SUMMARY

The publication of the first *Erwinia amylovora* genome has greatly accelerated and advanced our understanding of the fire blight organism. With the availability of multiple genomes, it quickly became clear that chromosomal diversity is relatively small, and that most of the pan-genome variance is attributable to plasmids. In addition to gaining a more detailed view of the known virulence factors, genomics has enabled new breakthrough studies of virulence regulation mechanisms. Furthermore, several niche adaptation and ecological fitness factors, though not directly influencing virulence, have been studied in greater detail, providing novel insights into the physiology and ecology of the bacterium. Additionally, application of genome data has yielded improved diagnostics and enabled population studies at different geographic scales.

**Keywords:** intraspecies diversity, niche adaptation, type III secretion system, diagnostics.

INTRODUCTION

Fire blight, the first plant disease to be described to be caused by a bacterium, is able to kill complete pome fruit orchards within a single growing season (Vanneste, 2000). After being detected first in North America, the disease has subsequently spread to New Zealand, Europe, North Africa, the Middle East, Russia, Central Asia and South Korea (Bonn and van der Zwet, 2000; Djaimurzina et al., 2014; Myung et al., 2016). This rapid and seemingly uncontrollable spread may reflect the relative weakness of its hosts to the disease, but also the strengths of the pathogen (Malnoy et al., 2012).

The yearly economic losses to commercial pome fruit orchards caused by fire blight infections are enormous (Duffy et al., 2005). A major line of research has thus been to start studying the genetics of its causative agent, *Erwinia amylovora*, with the aim to understand why the organism is this effective as a pathogen. Whereas in the past individual virulence factors were identified that largely or partly explained the pathogenicity (Oh and Beer, 2005), the era of genomics after the introduction of next generation sequencing devices has given us the chance of expanding this research to the genome scale. In contrast, efforts made to study the effect of the pathogen on the apple tree (Baldo et al., 2010; Kamber et al., 2016; Sarowar et al., 2011; Vrancken et al., 2013) are still in its early phase of understanding, mainly due to the complexity of the apple tree as a system.

Although another *E. amylovora* genome was sequenced as well, the first genome of *E. amylovora* to be published was the genome of the French isolate CFBP 1430 (Smits et al., 2010b). Though of most economical importance, *E. amylovora* was not the first species in the genus to be fully sequenced, as the genomes of *Erwinia tasmaniensis* Et1/99 (Kube et al., 2008) and *Erwinia pyrifoliae* DSM 12163^T^ (Smits et al., 2010a) were published before. This review intends to update current knowledge of the biology of *E. amylovora* that has been obtained since its genome was published in 2010. We will particularly discuss genomic insight of this pathogen in the larger context of its ecology, evolution and population genetics.

INTRA- AND INTERSPECIES DIVERSITY OF *E. AMYLOVORA* AT THE GENOMIC LEVEL

*E. amylovora* genomics. Currently, fifteen genome sequences of *E. amylovora* strains are publically available, of which three isolates are reported to specifically be pathogenic on *Rubus* spp. (Mann et al., 2013; McManus and Jones, 1995; Smits et al., 2014b; Starr et al., 1951). Many further genome sequences are still unpublished, including...
additional Rubus isolates. Though largely biased to Spiraeoideae strains belonging to the economically and more worldwide spread CRISPR group I (Rezzonico et al., 2011), this collection of genomic data allows the conclusion that there is only a limited diversity within this cosmopolitan CRISPR group, with sequence identities between two strains exceeding 99.99% (Smits et al., 2010b) as demonstrated by the presence of less than 2000 SNPs in the complete dataset of CRISPR I strains (T.H.M. Smits, unpublished information). On the other hand, the diversity among the genomes of Rubus-infecting strains is substantially higher (Mann et al., 2013), while also more diverse sequences of Spiraeoideae isolates have been obtained (T.H.M. Smits, unpublished information).

A major factor contributing to genomic diversity in Spiraeoideae-infecting isolates of E. amylovora is the presence of different plasmids in individual strains (Ismail et al., 2014; Llop et al., 2012). This has an influence on the total pan-genome of E. amylovora, leading to the impression that the pan-genome of the species is open. On the other hand, comparing only the chromosomal sequences, the pan-genome of Spiraeoideae-infecting isolates appears nearly closed (Mann et al., 2013).

**Intraspecies diversity of E. amylovora.** Many studies have already observed that it is relatively easy to separate Rubus-infecting isolates, currently reported only from USA and Canada (Braun and Hildebrand, 2005; Heimann and Worf, 1985; Ries and Otterbacher, 1977; Starr et al., 1951). Molecular studies based on housekeeping genes (Rezzonico et al., 2012b), rep-PCR, ribotyping (McManus and Jones, 1995), ITS (McGhee et al., 2002b) and CRISPR analysis (Rezzonico et al., 2011) allowed a fast assay to identify new strains. However, despite the long history of E. amylovora isolation, only few Rubus-infecting isolates are currently described and available (Rezzonico et al., 2012b).

Using the genome sequences, first efforts were undertaken to resolve the genetic base of the differential host range of Rubus-infecting isolates. Rezzonico et al. (2012b) have observed three major groups of Rubus-infecting isolates, each of them currently supported by a genome sequence (Mann et al., 2013). When comparing the genomes to those of Spiraeoideae-infecting isolates, several small-scale differences were observed. The gene cluster encoding lipopolysaccharide biosynthesis genes in Rubus-infecting isolates is substantially different from that of Spiraeoideae isolates (Rezzonico et al., 2012b). Furthermore, Rubus-infecting isolates contain a larger IT region (Mann et al., 2012). Analysis of the Hrp pathogenicity cluster showed that there are differences in single genes that exceed the average genome variation. Especially the eop1 gene, which shares 99% sequence identity within the Rubus-infecting isolates, has only 67% sequence identity to that of Spiraeoideae isolates. Deletion of eop1 from both Spiraeoideae-infecting and Rubus-infecting strains does not affect virulence. However, the addition of the Rubus eop1 to Spiraeoideae-infecting strains reduced their virulence on apple shoots, indicating that eop1 can function as a host specificity determinant (Asselin et al., 2011).
The Rubus-infecting isolates contain a 20 kb NRPS/PKS gene cluster close to the CRISPR regions (Rezzonico et al., 2011), which has eroded in the genome of Spiraeoideae-infecting isolates (Mann et al., 2013).

**Interspecies comparisons.** When comparing the genomes of *E. amylovora* to those of the closely related pathoadapted *Erwinia* spp. (*E. pyrifoliae*, *E. tasmaniensis* and *Erwinia piriflorinigrans*) (Kamber et al., 2012; Smits et al., 2013), it was observed that *E. amylovora* has the smallest genome size (3.8 Mb; Fig. 1) (Smits et al., 2010b). This is even more evident when comparing to the genome of the epiphyte *Erwinia billingiae* Eb661 (Kube et al., 2010). Although members of the closely related genus *Pantoea* may also have chromosomes as small as 4 Mb (Kamber et al., 2012), they appear not to have undergone a pathoadaptation process. On the other hand, the inclusion of several large gene clusters encoding proven and putative pathogenicity factors (Table 1), like diverse secretion systems and a second flagellum in the genomes of the pathoadapted species (Kamber et al., 2012; Smits et al., 2010b; Zhao and Qi, 2011), clearly shows that significant genome reduction must have taken place during the pathoadaptation process, as the size of the additional gene clusters sums up to more than 1 Mb, while the overall size of the genomes did not increase. Based on the comparison of the core genomes of *Pantoea* spp. and *E. billingiae* Eb661 (Kamber et al., 2012), it was hypothesized that the latter species has not undergone pathoadaptation, supporting the conclusion of a more generalist behavior for this species (Mergaert et al., 1999). With the addition of more genomes of this genus including the recently described novel species *Erwinia gerundensis* (Rezzonico et al., 2016), it was observed that pathoadaptation might have been induced several times during the shaping of the genus, as other species like *Erwinia tracheiphila* also show a pathoadaptation pattern (Shapiro et al., 2016).

**Plasmids.** In recent years, the sequencing of multiple genomes has led to the discovery of many new plasmids (Table 2). Whereas the plasmid pEA29 is almost ubiquitous (Bühlmann et al., 2014; Mann et al., 2013), other plasmids appear to be accessory material. Up to now, none of the plasmids has been shown to have an influence on the virulence of *E. amylovora* strains (Ismail et al., 2014; Llop et al., 2011). The novel plasmids are most probably derived from the general environmental plasmid pool, which has also been used as base for plasmids conferring multidrug antibiotic resistance to clinical isolates (Mann et al., 2013; Zhang et al., 2016; Zhu et al., 2009).

Plasmid pEA29 only shows diversity at few loci. It contains the short sequence repeat that was used for early strain typing (Table 1) (Kim and Geider, 1999; Schnabel and Jones, 1998), but also a second repeat, which allows the separation of larger populations (Bühlmann et al., 2014). Within the genome dataset, one isolate was included to have the streptomycin (Sm) resistance cassette in the transposon Tn3593 on pEA29 (Chiou and Jones, 1993). Additional diversity is present in the pEA29 plasmids from Rubus-infecting strains (McGhee et al., 2002a).

New plasmids have been identified within the frame of the genome sequencing, like pEA30 and several small plasmids from Rubus-infecting isolates (Mann et al., 2013) (Table 2). The plasmid pEI70, currently only detected in Europe, contains a non-integrative ICE-element (Llop et al., 2011). Plasmid pE6A8, isolated from Poland and Belgium (Ismail et al., 2014), is member of a plasmid family that includes the large plasmid pEA72 from strain ATCC 49946 (Sebaihia et al., 2010), but also a 78 kb plasmid found in a Mexican and a Californian strain (Ismail et al., 2014; Smits et al., 2014a, 2014b). Within the set of unpublished sequences, many other plasmids were found, that will be described in the near future (T.H.M. Smits, unpublished information).

**Niche Adaptation Factors**

Bacteria are able to adapt to environmental conditions at a certain location. This niche specialization process can be supported by the loss of functionalities involved in global survival in the environment (Moran, 2002), and

---

**Table 1.** Selected factors analyzed by comparative genomic approaches in a variety of genome-sequenced *Erwinia* spp. Only a single genome sequence is reported for each species. A “+” indicates the presence of the factor, a “−” indicates absence. The exopolysaccharides (EPS) are differentiated as amylovoran-like (AMS) or stewartan-like (CPS), while CRISPR systems are differentiated as *Escherichia coli*-type cas genes (Ecoli) or *Yersinia pestis*-type cas genes (Ypest).

<table>
<thead>
<tr>
<th>Species and strain</th>
<th>Hrp T3SS</th>
<th>Inv/Spa T3SS</th>
<th>EPS-type</th>
<th>Desferrioxamine</th>
<th>CRISPR</th>
<th>pEA29-like plasmid</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>E. amylovora</em> CFBP 1430</td>
<td>+</td>
<td>+</td>
<td>AMS</td>
<td>+</td>
<td>Ecoli</td>
<td>+</td>
<td>(Smits et al., 2010b)</td>
</tr>
<tr>
<td><em>E. pyrifoliae</em> DSM 12163T</td>
<td>+</td>
<td>+</td>
<td>AMS</td>
<td>+</td>
<td>Ecoli, Ypest</td>
<td>+</td>
<td>(Smits et al., 2010a)</td>
</tr>
<tr>
<td><em>E. tasmaniensis</em> Et1/299T</td>
<td>+</td>
<td>+</td>
<td>CPS</td>
<td>+</td>
<td>Ypest</td>
<td>−</td>
<td>(Kube et al., 2008)</td>
</tr>
<tr>
<td><em>E. piriflorinigrans</em> CFBP 5888T</td>
<td>+</td>
<td>+</td>
<td>AMS</td>
<td>+</td>
<td>Ypest</td>
<td>+</td>
<td>(Smits et al., 2013)</td>
</tr>
<tr>
<td><em>E. persticiae</em> CFBP 3622T</td>
<td>−</td>
<td>−</td>
<td>CPS</td>
<td>−</td>
<td>−</td>
<td>+</td>
<td>(Smits et al., unpublished)</td>
</tr>
<tr>
<td><em>E. billingiae</em> Eb661T</td>
<td>−</td>
<td>−</td>
<td>CPS</td>
<td>−</td>
<td>−</td>
<td>−</td>
<td>(Kube et al., 2010)</td>
</tr>
<tr>
<td><em>E. toletana</em> DAPP-PG735</td>
<td>−</td>
<td>−</td>
<td>CPS</td>
<td>−</td>
<td>−</td>
<td>−</td>
<td>(Passos da Silva et al., 2013)</td>
</tr>
<tr>
<td><em>E. tracheiphila</em> PSU1</td>
<td>+</td>
<td>+</td>
<td>CPS</td>
<td>−</td>
<td>−</td>
<td>−</td>
<td>(Shapiro et al., 2016)</td>
</tr>
<tr>
<td><em>E. gerundensis</em> EM595T</td>
<td>−</td>
<td>−</td>
<td>CPS</td>
<td>−</td>
<td>−</td>
<td>(+)</td>
<td>(Rezzonico et al., 2016)</td>
</tr>
</tbody>
</table>
may also include the acquisition of novel genetic material (Casadevall, 2008). Niche adaptation factors may include virulence factors (i.e. factors that are involved in the gain of access to and survival in host systems, factors that damage the host or factors that cause dysregulation of the host cellular function) and environmental fitness factors (Hill, 2012). The latter group includes products and strategies involved in attachment, macro and micronutrient acquisition, colonization and microbe-host communication strategies. The available genome sequences of \textit{E. amylovora} have allowed the examination of individual niche adaptation factors, involved in virulence, regulation and adaptation. Below, we summarize recent findings.

The Hrp Type III Secretion System (T3SS). Most Gram-negative bacterial pathogens rely on type III secretion systems (T3SS) for the interaction with their host (Tampakaki et al., 2010). The T3SS forms a needle structure that is able to penetrate the host cell wall and to deliver effector proteins directly into the cytosol of the eukaryotic host cell. The T3SSs in \textit{Yersinia}, \textit{Escherichia}, \textit{Salmonella} and \textit{Shigella} have been extensively studied as they have potentially devastating health effects on human life (Naum et al., 2011). In non-host plants, the T3SSs of plant pathogens cause a hypersensitive response, a collapse of plant tissue several hours after inoculation followed by tissue necrosis as a result of programmed cell death, as a reaction to incompatible pathogens (Tampakaki et al., 2010).

In \textit{E. amylovora}, the Hrp T3SS is one of the most important virulence factors and is absolutely required for pathogenicity (Malnoy et al., 2012; Oh and Beer, 2005). Comparable systems are also present in other pathoadapted \textit{Erwinia} spp. (Table I) (Kamber et al., 2012) and the plant pathogen \textit{Pantoea stewartii} subsp. \textit{stewartii} DC283 (Roper, 2011), that might have a common origin (Kirzinger et al., 2015). In contrast to previous reports (Oh and Beer, 2005), novel insights from genomics have shown that the Hrp T3SS was independently introduced from the so-called IT region (Mann et al., 2012). This latter region constitutes a remnant of an Integrative Conjugative Element (ICE), for which both border regions could be identified outside the Hrp T3SS region in the genome of \textit{E. piriflorinigrans} CFBP 5888T (Smits et al., 2013).

Many plant pathogens have a whole suite of T3SS effectors, that are selected based on the individual host range of the organism (Grant et al., 2006; Hajri et al., 2009). \textit{E. amylovora} has, in contrast, only eight effectors. The genes of four effectors are located within the Hrp T3SS gene cluster (Smits et al., 2010): harpins HrpN and HrpW, the host-range limiting factor Eop1 and the essential effector DspA/E (Asselin et al., 2011; Malnoy et al., 2012; Oh and Beer, 2005; Smits et al., 2010b). The four remaining effectors are spread over the genome, and are, in contrast to the previous set, only present in \textit{E. amylovora} while not in the other pathoadapted \textit{Erwinia} species (Malnoy et al., 2011a). One of the latter proteins, the effector AvrRpt2, is a cysteine protease with most similarity to the \textit{Pseudomonas syringae} AvrRpt2 (Zhao et al., 2006). It was shown that a single nucleotide polymorphism in the \textit{E. amylovora} avrRpt2 gene is responsible to overcome the \textit{Malus × robusta} 5 resistance to fire blight (Vogt et al., 2013), yet it does only affect but not break the resistance of \textit{Malus fusca} (Emeriewen et al., 2015). Genomic sequencing has shown that this mutation is present in a few natural isolates from the USA and Mexico (Smits et al., 2014a, 2014b; Vogt et al., 2013).

### Table 2. Plasmids described in \textit{E. amylovora} strains.

<table>
<thead>
<tr>
<th>Plasmid name</th>
<th>Original strain(s)</th>
<th>Size (kb)</th>
<th>Description</th>
<th>Replicon type</th>
<th>Mobilization type</th>
<th>Reference(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>pEA29</td>
<td>Nearly all strains</td>
<td>28</td>
<td>Standard plasmid, \text{thrOSGF}</td>
<td>IncF</td>
<td>None</td>
<td>(McGhee and Jones 2000; Smits et al., 2010b)</td>
</tr>
<tr>
<td>pEA34</td>
<td>NWI/1</td>
<td>35</td>
<td>pEA29 with Tn5393, \text{strAB}</td>
<td>IncF</td>
<td>None</td>
<td>(Chiu and Jones 1999; Smits et al., unpublished)</td>
</tr>
<tr>
<td>pEA68</td>
<td>692</td>
<td>68</td>
<td>mobAB, pil, tra</td>
<td>IncFIA</td>
<td>P13</td>
<td>(Ismail et al., 2014)</td>
</tr>
<tr>
<td>pEA72</td>
<td>ATCC 49496</td>
<td>71</td>
<td>mobAB, pil, tra</td>
<td>IncFIA</td>
<td>P13</td>
<td>(Sebahia et al., 2010)</td>
</tr>
<tr>
<td>pEA78</td>
<td>LA637</td>
<td>78</td>
<td>mobAB, pil, tra</td>
<td>IncFIA</td>
<td>P13</td>
<td>(Ismail et al., 2014; Smits et al., 2014b)</td>
</tr>
<tr>
<td>pEL60</td>
<td>LEBb66</td>
<td>60</td>
<td>mobAB, tra</td>
<td>IncL/M</td>
<td>P13</td>
<td>(Foster et al., 2004)</td>
</tr>
<tr>
<td>pEI70</td>
<td>ACW856-400</td>
<td>65</td>
<td>ICE-like plasmid</td>
<td>IncL/M</td>
<td>P6</td>
<td>(Foster et al., 2004)</td>
</tr>
<tr>
<td>pEU30</td>
<td>UTRF2</td>
<td>30</td>
<td>virB</td>
<td>IncU</td>
<td>P4</td>
<td>(Mann et al., 2013)</td>
</tr>
<tr>
<td>pEA30</td>
<td>Ea495</td>
<td>30</td>
<td>Identical to RSF1010, \text{strAB}, su2</td>
<td>IncQ</td>
<td>Q1</td>
<td>(Palmer et al., 1997)</td>
</tr>
<tr>
<td>pEA87</td>
<td>CA3R</td>
<td>8.7</td>
<td>Cryptic plasmid, not found with resequencing the strain</td>
<td>CoEl1</td>
<td>P5</td>
<td>(Mann et al., 2013)</td>
</tr>
<tr>
<td>pEAR3.3</td>
<td>ATCC BAA-2158</td>
<td>5.3</td>
<td></td>
<td>CoEl1</td>
<td>P5</td>
<td>(Mann et al., 2013)</td>
</tr>
<tr>
<td>pEAR4.2</td>
<td>ATCC BAA-2158</td>
<td>4.2</td>
<td></td>
<td>Unknown</td>
<td>None</td>
<td>(Mann et al., 2013)</td>
</tr>
<tr>
<td>pEA17</td>
<td>IH-3</td>
<td>1.7</td>
<td></td>
<td>CoEl1</td>
<td>None</td>
<td>(McGhee et al., 2002b)</td>
</tr>
<tr>
<td>pEA2.8</td>
<td>IL-5</td>
<td>2.8</td>
<td>Cryptic plasmid, not found with resequencing the strain</td>
<td>CoEl1</td>
<td>None</td>
<td>(Mann et al., 2013; McGhee et al., 2002b)</td>
</tr>
</tbody>
</table>
formation (Koczan et al., 2009) and defense against antimicrobial compounds (Ordax et al., 2010). Deletion of the biosynthetic ans genes or the regulatory rcs genes in *E. amylovora* leads to a loss of virulence (Bernhard et al., 1996; Koczan et al., 2009; Wang et al., 2009). The mono-mer structures of amylovoran and the related EPS stewartan of *P. stewartii* subsp. *stewartii* show some significant differences (Schollmeyer et al., 2012), that are also reflected in the genetic complement of these two strains and their sequenced relatives (Table 1) (Kamber et al., 2012). Recent studies have shown that two additional gene clusters are present in *E. amylovora* that are required for the export of alternative monomers to the periplasm (Langlotz et al., 2011; Wang et al., 2012c). Based on a complementation study, the additional genes were shown to be reduced active in *E. amylovora*, whereas these genes are required for the major structure of stewartan (Wang et al., 2012c).

**Regulation of virulence factors.** A regulatory cascade is responsible for the regulation of the Hrp cluster of *E. amylovora* (Khan et al., 2012; Zhao, 2014). Molecular genetic studies of *E. amylovora* pathogenesis demonstrated that expression of the hrp-T3SS genes is activated by the master regulator HrpL, a member of the exocytosplasmic functions (ECF) subfamily of sigma factors (Wei and Beer, 1995). Subsequently, HrpL binds to the hrp box in the promoter regions of structural genes and effectors of the T3SS and allows transcription of these genes (Oh and Beer, 2005; Wei et al., 2000). The regulon of HrpL includes all hrp and hrc genes and five putative T3SS effectors. As eight genes displayed apparent indirect HrpL regulation, it was suggested that the HrpL regulon of *E. amylovora* encompasses more than just T3SS regulation and that HrpL may communicate with other signaling networks to coordinate gene expression during pathogenesis (McNally et al., 2012). In *E. amylovora*, hrlP transcription is in turn controlled by sigma factor 54 (RpoN), which along with its modulation protein YbhH and an integration host factor (IHF), interacts with enhancer binding protein (EBP) HrpS to trigger the onset of T3SS (Ancona et al., 2014; Lee and Zhao, 2016; Lee et al., 2016). Furthermore, it has been recently reported that the linear nucleotide second messengers, guanosine tetraphosphate (ppGpp) and guanosine pentaphosphate (pppGpp), are also essential for the expression of the T3SS and virulence by activating the RpoN-HrpL alternative sigma factor cascade (Ancona et al., 2015; Zhao and Sundin, 2017).

Two-component signal transduction systems (TCSTs), consisting of a histidine kinase (HK) and a response regulator (RR), represent a major paradigm for signal transduction in prokaryotes. TCSTs play critical roles in sensing and responding to environmental conditions, and in bacterial pathogenesis. Genome-wide screening of two-component system identified four groups of mutants which exhibited varying levels of amylovoran production in *vitro* (Zhao et al., 2009b). Among them, the Rcs phosphorelay system is an essential regulatory system for pathogenicity and gene expression of virulence factors, especially of amylovoran (Wang et al., 2009, 2011). Microarray studies identified a large set of genes that are regulated by this system. While RcsB acts as a positive regulator, the sensor kinase RcsC positively controls ans cluster expression in planta but negatively in *vitro* (Wang et al., 2012b). In addition, the EnvZ/OmpR and GrrS/GrrA (GacS/GacA) systems, two widely distributed TCSTs in gammaproteobacteria, negatively co-regulate amylovoran biosynthesis and T3SS in *E. amylovora* (Li et al., 2014). It has recently been demonstrated in *E. amylovora* that negative regulation of virulence by GrrS/GrrA acts through the non-coding small regulatory RNA csrB/rsmB sRNA, which binds to the RNA-binding protein CsrA/RsmA and neutralizes its positive effect on T3SS gene expression and amylovoran production (Ancona et al., 2016). Thus, CsrA plays a central role in *E. amylovora* virulence and in positive regulation of T3SS and amylovoran (Ancona et al., 2016).

Genetic screening also identified other global regulatory genes for amylovoran biosynthesis, including AmyR, Lon protease, and H-NS (Eastgate et al., 1995; Hildebrand et al., 2006; Wang et al., 2012a). The novel regulator AmyR (EAMY_1304) plays a role of a negative regulator in the regulation of virulence factors including amylovoran biosynthesis (Wang et al., 2012a). Mutation of the amyR gene leads to a mucoid phenotype. In a background of amylovoran-overproducing mutants in other two-component signal transduction genes, overexpression of amyR leads to a strong reduction of amylovoran production (Wang et al., 2012a), showing that this regulator acts at a different level.

Two other global pathways controlling virulence regulation in *E. amylovora* have recently been discovered. The second messenger compound cyclic di-GMP (c-di-GMP) also extensively regulates the major pathogenesis systems in *E. amylovora* including the T3SS, amylovoran biosynthesis, biofilm formation, and motility (Edmunds et al., 2013). Cyclic di-GMP impacts transcriptional regulation of target genes through binding to protein regulators or to riboswitches; in addition, binding to c-di-GMP can functionally activate proteins (Römling et al., 2013). Intracellular levels of c-di-GMP in *E. amylovora* are controlled through the actions of five guanylate cyclase enzymes that synthesize c-di-GMP and three phosphodiesterase enzymes that degrade the molecule (Edmunds et al., 2013). High intracellular levels of c-di-GMP in *E. amylovora* promote amylovoran synthesis and biofilm formation while simultaneously inhibiting the T3SS.

The second global regulatory pathway involves non-coding regulatory small RNAs (sRNAs) that require the RNA chaperone Hfq for stability and functional activation. An *E. amylovora* hfq mutant exhibits severely-reduced virulence, most likely due to a reduction in translocation of DspA/E into host cells and other impacts on proteins secreted by the T3SS (Zeng et al., 2013). In addition, amylovoran production is significantly reduced in the hfq
mutant, and cells exhibit a hyper-attaching phenotype (Zeng et al., 2013). *E. amylovora* encodes at least 40 Hfq-dependent sRNAs, and deletion of at least three of these (arcZ, trybA, brs21, rprA) results in significant reductions in virulence in an immature pear model (Zeng et al., 2013). Additional Hfq-dependent sRNAs including brs6 and omrAB significantly affect motility and amylovoran production (Zeng and Sundin, 2014).

It is clear that the regulatory networks for *E. amylovora* virulence factors are highly complex and may be far more complex than currently known. Understanding how all of the varied regulatory pathways of *E. amylovora* intersect and respond to environmental signals is the next challenge in the development of a complete model of pathogenesis in this organism.

**Inv-Spa-type T3SSs.** Additionally to the Hrp T3SS, *E. amylovora* also contains two Inv/Spa-type T3SSs (Smits et al., 2010b), a situation also observed in *E. piriflorinigrans* CFBP 5888 (Table 1) (Smits et al., 2013). Only one of these is present in *E. pyrifolae*, while the second is also partially present in *E. tasmaniensis* Et1/99 (Kamber et al., 2012). As these T3SS were proven to not play a role in host plant pathogenicity (Zhao et al., 2009a) and resembled the Inv/Spa T3SS of the insect symbiont *Sodalis glossinidius* str. *S. mossi* (Dale et al., 2001; Kirzinger et al., 2015), they were hypothesized to play a potential role in the interaction with insects (Smits et al., 2010b). This hypothesis is supported by a recent study describing that the Inv/Spa-type T3SS in *P. stewartii* subsp. *stewartii* DC283 is required for persistence in its flea beetle vector *Chetaocrema pulicaria* (Correa et al., 2012). Although a close vector relationship is not recognized with *E. amylovora*, it would be worth examining the potential roles for Inv/Spa T3SSs in the *E. amylovora* life cycle.

**Autoinduction and quorum-sensing.** In Gram-negative bacteria, two major quorum-sensing systems are known, that are defined by the chemical structure of the autoinducer (AI) signal molecule (Miller and Bassler, 2001). For both systems, a conclusive support of their presence is lacking for *E. amylovora*. Previous reports on the presence of the gene pair expRI, encoding the enzyme responsible for the production of the AI-1 signal N-acylhomoserine lactone (AHL) and its cognate receptor in *E. amylovora* (Molina et al., 2005; Venturi et al., 2004), were not confirmed by the annotation of multiple genomes within the species (Mann et al., 2013; Smits et al., 2010b).

Whereas the genomes of *E. tasmaniensis* Et1/99 and *E. piriflorinigrans* CFBP 5888 encode orthologs of the *Pectobacterium carotovorum* expRI genes (Sabag-Daigle and Ahmer, 2012; Smits et al., 2011a, 2013), the absence of the expRI genes in *E. amylovora* can be explained by the replacement of the region between the smpB and tsx genes, which contains the quorum-sensing genes in *E. tasmaniensis* Et1/99 and *E. piriflorinigrans* CFBP 5888, through a flagellar gene cluster in *E. amylovora* CFBP 1430 and *E. pyrifolae* DSM 12163 (Smits et al., 2011a, 2010b). Indeed, a BlastN search with a previously sequenced DNA fragment (GenBank Accession No. AJ841286) purportedly spanning the expRI genes in *E. amylovora* (Venturi et al., 2004) resulted in high sequence identity within the genomes of *Pectobacterium* spp., but no hits to *E. amylovora*. Furthermore, in silico analysis on all the sequenced *E. amylovora* genomes did not yield any viable annealing site neither for the EAM1/EAM2 primer pair described in the latter work nor for the AHLea-fw/-rev primer pair designed by (Molina et al., 2005), yet both primer sets showed a perfect match with the complete genome sequence of *Pectobacterium carotovorum* SCC3193 (GenBank Accession No. CP003415). This suggests that in both studies, the positive amplification may have been the outcome of a contamination with *Pectobacterium* spp. DNA in the template used for the PCR. In addition to these inconsistencies on a genomic basis, phenotypic tests gave rise to contradictory results. The purported AHL signal was identified to be either N-(3-oxo-hexanoyl)-homoserine lactone (3-oxo-C6-HSL) or N-(3-hydroxy-hexanoyl)-homoserine lactone (3-OH-C6-HSL) in the earlier paper using thin layer chromatography and reporter strain *E. coli* JM109 (pSB401) (Venturi et al., 2004). Yet neither aforementioned HSLs could be detected by reporter strain *Chromobacterium violaceum* CV026 (Burton et al., 2005), which yielded a positive signal in the experiments performed by Molina et al. (2005). To add to the confusion, the above results could not be confirmed by a subsequent work that failed to detect any endogenous AHL production either using *C. violaceum* CV026 or by in trans expression of a green fluorescent protein after transformation of *E. amylovora* with reporter plasmid pJB1A32 (Jakovljevic et al., 2008).

Despite the confusion about the nature of the AHL produced, a gene encoding a LuxR-type transcriptional activator with sequence identity to the cell division control protein SdiA of *E. coli* and *Salmonella typhimurium* (Wang et al., 1991) could be confirmed in the genome of *E. amylovora* CFBP 1430 (Smits et al., 2010b). Directly downstream of sdiA, a luxI homolog encoding a putative AI-1 synthesis protein that is related to the AHL synthase PhzI of *Pseudomonas chlororaphis* was found (Smits et al., 2010b). Phylogenetic analysis suggests that PhzI may be the long-lost cognate signal synthase for SdiA. The sdiA/phzI gene pair is present in all sequenced *Erwinia* and *Pantoea* spp., whereas a deletion event has apparently removed the luxI homolog from the remaining members of the Enterobacteriaceae analyzed so far (Sabag-Daigle and Ahmer, 2012). Deletion of both sdiA and phzI in *E. amylovora* showed that these genes are not involved in virulence, amylovoran production and motility (Zhao et al., 2009b). The identity between *E. amylovora* PhzI and *Pantoea ananatis* RhI (KKW51348) is 65% at protein level. In the latter species, both C6-HSL and 3-oxo-C6-HSL were shown to be produced and detected by the parallel EanI/EanR
system (Morohoshi et al., 2007), which is not present in *E. amylovora*, and mutation of RhlI was found not to affect the ability of *P. ananatis* to produce short-chain (C4 to C8) HSLs (Sibanda et al., 2016). Altogether, these results may point to the fact that the autoinducer produced by RhlI in *E. amylovora* is not 3-oxo-C6-HSL, but a not yet identified long-chain HSL. In any case, the question remains unsettled given the contradictory results obtained so far.

In several *Vibrio* spp., the LuxS protein was shown to catalyze the production of the AI-2 autoinducer molecule controlling a second quorum sensing system (Schraud et al., 2001). Since the luxS gene was found to be widespread among vast subgroups of the bacterial kingdom, it was hypothesized that AI-2 may constitute the basis of a universal microbial chemical language among different species. However, LuxS has also a primary metabolic role in the activated methyl cycle (AMC), which is involved in the generation of S-adenosyl-L-methionine (SAM), the major methyl donor in the cell, and the recycling of methionine by detoxification of S-adenosyl-L-homocysteine (SAH) (Winzer et al., 2003). The direct AI-2 precursor 4,5-dihydroxy-2,3-pentadione (DPD) is formed as a by-product of the AMC and in Vibroniales, this signal is detected by the two-component sensor kinase LuxPQ. In some species belonging to the families of Enterobacteriaceae, Pasteurellaceae and Bacillaceae (Rezzonico and Duffy, 2008; Xavier and Bassler, 2005), the task of detecting the AI-2 signal is carried out by an ABC-transporter, the Lsr-receptor complex. Most of the species carrying the latter are either pathogens or endosymbions of animals, or live in closed ecosystems in tight association with plants or fungi (Rezzonico et al., 2012a). A thorough approach based on bioinformatics was not able to detect any ortholog of known AI-2 receptors in the genomes of currently sequenced *Erwinia* spp. (Rezzonico and Duffy, 2008; Rezzonico et al., 2012a). This reinforced the results of a previous work (Rezzonico and Duffy, 2007), where no significant AI-2 production was detected in wild-type strains of *E. amylovora* and where mutational analysis could not identify any quorum-sensing related effects in co-culture experiments of wild-type and luxS negative strains. The phenotype of the luxS-mutant strain showed rather that the primary role of LuxS is limited to the AMC and the methionine metabolism (Rezzonico and Duffy, 2007; Smits et al., 2010b). Additionally, after deletion of the luxPQ genes in *E. amylovora*, the deletion mutant was not showing a different phenotype for virulence, amylovoran biosynthesis and motility than the wild-type strain (Zhao et al., 2009b).

**Type VI secretion systems.** Unlike first reports (Hood et al., 2010; Schell et al., 2007), the type VI secretion systems (T6SS) are now regarded as interaction factors rather than virulence factors (Jani and Cotter, 2010; Schwar et al., 2010). *E. amylovora* contains a complete (T6SS-1) and a partial (T6SS-2) T6SS gene cluster (Smits et al., 2010b), a situation also found in other genome-sequenced *Erwinia* and *Pantoea* spp. (De Maayer et al., 2011). *E. amylovora* contains a third T6SS gene cluster (T6SS-3), which only has counterparts in *E. billingiae* Eb661 and several individual strains of *Pantoea* species, including *Pantoea agglomerans* E325 (De Maayer et al., 2011; Sarri et al., 2012). In different liquid media, all three T6SS of *E. amylovora* CFBP 1430 were shown to be transcribed (Kamber et al., 2011).

When deleting the serine kinase (EAMY_3011) in the T6SS-1 of *E. amylovora* Ea1189, there was no influence on virulence, but in contrast to the wild-type strain, motility was irregular (Zhao et al., 2009b). Similar results were obtained when deleting two of the T6SS core genes in *E. amylovora* CFBP 1430, independent if the deletion is in T6SS-1 or T6SS-3 (Kamber et al., 2017). Transcriptome analysis showed differential expression of membrane-related functions, including the T3SSs, iron acquisition, chemotaxis, flagellar, and fimbrial genes, but the deletion of either or both T6SS clusters had only a minor effect on the virulence in planta (Kamber et al., 2017).

**Pathogen self-defense.** There is a dynamic interaction between pathogenic microorganism and their competitors (Duffy et al., 2003). Phytopathogens are able to defend themselves against influences by other bacteria and to the reaction of the plant to their presence. Although largely unexploited, self-defense is a common mechanism for phytopathogens. Especially the reaction of the phytopathogen to biological control organisms has received little attention.

On apple flowers, the main entry site for *E. amylovora*, many bacterial species can be found, several of which are close relatives to the pathogen (Pusey et al., 2009). The members of the genus *Pantoea* are known to be effective agents of biocontrol, which is exerted through several mechanisms: competition for space and nutrients (Smits et al., 2011b; Wilson and Lindow, 1994), acidification of the habitat (Pusey et al., 2008, 2011) or antibiotic production (Stockwell et al., 2002; Vanneste et al., 1992). It is not exactly known how *E. amylovora* reacts to these mechanisms.

**Coexistence on flowers is possible when the available nutrient sources are shared.** Organisms living in a similar niche also will have a similar nutrient utilization pattern. Competition for nutrient sources by biocontrol strains as a mechanism has been hypothesized to depend upon a high niche overlap index (NOI) (Wilson and Lindow, 1994). There is a large overlap in substrate utilization patterns between *E. amylovora* and biocontrol strain *Pantoea vagans* C9-1, with a NOI of 0.96 (Smits et al., 2011b). Compared to the *Pseudomonas fluorescens* A506, the NOI was only 0.71 (Stockwell et al., 2010). In fact, *E. amylovora* is able to compete efficiently with some *Pantoea* strains when it is present on the flower before the biocontrol strain is applied (Giddens et al., 2003).

**Multidrug resistance transporters.** Bacterial multidrug efflux pumps, primarily mediated by secondary transporters that typically utilize the proton motive force,
are common protection tools as a reaction to foreign compounds (Nikaido and Pagès, 2012). In E. amylovora, several multidrug efflux transporters have been characterized up to now. The AcrAB transport system (EAMY_1008 – EAMY_1009) is a protein complex that protects the pathogen from naturally occurring phytoalexins (Burse et al., 2004a). Mutation of acrB in E. amylovora dramatically reduced tolerance to apple phytoalexins as well as virulence on apple plants. AcrAB confers ecological benefits during colonization of the host plant and, thus, influences the virulence of E. amylovora on apple. The transporter NorM (EAMY_1684) confers resistance to several hydrophobic cationic antibiotics (Burse et al., 2004a). Mutation of norM in E. amylovora significantly reduced tolerance to toxins produced by epiphytic P. agglomerans strains isolated from apple and quince blossoms. Therefore, NorM is involved in the counteraction of biocontrol mechanisms on the host plant. The RND family efflux pump AcrD (EAMY_2508) displays only a limited range of amphiphilic compounds that are exported, and there is no effect on virulence (Pletzer and Weingart, 2014a). Two three-component RND transporters, MdtABC (EAMY_2262 – EAMY2264) and MdtUVW (EAMY_0683 – EAMY_0685), are involved in resistance to plant antimicrobials like tannin or flavonoids (Pletzer and Weingart, 2014b). Mutants are impaired in their virulence, as their ability to multiply in apple root stocks was reduced. The different multidrug resistance genes are under the regulation of different sets of global stress regulators, including BaeR and CpxR (Pletzer et al., 2014, 2015), indicating that the ecological role is coupled to environmental signals.

**Streptomycin resistance.** The use of the aminoglycoside streptomycin has long been the only alternative to effectively reduce fire blight (Stockwell and Duffy, 2012). E. amylovora wild type strains are highly sensitive to this antibiotic, and are unable to grow on media amended with concentrations of streptomycin as low as 8 μg ml⁻¹ (Chiou and Jones, 1995). Nevertheless, two different mechanisms of resistance have developed in areas that allowed a relatively uncontrolled use of the antibiotic (McManus and Jones, 1994).

Chromosomal resistance is acquired due to spontaneous single nucleotide mutations in the rpsL gene, which encodes the S12 protein included in the 30S small ribosomal subunit, and result in bacteria that are capable to grow at very high streptomycin concentration exceeding 16,000 μg ml⁻¹ (M. Escursell, unpublished information). These mutations cause a substitution of the existing lysine either at position 43 or 88 of the protein that prevents inhibitory binding of streptomycin while preserving the functionality of the ribosome (Chiou and Jones, 1995). In vitro competition assays have demonstrated that among all the possible mutations examined, only the substitution of lysine by arginine at position 43 (K43R) results in a resistant variant with fitness comparable to that of the wild-type under conditions mimicking the flower environment when the selective pressure is relieved, which explains its prevalence in the field (M. Escursell, unpublished information).

The other resistance mechanism consists in the transposable element Tn3393 containing the gene cluster strA-strB, which encodes a phosphotransferase that enzymatically inactivates streptomycin. This transposon is common to several Gram-negative bacteria and in E. amylovora was found to be incorporated either in plasmid pEa29 or pEa34 (Chiou and Jones, 1993). These strains display a lower resistance to streptomycin (approximately 1000 μg ml⁻¹) with respect to the one carrying the chromosomal mutation and are thus termed medium resistant (Chiou and Jones, 1995). Tn3393-mediated streptomycin resistance has now been reported in three geographically distinct locations in the United States (Förster et al., 2015; McGhee et al., 2011; Tancos and Cox, 2016; Tancos et al., 2016). Concerns that this type of resistance may be horizontally transferred to the target bacteria via formulations contaminated by the resistance genes of the antibiotic producer organism used to prepare them have been cleared by showing their absence in a range of commercial products (Rezzonico et al., 2008).

**Biosynthesis of sulfur-containing compounds.** In recent years, several reports on the biosynthesis of sulfur-containing compounds have been published. Whereas the function of glutathione is well-known as a redox compound, the newly identified compound ovothiol A could very well have a similar function in E. amylovora (Seebeck, 2013). The compound has a very acidic thiol group that can serve as a one-electron donor and has a redox potential in the range of that of protein disulfide isomerasers. Therefore, it might be an efficient scavenger for radicals and peroxides (Braunshausen and Seebeck, 2011), which are the result of the primary defense of apple after challenge with the pathogen. The gene encoding the first step in ovothiol A biosynthesis has been identified as ovoA (EAMY_0003), and encodes an iron(III)- and oxygen-dependent sulfoxide synthase (Mashabela and Seebeck, 2013).

The biosynthesis of the cytotoxic compound 6-thioguanine has been reported already long ago for E. amylovora (Feistner and Staub, 1986), but it is only recently that the biosynthetic cluster (EAMY_1020 – EAMY_1024) has been characterized (Coyne et al., 2013; Wensing et al., 2014). Its role in virulence is contradictory: Coyne et al. have defined a key role for 6-thioguanine in pathogenicity (Coyne et al., 2013), whereas Wensing et al. reported that the compound was not affecting virulence (Wensing et al., 2014). The exact biochemical pathway of 6-thioguanine biosynthesis still remains to be elucidated.

**Siderophores.** Although an essential nutrient for living organisms, iron is only limited available in plants. To obtain sufficient iron for microbial growth, many bacteria
have the ability to produce low-molecular compounds that are able to chelate iron at high affinity (Expert et al., 1996), while other bacteria have the ability to capture these iron siderophores for their own iron metabolism (Deiss et al., 1998). As competition for iron influences the virulence both in human and plant pathogens, siderophores are regarded as a virulence factor (Franza and Expert, 2013).

*E. amylovora* produces the siderophore desferrioxamine E (Feistner et al., 1993) as a reaction to low-iron conditions. Initially also interpreted as a virulence factor (Expert et al., 1996), it soon showed out that desferrioxamine uptake does not affect virulence towards apple, but that desferrioxamine uptake mutants are less able to colonize floral tissues and to cause necrosis (Dellagi et al., 1998). This indicates that desferrioxamine biosynthesis and uptake rather play a general role in colonization, and are as such rather to be regarded as niche adaptation factors in *E. amylovora*.

With the description of the *E. amylovora* genome, the biosynthetic pathway for desferrioxamine E could be fully elucidated (Smits et al., 2010b). Comparative genomic studies have shown that several genome-sequenced species of *Erwinia* and *Pantoea* have the dfoJAC gene cluster for biosynthesis of desferrioxamine E (Smits and Duffy, 2011). As the iron ferrioxamine uptake is also present in these organisms, they would be able to compete for iron-ferrioxamine with *E. amylovora*. With the sequencing of more genomes, it has become evident that not all *Erwinia* species have this potential (Table 1).

**CRISPR and phage defense.** The CRISPR system confers acquired heritable immunity against mobile nucleic acid elements in prokaryotes, limiting phage infection and horizontal gene transfer of plasmids (van der Oost et al., 2009). CRISPR regions are composed by highly conserved direct DNA repeats that are interspersed by unique, similarly sized spacers acquired when the cell is challenged with foreign nucleic acids. Together with the associated Cas proteins, the system confers immunity against infection by foreign DNA or RNA by a mechanism based on the strict identity between the incorporated CRISPR spacer and the nucleic acid target (Brouns et al., 2008). Several spacers sequences pointing to known plasmids or phages of Enterobacteriaceae were found in the CRISPR regions of a series of worldwide isolates of *E. amylovora* (McGhee and Sundin, 2012; Rezzonico et al., 2011). Nevertheless, several novel phages isolated from fruit production environments in Switzerland displayed variable host ranges with respect to the different *E. amylovora* strains analyzed (Born et al., 2011), even if the latter carried identical CRISPR profiles (Rezzonico et al., 2011). This suggests that there is at least an additional mechanism that regulates phages resistance in *E. amylovora*. Given the high genetic homogeneity of the isolates analyzed, this mechanism could reside more in the regulation of certain genes rather than in a specific interaction with the phages, as demonstrated by the preference of certain viruses for either high or low exopolysaccharide producing bacterial hosts (Roach et al., 2013), which can be mediated for example through the production of a phage depolymerase that enables enzymatic bacterial capsule removal (Born et al., 2014).

**APPLIED GENOMICS FOR PHYTOSANITARY CONTROL STRATEGIES**

**Detection and diagnostics.** The current standard for detection of *E. amylovora* in Europe is described in the EPPO protocol, a collection of established and validated protocols (European and Mediterranean Plant Protection Organization – EPPO, 2004). Since then, a large variety of methods improving the detection was developed. A lateral-flow immunostrip assay was developed for specific detection of the fire blight bacterium in symptomatic plant material that can be applied directly on the field (Braun-Kiewnick et al., 2011). Other methods use molecular techniques. Optimization of standard PCR protocols [reviewed in Powney et al. (2011)] and use of quantitative PCR methods [reviewed in Dreo et al. (2012) and Pirc et al. (2009)] are now available and validated. A recent addition to this is the use of loop-mediated isothermal amplification (LAMP), for which three different primer sets based on different targets are developed (Bühlmann et al., 2013; Moradi et al., 2012; Temple and Johnson, 2011).

An important factor is the design of primers for such method. With the availability of genome sequences of a broader range of *Erwinia* spp. and other genera, it is now possible to design assays that are highly specific for *E. amylovora* (Bühlmann et al., 2013). This avoids the detection of closely related species yielding false positive reactions. A good example is the widely used primer set for detection of the SSR on plasmid pEA29 of *E. amylovora* (Llop et al., 2000), the target of which is also present in the plasmid pEP36 of *E. pyrifoliae* (Jock et al., 2003). Even the highly specific LAMP primer set designed by Moradi et al. (2012) using six primers is able to detect *E. pyrifoliae* (Bühlmann et al., 2013); a fact that could easily be avoided by checking the primer set against the genome of the latter species (Smits et al., 2010a). It is thus important to use the maximum of information contained in a genome and to combine it with the maximum number of closely related species to validate the specificity of a primer set.

**Monitoring and epidemiology.** Efficient phytosanitary and control measures must be built upon epidemiological understanding of the fire blight disease cycle. Measures aimed at the protection of commercial pome fruit objects entail identification and containment of inoculum reservoirs, both from existing orchards as well as from alternate host plants from the environment. Recently, new genomics-insights have presented an opportunity to confirm identity of otherwise indistinguishable pathogen isolates (Smits et al., 2010b). The development of this knowledge into
source-tracking methods (Bühlmann et al., 2013; Rezzonico et al., 2011) has the potential to improve the successful implementation of phytosanitary strategies by focusing on removal of only true inoculum sources, while preventing unneeded clearing of old-growth landscape trees, native forest species or unrelated orchards (Gusberti et al., 2015).

Fire blight infection during the blossoming period is a main driver of epidemic development and spread. Application of control products to flowers reduces infection, limits pathogen inoculum build-up, and reduces available inoculum for vector dissemination (Johnson and Stockwell, 1998). Correct timing is critical for efficacy and can be implemented by performing the applications in accordance with established fire blight forecasting models such as Maryblyt or Cougarblight (Smith and Pusey, 2010; van der Zweet et al., 1994). Through the input of weather and phenological variables, it is possible to predict pathogen population growth and inform when potential infection days are likely to occur. However, in current models the presence of the pathogen in a given location is only presumed, so that the use of detection methods for routine flower monitoring to quantify the actual level of inoculum has the potential to reduce unneeded treatments thereby improving the efficacy of control applications (V. Stockwell, unpublished information).

Closely related pathogens that cause symptoms similar to fire blight but are not included in current phytosanitary regulations have recently been described from East Asia (E. pyrifoliæ and Erwinia uzenensis) (Kim et al., 1999; Mat-suura et al., 2012) and Spain (E. piriflorinigrans) (Palacio-Bielsa et al., 2012). Whether these occur more widely in Europe or pose an emerging phytosanitary threat is largely unstudied, a lack due to of efficient methods for detecting these bacteria in plant samples.

**Source-tracking and regulatory strategies.** *E. amylovora* is a relatively genetically homogenous bacterial species: only limited genotype groupings could be obtained across large geographic areas using the most currently available molecular methods like AFLP (Rico et al., 2004), rep-PCR (Rico et al., 2008), PFGE (Jock et al., 2002) or ribotyping (Donat et al., 2007; McManus and Jones, 1995). These methods are thus inadequate for discrimination among isolates within local populations or for source-tracking of inoculum reservoirs and outbreak origins. Recent diversity studies based on the genome sequence of *E. amylovora* CFBP 1430 (Smits et al., 2010b) have exploited the local diversity in clustered regularly interspaced short palindromic repeats (CRISPR). Sequence analysis of CRISPR repeat regions (CRR) showed that there is a larger diversity in *Rubus*-infecting isolates, but that also diversity exists (in the form of three major CRISPR groups) among Spiraeoideae-infecting isolates from North America (McGhee and Sundin, 2012; Rezzonico et al., 2011), which is the geographic center of origin of the disease. However, there are hardly differences within European or Mediterranean isolates, leading to the conclusion that there was an evolutionary bottleneck in the spreading of *E. amylovora* from North America, with only strains belonging to CRISPR group I (typical of the U.S. East coast) having spread to Europe and New Zealand in the 20th century (Rezzonico et al., 2011). A major disadvantage of this technique is the extensive amount of labor needed for sequencing the complete CRRs, which can individually be as long as 6 kb and that may not contain any related repeat, thus requiring the design of ever-new primers to carry out a time-consuming spacer crawling strategy.

A possible alternative to CRISPR is the use of variable number tandem repeats (VNTRs), which are locations in the genome that are composed from short sequences that show variations in length between individuals. In *E. amylovora*, comparison between the genomes of strains CFBP 1430 and ATCC 49946 (Ea273) (Dreo et al., 2011; Smits et al., 2010b) revealed six suitable VNTRs that potentially show enough variability to be employed to analyze the diversity within local isolates of *E. amylovora*. With respect to CRISPRs, automation of multiple locus VNTR analysis (MLVA) is quite straightforward, e.g. in form of labeled fragment analysis, and can be applied to larger sets of isolates providing insights into diversity, biogeography and phylogeny of the pathogen (Bühlmann et al., 2014).

Single nucleotide polymorphisms (SNPs) have the potential to offer the most detailed fingerprint for each strain and are thus the most promising features that can be used as unequivocal phylogenetic markers to pinpoint diversity in genetically monomorphic bacteria (Achtman, 2008; Comas et al., 2009; Holt et al., 2010). As the costs of whole-genome sequencing and re-sequencing techniques is constantly decreasing (Logares et al., 2012), it is foreseeable that the complete variome of hundreds of bacterial isolates will become readily accessible soon, thus making the mapping of SNPs the method of choice to study diversity within a species. In *E. amylovora*, a point mutation in the *galE* gene was able to distinguish isolates of North American origin from those of European/Mediterranean origin, but the analysis was based on a single SNP position and the number of strains considered was very limited (Sauer et al., 2008). The first data available from sequencing 140 isolates of *E. amylovora* are promising, and will be published in due time (T.H.M. Smits, unpublished information), setting the stage for diversity studies through extensive SNP analysis.

**CHALLENGES GOING FORWARD**

The entry into the genomics era has facilitated major advancement on the knowledge of *E. amylovora* and related *Erwinia* spp. This has not only led to an improved understanding of the virulence factors (Kamber et al., 2012; Smits et al., 2011a), but also now allows for complete genome-wide regulation and proteomic studies.
(Holtappels et al., 2016; McNally et al., 2012; Sarowar et al., 2011; Wang et al., 2012b; Wu et al., 2013). The development of highly specific assays to detect E. amylovora in field samples (Bühlmann et al., 2013) and of source tracking assays (Bühlmann et al., 2014; McGhee and Sundin, 2012; Rezzonico et al., 2011) was only possible due to the knowledge gathered from the first fully sequenced genomes. Still, there is a lack of information on the epidemiology of E. amylovora, for which the newly developed methods may deliver the molecular fundamentals. This increased knowledge will, together with improved understanding of the pathogen’s Achilles heel, ultimately lead to novel control strategies.

ACKNOWLEDGEMENTS

This work was supported by the EU FP7 ERA-Net EUPHORESCO project ‘PhytoFire’ (FR), the Swiss Federal Office of Agriculture project ‘Achilles’ within the frame-work of the Swiss ProfiCrops Integrated Project Feuer-brand (THMS), the department of Life Sciences and Fac-ility Management of ZHAW Wädenswil (BD), the Agricul-ture and Food Research Initiative Competitive Grants Program Grants no. 2016-67013-24812 (GWS) and 2015-67013-23068 (GWS) from the USDA National Institute of Food and Agriculture, and Michigan State University AgBioResearch.

REFERENCES


Chiou C.-S., Jones A.L., 1993. Nucleotide sequence analysis of a transposon (Tn3393) carrying streptomycin resistance

Journal of Plant Pathology (2017), 99 (Special issue), 7-23

BMC Plant Biology


Received February 2, 2017
Accepted May 10, 2017