable to hypothesize that the attachment of OY-AMP to host microfilament aids invasion of the host insect by the phytoplasma, and to speculate that the OY-Amp may constitute a novel class of bacterial effector protein.

10.8.5 Phospholipase

The HVR of a SVM in the chromosome of Malaysian periwinkle yellows (MPY) phytoplasma harbors a gene encoding a membrane-targeted, putatively secreted phospholipase (PL). Presence of the PL-encoding gene in the SVM's HVR is consistent with the hypothesis that the gene was acquired by MPY through horizontal transfer (Jomantiene et al. 2007). Both OYM and AYWB phytoplasmas, which are closely related to MPY phytoplasma, lack a putative PL-encoding sequence, reinforcing the hypothesis of horizontal acquisition of the MPY PL gene. Phospholipases are lipolytic enzymes that act upon phospholipids, one of the major constituents of eukaryotic cell membranes and that can be responsible for significant damage to host membranes in bacterial infections (Banerji et al. 2008; Istivan and Coloe 2006; Sitkiewitcz et al. 2006). Recently, work was initiated to test the hypothesis that the MPY PL protein is an active enzyme; following heterologous expression of the MPY PL gene in bacterial and yeast hosts, in vitro functional analyses demonstrated the expressed protein's lipolytic activity (Gedvilaite et al. 2014). Since some phytoplasmal disease symptoms could possibly be accounted for, at least in part, by damage and/or degradation of host cell membranes, it is possible that the MPY phytoplasma PL acts as a pathogenicity factor in the plant, and/or insect, host.

10.9 Concluding Remarks

Much has been learned from analyses of available phytoplasma genome sequence data, but these advances have opened only a small window on the genomic diversity likely to be found among extant phytoplasmas. Lineage-specific adaptation to a vastly broad range of plant host–insect vector ecosystem niches worldwide undoubtedly gave rise to many more phytoplasma species than currently known. To date, the genomes of only four phytoplasmas, classified in just three major groups, have been completely sequenced, whereas 32 groups and 36 '*Candidatus* Phytoplasma' species have already been delineated. While gene loss and horizontal acquisition constitute the major driving forces behind the evolution of phytoplasma genomes,

rapid changes in vertically inherited genes also play an important role in radiation of phytoplasma lineages. The rapidly evolving nature of vertically inherited phytoplasmal genes is evident and deserves attention in future studies. These genome-based observations are predictive of significant divergent evolution among phytoplasmas.

Yet, a common genomic thread unites all phytoplasmas. Evidence now indicates that, through the formation of phage-based genomic islands, SVMs, the phytoplasma progenitor acquired new capabilities. Repeated and targeted chromosomal integration of phage genomes, and further gene acquisition through targeted insertion of mobile gene cassette-like elements, shaped the phytoplasma genome. Fusion of progenitor genome with phage genome(s) thus provided platforms for horizontal gene transfers enabling transkingdom parasitism and pathogenicity. Remnants of these events remain in all studied phytoplasma genomes. While analysis of the four completely sequenced phytoplasma genomes reveals footprints of other events that also contributed to genome evolution and the radiation of species, it is apparent that evolutionary emergence of the phytoplasma clade is in large part attributable to a singularly critical event, genome fusion of progenitor and phage.

Note added in proof:

While this chapter was in press, the following paper appeared: Andersen et al. 2013. BMC Genomics 14, 529. doi:10.1186/1471-2164-14-529. The information in that paper does not alter the conclusions expressed in this chapter.

References

Abreu IA, Cabelli DE (2010) Superoxide dismutases-a review of the metal-associated mechanistic variations. Biochim Biophys Acta 1804:263–274

Al-Saady NA, Khan AJ, Calari A, Al-Subhi AM, Bertaccini A (2008) '*Candidatus* Phytoplasma omanense', associated with witches'-broom of *Cassia italica* (Mill.) Spreng. in Oman. Int J Syst Evol Microbiol 58:461–466

Arashida R, Kakizawa S, Hoshi A, Ishii Y, Jung HY, Kagiwada S, Yamaji Y, Oshima K, Namba S (2008a) Heterogeneic dynamics of the structures of multiple gene clusters in two pathogenetically different lines originating from the same phytoplasma. DNA Cell Biol 27:209–217

Arashida R, Kakizawa S, Ishii Y, Hoshi A, Jung HY, Kagiwada S, Yamaji Y, Oshima K, Namba S (2008b) Cloning and characterization of the antigenic membrane protein (Amp) gene and in situ detection of Amp from malformed flowers infected with Japanese hydrangea phyllody phytoplasma. Phytopathology 98:769–775

Aurass P, Schlegel M, Metwally O, Harding CR, Schroeder GN, Frankel G, Flieger A (2013) The Legionella pneumophila Dot/Icm-secreted effector PlcC/CegC1 together with PlcA and PlcB promotes virulence and belongs to a novel zinc metallophospholipase C family present in bacteria and fungi. J Biol Chem (Epub ahead of print)

Bagadia PG, Polashock J, Bottner-Parker KD, Zhao Y, Davis RE, Lee I-M (2013) Characterization and molecular differentiation of 16SrI-E and 16SrIX-E phytoplasmas associated with blueberry stunt disease in New Jersey. Mol Cell Probes 27:90–97

Bai X, Zhang J, Ewing A, Miller SA, Jancso Radek A, Shevchenko DV, Tsukerman K, Walunas T, Lapidus A, Campbell JW, Hogenhout SA (2006) Living with genome instability: the adaptation of phytoplasmas to diverse environments of their insect and plant hosts. J Bacteriol 188:3682–3696

Bai X, Correa VR, Toruño TY, Ammar el D, Kamoun S, Hogenhout SA (2009) AY-WB phytoplasma secretes a protein that targets plant cell nuclei. Mol Plant Microbe Interact 22:18–30

Banerji S, Aurass P, Flieger A (2008) The manifold phospholipases A of *Legionella pneumophila*—identification, export, regulation, and their link to bacterial virulence. Int J Medical Microbiol 298:169–181

Barbara DJ, Morton A, Clark MF, Davies DL (2002) Immunodominant membrane proteins from two phytoplasmas in the aster yellows clade (chlorante aster yellows and clover phyllody) are highly divergent in the major hydrophilic region. Microbiology 148:157–167

Barré A, de Daruvar A, Blanchard A (2004) MolliGen, a database dedicated to the comparative genomics of Mollicutes. Nucleic Acids Res 32:D307–D310

Belenky P, Bogan KL, Brenner C (2007) NAD^+ metabolism in health and disease. Trends Biochem Sci 32(1):12–19

Bertamini M, Grando MS, Muthuchelian K, Nedunchezhian N (2002a) Effect of phytoplasmal infection on photosystem II efficiency and thylakoid membrane protein changes in field grown apple *(Malus pumila)* leaves. Physiol Mol Plant Pathol 61:349–356

Bertamini M, Nedunchezhian N, Tomasi F, Grando MS (2002b) Phytoplasma [Stolbur-subgroup (Bois Noir-BN)] infection inhibits photosynthetic pigments, ribulose-1,5-bisphosphate carboxylase and photosynthetic activities in field grown grapevine (*Vitis*

vinifera L. cv. Chardonnay) leaves. Physiol Mol Plant Pathol 61:357–366

Bessman MJ, Frick DN, O'Handley SF (1996) The MutT proteins or "Nudix" hydrolases, a family of versatile, widely distributed, "housecleaning" enzymes. J Biol Chem 271:25059–25062

Bi J, Wang H, Xie J (2011) Comparative genomics of NAD(P) biosynthesis and novel antibiotic drug targets. J Cell Physiol 226:331–340

Bizarro CV, Schuck DC (2007) Purine and pyrimidine nucleotide metabolism in Mollicutes. Genet Mol Biol 30(1):190–201

Boulila-Zoghlami L, Gallusci P, Holzer FM, Basset GJ, Djebali W, Chaïbi W, Walling LL, Brouquisse R (2011) Up-regulation of leucine aminopeptidase-A in cadmium-treated tomato roots. Planta 234:857–863

Brüssow H, Canchaya C, Hardt WD (2004) Phages and the evolution of bacterial pathogens: from genomic rearrangements to lysogenic conversion. Microbiol Mol Biol Rev 68:560–602

Buchanan BB, Holmgren A, Jacquot JP, Scheibe R (2012) Fifty years in the thioredoxin field and a bountiful harvest. Biochim Biophys Acta 1820:1822–1829

Bukata L, Altabe S, de Mendoza D, Ugalde RA, Comerci DJ (2008) Phosphatidylethanolamine synthesis is required for optimal virulence of *Brucella abortus*. J Bacteriol 190:8197–8203

Cai H, Wei W, Davis RE, Chen H, Zhao Y (2008) Genetic diversity among phytoplasmas infecting *Opuntia* species: virtual RFLP analysis identifies new subgroups in the peanut witches'-broom phytoplasma group. Int J Syst Evol Microbiol 58:1448–1457

Caldara M, Dupont G, Leroy F, Goldbeter A, De Vuyst L, Cunin R (2008) Arginine biosynthesis in *Escherichia coli*: experimental perturbation and mathematical modeling. J Biol Chem 283:6347–6358

Cappiello M, Lazzarotti A, Buono F, Scaloni A, D'Ambrosio C, Amodeo P, Méndez BL, Pelosi P, Del Corso A, Mura U (2004) New role for leucyl aminopeptidase in glutathione turnover. Biochem J 378:35–44

Carraro L, Ermacora P, Loi N, Osler R (2004) The recovery phenomenon in apple proliferation-infected apple trees. J Plant Pathol 86:141–146

Carroll RK, Robison TM, Rivera FE, Davenport JE, Jonsson IM, Florczyk D, Tarkowski A, Potempa J, Koziel J, Shaw LN (2012) Identification of an intracellular M17 family leucine aminopeptidase that is required for virulence in *Staphylococcus aureus*. Microbes Infect 14:989–999

Cenens W, Mebrhatu MT, Makumi A, Ceyssens PJ, Lavigne R, Van Houdt R, Taddei F, Aertsen A (2013) Expression of a novel P22 ORFan gene reveals the phage carrier state in *Salmonella typhimurium*. PLoS Genet 9(2):e1003269

Cettul E, Firrao G (2011) Development of phytoplasma-induced flower symptoms in *Arabidopsis thaliana*. Physiol Mol Plant Pathol 76:204–211

Chen YL, Montedonico AE, Kauffman S, Dunlap JR, Menn FM, Reynolds TB (2010) Phosphatidylserine synthase and phosphatidylserine decarboxylase are essential for cell wall integrity and virulence in *Candida albicans*. Mol Microbiol 75:1112–1132

Chimalapati S, Cohen JM, Camberlein E, MacDonald N, Durmort C, Vernet T, Hermans PW, Mitchell T, Brown JS (2012) Effects of deletion of the *Streptococcus pneumoniae* lipoprotein diacylglyceryl transferase gene lgt on ABC transporter function and on growth in vivo. PLoS ONE 7(7):e41393

Clarke S, Banfield K. (2001) S-adenosylmethionine-dependent methyltransferases. In: Carmel R, Jacobsen DW (eds) Homocysteine in health and disease. New York, Cambridge University Press, pp 63–79

Conde-Alvarez R, Grilló MJ, Salcedo SP, de Miguel MJ, Fugier E, Gorvel JP, Moriyón I, Iriarte M (2006) Synthesis of phosphatidylcholine, a typical eukaryotic phospholipid, is necessary for full virulence of the intracellular bacterial parasite *Brucella abortus*. Cell Microbiol 8:1322–1335

Conrady DG, Wilson JJ, Herr AB (2013) Structural basis for Zn2+-dependent intercellular adhesion in staphylococcal biofilms. Proc Natl Acad Sci USA 110:E202–E211

Contaldo N, Bertaccini A, Paltrinieri S, Windsor HM, Windsor GD (2012) Axenic culture of plant pathogenic phytoplasmas. Phytopathologia Mediterranea 51:607–617

Cumby N, Edwards AM, Davidson AR, Maxwell KL (2012) The bacteriophage HK97 gp15 moron element encodes a novel superinfection exclusion protein. J Bacteriol 194:5012–5019

Curković-Perica M, Lepedus H, Seruga-Musić M (2007) Effect of indole-3-butyric acid on phytoplasmas in infected *Catharanthus roseus* shoots grown *in vitro*. FEMS Microbiol Lett 268:171–177

Darling AC, Mau B, Blattner FR, Perna NT (2004) Mauve: multiple alignment of conserved genomic sequence with rearrangements. Genome Res 14:1394–1403

Davis RE, Lee I-M (1982) Pathogenicity of spiroplasmas, mycoplasmalike organisms, and vascular-limited fastidious walled bacteria. In: Mount M, Lacy G (eds) Phytopathogenic Prokaryotes, vol 1. Academic Press, New York, pp 491–513

Davis RE, Sinclair WA (1998) Phytoplasma identity and disease etiology. Phytopathology 88:1372–1376

Davis RE, Worley JF (1973) Spiroplasma: motile, helical microorganism associated with corn stunt disease. Phytopathology 63:403–408

Davis RE, Worley JF, Whitcomb RF, Ishijima R, Steere RL (1972) Helical filaments produced by a mycoplasmalike organism associated with corn stunt disease. Science 176:521–523

Davis RE, Dally EL, Gundersen DE, Lee I-M, Habili N (1997) '*Candidatus* Phytoplasma australiense', a new phytoplasma taxon associated with Australian grapevine yellows. Int J Syst Bacteriol 47:262–269

Davis RE, Jomantiene R, Zhao Y, Dally EL (2003) Folate biosynthesis pseudogenes, *ψfolP* and *ψfolK*, and an O-sialoglycoprotein endopeptidase gene homolog in the phytoplasma genome. DNA Cell Biol 22:697–706

Davis RE, Jomantiene R, Zhao Y (2005) Lineage-specific decay of folate biosynthesis genes suggests ongoing host adaptation in phytoplasmas. DNA Cell Biol 24:832–840

Davis RE, Jomantiene R, Zhao Y (2007) Dynamic structures in phytoplasma genomes: sequence variable mosaics (SVMs) of clustered genes. Bull Insectol 60:119–120

Davis RE, Zhao Y, Dally EL, Lee IM, Jomantiene R, Douglas SM (2013) '*Candidatus* Phytoplasma pruni', a novel taxon associated with X-disease of stone fruits, *Prunus* spp.: multilocus characterization based on 16S rRNA, *secY*, and ribosomal protein genes. Int J Syst Evol Microbiol 63:766–776

Diener TO (1971) Potato spindle tuber "virus". IV. A replicating, low molecular weight RNA. Virology 45:411–428

Ding Y, Wu W, Wei W, Davis RE, Lee I-M, Hammond RW, Sheng J, Shen L, Jinag Y, Zhao Y (2013) Potato purple top phytoplasma-induced disruption of gibberellin homeostasis in tomato plants. Ann Appl Biol 162:131–139

Doi YM, Teranaka M, Yora K, Asuyama H (1967) Mycoplasma or PLT-group-like microorganisms found in the phloem elements of plants infected with mulberry dwarf, potato witches' broom, aster yellows, or paulownia witches' broom. Ann Phytopathol Soc Jpn 33:259–266

Durante G, Casati P, Clair D, Quaglino F, Bulgari D, Boudon-Padieu E, Bianco PA (2012) Sequence analyses of S10-*spc* operon among 16SrV group phytoplasmas: phylogenetic relationships and identification of discriminating single nucleotide polymorphisms. Ann Appl Biol 161:234–246

Eadsforth TC, Gardiner M, Maluf FV, McElroy S, James D, Frearson J, Gray D, Hunter WN (2012) Assessment of Pseudomonas aeruginosa N5,N10-methylenetetrahydrofolate dehydrogenase-cyclohydrolase as a potential antibacterial drug target. PLoS ONE 7(4):e35973

Frago S, Martínez-Júlvez M, Serrano A, Medina M (2008) Structural analysis of FAD synthetase from *Corynebacterium ammoniagenes*. BMC Microbiol 8:160

Fraser CM, Gocayne JD, White O, Adams MD, Clayton RA, Fleischmann RD, Bult CJ, Kerlavage AR, Sutton G, Kelley JM, Fritchman RD, Weidman JF, Small KV, Sandusky M, Fuhrmann J, Nguyen D, Utterback TR, Saudek DM, Phillips CA, Merrick JM, Tomb JF, Dougherty BA, Bott KF, Hu PC, Lucier TS, Peterson SN, Smith HO, Hutchison CA 3rd, Venter JC (1995) The minimal gene complement of *Mycoplasma genitalium*. Science 270:397–403

Garavaglia S, Galizzi A, Rizzi M (2003) Allosteric regulation of *Bacillus subtilis* NAD kinase by quinolinic acid. J Bacteriol 185:4844–4850

Gedvilaite A, Jomantiene R, Dabrisius J, Norkiene M, Davis RE (2014) Functional analysis of a lipolytic protein encoded in phytoplasma phage based genomic island. Microbiol Res 169:388–394

Ghosh SK, Raychaudhuri SP, Varma A, Nariani TK (1971) Isolation and culture of mycoplasma associated with citrus greening disease. Curr Sci 40:299–300

Glass JI, Assad-Garcia N, Alperovich N, Yooseph S, Lewis MR, Maruf M, Hutchison CA 3rd, Smith HO, Venter JC (2006) Essential genes of a minimal bacterium. Proc Natl Acad Sci USA 103:425–430

Gundersen DE, Lee IM, Rehner SA, Davis RE, Kingsbury DT (1994) Phylogeny of mycoplasmalike organisms (phytoplasmas): a basis for their classification. J Bacteriol 176:5244–5254

Hampton RO, Stevens JG, Allen TC (1969) Mechanically transmissible mycoplasma from naturally infected peas. Plant Dis Reporter 53:499–503

Hanschmann EM, Godoy JR, Berndt C, Hudemann C, Lillig CH (2013) Thioredoxins, glutaredoxins, and peroxiredoxins-molecular mechanisms and health significance: from cofactors to antioxidants to redox signaling. Antioxid Redox Signal, Mar 28, 2013 (Epub ahead of print)

Hantke K (2005) Bacterial zinc uptake and regulators. Curr Opin Microbiol 8:196–202

Himeno M, Neriya Y, Minato N, Miura C, Sugawara K, Ishii Y, Yamaji Y, Kakizawa S, Oshima K, Namba S (2011) Unique morphological changes in plant pathogenic phytoplasma-infected petunia flowers are related to transcriptional regulation of floral homeotic genes in an organ-specific manner. Plant J 67:971–979

Hoshi A, Oshima K, Kakizawa S, Ishii Y, Ozeki J, Hashimoto M, Komatsu K, Kagiwada S, Yamaji Y, Namba S (2009) A unique virulence factor for proliferation and dwarfism in plants identified from a phytopathogenic bacterium. Proc Natl Acad Sci USA 106:6416–6421

Hsiao W, Wan I, Jones SJ, Brinkman FS (2003) IslandPath: aiding detection of genomic islands in prokaryotes. Bioinformatics 19:418–420

IRPCM Phytoplasma/Spiroplasma Working Team-Phytoplasma Taxonomy Group (2004) '*Candidatus* Phytoplasma', a taxon for the wall-less, non-helical prokaryotes that colonize plant phloem and insects. Int J Syst Evol Microbiol 54:1243–1255

Ishiie T, Doi Y, Yora K, Asuyama H (1967) Suppressive effects of antibiotics of tetracycline group on symptom development of mulberry dwarf disease. Ann Phytopathol Soc Jpn 33:267–275

Istivan TS, Coloe PJ (2006) Phospholipase A in Gram-negative bacteria and its role in pathogenesis. Microbiology 152:1263–1274

Javid-Majd F, Blanchard JS (2000) Mechanistic analysis of the *argE*-encoded N-acetylornithine deacetylase. Biochemistry 39(6):1285–1293

Jomantiene R, Davis RE (2006) Clusters of diverse genes existing as multiple, sequence-variable mosaics in a

phytoplasma genome. FEMS Microbiol Lett 255:59–65

Jomantiene R, Zhao Y, Davis RE (2007) Sequence-variable mosaics: composites of recurrent transposition characterizing the genomes of phylogenetically diverse phytoplasmas. DNA Cell Biol 26:557–564

Jomantiene R, Zhao Y, Lee I-M, Davis RE (2011) Phytoplasmas infecting sour cherry and lilac represent two distinct lineages having closes evolutionary affinities with clover phyllody phytoplasma. Eur J Plant Pathol 130:97–107

Jorasch P, Wolter FP, Zähringer U, Heinz E (1998) A UDP glucosyltransferase from *Bacillus subtilis* successively transfers up to four glucose residues to 1,2-diacylglycerol: expression of *ypfP* in *Escherichia coli* and structural analysis of its reaction products. Mol Microbiol 29:419–430

Kakizawa S, Oshima K, Nishigawa H, Jung HY, Wei W, Suzuki S, Tanaka M, Miyata S, Ugaki M, Namba S (2004) Secretion of immunodominant membrane protein from onion yellows phytoplasma through the Sec protein-translocation system in *Escherichia coli*. Microbiology 150:135–142

Karlin S (2001) Detecting anomalous gene clusters and pathogenicity islands in diverse bacterial genomes. Trends Microbiol 9:335–343

Kawai S, Murata K (2008) Structure and function of NAD kinase and NADP phosphatase: key enzymes that regulate the intracellular balance of NAD(H) and NADP(H). Biosci Biotechnol Biochem 72:919–930

Klimasauskas S, Weinhold E (2007) A new tool for biotechnology: AdoMet-dependent methyltransferases. Trends Biotechnol 25:99–104

Kobayashi M, Shimizu S (1999) Cobalt proteins. Eur J Biochem 261:1–9

Kube M, Schneider B, Kuhl H, Dandekar T, Heitmann K, Migdoll AM, Reinhardt R, Seemüller E (2008) The linear chromosome of the plant-pathogenic mycoplasma '*Candidatus* Phytoplasma mali'. BMC Genom 9:306

Kube M, Mitrovic J, Duduk B, Rabus R, Seemüller E (2012) Current view on phytoplasma genomes and encoded metabolism. Sci World J 2012:185942

Lee I-M, Hammond RW, Davis RE, Gundersen DE (1993) Universal amplification and analysis of pathogen 16S rDNA for classification and identification of mycoplasmalike organisms. Phytopathology 83:834–842

Lee I-M, Gundersen-Rindal DE, Davis RE, Bartoszyk IM (1998) Revised classification scheme of phytoplasmas based on RFLP analysis of 16S rRNA and ribosomal protein gene sequences. Int J Syst Bacteriol 48:1153–1169

Lee I-M, Davis RE, Gundersen-Rindal DE (2000) Phytoplasma: phytopathogenic mollicutes. Annu Rev Microbiol 54:221–255

Lee I-M, Gundersen-Rindal DE, Davis RE, Bottner KD, Marcone C, Seemüller E (2004) '*Candidatus* Phytoplasma asteris', a novel phytoplasma taxon associated with aster yellows and related diseases. Int J Syst Evol Microbiol 54:1037–1048

Lee I-M, Bottner KD, Secor G, Rivera-Varas V (2006) '*Candidatus* Phytoplasma americanum', a phytoplasma associated with a potato purple top wilt disease complex. Int J Syst Evol Microbiol 56:1593–1597

Lee I-M, Bottner-Parker KD, Zhao Y, Villalobos W, Moreira L (2011) '*Candidatus* Phytoplasma costaricanum' a new phytoplasma associated with a newly emerging disease in soybean in Costa Rica. Int J Syst Evol Microbiol 61:2822–2826

Lepka P, Stitt M, Moll E, Seemüller E (1999) Effect of phytoplasmal infection on concentration and translocation of carbohydrates and amino acids in periwinkle and tobacco. Physiol Mol Plant Pathol 55:59–68

Li L, Storm P, Karlsson OP, Berg S, Wieslander A (2003) Irreversible binding and activity control of the 1,2-diacylglycerol 3-glucosyltransferase from *Acholeplasma laidlawii* at an anionic lipid bilayer surface. Biochemistry 42:9677–9686

Liefting LW, Veerakone S, Clover RG, Ward LI (2011) An update on phytoplasma diseases in New Zealand. Bull Insectol 64:S93–S94

Lin SC, Lee CS, Chin RJ (1970) Isolation and cultivation of, and inoculation with a mycoplasma causing white leaf disease of sugarcane. Phytopathology 60:795–797

Lombardo G, Pignattelli P (1970) Cultivation in a cell-free medium of a mycoplasma-like organism from *Vinca rosea* with phyllody symptoms of the flowers. Ann Microbiol 20:83–88

Lu ZJ, Markham GD (2002) Enzymatic properties of S-adenosylmethionine synthetase from the archaeon *Methanococcus jannaschii*. J Biol Chem 277:16624–16631

MacLean AM, Sugio A, Makarova OV, Findlay KC, Grieve VM, Tóth R, Nicolaisen M, Hogenhout SA (2011) Phytoplasma effector SAP54 induces indeterminate leaf-like flower development in *Arabidopsis* plants. Plant Physiol 157:831–841

Magni G, Di Stefano M, Orsomando G, Raffaelli N, Ruggieri S (2009) NAD(P) biosynthesis enzymes as potential targets for selective drug design. Curr Med Chem 16:1372–1390

Marcone C, Neimark H, Ragozzino A, Lauer U, Seemüller E (1999) Chromosome sizes of phytoplasmas composing major phylogenetic groups and subgroups. Phytopathology 89:805–810

McLennan AG (2006) The Nudix hydrolase superfamily. Cell Mol Life Sci 63:123–143

Ménétret JF, Schaletzky J, Clemons WM Jr, Osborne AR, Skånland SS, Denison C, Gygi SP, Kirkpatrick DS, Park E, Ludtke SJ, Rapoport TA, Akey CW (2007) Ribosome binding of a single copy of the SecY complex: implications for protein translocation. Mol Cell 28:1083–1092

Meyer Y, Buchanan BB, Vignols F, Reichheld JP (2009) Thioredoxins and glutaredoxins: unifying elements in redox biology. Annu Rev Genet 43:335–367

Mina JG, Pan SY, Wansadhipathi NK, Bruce CR, Shams-Eldin H, Schwarz RT, Steel PG, Denny PW (2009) The *Trypanosoma brucei* sphingolipid synthase, an essential enzyme and drug target. Mol Biochem Parasitol 168:16–23

Moore CM, Helmann JD (2005) Metal ion homeostasis in *Bacillus subtilis*. Curr Opin Microbiol 8:188–195

Morton A, Davies DL, Blomquist CL, Barbara DJ (2003) Characterization of homologues of the apple proliferation immunodominant membrane protein gene from three related phytoplasmas. Mol Plant Pathol 4:109–114

Muller JP (1999) Effects of pre-protein overexpression on SecB synthesis in *Escherichia coli*. FEMS Microbiol Lett 176:219–227

Murray RG, Schleifer KH (1994) Taxonomic notes: a proposal for recording the properties of putative taxa of procaryotes. Int J Syst Bacteriol 44:174–176

Murray RG, Stackebrandt E (1995) Taxonomic note: implementation of the provisional status Candidatus for incompletely described procaryotes. Int J Syst Bacteriol 45:186–187

Nagiec MM, Nagiec EE, Baltisberger JA, Wells GB, Lester RL, Dickson RC (1997) Sphingolipid synthesis as a target for antifungal drugs. Complementation of the inositol phosphorylceramide synthase defect in a mutant strain of *Saccharomyces cerevisiae* by the AUR1 gene. J Biol Chem 272:9809–9817

Nejat N, Vadamalai G, Davis RE, Harrison NA, Sijam K, Dickinson M, Abdullah SNA, Zhao Y (2013) '*Candidatus* Phytoplasma malaysianum', a novel taxon associated with virescence and phyllody of Madagascar periwinkle (*Catharanthus roseus*). Int J Syst Evol Microbiol 63:540–548

O'Farrell HC, Rife JP (2012) *Staphylococcus aureus* and *Escherichia coli* have disparate dependences on KsgA for growth and ribosome biogenesis. BMC Microbiol 12:244

Oshima K, Kakizawa S, Nishigawa H, Jung HY, Wei W, Suzuki S, Arashida R, Nakata D, Miyata S, Ugaki M, Namba S (2004) Reductive evolution suggested from the complete genome sequence of a plant-pathogenic phytoplasma. Nat Genet 36:27–29

Pailler J, Aucher W, Pires M, Buddelmeijer N (2012) Phosphatidylglycerol: prolipoprotein diacylglyceryl transferase (Lgt) of *Escherichia coli* has seven transmembrane segments, and its essential residues are embedded in the membrane. J Bacteriol 194:2142–2151

Patra T, Koley H, Ramamurthy T, Ghose AC, Nandy RK (2012) The Entner-Doudoroff pathway is obligatory for gluconate utilization and contributes to the pathogenicity of *Vibrio cholerae*. J Bacteriol 194:3377–3385

Pérez-Donoso AG, Sun Q, Roper MC, Greve LC, Kirkpatrick B, Labavitch JM (2010) Cell wall-degrading enzymes enlarge the pore size of intervessel pit membranes in healthy and *Xylella fastidiosa*-infected grapevines. Plant Physiol 152:1748–1759

Pinna GG, Spano A, Carru C, Pes GM, Delogu G, Miscali AR, Zanetti S, Sisini A (1991) Peptidase activity and toxicity of strains of *Pseudomonas aeruginosa*. Boll Soc Ital Biol Sper 67:287–294

Pollak N, Dölle C, Ziegler M (2007) The power to reduce: pyridine nucleotides - small molecules with a multitude of functions. Biochem J 402:205–218

Pracros P, Renaudin J, Eveillard S, Mouras A, Hernould M (2006) Tomato flower abnormalities induced by stolbur phytoplasma infection are associated with changes of expression of floral development genes. Mol Plant Microbe Interact 19:62–68

Pratt S, Wansadhipathi-Kannangara NK, Bruce CR, Mina JG, Shams-Eldin H, Casas J, Hanada K, Schwarz RT, Sonda S, Denny PW (2013) Sphingolipid synthesis and scavenging in the intracellular apicomplexan parasite, *Toxoplasma gondii*. Mol Biochem Parasitol 187:43–51

Quaglino F, Zhao Y, Bianco P, Wei W, Casati P, Durante G, Davis RE (2009) New 16Sr subgroups and distinct SNP lineages among grapevine Bois noir phytoplasma populations. Ann Appl Biol 154:279–289

Quaglino F, Zhao Y, Casati P, Bulgari D, Bianco PA, Wei W, Davis RE (2013) 'Candidatus Phytoplasma solani', a novel taxon associated with stolbur and bois noir related diseases of plants. Int J Syst Evol Microbiol (Epub ahead of print 18 Jan 2013)

Rapoport TA (2007) Protein translocation across the eukaryotic endoplasmic reticulum and bacterial plasma membranes. Nature 450:663–669

Razin S, Yogev D, Naot Y (1998) Molecular biology and pathogenicity of mycoplasmas. Microbiol Mol Biol Rev 62:1094–1156

Recchia GD, Hall RM (1997) Origins of the mobile gene cassettes found in integrons. Trends Microbiol 5:389–394

Rodionov DA, Hebbeln P, Gelfand MS, Eitinger T (2006) Comparative and functional genomic analysis of prokaryotic nickel and cobalt uptake transporters: evidence for a novel group of ATP-binding cassette transporters. Bacteriol 188:317–327

Saccardo F, Martini M, Palmano S, Ermacora P, Scortichini M, Loi N, Firrao G (2012) Genome drafts of four phytoplasma strains of the ribosomal group 16SrIII. Microbiology 158:2805–2814

Seemüller E, Marcone C, Lauer U, Ragozzino A, Göschl M (1998) Current status of molecular classification of the phytoplasmas. J Plant Pathol 80:3–26

Shen BW, Dyer DH, Huang JY, D'Ari L, Rabinowitz J, Stoddard BL (1999) The crystal structure of a bacterial, bifunctional 5,10-methylene-tetrahydrofolate dehydrogenase/cyclohydrolase. Protein Sci 8:1342–1349

Sitkiewicz I, Nagiec MJ, Sumby P, Butler SD, Cywes-Bentley C, Musser JM (2006) Emergence of a bacterial clone with enhanced virulence by acquisition of a phage encoding a secreted phospholipase A2. Proc Natl Acad Sci USA 103:16009–16014

Sparrow CP, Raetz CR (1985) Purification and properties of the membrane-bound CDP-diglyceride synthetase from *Escherichia coli*. J Biol Chem 260:12084–12091

Stackebrandt E (2007) Forces shaping bacterial systematics. Microbe 2:283–288

Su YT, Chen JC, Lin CP (2011) Phytoplasma-induced floral abnormalities in *Catharanthus roseus* are associated with phytoplasma accumulation and transcript repression of floral organ identity genes. Mol Plant Microbe Interact 24:1502–1512

Sugio A, Kingdom HN, MacLean AM, Grieve VM, Hogenhout SA (2011) Phytoplasma protein effector SAP11 enhances insect vector reproduction by manipulating plant development and defense hormone biosynthesis. Proc Natl Acad Sci USA 108:E1254–E1263

Suzuki S, Oshima K, Kakizawa S, Arashida R, Jung HY, Yamaji Y, Nishigawa H, Ugaki M, Namba S (2006) Interaction between the membrane protein of a pathogen and insect microfilament complex determines insect-vector specificity. Proc Natl Acad Sci USA 103:4252–4257

Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S (2011) MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Mol Biol Evol 28:2731–2739

Tran-Nguyen LT, Kube M, Schneider B, Reinhardt R, Gibb KS (2008) Comparative genome analysis of '*Candidatus* Phytoplasma australiense' (subgroup tuf-Australia I; rp-A) and '*Ca.* Phytoplasma asteris' Strains OY-M and AY-WB. J Bacteriol 190:3979–3991

Trumbo P, Yates AA, Schlicker S, Poos M (2001) Dietary reference intakes: vitamin A, vitamin K, arsenic, boron, chromium, copper, iodine, iron, manganese, molybdenum, nickel, silicon, vanadium, and zinc. J Am Diet Assoc 101:294–301

Van Vu B, Itoh K, Nguyen QB, Tosa Y, Nakayashiki H (2012) Cellulases belonging to glycoside hydrolase families 6 and 7 contribute to the virulence of *Magnaporthe oryzae*. Mol Plant Microbe Interact 25:1135–1141

Wei W, Cai H, Chen H, Davis RE, Zhao Y (2007a) First report of a natural infection of *Opuntia* sp. by a '*Candidatus* Phytoplasma asteris'-related phytoplasma in China. Plant Dis 91:461

Wei W, Davis RE, Lee I-M, Zhao Y (2007b) Computer-simulated RFLP analysis of 16S rRNA genes: identification of ten new phytoplasma groups. Int J Syst Evol Microbiol 57:1855–1867

Wei W, Lee I-M, Davis RE, Suo X, Zhao Y (2007c) Virtual RFLP analysis of 16S rDNA sequences identifies new subgroups in the clover proliferation phytoplasma group. Bull Insectol 60:349–350

Wei W, Davis RE, Jomantiene R, Zhao Y (2008a) Ancient, recurrent phage attacks and recombination events shaped dynamic sequence-variable mosaic structures at the root of phytoplasma genome evolution. Proc Natl Acad Sci USA 105:11827–11832

Wei W, Lee I-M, Davis RE, Suo X, Zhao Y (2008b) Automated RFLP pattern comparison and similarity coefficient calculation for rapid delineation of new and distinct phytoplasma 16Sr subgroup lineages. Int J Syst Evol Microbiol 58:2368–2377

Win NK, Lee SY, Bertaccini A, Namba S, Jung HY (2013) '*Candidatus* Phytoplasma balanitae' associated with witches' broom disease of *Balanites triflora*. Int J Syst Evol Microbiol 63:636–640

Yahr TL, Wickner WT (2000) Evaluating the oligomeric state of SecYEG in preprotein translocase. EMBO J 19:4393–4401

Zhao Y, Wang H, Hammond RW, Jomantiene R, Liu Q, Lin S, Roe BA, Davis RE (2004) Predicted ATP-binding cassette systems in the phytopathogenic mollicute *Spiroplasma kunkelii*. Mol Genet Genomics 271:325–338

Zhao Y, Davis RE, Lee I-M (2005) Phylogenetic positions of '*Candidatus* Phytoplasma asteris' and *Spiroplasma kunkelii* as inferred from multiple sets of concatenated core housekeeping proteins. Int J Syst Evol Microbiol 55:2131–2141

Zhao Y, Sun Q, Wei W, Davis RE, Wu W, Liu Q (2009a) '*Candidatus* Phytoplasma tamaricis', a novel taxon discovered in witches'-broom diseased salt cedar (*Tamarix chinensis* Lour.) Int. J. Syst. Evol. Microbiol 59:2496–2504

Zhao Y, Wei W, Lee I-M, Shao J, Suo X, Davis RE (2009b) Construction of an interactive online phytoplasma classification tool, *i*PhyClassifier, and its application in analysis of the peach X-disease phytoplasma group (16SrIII). Int J Syst Evol Microbiol 59:2582–2593

Zhao Y, Wei W, Davis RE, Lee I-M (2010) Recent advances in 16S rRNA gene-based phytoplasma differentiation, classification and taxonomy. In: Weintraub P, Jones P (eds) Phytoplasmas: genomes, plant hosts and vector. CABI Publishing, Wallingford, pp 64–92

Ziegler H (1975) Nature of substances in phloem. In: Zimmermann MH, Milburn JA (eds) Encyclopedia of plant physiology, transport in plants, vol 1. Springer, Berlin, pp 57–100

Index

D. C. Gross et al. (eds.), *Genomics of Plant-Associated Bacteria*,
DOI: 10.1007/978-3-642-55378-3, © Springer-Verlag Berlin Heidelberg 2014