REVIEW

Genomics and current genetic understanding of *Erwinia amylovora* and the fire blight antagonist *Pantoea vagans*

Tim Kamber · Theo H. M. Smits · Fabio Rezzonico ·
Brion Duffy

Abstract The bacterial plant pathogen *Erwinia amylovora* causes fire blight, a major disease threat to pome fruit production worldwide with further impact on a wide-range of *Rosaceae* species. Important factors contributing to the development of the disease were discovered in the last decades. Comparative genomics of the genera *Erwinia* and *Pantoea* is coming into focus with the recent availability of complete genome sequences. Insights from comparative genomics now position us to answer fundamental questions regarding the evolution of *E. amylovora* as a successful pathogen and the critical elements for biocontrol activity of *Pantoea* spp. This trove of new data promises to reveal novel determinants and to understand interactive pathways for virulence, host range and ecological fitness. The ultimate aim is now to apply genomics and identify the pathogen Achilles heels and antagonist mechanisms of action as targets for designing innovative control strategies for fire blight.

Keywords Comparative genomics ·
Type III secretion systems · Exopolysaccharides ·
Virulence factor · Biocontrol

T. Kamber · T. H. M. Smits · F. Rezzonico · B. Duffy (✉)
Plant Protection Division, Agroscope Changins-Wädenswil
ACW, Swiss National Competence Center for Fire Blight,
8820 Wädenswil, Switzerland
e-mail: duffy@acw.admin.ch

Introduction

The enterobacterium *Erwinia amylovora* is the causal agent of the fire blight disease, threatening global pome fruit production (i.e., apple, pear) and a wide-variety of *Rosaceae* (Spiraeoideae, Maloideae, *Rubus*) species, including ecological cornerstone species (e.g., forest, landscape and rural ecosystems). The pathogen first was described in the late 1790s and originated in North America. From there it has relatively recently dispersed to New Zealand in the late 1910s, the United Kingdom and Northern Europe in the late 1950s and the Middle East in the mid-1960s (Bonn and van der Zwet 2000). Since the first reports of fire blight in Europe, the pathogen has continued to spread across the continent (Jock et al. 2002) and now threatens Central Asia, the germplasm region of origin for apple and pear. The quarantine status of *E. amylovora* in many countries imposes further economic losses from phytosanitary control measures and as a highly charged trade (Calvin and Krissoff 1998).

The general epidemiology of fire blight is well understood and is the basis for current control strategies. *E. amylovora* lacks enzymatic means for penetrating healthy host tissues and infects plants through natural openings (e.g., floral nectaries, leaf hydathodes) and via wounds (e.g., from hail or insect damage). Once inside a host, the pathogen can spread in the plant through the vascular system (Billing 2011), and aggressive sanitation is the key to remove inoculum reservoirs (e.g., tree removal) and prevent further advance within infected hosts (e.g., pruning well beyond visible disease symptoms). Dependent on the infected plant part, the disease develops as flower, shoot or rootstock blight. Typical symptoms are flower necrosis, fruit rot, shepherd's crook in shoots, bacterial ooze and cankers in woody tissue. The disease develops

under defined weather conditions that enabled the deployment of fire blight forecast models. Pathogen establishment and rapid infection of flowers (i.e., blossom blight) is the main infection court and epidemic driver of fire blight and thus the primary focus for preventative disease control efforts.

Control options against blossom blight include antibiotic and biocontrol agent application during bloom to reduce the epiphytic population of *E. amylovora*. Resistance to streptomycin (Chiou and Jones 1995), the most effective antibiotic against *E. amylovora*, and regulatory restriction on antibiotic use in plant agriculture (McManus et al. 2002) demands development of novel control measures with comparable efficacy. Natural epiphytic bacteria of the closely related species *Pantoea agglomerans* and *Pantoea vagans* have proven to be among the most reliable and effective antagonists of *E. amylovora* when applied during bloom-time (Stockwell et al. 2010). *P. agglomerans* strains have been isolated from various environments (e.g., soil, plant, water) (Gavini et al. 1989; Rezzonico et al. 2009), reflecting their potential to successfully compete with indigenous microbial populations. The growth inhibition induced by *P. agglomerans* strains might result from nutrient competition, active site exclusion and antibiotic production or the combination of these processes (Kearns and Mahanty 1998; Vanneste et al. 1992; Wodzinski et al. 1994; Pusey et al. 2008).

Although important insights have been acquired regarding this important phytopathogenic bacterium (Table 1), much remains uncertain about the evolutionary genetics of *E. amylovora*. The recent sequencing of eight *Erwinia* genomes (i.e., 3 *E. amylovora*, 2 *Erwinia pyrifoliae*, *Erwinia* sp., *Erwinia billingiae* and *Erwinia tasmaniensis*) (Sebaihia et al. 2010; Smits et al. 2010a, b; Kube et al. 2008, 2010; Park et al. 2011) and three genomes of the closely related genus *Pantoea* (i.e., *P. vagans*, *P. agglomerans* and *Pantoea ananatis*) (Smits et al. 2010c; De Maayer et al. 2010) provides a solid genomics foundation to infer the species evolution within these genera.

Erwinia amylovora genomics

The genome of the *Crataegus* isolate *E. amylovora* CFBP 1430 was sequenced, consisting of a 3.8 Mb chromosome and the 28 kb plasmid pEA29 (Smits et al. 2010b). In total, 3,736 CDS were automatically assigned using GenDB (Meyer et al. 2003) and manually curated. The annotated genome allowed the identification of known virulence factor genes [e.g., hypersensitivity response and pathogenicity (*hrp*), amylovoran biosynthesis gene cluster] of *E. amylovora* in the genome (Oh and Beer 2005), but also included the genes encoding several new putative factors,

like two additional Inv/Spa-type type III secretion systems (T3SSs), a second flagellum and the complete desferrioxamine E biosynthesis cluster (Table 1) (Smits et al. 2010b).

Currently, two further *E. amylovora* genome sequences are available: *Malus* isolate Ea273 (ATCC 49946) and the *Rubus* isolate ATCC BAA-2158 (Powney et al. 2011). Comparison of the Ea273 genome to that of CFBP 1430 clearly shows that two large rearrangements must have occurred within the rRNA regions (Fig. 1) (Smits et al. 2010b) that could explain differences in the PFGE patterns observed before (Zhang and Geider 1997; Jock et al. 2002). Several other, relatively small differences mainly in the ITS regions rendered the chromosome of Ea273 in total 301 bp larger than that of CFBP 1430. Nevertheless, the both genomes share over 99.99% sequence identity over the complete length, indicating only minimal evolution since the geographical dispersal. Plasmid pEA72, identified in the genome of *E. amylovora* Ea273 (Sebaihia et al. 2010), is absent in *E. amylovora* CFBP 1430.

The draft genome of the closely related but genetically distinct *E. amylovora* strain ATCC BAA-2158 with restricted pathogenicity to *Rubus* spp. was recently published (Powney et al. 2011). Also here, collinearity was obtained over large regions of the chromosome. This strain carries, in addition to pEA29, two small plasmids (pEAR4.3, pEAR5.2) (Table 2). A total of 373 singletons in this strain may give indications towards the restricted host range of this strain (Powney et al. 2011).

An obvious difference between the currently sequenced *E. amylovora* genomes is the presence of (cryptic) plasmids (Smits et al. 2010b; Sebaihia et al. 2010; Powney et al. 2011). Plasmids appear to be a major factor influencing the pan-genome of *E. amylovora*. However, although several plasmids of different sizes have been detected in isolates of this species (Chiou and Jones 1991; Foster et al. 2004; Laurent et al. 1989; McGhee et al. 2002; Steinberger et al. 1990), the knowledge on this extra-chromosomal material is limited to few strains and plasmids.

Erwinia inter-species genomics

Another five *Erwinia* genomes were recently sequenced, namely two *E. pyrifoliae* strains (DSM 12163[T] and Ep1/96) (Smits et al. 2010a; Kube et al. 2010), *Erwinia* sp. Ejp617 (Park et al. 2011), *E. tasmaniensis* Et1/99 (Kube et al. 2008) and *E. billingiae* Eb661 (Kube et al. 2010) and are available for inter-species comparisons.

E. pyrifoliae, a close relative of *E. amylovora,* is primarily a pathogen of Asian or Nashi pear (*Pyrus pyrifolia*) with a restricted geographical distribution to East Asia (Kim et al. 1999). Disease symptoms caused by *E. pyrifoliae* are almost indistinguishable to those of *E. amylovora*

Table 1 Genes and gene clusters assessed on potential impact on pathogenicity

	Gene(s)	Impact on pathogenicity	Assessment	References	Locus tag[a]
Type 1 secretion system	prtADEF	No	Immature pear fruits, apple seedlings	Zhang et al. (1999)	EAMY_3577-3581
Type 2 secretion system	outCDEFFHJIKLMNOS-chiV	No	Immature pear fruits, apple seedlings	Zhao et al. (2009)	EAMY_2865-2878
Type 3 secretion system	hrpN	Yes	Immature pear fruits, no HR in tobacco leaves	Wei et al. (1992)	EAMY_0552
	dspA/E	Yes	Immature pear fruits, apple and cotoneaster shoot	Bogdanove et al. (1998)	EAMY_0557
	hrpW	No	Immature pear fruits and apple and pear seedlings	Kim and Beer (1998)	EAMY_0556
	hrpK	No	Immature pear fruits, apple shoots, HR tobacco plants	Oh et al. (2005)	EAMY_0519
	hrpA	Yes	Apple seedlings, no HR tobacco leaves	Jin et al. (2001)	EAMY_0542
	hrpJ	Yes	Immature pear fruits, reduced HR tobacco leaves	Nissinen et al. (2007)	EAMY_0535
	orfB (eopB, eop1)	No	Immature pear fruit assay	Asselin et al. (2006)	EAMY_0554
Additional T3SS effectors	eop2 (hopAK1)	?		Nissinen et al. (2007)	EAMY_0653
	eop3	?		Nissinen et al. (2007)	EAMY_2270
	hopPtoC	No	Immature pear	Zhao et al. (2005)	EAMY_0744
	avrRpt2	Reduced	Immature pear	Zhao et al. (2006)	EAMY_3175
Chaperones	dspB/F	Reduced	Pear seedlings	Gaudriault et al. (2002)	EAMY_0558
	orfA	No	Immature pear fruits	Asselin et al. (2006)	EAMY_0553
T3SS associated	hsvABC	Yes apple/no pear	Immature pear fruits, apple shoots, HR in tobacco	Oh et al. (2005)	EAMY_0520-0522
Regulators	hrpY	Yes	Immature pear fruits, no HR in tobacco	Wei et al. (2000)	EAMY_0538
	hrpX	Reduced	Immature pear fruits, no HR in tobacco	Wei et al. (2000)	EAMY_0537
	hrpS	Yes	Immature pear fruits, no HR in tobacco	Wei and Beer (1993)	EAMY_0539
	hrpL	Yes	Immature pear fruits, no HR in tobacco	Wei and Beer (1995)	EAMY_0536
Inv/Spa-type	PAI-2	No	Immature pear fruits/apple seedlings	Zhao et al. (2009)	EAMY_0771-0792
	PAI-3	No	Immature pear fruits/apple seedlings	Zhao et al. (2009)	EAMY_1573-1593
Iron uptake, siderophores	dfoJAC	Yes flowers/no shoots	Apple seedlings/apple flower	Dellagi et al. (1998)	EAMY_3238-3240
	foxR	Yes	Apple seedlings/apple flower	Dellagi et al. (1998)	EAMY_3241
Metabolism					
Sorbitol	srlAEBDMR	Yes apple/no pear	Immature pear/apple shoots	Aldridge et al. (1997)	EAMY_3071-3076
Exopolysaccharide					
Levansucrase	lsc	Reduced	Pear seedlings	Geier and Geider (1993)	EAMY_3695
Amylovoran	amsABCDEFGHI	Yes	Pear slices/pear seedlings	Bellemann and Geider (1992)	EAMY_2241-2253
Regulator	rscBCD	Yes	Immature pear fruits	Wang et al. (2009)	EAMY_2342-2343

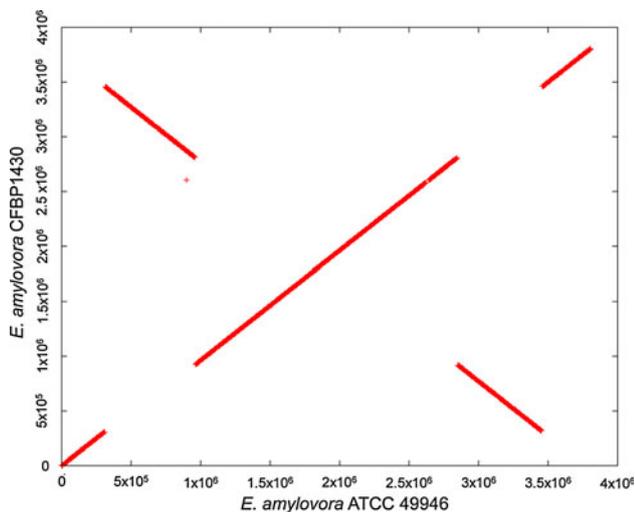[a] Locus tags derive from *E. amylovora* CFBP 1430 (accession number FN434113)

Fig. 1 Synteny plot of *E. amylovora* strains CFBP 1430 and ATCC 49946 generated using EDGAR (Blom et al. 2009). The position of each CDS given on the *X* axis is plotted against the position of its ortholog in the other chromosome given on the *Y* axis. Identical gene organization on the chromosomes results in a diagonal plot, inversions and chromosomal rearrangements are plotted perpendicular to it

(Rhim et al. 1999). Another pathogenic *Erwinia* species, causing bacterial shoot blight of pear in Japan, was first described as *E. amylovora* due to similar disease symptoms and was later found to be closer related to *E. pyrifoliae* than to *E. amylovora* (Mizuno et al. 2000; Matsuura et al. 2007; Geider et al. 2009). *E. tasmaniensis* strains were isolated from apple and pear as non-pathogenic epiphytic bacteria (Geider et al. 2006; Powney et al. 2011). The non-pathogenic *E. billingiae* was isolated as non-pigmented *E. herbicola* and later reclassified (Billing and Baker 1963; Mergaert et al. 1999).

Chromosomal collinearity of *E. amylovora* to the closely related *E. pyrifoliae* strains, *Erwinia* sp. Ejp617 and *E. tasmaniensis* was observed, but chromosomal large-scale rearrangements were detected. These ecologically distinct species harbor distinct plasmids with low sequence similarity. The genome sizes of the above described *Erwinia* species range from approximately 3.8–4 Mb with the exception of *E. billingiae* having a genome size of 5.4 Mb (Table 2). The genome sizes of the sequenced pathogenic *Erwinia* spp. are markedly smaller as compared to other enterobacterial genomes, likely the effect of genome erosion (Georgiades and Raoult 2011). The loss of epiphytic fitness factors due to genome size reduction and the acquisition of the Hrp-type T3SS and genes needed for the biosynthesis of the exopolysaccharides levan and amylovoran are potentially a result of adaption to the pathogenic lifestyle.

Differential gene content can be displayed in Venn diagrams, calculated by reciprocal BLAST of all CDS to

all the input sequences. Areas displayed in such a Venn diagram represent a subset of the compared genomes and the number of genes is indicated by numbers. The core genome (Tettelin et al. 2005) consists of CDS shared by all genome sequences, which usually includes metabolic and cellular functions. CDS shared by two or several species are present in overlapping regions. Areas not shared by any other pool of CDS (singletons) include accessory genes which may provide additional functions (virulence factors, host-range determinants, metabolic processes) and contribute to species variability (Blom et al. 2009).

The core genome of *E. amylovora* strain CFBP 1430, *E. tasmaniensis* Et1/99, *E. pyrifoliae* DSM 12163^T and the non-pathogenic *E. billingiae* Eb661 displays 2414 genes shared between these species. The genes only shared by *E. amylovora* CFBP 1430, *E. tasmaniensis* Et1/99 and *E. pyrifoliae* DSM 12163^T, in different combinations, include the most important virulence factors, e.g., T3SSs and the exopolysaccharides amylovoran and levan (Table 3). These genes are absent in *E. billingiae* Eb661 and might therefore represent the "pathogenic core" of disease eliciting *Erwinia* species (Fig. 2a). Virulence determinants setting host-range and specificity are most likely included in the singletons for the broad host range *E. amylovora*, and absent in the genomes of *Erwinia* species with restricted host-ranges (Fig. 2a).

Selected features clarified using comparative genomics

Type III secretion systems

T3SSs are part of the "pathogenic core" and are absent in *E. billingiae* Eb661. The Hrp T3SS genes were identified in *E. pyrifoliae* DSM 12163^T (Smits et al. 2010a) and *E. tasmaniensis* Et1/99 (Kube et al. 2008) showing differences in gene content compared to *E. amylovora* CFBP 1430 (Smits et al. 2010b). *E. tasmaniensis* Et1/99 lacks the HAE region present in the other two species, as well as ORFU1 and ORFU2 that are only present in *E. amylovora* CFBP 1430 (Fig. 3a).

Proteins of *E. amylovora* secreted by the Hrp-T3SS have been demonstrated to be essential for pathogenicity on host-plants (Table 1) (Oh and Beer 2005). The Hrp T3SS gene cluster is located on pathogenicity island 1 (PAI-1) and consists of the *hrp/hrc* region, flanked by the Hrp effectors and elicitors (HEE) region and the Hrp-associated enzymes (HAE) region. The Hrp/Hrc region contains regulatory genes as well as genes encoding for secreted proteins. DspA/E and HrpN encoded by genes in the HEE are secreted proteins essential pathogenicity factors of *E. amylovora* (Gaudriault et al. 1997; Wei et al. 1992).

Table 2 Summary of sequenced genomes of *Erwinia* and *Pantoea*

Species/strain	Host/origin	Status	Genome	Size (Mb)	CDS	References
E. amylovora CFBP 1430	*Crataegus sp.* (Hawthorn), France	Finished	Chromosome	3.806	3,706	Smits et al. (2010b)
			pEA29	0.028	28	
E. amylovora ATCC 49946	*Malus domestica* (Apple), USA	Finished	Chromosome	3.806	3,483	Sebaihia et al. (2010)
			pEA72	0.071	87	
			pEA29	0.028	28	
E. amylovora ATCC BAA-2158	*Rubus* sp. (Blackberry), USA	Draft	Chromosome	3.81	3,831	Powney et al. (2011)
			(29 contigs)			
			pEA29	0.028	28	
			pEAR5.2	0.005	6	
			pEAR4.3	0.004	4	
E. pyrifoliae DSM 12163^T	*Pyrus pyrifolia* (Asian pear, Nashi), South Korea	Finished	Chromosome	4.026	3,986	Smits et al. (2010a)
			pEP36	0.036	40	
			pEP5	0.005	7	
			pEP3	0.003	4	
			pEP2.6	0.003	1	
E. pyrifoliae Ep1/96	*Pyrus pyrifolia* (Asian pear, Nashi), South Korea	Finished	Chromosome	4.026	3,645	Kube et al. (2010)
			pEP36	0.036	37	
			pEP5	0.005	5	
			pEP3	0.003	4	
			pEP2.6	0.003	6	
Erwinia sp. Ejp617	*Pyrus pyrifolia* (Asian pear, Nashi), Japan	Finished	Chromosome	3.909	3,600	Park et al. (2011)
			pEJP30.8	0.031	34	
			pEJP6.4	0.006	0	
			pEJP5.2	0.005	6	
			pEJP3.2	0.003	0	
			pEJP2.6	0.003	0	
E. tasmaniensis Et1/99	*M. domestica* (Apple), Australia	Finished	Chromosome	3.883	3,427	Kube et al. (2008)
			pET49	0.049	61	
			pET46	0.046	39	
			pET45	0.045	46	
			pET35	0.035	42	
			pET09	0.009	7	
E. billingiae Eb661	*P. communis* (Pear), UK	Finished	Chromosome	5.1	4,587	Kube et al. (2010)
			pEB170	0.17	220	
			pEB102	0.102	114	
P. vagans C9-1	*M. domestica* (Apple), USA	Finished	Chromosome	4.025	3,665	Smits et al. (2010c)
			pPag1	0.168	162	
			pPag2	0.166	229	
			pPag3	0.53	535	
P. agglomerans E325	*M. domestica* (Apple), USA	Draft	Genome	4.774	4,495	Smits and Duffy (unpublished)
P. ananatis LMG 20103	*Eucalyptus grandis* × *E. nitens* Hybrid A. (Eucalyptus), South Africa	Finished	Chromosome	4.69	4,272	De Maayer et al. (2010)

**Table 3** Selected factors analyzed by comparative genomic approaches

Species (strain)	T3SS			Exopolysaccharides			Flagellar		Siderophores			Sorbitol
	Hrp	Inv/Spa-1	Inv/Spa-2	AMS	CPS	*lsc*	Flg1	Flg2	*foxR/A*	*dfoJAC*	Ent	Srl
E. amylovora (CFBP 1430, ATCC 49946)	+	+	+	+	−	+	+	+	+	+	−	+
E. pyrifoliae (DSM 12163[T], Ep1/96)	+	−	+	+	−	−	+	+	+	+	−	+
Erwinia sp. (Ejp617)	+	−	+	+	−	−	+	+	+	+	−	+
E. tasmaniensis (Et1/99)	+[a]	+[b]	+	−	+	+	+	−	+	+	−	−
E. billingiae (Eb661)	−	−	−	−	+	−	+	−	+	+	−	+
P. vagans (C9-1)	−	−	−	−	+	−	+	−	+	+	+	+
P. agglomerans (E325)	−	−	−	−	+	−	+	−	+	+	+	−
P. ananatis (LMG 20103)	−	−	−	−	+	−	+	−	+	+	−	+

Presence or absence of genes and gene cluster are indicated by the symbols + and −, respectively

[a] HAE absent

[b] Partial

Fig. 2 Venn diagram of *Erwinia* spp. (**a**) and *Pantoea* spp. including *E. billingiae* (**b**) generated using EDGAR. The numbers of CDS is indicated. Overlapping areas indicate shared CDS. The "pathogenic" (Fig. 2a)—and "biocontrol" core (Fig. 2b), respectively, are indicated by *dotted lines*
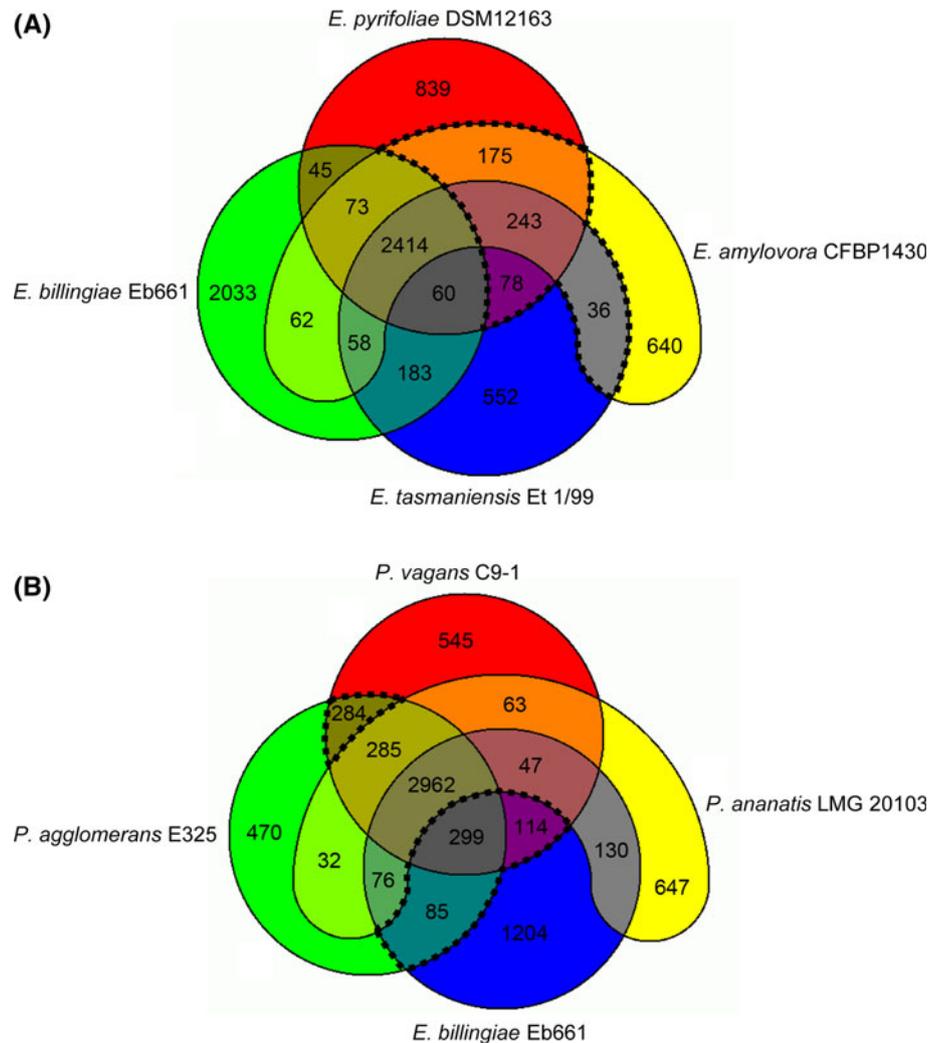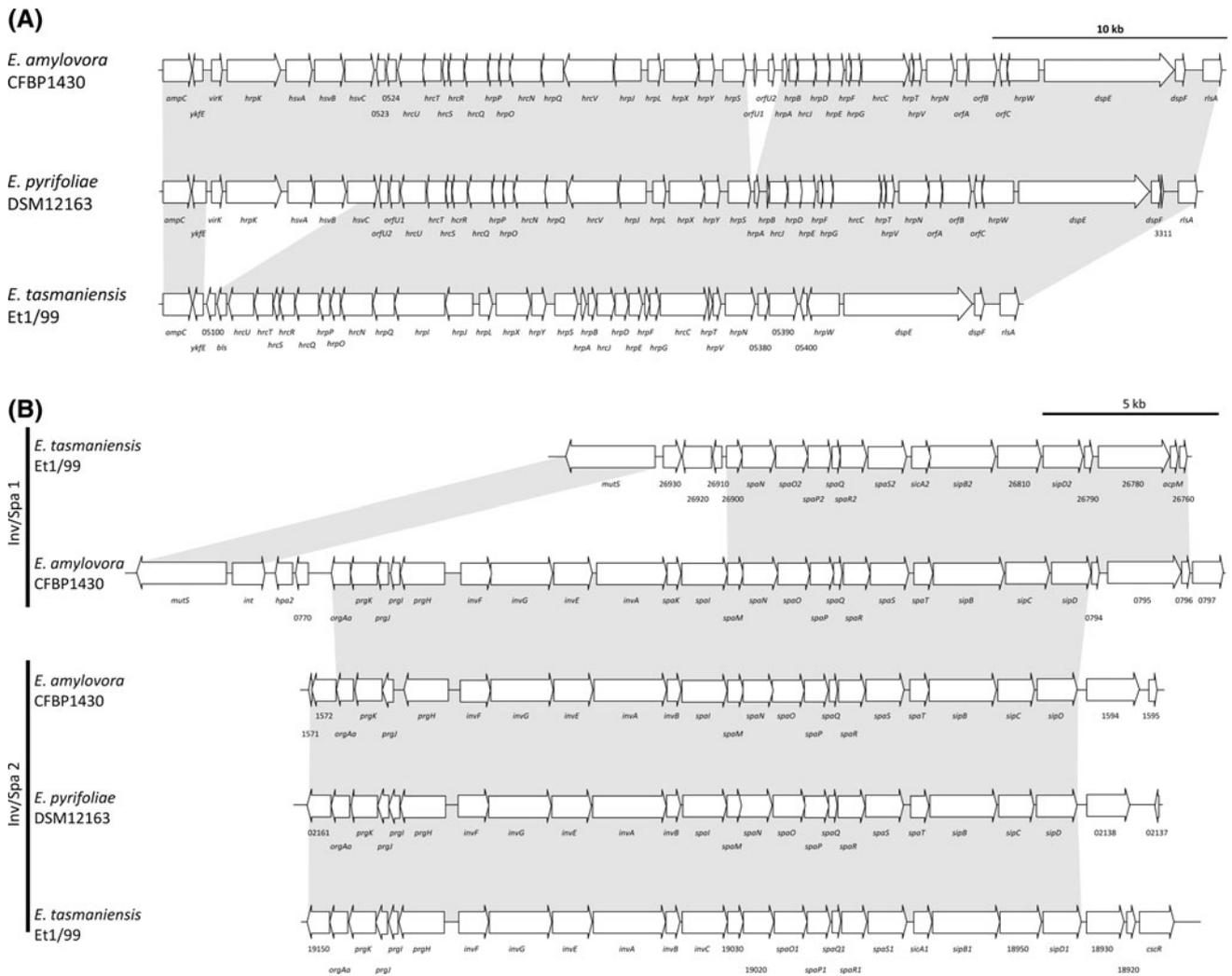
Fig. 3 Comparison of the Hrp (**a**) and Inv/Spa-type (**b**) Type III Secretion Systems in *E. amylovora* CFBP 1430, *E. pyrifoliae* DSM 12163[T] and *E. tasmaniensis* Et1/99. Related genes are shaded *grey*

The products of the hrp-associated systemic virulence genes (*hsv*) encoded in the HAE region are required for systemic infection of host-plants (Oh et al. 2005).

Additional T3SSs (Inv/Spa-1, Inv/Spa-2) were identified in the genomes of the pathogenic *Erwinia* spp. that differ in gene content. While the *inv/spa-2* gene cluster is present in *E. amylovora* CFBP 1430, *E. pyrifoliae* DSM 12163[T] and *E. tasmaniensis* Et1/99, the *inv/spa-1* gene cluster is absent in *E. pyrifoliae* DSM 12163[T] and only partially present in *E. tasmaniensis* Et1/99 (Fig. 3b). The *inv/spa*-type T3SSs are located in low G + C regions on the chromosome of the three *Erwinia* spp. and might therefore be an acquired trait. The *inv/spa*-type T3SSs are similar to the *Salmonella* pathogenicity island SPI1 T3SS of *Salmonella typhimurium* LT-2 (McClelland et al. 2001) and the *inv/spa* T3SS of the insect endosymbiont *Sodalis glossinidius* (Dale et al. 2001) and are not directly implicated in virulence on host-plants (Zhao et al. 2009).

Exopolysaccharides

The *E. amylovora* CFBP 1430 and *E. pyrifoliae* DSM 12163[T] exopolysaccharide gene clusters differ from the respective clusters in *E. tasmaniensis* Et1/99 and *E. billingiae* Eb661 by an exchange of two glycosyltransferases (Fig. 4) resulting potentially in the production of amylovoran. Amylovoran biosynthesis is a specific virulence factor in *E. amylovora* and *E. pyrifoliae* reflected in the fact that deletion or mutagenesis of specific genes renders the pathogens avirulent (Bellemann and Geider 1992; Kim et al. 2002). The exopolysaccharide gene clusters, producing CPS of the two non-amylovoran producing species and also of *Pantoea* spp., might represent the ancestral state.

The additional exopolysaccharide levan is only produced by *E. amylovora* and *E. tasmaniensis*, whereas not by *E. billingiae* and *E. pyrifoliae*. The gene encoding the levansucrase, responsible for the synthesis of levan, most

Fig. 4 Comparison of the exopolysaccharide gene clusters of *Erwinia* spp. and *Pantoea* spp. Related genes are *shaded*
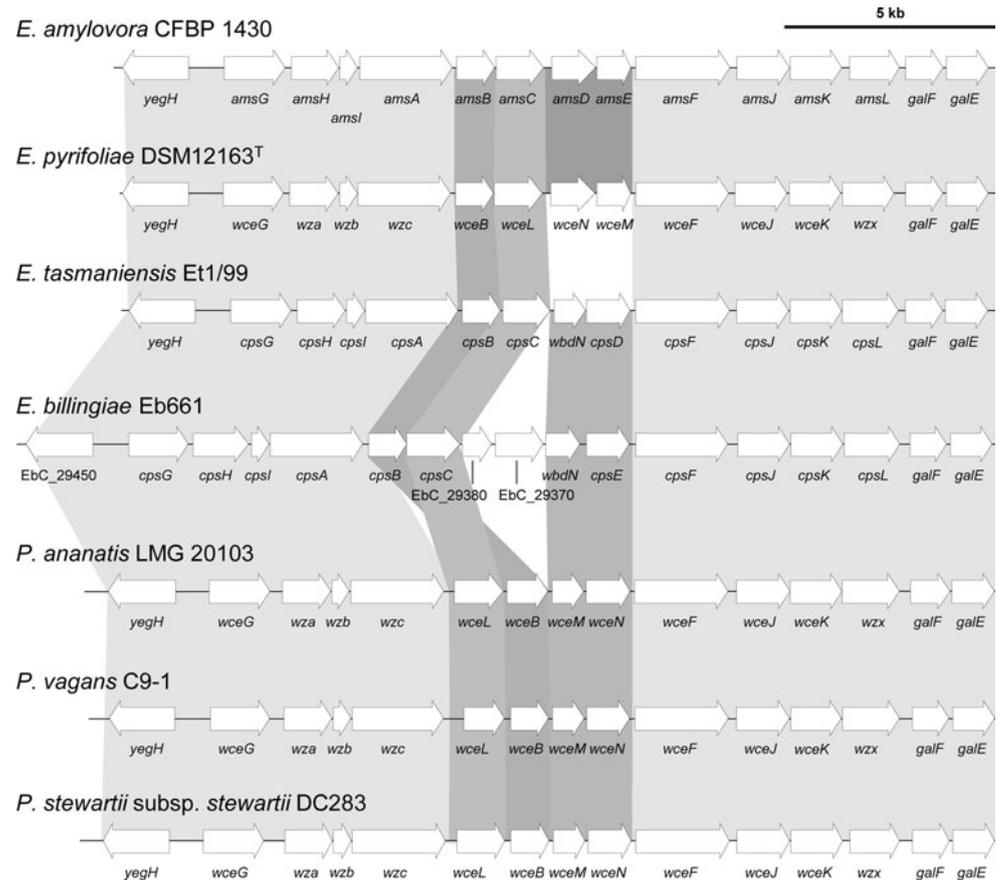
likely was acquired by a common ancestor of *E. amylovora*, *E. tasmaniensis* and *E. pyrifoliae*. The gene is retained by *E. amylovora* and *E. tasmaniensis*, whereas lost by *E. pyrifoliae*.

Siderophores

All so far sequenced genomes of *Erwinia* spp. contain the desferrioxamine E siderophore biosynthesis gene cluster (Kube et al. 2010; Smits et al. 2010a, b), whereas the enterobactin gene cluster, producing the catecholate siderophore enterobactin found in the genomes of many enterobacteria is absent.

Iron is an essential nutritional factor, required as cofactor for many proteins. In iron deprived environments high-affinity iron uptake siderophores are secreted to the environment to gain access to this limited factor by removing it from minerals and organic substances. *Erwinia* spp. produce the hydroxamate siderophore desferrioxamine E (Feistner et al. 1993; Kachadourian et al. 1996) and the specific TonB-dependent ferrioxamine receptor FoxR, both involved in iron uptake. Mutation of these genes leads to colonization defects of *E. amylovora* on flowers (Dellagi et al. 1998), whereas DFO E might be protective to oxidative conditions (Venisse et al. 2003).

Phylogenomic applications

A core genome tree was constructed (Fig. 5) which displays the phylogeny of the genus *Erwinia* which is in accordance to trees based on *gyrB* sequences. On both trees, *E. billingiae* groups to the *Erwinia* spp. and is the most distantly related *Erwinia* species sequenced until now. The position of *E. billingiae* seems to be close to the genus delineation between *Erwinia* and *Pantoea*. The
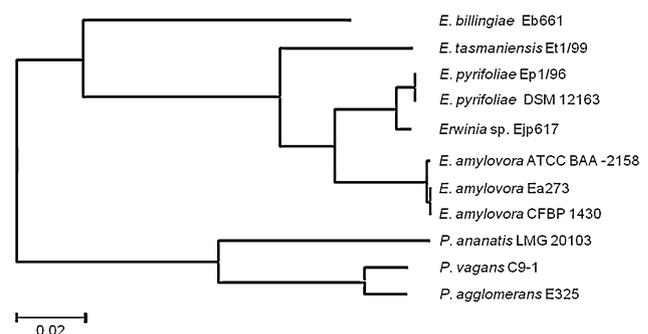

Fig. 5 Phylogenetic tree based on the 2022 core genes of *Erwinia* and *Pantoea* spp. generated using EDGAR. Percent divergence is indicated by the *scale bar*

marked difference in genome size, approximately 5.4 Mb *E. billingiae* Eb661 versus the nearly 4 Mb for the other *Erwinia* spp., might be the result of genome size reduction during specification to plant pathogenicity in the latter species. The Hrp, Inv/Spa-1, Inv/Spa- 2 T3SSs, absent in the genome of *E. billingiae*, might have been acquired by the pathogenic *Erwinia* ancestor, prior the divergence of *E. amylovora*, *E. pyrifoliae* and *E. tasmaniensis*.

Erwinia billingiae

The genera *Erwinia* and *Pantoea* are closely related (Hauben et al. 1998), which is supported by the core genome tree (Fig. 5) and the fact that gene synteny across the genus border is retained for large regions. *E. billingiae* first was isolated and described as non-pigmented *E. herbicola* (Billing and Baker 1963) [now *P. agglomerans* (Gavini et al. 1989)] and later reassigned to *E. billingiae* (Mergaert et al. 1999).

The exopolysaccharides amylovoran and levan are not produced by *E. billingiae*, but a capsular polysaccharide (CPS) similar to that of *Pantoea* spp. is formed. Compared to *Pantoea* spp., the CPS gene cluster in *E. billingiae* Eb661 has an inversion of two genes that is distinct for the members of the genus *Erwinia*. The number of CDS shared by other *Erwinia* species and *E. billingiae* (Fig. 2a) in the core genome is smaller (approximately 500 CDS) than the core genome calculated when *E. billingiae* is included with *Pantoea* species (Fig. 2b). The genes shared with *Pantoea* include many carbohydrate uptake and utilization pathways and phosphonate utilization; genes absent in other *Erwinia* spp. These genes, possibly involved in the epiphytic fitness, were lost in the course of genome size reduction towards a more specific plant-associated or pathogenic lifestyle.

Pantoea biocontrol agent genomics

The genome of the biocontrol agent *P. vagans* C9-1 (Ishimaru et al. 1988; Rezzonico et al. 2009, 2010) was recently sequenced, consisting of a 4,025 Mb circular chromosome and the three plasmids pPag1, pPag2 and pPag3 (Smits et al. 2010c). A total of 4,619 CDS were assigned using GenDB (Meyer et al. 2003) and manually annotated. Genome sequencing of *P. agglomerans* E325 (Pusey et al. 2008; Pusey 1997), a second biocontrol agent, is in progress (Smits and Duffy, unpublished). The genome sequences of both *Pantoea* spp. lack known enterobacterial virulence determinants such as T3SSs, toxins and pectolytic enzymes. A large repertoire of carbohydrate metabolic pathways, epiphytic fitness genes (e.g., AI-1 quorum sensing genes, IAA and carotenoid biosynthesis) and the biosynthetic genes for the antibacterial metabolite pantocin

A were identified in the genome sequence of *P. vagans* C9-1 (Smits et al. 2010c).

Pantoea comparative genomics

Additional to the two biocontrol strains, the genome of *P. ananatis* LMG 20103, the causative agent of *Eucalyptus* blight and dieback, was sequenced (De Maayer et al. 2010) and was therefore included in comparative genomics analysis. Although being described as a plant pathogen (Goszczynska et al. 2006, 2007), its genome sequence lacks T3SSs, a major virulence factor in many plant-associated pathogens, rendering this organism an unusual plant pathogen (Coutinho and Venter 2009).

The calculated core genome of *P. ananatis* LMG 20103, *P. vagans* C9-1, *P. agglomerans* E325 and *E. billingiae* Eb661 (Fig. 2b) includes, additionally to the genes of the *Erwinia* core genome, carbohydrate metabolic pathways for maltose, rhamnose, glucarate, xylose, uronate, L-lactate, acetate as well as utilization of phosphonates, and many transporters (all types). Genes absent in the genome of the plant pathogen *P. ananatis* LMG 20103, but shared by the two other *Pantoea* and/or *E. billingiae* might represent the "biocontrol core" (Fig. 2b) due to the fact that genes potentially contributing to epiphytic fitness (e.g., nitrate assimilation, enterobactin synthesis genes) are common to these groups. Other factors implied in biocontrol efficacy, such as antibiotic biosynthesis, are not shared between the *Pantoea* spp. since they produce different antibiotics (Pusey et al. 2008; Coutinho and Venter 2009; Ishimaru et al. 1988). The antibiotic pantocin A biosynthesis genes, for example are only present in *P. vagans* C9-1, whereas absent in *P. ananatis* LMG 20103, *P. agglomerans* E325 and *E. billingiae* Eb661. The biosynthesis genes of *P. vagans* C9-1 are located on a low-G + C genomic island of about 29 kb, which was probably acquired by horizontal gene transfer (Smits et al. 2010c, d). The exopolysaccharide gene clusters of the *Pantoea* spp. are similar, whereas the *E. billingiae* cluster differs by inversion of two genes (Fig. 4).

Perspectives

The available sequenced *Erwinia* genomes from different hosts enable the identification of virulence, host-specificity and metabolic determinants involved in pathogenicity by comparative genomic analysis. The analyses could yield the information needed to develop novel control measures for the fire blight disease. Genome sequencing and analysis of *Pantoea* spp. will reveal their potential and, for the already successfully used biocontrol agents, uncover the factors (metabolism, antibiotic production) responsible for

effective biocontrol. As more *Erwinia* and *Pantoea* genomes get sequenced, these can be used to consolidate the current data as well as refine evolutionary aspects.

References

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

Asselin JE, Oh CS, Nissinen RM, Beer SV (2006) The secretion of EopB from *Erwinia amylovora*. Acta Hort 704:409–415

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

Billing E (2011) Fire blight. Why do views on host invasion by *Erwinia amylovora* differ? Plant Pathol 60:178–189

Billing E, Baker LAE (1963) Characteristics of *Erwinia*-like organisms found in plant material. J Appl Bacteriol 26:58–65

Blom J, Albaum SP, Doppmeier D, Puhler A, Vorholter FJ, Zakrzewski M, Goesmann A (2009) EDGAR: a software framework for the comparative analysis of prokaryotic genomes. BMC Bioinformatics 10:154

Bogdanove AJ, Kim JF, Wei ZM, Kolchinsky P, Charkowski AO, Conlin AK, Collmer A, Beer SV (1998) 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

Bonn WG, van der Zwet T (2000) Distribution and economic importance of fire blight. In: Vanneste JL (ed) Fire blight: the disease and its causative agent, *Erwinia amylovora*. CAB International, Wallingford, pp 37–53

Calvin L, Krissoff B (1998) Technical barriers to trade: a case study of phytosanitary barriers and US-Japanese apple trade. J Agr Resour Econ 23:351–366

Chiou CS, Jones AL (1991) The analysis of plasmid-mediated streptomycin resistance in *Erwinia amylovora*. Phytopathology 81:710–714

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

Coutinho TA, Venter SN (2009) *Pantoea ananatis*: an unconventional plant pathogen. Mol Plant Pathol 10:325–335

Dale C, Young SA, Haydon DT, Welburn SC (2001) The insect endosymbiont *Sodalis glossinidius* utilizes a type III secretion system for cell invasion. Proc Natl Acad Sci USA 98:1883–1888

De Maayer P, Chan WY, Venter SN, Toth IK, Birch PR, Joubert F, Coutinho TA (2010) Genome sequence of *Pantoea ananatis* LMG20103, the causative agent of *Eucalyptus* blight and dieback. J Bacteriol 192:2936–2937

Dellagi A, Brisset MN, Paulin JP, Expert D (1998) Dual role of desferrioxamine in *Erwinia amylovora* pathogenicity. Mol Plant Microb Interact 11:734–742

Feistner GJ, Stahl DC, Gabrik AH (1993) Proferrioxamine siderophores of *Erwinia amylovora*—a capillary liquid-chromatographic electrospray tandem mass-spectrometric study. Org Mass Spectrom 28:163–175

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 the Hrp secretion pathway in a DspB-dependent way. Mol Microbiol 26:1057–1069

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

Gavini F, Mergaert J, Beji A, Mielcarek C, Izard D, Kersters K, Deley J (1989) Transfer of *Enterobacter agglomerans* (Beijerinck 1888) Ewing and Fife 1972 to *Pantoea* gen. nov. as *Pantoea agglomerans* comb. nov. and description of *Pantoea dispersa* sp. nov. Int J Syst Bacteriol 39:337–345

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

Geider K, Auling G, Jakovljevic V, Völksch B (2009) A polyphasic approach assigns the pathogenic *Erwinia* strains from diseased pear trees in Japan to *Erwinia pyrifoliae*. Lett Appl Microbiol 48:324–330

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

Georgiades K, Raoult D (2011) Genomes of the most dangerous epidemic bacteria have a virulence repertoire characterized by fewer genes but more toxin-antitoxin modules. PLoS One 6:e17962

Goszczynska T, Moloto VM, Venter SN, Coutinho TA (2006) Isolation and identification of *Pantoea ananatis* from onion seed in South Africa. Seed Sci Technol 34:655–668

Goszczynska T, Botha WJ, Venter SN, Coutinho TA (2007) Isolation and identification of the causal agent of brown stalk rot, a new disease of maize in South Africa. Plant Dis 91:711–718

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

Ishimaru CA, Klos EJ, Brubaker RR (1988) Multiple antibiotic production by *Erwinia herbicola*. Phytopathology 78:746–750

Jin QL, Hu WQ, 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

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

Kachadourian R, Dellagi A, Laurent J, Bricard L, Kunesch G, Expert D (1996) Desferrioxamine-dependent iron transport in *Erwinia amylovora* CFBP1430: cloning of the gene encoding the ferrioxamine receptor FoxR. Biometals 9:143–150

Kearns LP, Mahanty HK (1998) Antibiotic production by *Erwinia herbicola* Eh1087: its role in inhibition of *Erwinia amylovora* and partial characterization of antibiotic biosynthesis genes. Appl Environ Microbiol 64:1837–1844

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 Bacteriol 180:5203–5210

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 49:899–905

Kim WS, Schollmeyer M, Nimtz M, Wray V, Geider K (2002) Genetics of biosynthesis and structure of the capsular exopolysaccharide from the Asian pear pathogen *Erwinia pyrifoliae*. Microbiology 148:4015–4024

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, Migdoll AM, Gehring I, Heitmann 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 Genomics 11:393

Laurent J, Barny MA, Kotoujansky A, Dufriche P, Vanneste JL (1989) Characterization of a ubiquitous plasmid in *Erwinia amylovora*. Mol Plant Microb Interact 2:160–164

Matsuura T, Shinohara H, Inoue Y, Azegami K, Tsushima S, Tsukamoto T, Mizuno A (2007) *Erwinia* isolates from the bacterial shoot blight of pear in Japan are closely related to *Erwinia pyrifoliae* based on phylogenetic analyses of *gyrB* and *rpoD* genes. J Gen Plant Pathol 73:53–58

McClelland M, Sanderson KE, Spieth J, Clifton SW, Latreille P, Courtney L, Porwollik S, Ali J, Dante M, Du F, Hou S, Layman D, Leonard S, Nguyen C, Scott K, Holmes A, Grewal N, Mulvaney E, Ryan E, Sun H, Florea L, Miller W, Stoneking T, Nhan M, Waterston R, Wilson RK (2001) Complete genome sequence of *Salmonella enterica* serovar Typhimurium LT2. Nature 413:852–856

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

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

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

Meyer F, Goesmann A, McHardy AC, Bartels D, Bekel T, Clausen J, Kalinowski J, Linke B, Rupp O, Giegerich R, Pühler A (2003) GenDB—an open source genome annotation system for prokaryote genomes. Nucleic Acids Res 31:2187–2195

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

Nissinen RM, Ytterberg AJ, Bogdanove AJ, Van Wijk KJ, 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

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, 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

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

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

Pusey PL (1997) Crab apple blossoms as a model for research on biological control of fire blight. Phytopathology 87:1096–1102

Pusey PL, Stockwell VO, Rudell DR (2008) Antibiosis and acidification by *Pantoea agglomerans* strain E325 may contribute to suppression of *Erwinia amylovora*. Phytopathology 98:1136–1143

Rezzonico F, Smits THM, Montesinos E, Frey JE, Duffy B (2009) Genotypic comparison of *Pantoea agglomerans* plant and clinical strains. BMC Microbiol 9:204

Rezzonico F, Vogel G, Duffy B, Tonolla M (2010) Application of whole-cell matrix-assisted laser desorption ionization-time of flight mass spectrometry for rapid identification and clustering analysis of *Pantoea* species. Appl Environ Microbiol 76:4497–4509

Rhim SL, 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

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

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 Genomics 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, Kamber T, Goesmann A, Ishimaru CA, Stockwell VO, Frey JE, Duffy B (2010c) Genome sequence of the biocontrol agent *Pantoea vagans* strain C9–1. J Bacteriol 192:6486–6487

Smits THM, Rezzonico F, Pelludat C, Goesmann A, Frey JE, Duffy B (2010d) Genomic and phenotypic characterization of a nonpigmented variant of *Pantoea vagans* biocontrol strain C9-1 lacking the 530-kb megaplasmid pPag3. FEMS Microbiol Lett 308:48–54

Steinberger EM, Cheng GY, Beer SV (1990) Characterization of a 56-kb plasmid of *Erwinia amylovora* Ea322: its noninvolvement in pathogenicity. Plasmid 24:12–24

Stockwell VO, Johnson KB, Sugar D, Loper JE (2010) Control of fire blight by *Pseudomonas fluorescens* A506 and *Pantoea vagans* C9-1 applied as single strains and mixed inocula. Phytopathology 100:1330–1339

Tettelin H, Masignani V, Cieslewicz MJ, Donati C, Medini D, Ward NL, Angiuoli SV, Crabtree J, Jones AL, Durkin AS, Deboy RT, Davidsen TM, Mora M, Scarselli M, Margarit y Ros I, Peterson JD, Hauser CR, Sundaram JP, Nelson WC, Madupu R, Brinkac LM, Dodson RJ, Rosovitz MJ, Sullivan SA, Daugherty SC, Haft DH, Selengut J, Gwinn ML, Zhou L, Zafar N, Khouri H, Radune D, Dimitrov G, Watkins K, O'Connor KJ, Smith S, Utterback TR, White O, Rubens CE, Grandi G, Madoff LC, Kasper DL, Telford JL, Wessels MR, Rappuoli R, Fraser CM (2005) Genome analysis of multiple pathogenic isolates of *Streptococcus agalactiae*: implications for the microbial "pan-genome". Proc Natl Acad Sci USA 102:13950–13955

Vanneste JL, Yu J, Beer SV (1992) Role of antibiotic production by *Erwinia herbicola* Eh252 in biological control of *Erwinia amylovora*. J Bacteriol 174:2785–2796

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

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

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

Wei ZM, Beer SV (1995) *hrpL* activates *Erwinia amylovora* hrp gene transcription and is a member of the ECF subfamily of σ 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 ZM, 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 Microb Interact 13:1251–1262

Wodzinski RS, Umholtz TE, Rundle JR, Beer SV (1994) Mechanisms of inhibition of *Erwinia amylovora* by *Erw. Herbicola* in vitro and in vivo. J Appl Bacteriol 76:22–29

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

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

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 DP (2009) Construction and analysis of pathogenicity island deletion mutants of *Erwinia amylovora*. Can J Microbiol 55:457–464