Genes Genom (2017) 39:945–955
DOI 10.1007/s13258-017-0559-y

Online ISSN 2092-9293
Print ISSN 1976-9571

RESEARCH ARTICLE

Complete genome sequence and pathogenic genes analysis of *Pectobacterium atroseptica* JG10-08

Dai Zhang[1] · Yuan Zhou[1] · Dongmei Zhao[1] · Jiehua Zhu[1] · Zhihui Yang[1] · Mingming Zhu[1]

Received: 9 August 2016 / Accepted: 30 April 2017 / Published online: 30 May 2017

Abstract *Pectobacterium atroseptica* is known as a rod-shaped gram-negative bacterial pathogen associated with the blackleg of potato. *P. atroseptica* has been widely identified as the predominant agent causing tuber rot in temperate regions, a disease that leads to severe economic losses to potato industry. In this study, we provide the complete genome sequence of *P. atroseptica* JG10-08, which revealed that *P. atroseptica* strain JG10-08 carries a single 5,004,926 bp chromosome with 51.15% G+C content and harbors 4252 predicted coding genes. Phylogenetic analysis based on the genome sequences showed a close evolutionary relationship between *P. atroseptica* and *Pectobacterium wasabiae*. We discovered total 168 genes were potentially related to pathogenesis including 9 strain-specific genes encoding toxins on the genome of JG10-08. Further comparison with other species in *Pectobacterium* revealed a better understanding of pathogenic factors, especially secretion systems in *P. atroseptica* JG10-08. Collectively, the results of this research provide a solid foundation for discovering the underlying pathogenic mechanisms of *P. atroseptica* and offer the information to develop more effective strategies against blackleg of potatoes.

Keywords *Pectobacterium atroseptica* · Genome sequence · Blackleg of potato · Tuber rot · Pathogenic factors

✉ Jiehua Zhu
zhujiehua356@126.com

[1] College of Plant Protection, Agricultural University of Hebei, 2596 Lekai South Street, Lianchi District, Baoding, Hebei, China

Introduction

Potato (*Solanum tuberosum*) is considered as the fourth main food crop in the world and widely cultivated because no special growth environment is required (Larkin and Honeycutt 2006). Potato is surpassed only by rice (*Oryza sativa*), wheat (*Triticum aestivum*) and maize (*Zea mays*) in terms of crop production and cultivated areas (King and Slavin 2013). However, various types of pathogens including bacteria, virus, fungi and pests can cause severe economic losses to potato yields (Hagland 2011). Blackleg of potato, representing one of most important bacterial diseases, is caused by several bacterial species belonging to the soft rot *Enterobacteriaceae* (SRE) family. It has been documented that *Pectobacterium carotovorum* subsp. *brasiliense* (*Pcb*) (Duarte et al. 2004), *Pectobacterium atrosepticum* (*Pa*) (Gardan et al. 2003), *Pectobacterium carotovorum* subsp. *carotovorum* (*Pcc*), *Pectobacterium wasabiae* (*Pwa*) (Pitman et al. 2008) and several *Dickeya* spp. are identified as causal agents for potato blackleg in the field (Toth et al. 2011; van der Wolf et al. 2014). However, *P. atrosepticum* (formerly named *Erwinia carotovora* subsp. *atrosepticum*) (Hauben et al. 1998), only restricted to potato, is regarded as the predominant blackleg pathogen occurred in temperate regions, which is mostly characterized by symptoms such as black rot lesion (Gardan et al. 2003). Currently, no methods are effective to control the disease caused by *P. atrosepticum*. There are no available chemical agents to prevent the spread of these pathogens. In addition, planting patterns and storage conditions are not efficient against the disease (Czajkowski et al. 2011; Yaganza et al. 2014).

Except of *Pectobacterium* and *Dickeya*, the bacterial family *Enterobacteriaceae* also includes many well-studied human pathogens such as *Shigella, Yersinia*, and

Salmonella, as well as the model species *Escherichia coli*. Many genome sequences have been reported from this family (Parkhill et al. 2001a, b; Perna et al. 2001; Welch et al. 2002; Duchaud et al. 2003). Especially, the genome sequence analysis of *E. coli* K-12 offers crucial information for genome development, because this strain is widely referred as the standard model strain in almost all areas of biological researches (Blattner et al. 1997). While the genome analysis of bacterium in *Enterobacteriaceae* has provided the comprehensive data for scientific researches, limited genetic information of *Pectobacterium* genus, playing a key role in *Enterobacteriaceae*, has been collected.

In present, totally seven genome sequences of *Pectobacterium* are now available and enable us to perform the comparative genomics analysis on these species including two strains of *P. carotovorum* (PC1 and PCC21), two strains of *P. atrosepticum* (SCRI1043 and CFBP6276) (Bell et al. 2004a; Kwasiborski et al. 2013), three strains of *P. wasabiae* (WPP163, CFBP3304 and SCC3193) (Koskinen et al. 2012; Nykyri et al. 2012). In previous researches on these strains, pathogenic genes encoding effector proteins and the secretion system have an effect on plant disease. Moreover, soft rot pathogenesis basically relies on the plant cell wall degrading enzymes (PCWDEs) that can lead to extensive tissue maceration (Expert 1999; Franza et al. 2002). However, the mechanism may be more complex and subtle than previous theoretical results (Mulholland et al. 1993; Jones et al. 1999; Kang et al. 2002; Toth et al. 2003). Here we present the complete genome sequence of *P. atrosepticum* JG10-08 to reveal its molecular pathogenesis, which could supplement further references for pathogenic mechanisms. This research could also provide a key source of new genetic materials and present a target for biological therapy.

To date, blackleg disease of potato is becoming prevalent in potato-growing regions in China. However, the genome sequence of *P. atrosepticum* strain isolated from China remains unknown. Here we present the complete genome sequence of *P. atrosepticum* JG10-08 obtained from infected potato tubers, performed the comparative genomics analysis with those of five annotated SRE and identified several strain-specific genes that potentially contribute to virulence.

Materials and methods

Bacterial strain and genomic DNA extraction

Pectobacterium atroseptica JG10-08 was isolated from the potato tuber with blackleg disease symptoms in Heibei, China. This strain was typically incubated in Luria–Bertani (LB) liquid medium at 25 °C for 48 h. Genomic DNA was extracted from cultured bacteria using CTAB method (Hyman et al. 2000).

Sequencing the whole genome

The genome sequencing of *P. atrosepticum* JG10-08 was performed by Illumina HiSeq2000 platform (2×100 bp) and the total sequencing coverage was 95-fold. The obtained high-quality paired-end reads were assembled by using SOAP denovo (http://soap.genomics.org.cn) and SOAP GapCloser was also applied to close the gaps after assembly (Luo et al. 2012). Subsequently, a draft genome with 25 scaffolds was generated. The gaps between scaffolds were closed via PCR and Sanger sequencing.

Glimmer was used to determine and functionally categorize the predicted protein-coding sequences (Delcher et al. 1999). The genome sequences were annotated through Rapid Annotations using Subsystem Technology (RAST) (Aziz et al. 2008), identified by performing manual NCBI BLAST searches, and compared with the coding sequences (CDSs) of the genome of *P. atrosepticum* SCRI1043, *P. carotovorum subsp. carotovorum* PC1, *P. carotovorum subsp. carotovorum* PCC21, *P. wasabiae* SCC3193 and *P. wasabiae* WPP163. Comparative genome analyses for functions were conducted using Clusters of Orthologous Group of proteins (COGs).

Phylogenetic analyses

Phylogenetic relationships were analyzed based on both the genome sequences and 16S rRNA genes among six *Pectobacterium* species and one *Yersinia* species, including *P. atrosepticum* JG10-08 (GenBank accession no. CP007744.1), *P. atrosepticum* SCRI1043 (BX950851.1), *P. carotovorum* subsp. *carotovorum* PC1 (CP001657.1), *P. carotovorum* subsp. *carotovorum* PCC21 (CP003776.1), *P. wasabiae* SCC3193 (CP003415.1), *P. wasabiae* WPP163 (CP001790.1) and *Yersinia pestis* CO92 (CP009973.1). The ClustalW software was used to align the gene sequences after trimming to remove ambiguously aligned regions. The phylogenetic tree was performed using the neighbor-joining method of Phylip.

According to the results of phylogenetic analyses, strain SCRI1043, PCC21 and WPP163 were chosen to compare with the strain JG10-08 on synteny using Mauve. Next, the similarity of locally collinear blocks (LCB) was obtained through Gene Nees.

Analysis of related pathogenic genes

The genome sequences and the function annotation database of *P. atrosepticum* SCRI1043, *P. carotovorum* subsp. *carotovorum* PC1, *P. carotovorum* subsp. *carotovorum*

PCC21, *P. wasabiae* SCC3193 and *P. wasabiae* WPP163 were downloaded by the NCBI website. Through the sequence analysis, a large number of main pathogenic genes could be obtained. According to the annotated functions of genes, we predicted the relative genes coding the toxins, PCWDEs and six secretion systems for these five pathogenic bacteria and JG10-08.

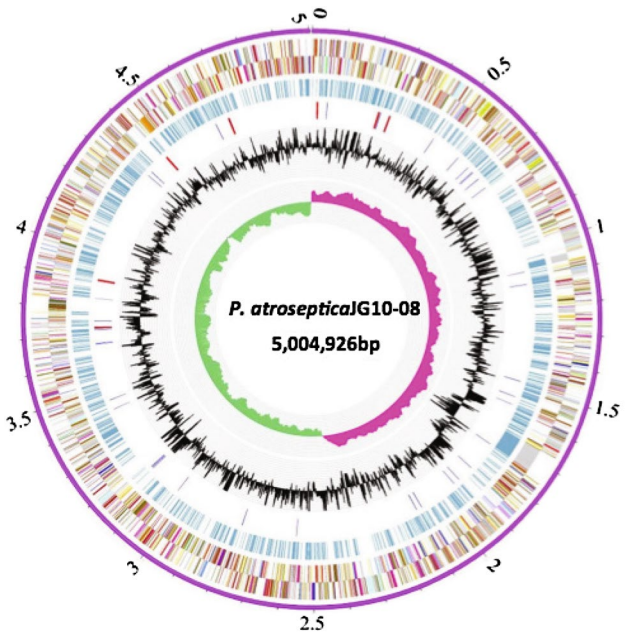

Fig. 1 Circular representation of the *P. atrosepticum* JG10-08 genome. The *outer scale* shows the genome sequences. From the outside to inside circle, *circle 1* and *2* indicate COG annotated coding sequences. *Circle 3* shows KEGG enzymes. *Circle 4* shows RNA genes. *Circle 5* represents the GC content (%) of the *P. atrosepticum* JG10-08, and GC skew are shown in *circle 6*

Results and discussion

General genomic features

The genome of *P. atrosepticum* JG10-08 contains a 5,004,926 bp long circular chromosome with no plasmid (Fig. 1). The average G+C content of the genome is 51.15%, which is similar to that of *P. wasabiae* (50.53%), but is lower than that of *P. carotovorum* (52.02%) (Table 1).

A total of 4672 CDSs were annotated on the *P. atrosepticum* JG10-08 genome using Glimmer 2.0. This finding indicated that the predicted CDSs account for 86.1% of total genome length (average, 941.9 bp). These annotated genes are transcribed in the positive and negative directions from the perspective of the direction of DNA replication, respectively. The *P. atrosepticum* genome encodes 76 tRNA operons and 39 rRNA genes, which are distributed in seven sets of 16S-26S-5S rRNA operon regions.

To further determine the difference in functions encoded by these 4672 genes, we analyzed the data using clusters of orthologous group of proteins (COGs). Our results revealed that 4252 (91%) predicted genes of *P. atrosepticum* were assigned by the COG categories (Table 2). Among these assigned genes, 43.27% of the genes are related to metabolism, 21.05% to cellular processes and signaling, and 16.56% to information storage and processing. However, 19.12% of the genes cannot be assigned in COG categories because their features and functions remain obscure. The COGs category is extremely essential for classifying and illustrating the annotated gene data of a complete genome as evidenced by predicting the functions of protein families (Tatusov et al. 1997).

Comparison of the genome sequences of *P. atrosepticum* JG10-08 with those of other *Pectobacterium* spp

To examine the relationships of *P. atrosepticum* JG10-08 with previously-sequenced strains within *Pectobacterium*

Table 1 The genomic features of sequenced genomes of *Pectobacterium* strains

Species	Bacterium	Length (Mb)	GC%	Average GC%
Pectobacterium atrosepticum	JG10-08	5.00	51.15	50.34
	SCRI1043	5.06	51.00	
	SFBP6276	4.87	51.09	
Pectobacterium wasabiae	WPP163	5.06	50.50	50.53
	CFBP3304	5.15	50.60	
	CFIA1002	5.01	50.60	
	SCC3193	5.16	50.40	
Pectobacterium carotovorum	PC1	4.86	51.90	52.02
	PCC21	4.84	52.20	
	PBR1692	4.92	51.90	
	WPP14	4.82	52.00	

Table 2 Distribution of genes associated with the 25 general COG functional categories

Category	Code	Description	Annotated gene	
			No.	%
Metabolism	C	Energy production and conversion	241	5.67
	E	Amino acid transport and metabolism	456	10.72
	F	Nucleotide transport and metabolism	81	1.90
	G	Carbohydrate transport and metabolism	349	8.21
	H	Coenzyme transport and metabolism	157	3.69
	I	Lipid transport and metabolism	108	2.54
	P	Inorganic ion transport and metabolism	326	7.67
	Q	Secondary metabolites biosynthesis, transport and catabolism	122	2.87
Cellular processes and signaling	D	Cell cycle control, cell division, chromosome partitioning	41	0.96
	M	Cell wall/membrane/envelope biogenesis	241	5.67
	N	Cell motility	110	2.59
	O	Posttranslational modification, protein turnover, chaperones	141	3.32
	T	Signal transduction mechanisms	200	4.70
	U	Intracellular trafficking, secretion, and vesicular transport	118	2.78
	V	Defense mechanisms	44	1.03
	W	Extracellular structures	0	0
	Y	Nuclear structure	0	0
	Z	Cytoskeleton	0	0
Information storage and processing	J	Translation, ribosomal structure and biogenesis	181	4.26
	A	RNA processing and modification	1	0.02
	B	Chromatin structure and dynamics	0	0
	K	Transcription	344	8.09
	L	Replication, recombination and repair	178	4.19
Poorly characterized	R	General function prediction only	499	11.74
	S	Function unknown	314	7.38
	X	Not annotated	0	0
Total			4252	100

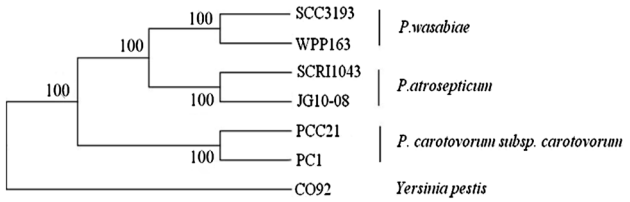

Fig. 2 Phylogenetic relationship of *P. atrosepticum* JG10-08 genome sequence with those of five other *Pectobacterium* strains and one *Y. pestis* strain. *P. atrosepticum* JG10-08 is closely related with *P. atrosepticum* SCRI1043

genera, we downloaded the genome sequences of five other *Pectobacterium* strains and one *Y. pestis* strain from NCBI, and performed the phylogenetic analysis (Fig. 2). Phylogenetic tree based on the whole genome revealed that *P. atrosepticum* has the closest relationship with *P. wasabiae*. The host range and survival conditions of *P. carotovorum* are more extensive than those of *P. atrosepticum* and *P. wasabiae*, probably leading to distant association with them. Unexpectedly, the relationships among different species in *Pectobacterium* based on phylogenetic analysis of 16S rRNA remains inaccurate (Fig. 3) because the strain SCC3193 was more distantly related with the strain WPP163. Previous findings based on the comparison of proteomes of all sequenced soft rot bacterium (Nykyri et al. 2012) showed that SCC3193 previously classified into *P. arotovorum* is grouped into *P. wasabiae*. Our result based on phylogenic tree of the genome sequences is in line with this finding. Thus, phylogenetic analysis at the whole genome level provided a strong support and an accurate classification for the species.

To further compare the genome structures of these sequenced strains within *Pectobacterium* genera, the whole genome sequences were compared using Mauve. We found that the location of genes in *P. atrosepticum* JG10-08 was different from that of *P. wasabiae* WPP163 and *P. carotovorum* PCC21. The aligned genes of *P. atrosepticum* JG10-08 are mostly oriented in forward direction relative to the genome sequences, highly similar to *P. atrosepticum*

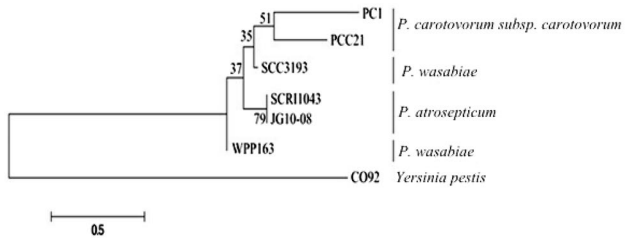

Fig. 3 Phylogenetic relationship of *P. atrosepticum* JG10-08 based on 16S rRNA gene with those of five other *Pectobacterium* strains and one *Y. pestis* strain. *P. atrosepticum* JG10-08 is closely related with *P. atrosepticum* SCRI1043

SCRI1043, while those of *P. wasabiae* WPP163 and *P. carotovorum* PCC21 are oriented in forward and reverse complementary direction (Fig. 4). In comparison to SCRI1043, we found there is no gene insertion or deletion of large fragments in *P. atrosepticum* JG10-08, only two loci of genes inversion occurred, as well as the frequent gene rearrangement. In addition, we calculated the sequence similarity of LCB among these four strains using Gene nees software. The genome arrangement of JG10-08 displays almost same synteny with SCRI1043 and differs by only 7.0% in the pairwise alignment. However, the JG10-08 differs by 47.7 and 49.4% from PCC21 and WPP163, respectively. The SCRI1043 differs by 48.3 and 49.1% from PCC21 and WPP163, respectively. The differences between *P. wasabiae* WPP163 and *P. carotovorum* PCC21 in the pairwise alignments are much larger than the differences between other strains. Taken together, the analysis of the pairwise alignment supports the previous finding that JG10-08 and SCRI1043 belong to the same species. Moreover, *P. atrosepticum* strains share more similarities with *P. carotovorum* than *P. wasabiae* strains. We assume that in the process of evolution, the genomes of strain JG10-08 and SCRI1043 are relatively stable, which results in close relationship. But, there are quite numbers of changes in LCB in *P. wasabiae* WPP163 and *P. carotovorum* PCC21 during species evolution.

Pathogenic factors

Total 168 genes associated with pathogenicity were identified in the genome of *P. atrosepticum* JG10-08. Among them, 25 genes encoded plant cell wall degrading enzymes (PCWDs), 22 genes were related to toxins and 121 genes were involved in secretion system (Table 3). Statistics analysis revealed that the number of genes encoding toxins among these three species of bacterium is similar. However, there was an extremely significant divergence in the number of genes encoding PCWDEs. Moreover, the number of genes involving in secretion system and pathogenicity in three species was relative in 1% confidence interval. However, there was significant difference between *P. carotovorum* (a) and *P. atrosepticum* (b) as well as between *P. carotovorum* (a) and *P. wasabiae* (b) at 5% different level.

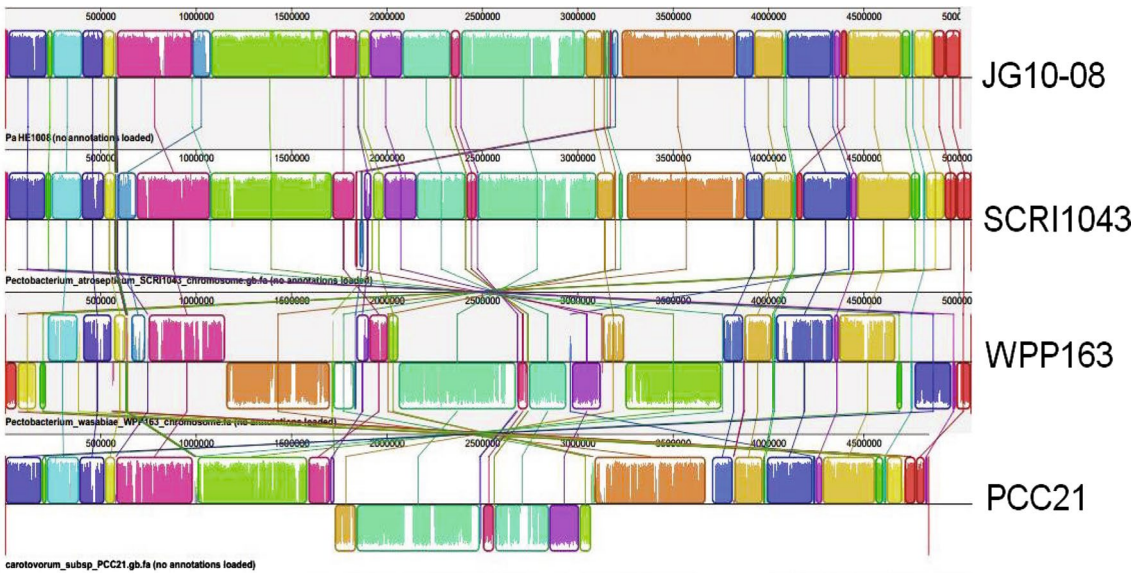

Fig. 4 Synteny analysis of *P. atrosepticum* JG10-08 and SCRI1043, *P. wasabiae* WPP163 and *P. carotovorum* PCC21 genomes. Pairwise alignments of genomes were generated using Mauve. The sequence similarity in the pairwise alignment of *P. atrosepticum* JG10-08 and SCRI1043 was 93.0%. The similarity between *P. atrosepticum* JG10-08 and *P. carotovorum* PCC21 was 52.3% and between *P. atrosepticum* JG10-08 and *P. wasabiae* WPP163 was 50.6%. The sequence similarity in the pairwise alignment of *P. atrosepticum* SCRI1043 and *P. carotovorum* PCC21 was 51.7%, between *P. atrosepticum* SCRI1043 and *P. wasabiae* WPP163 was 50.9% and between *P. wasabiae* WPP163 and *P. carotovorum* PCC21 was 46.3%

Table 3 The number of pathogenic genes distributed in six *Pectobacterium* strains

Virulence factors	Number of genes involving in virulence factors					
	Pectobacterium atrosepticum JG10-08	*Pectobacterium atrosepticum* SCRI 1043	*Pectobacterium wasabiae* WPP163	*Pectobacterium wasabiae* SCC3193	*Pectobacterium carotovorum* PC1	*Pectobacterium carotovorum* PCC21
Secretion system	121	124	106	140	201	193
PCWDs	25	25	16	16	41	41
Toxin	22	13	10	10	16	16
Total	168	162	132	166	258	250

The number of total pathogenic genes and genes involving in secretion of *P. carotovorum* was notably higher than that of *P. atrosepticum* and *P. wasabiae* (Table 4), which are consistent with the above observations on genetic relationship among these three species.

Genes encoding toxins

Compared with the *P. atrosepticum* SCRI1043, nine specific virulence genes were identified uniquely to *P. atrosepticum* JG10-08, including *YafW, YefM, YkfI, YoeB, RelE, RelB, StbD, StbE* and *Phd*. It has been reported that the toxin proteins encoded by *YafW, YefM, YkfI* and *YoeB* are able to directly induce plant diseases (Christensen et al. 2004; Harrison et al. 2009). The *RelE, RelB, StbD* and *StbE* genes play a key role in replication and stability of toxins (Takagi et al. 2005; Li et al. 2009; Unterholzner et al. 2013). The *Phd* gene encodes toxin protein, which is associated with the suppression of host defense system. Although the functions of these nine genes remain to be determined, our data provide a framework for investigating the pathogenic system of *P. atrosepticum*.

Plant cell wall-degrading enzymes are essential for pathogenesis of *P. atrosepticum*

The *P. atrosepticum* JG10-08 genome contains number of PCWDE genes whose products are released to extracellular space of host. These proteins play a crucial role in three distinct pathogenic ways including degradation, nutrition and feedback regulation (Yang et al. 2007). The pathogens benefit from the nutrients produced after degradation, and these degradation products accumulated in the host can induce bacterium to generate more enzymes. Therefore, the production of PCWDEs is the hallmark to soft rot pectobacteria infection. We identified a total of 25 known or putatively related genes encoding pectinases, cellulases and proteinases.

Pectinesterase plays a key role in the pathogenicity of *P. atrosepticum*. It can utilize the intermediate layer and pectin in cell wall to cause the death of the host tissues. Interestingly, there are two more pxo-polygalacturonate lyase-coding genes in *P. atrosepticum* JG10-08 compared with strain SCRI1043 (Table 5). Through n-blast software we found that these two genes started from the position of 2,349,204 and 2,630,527 bp on the genome, and the length is 1706 and 1631 bp, respectively. However, *P. atrosepticum* SCRI1043 carries these two genes without annotation. Thus, it is essential to determine the functions of them in the future study.

Endo polygalacturonate lyase (pel) secreted by bacteria is one of the most crucial virulence factors for plant cell wall degradation. The *pel* gene cluster consists of three different genes including *pelA, pelB* and *pelC* in *P. atrosepticum* JG10-08, which is the same as that in *P. atrosepticum* SCRI1043. The number of genes in pel clusters is significantly different among species and subspecies. For example, *pelA, pelB, pelC* and *pelD* were carried in Pcc, the pel

Table 4 The significant difference analysis in the total number of pathogenic genes of *Pectobacterium* genus

Species	Number of genes encoding the virulence factors			
	Secretion system	PCWDs	Toxin	Pathogenic genes
Pectobacterium carotovorum	197 a(A)	41 a(A)	16 a(A)	254 a(A)
Pectobacterium atrosepticum	122.5 b(A)	24.5 b(B)	17.5 a(A)	165 b(A)
Pectobacterium wasabiae	123 b(A)	16 c(C)	10 a(A)	149 b(A)

Lowercase letters (a, b, c) represent the difference in 5% confidence interval, uppercase letters (A, B, C) represent the extremely significant difference in 1% confidence interval. The same letter indicates that the difference is not significant, while different letters reveal the significant differences

Table 5 The distribution of different kinds of pectinesterase in *P. atrosepticum* strain JG10-08 and SCRI1043

Pectinesterase	No.	Substrate	No.	
			JG10-08	SCRI1043
Pectin acetylesterase	EC 3.1.1.6	Pectin	2	2
Pectin methyl esterase	EC 3.1.1.11	Pectin	2	2
Polygalacturonase	EC 3.2.1.15	Pectic acid	3	3
Pxo-polygalacturonase	EC 3.2.1.82	Pectic acid	1	1
Endo polygalacturonate lyase	EC 4.2.2.2	Pectic acid	8	8
Pxo-polygalacturonate lyase	EC 4.2.2.9	Pectic acid	3	1
Oligosaccharides of the enzyme	EC 4.2.2.6	D-Galacturonic acid	1	1
Pectin lyase	EC 4.2.2.10	Pectin	2	2
Total	–	–	22	20

The "EC" represents the serial number named by International Enzyme Commision. The first digit represents the classification name of the enzyme; the second digit represents the sub class; the third digit represents a small group in a subclass; the fourth digit represents the arranging number in the sub-subclass of the enzyme

cluster in *Dickeya dadantii* contains more than five genes (*pelA, pelB, pelC, pelD* and *pelE*), as well as at least four secondary genes (*pelI, pelL, pelZ* and *pelX*). Moreover, secondary *pel* genes are found to be involved in the protein expression. It is likely that the host range of *Pa* is single. But, infected plants by *Pcc* and *Dickeya dadantii* range widely.

Secretion system

In gram-negative bacteria, six types of secretion systems including I–VI are used by bacteria to export many extracellular enzymes and abundant of effector proteins (Fig. 5) (Lory 1998), which are also conserved in *P. atrosepticum* JG10-08. Through the secretion systems, effectors can be transported inside the plant cell where they can manipulate the host processes to facilitate bacterial infection (Holst et al. 1996). Type I secretion system (T1SS) mainly contains TolC, HlyB and HlyD proteins, and ABC systems (Davidson and Chen 2004). Type II and V secretion systems depend on Sec systems and known as a two-step transportation. Previous studies have proved that pectinases and cellulases are secreted via the type II secretion system (T2SS), and its inactivation renders *Pectobacterium* virulence (Reeves et al. 1993; Sandkvist 2001; de Chial et al. 2003; Filloux 2004). Type II system, mainly composed of common secretion and *Sec* proteins encoded by 15 *Gsp* gene clusters, is essential for the bacteria. *AidA* gene (3110 bp) is contained both in *P. atrosepticum* JG10-08 and SCRI1043. However, there is no evidence that this gene has been found in the other species in *Pectobacterium* based on nblast comparison. Type V system contains the auto transporter and two-partner secretion system (Henderson et al. 1998), and further experiments are needed to determine the predictable function of *AidA* gene related to adhesion in JG10-08. Bacteria also possess type VI system, which is recently discovered as a new secretion mechanism. Previous findings showed that two genes *vasK* and *vasH*

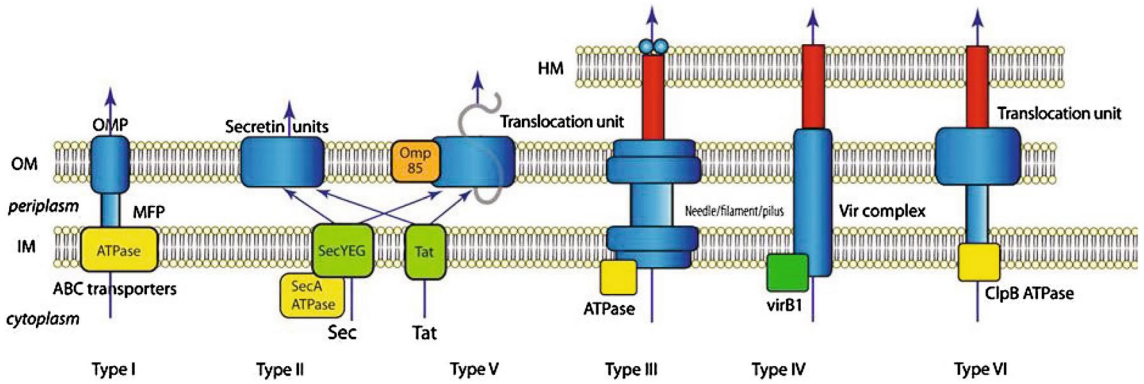

Fig. 5 The overview of secretion systems utilized by bacterium adapted from. *HM* host membrane, *OM* outer membrane, *IM* inner membrane, *OMP* outer membrane protein, *MFP* membrane fusion protein. ATPases was shown in *yellow* parts (Tseng et al. 2009)

involving in type VI secretion systems expressed during infection of potato tubers and the virulence of the mutants without these two genes is significantly reduced (Boyer et al. 2009; Leiman et al. 2009; Silverman et al. 2011).

The *Hrp* gene cluster encoding type III system is the most important feature in strain JG10-08. Compared with stain SCRI1043, JG10-08 shares some similarities in sequences. However, there are several loci containing the gene translocation (Fig. 6). In addition, the *HrpI* and *HrpD* gene are inverted in orientation. Moreover, the *HrpY* gene is absent from the genome of *P. atrosepticum* JG10-08. Although the gene translocation, inversion and deletion existed in the genome of *P. atrosepticum* JG10-08, there are no influences on gene expression and function. The reason is probably due to the fact that there are some relevant genes supplementing the functions of the deleted or disordered genes during the process of evolution. Besides of the *Hrp* genes, *P. atrosepticum* JG10-08 contains several other annotated gene clusters related to type III secretion function (Fig. 7), such as *Hrc, Yes, Esc*, and *Spa*, which is believed to contribute to the independent secretion systems and direct transportation of the effectors from cytoplasm to the cell surface (Cornelis and Van Gijsegem 2000; Jin et al. 2003).

The type IV secretion system (T4SS) are large protein complexes that can transport DNA, large molecular proteins and nuclear-protein compounds. *Agrobacterium tumefaciens* includes this system encoded by *VirB, Pic, Cpa, Tad* and *Rep* genes (Aguilar et al. 2010). A virB-type IV secretion system that is best known as the

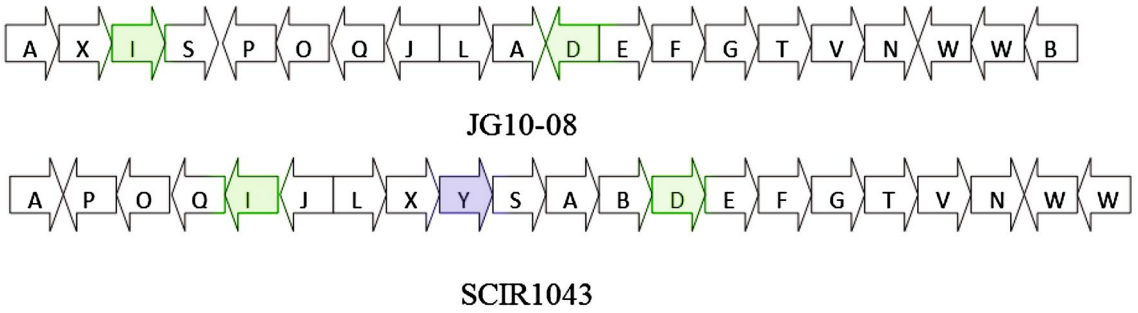

Fig. 6 Comparison of the Hrp gene cluster between *P. atrosepticum* JG10-08 and SCRI1043. Each gene is indicated by an *arrow*, and is ordered at the corresponding position. The length of *arrow* does not reflect the size of gene. The *light green arrows* mean reverse orientation of genes between *P. atrosepticum* JG10-08 and SCRI1043. A *purple arrow* represents the gene absent from the genome of *P. atrosepticum* JG10-08

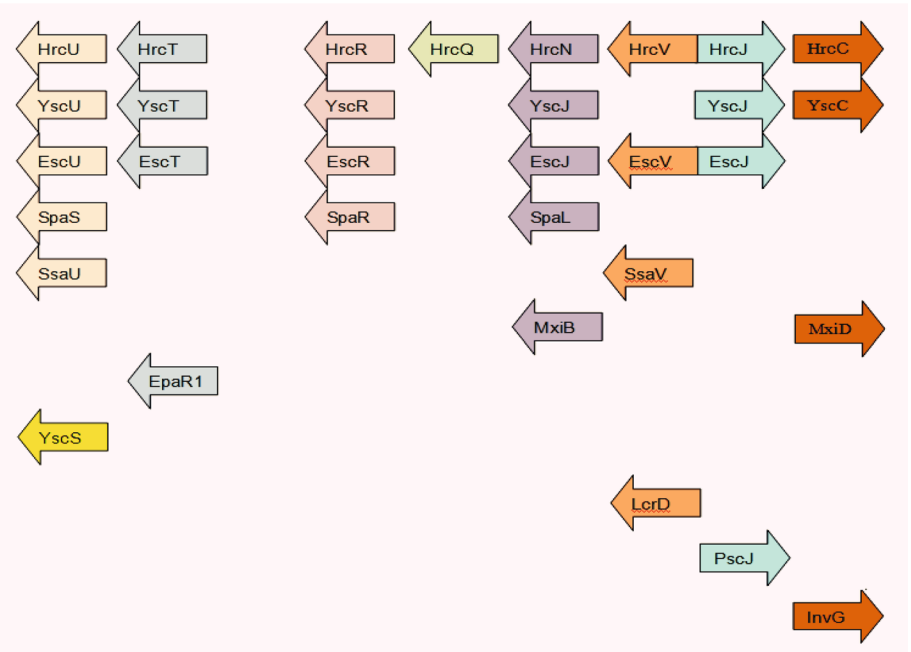

Fig. 7 Gene cluster analysis in *P. atrosepticum* JG10-08. Predicated genes and their orientation are shown by *arrows*. The length does not represent the corresponding size of gene. The *arrowheads* with the same *color* show that these genes are homologous

Fig. 8 Predictions of the *Pil* gene cluster in *P. atrosepticum* JG10-08. Predicated genes and their orientation are shown in *arrows*. The length of *arrow* does not represent the corresponding size of gene. *Green arrows* represent *PilA*, *PilB* and *PilC*, which are also present in *P. atrosepticum* SCRI 1043

Tiplasmid transferring system of *A. tumefaciens* (Bell et al. 2004b) is conserved in *P. atrosepticum*. However, *VirB7* gene is deleted in *P. atrosepticum* SCRI1043, and *VirB7* and *VirB3* genes are both deleted from *P. atrosepticum* JG10-08. Previous studies revealed that the lack of *VirB3* and *VirB7* genes is common to lots of bacterium. In addition, JG10-08 harbors the *Pil* gene (Fig. 8) involving in the assembly of pilin, which plays an important role in early infection and colonization (Henderson et al. 1998). It has been documented that nearly 40 *Pil* genes are identified in *Pseudomonas aeruginosa* and some of these genes contribute to virulence (Ishimoto and Lory 1992; Alm et al. 1996; Shikata et al. 2016), which strongly suggested that these genes in JG10-08 have the potential virulence functions. Thus, the presence of the type IV system in the *P. atrosepticum* JG10-08 genome implies that this system is required to infect the plants.

Conclusion

The GC content of JG10-08 is similar to that of *P. wasabiae* while slightly lower than that of *P. carotovorum*. The comparative genomic analysis based on whole-genome sequences of *P. atrosepticum* JG10-08 with *P. wasabiae* and *P. carotovorum* provided significant insights into their relationships in the process of evolution. Our results suggested the genes involving in metabolism account for a predominant proportion of all annotated genes, which reflects a high degree of adaptation and interaction of *P. atrosepticum* with the host. In addition, our studies highlight the importance of the bacterial secretion systems for pathogenesis, and indicate that bacteria are capable of synthesizing and transporting virulence factors. All in all, the results of this research will serve as an important source for the development of biological measures to control blackleg of potatoes.

Acknowledgements This work was funded by the Earmarked Fund for Modern Agro-industry Technology Research System (CARS-10-P12) and the Special Fund for Agro-scientific Research in the Public Interest (201303018).

Compliance with ethical standards

Research involving human and animal rights This article does not contain any studies with human subjects or animals performed by any of the authors.

Conflict of interest Dai Zhang declares that he/she has no conflict of interest. Yuan Zhou declares that he/she has no conflict of interest. Dongmei Zhao declares that he/she has no conflict of interest. Jiehua Zhu declares that he/she has no conflict of interest. Zhihui Yang declares that he/she has no conflict of interest. Mingming Zhu declares that he/she has no conflict of interest.

References

Aguilar J, Zupan J, Cameron TA, Zambryski PC (2010) Agrobacterium type IV secretion system and its substrates form helical arrays around the circumference of virulence-induced cells. Proc Natl Acad Sci USA 107: 3758–3763

Alm RA, Hallinan JP, Watson AA, Mattick JS (1996) Fimbrial biogenesis genes of *Pseudomonas aeruginosa*: PilW and pilX increase the similarity of type 4 fimbriae to the GSP protein-secretion systems and pilY1 encodes a gonococcal PilC homologue. Mol Microbiol 22:161–173

Aziz RK, Bartels D, Best AA, DeJongh M, Disz T, Edwards RA, Formsma K, Gerdes S, Glass EM, Kubal M et al (2008) The RAST server: rapid annotations using subsystems technology. BMC Genom 9:75

Bell KS, Sebaihia M, Pritchard L, Holden MT, Hyman LJ, Holeva MC, Thomson NR, Bentley SD, Churcher LJ, Mungall K et al (2004a) Genome sequence of the enterobacterial phytopathogen *Erwinia carotovora* subsp. *atroseptica* and characterization of virulence factors. Proc Natl Acad Sci USA 101:11105–11110

Bell KS, Sebaihia M, Pritchard L, Holden MTG, Hyman LJ, Holeva MC, Thomson NR, Bentley SD, Churcher LJC, Mungall K et al (2004b) Genome sequence of the enterobacterial phytopathogen *Erwinia carotovora* subsp. *atroseptica* and characterization of virulence factors. Proc Natl Acad Sci USA 101:11105–11110

Blattner FR, Plunkett G 3rd, Bloch CA, Perna NT, Burland V, Riley M, Collado-Vides J, Glasner JD, Rode CK, Mayhew GF et al (1997) The complete genome sequence of *Escherichia coli* K-12. Science 277:1453–1462

Boyer F, Fichant G, Berthod J, Vandenbrouck Y, Attree I (2009) Dissecting the bacterial type VI secretion system by a genome wide in silico analysis: what can be learned from available microbial genomic resources? BMC Genom 10:104

Christensen SK, Maenhaut-Michel G, Mine N, Gottesman S, Gerdes K, Van Melderen L (2004) Overproduction of the Lon protease triggers inhibition of translation in *Escherichia coli*: involvement of the yefM-yoeB toxin-antitoxin system. Mol Microbiol 51:1705–1717

Cornelis GR, Van Gijsegem F (2000) Assembly and function of type III secretory systems. Annu Rev Microbiol 54:735–774

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

Davidson AL, Chen J (2004) ATP-binding cassette transporters in bacteria. Annu Rev Biochem 73:241–268

de Chial M, Ghysels B, Beatson SA, Geoffroy V, Meyer JM, Pattery T, Baysse C, Chablain P, Parsons YN, Winstanley C et al (2003) Identification of type II and type III pyoverdine receptors from *Pseudomonas aeruginosa*. Microbiology 149:821–831

Delcher AL, Harmon D, Kasif S, White O, Salzberg SL (1999) Improved microbial gene identification with GLIMMER. Nucleic Acids Res 27:4636–4641

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

Duchaud E, Rusniok C, Frangeul L, Buchrieser C, Givaudan A, Taourit S, Bocs S, Boursaux-Eude C, Chandler M, Charles JF et al (2003) The genome sequence of the entomopathogenic bacterium *Photorhabdus luminescens*. Nat Biotechnol 21:1307–1313

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

Filloux A (2004) The underlying mechanisms of type II protein secretion. BBA Biomembranes 1694:163–179

Franza T, Michaud-Soret I, Piquerel P, Expert D (2002) Coupling of iron assimilation and pectinolysis in *Erwinia chrysanthemi* 3937. Mol Plant Microbe Interact 15:1181–1191

Gardan L, Gouy C, Christen R, Samson R (2003) Elevation of three subspecies of *Pectobacterium carotovorum* to species level: *Pectobacterium atrosepticum* sp nov., *Pectobacterium betavasculorum* sp nov. and *Pectobacterium wasabiae* sp nov. Int J Syst Evol Microbiol 53:381–391

Hagland M (2011) Potato perspectives. It's hard to know what the implications are for the future at the moment any technological or consumer innovation emerges. Healthcare Inform 28:8

Harrison JJ, Wade WD, Akierman S, Vacchi-Suzzi C, Stremick CA, Turner RJ, Ceri H (2009) The chromosomal toxin gene yafq is a determinant of multidrug tolerance for *Escherichia coli* growing in a biofilm. Antimicrob Agents Chemother 53:2253–2258

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

Henderson IR, Navarro-Garcia F, Nataro JP (1998) The great escape: structure and function of the autotransporter proteins. Trends Microbiol 6:370–378

Holst O, Ulmer AJ, Brade H, Flad HD, Rietschel ET (1996) Biochemistry and cell biology of bacterial endotoxins. FEMS Immunol Med Microbiol 16:83–104

Hyman LJ, Birch PR, Dellagi A, Avrova AO, Toth IK (2000) A competitive PCR-based method for the detection and quantification of *Erwinia carotovora* subsp. *atroseptica* On potato tubers. Lett Appl Microbiol 30:330–335

Ishimoto KS, Lory S (1992) Identification of Pilr, which encodes a transcriptional activator of the *Pseudomonas aeruginosa* pilin gene. J Bacteriol 174:3514–3521

Jin QL, Thilmony R, Zwiesler-Vollick J, He SY (2003) Type III protein secretion in *Pseudomonas syringae*. J Microb Infect 5:301–310

Jones HA, Lillard JW, Perry RD (1999) HmsT, a protein essential for expression of the haemin storage (Hms(+)) phenotype of *Yersinia pestis*. Microbiology 145:2117–2128

Kang YW, Liu HL, Genin S, Schell MA, Denny TP (2002) *Ralstonia solanacearum* requires type 4 pili to adhere to multiple surfaces and for natural transformation and virulence. Mol Microbiol 46:427–437

King JC, Slavin JL (2013) White potatoes, human health, and dietary guidance. Adv Nutr 4:393S–401S

Koskinen JP, Laine P, Niemi O, Nykyri J, Harjunpaa H, Auvinen P, Paulin L, Pirhonen M, Palva T, Holm L (2012) Genome sequence of *Pectobacterium* sp. strain SCC3193. J Bacteriol 194:6004

Kwasiborski A, Mondy S, Beury-Cirou A, Faure D (2013) Genome sequence of *Pectobacterium atrosepticum* strain cfbp6276, causing blackleg and soft rot diseases on potato plants and tubers. Genome Announc 1:e00374

Larkin RP, Honeycutt CW (2006) Effects of different 3-year cropping systems on soil microbial communities and rhizoctonia diseases of potato. Phytopathology 96:68–79

Leiman PG, Basler M, Ramagopal UA, Bonanno JB, Sauder JM, Pukatzki S, Burley SK, Almo SC, Mekalanos JJ (2009) Type VI secretion apparatus and phage tail-associated protein complexes share a common evolutionary origin. Proc Natl Acad Sci USA 106:4154–4159

Li GY, Zhang YL, Inouye M, Ikura M (2009) Inhibitory mechanism of *Escherichia coli* RelE-RelB toxin-antitoxin module involves a helix displacement near an mrna interferase active site. J Biol Chem 284:14628–14636

Lory S (1998) Secretion of proteins and assembly of bacterial surface organelles: shared pathways of extracellular protein targeting. Curr Opin Microbiol 1:27–35

Luo RB, Liu BH, Xie YL, Li ZY, Huang WH, Yuan JY, He GZ, Chen YX, Pan Q, Liu YJ et al (2012) SOAPdenovo2: an empirically improved memory-efficient short-read de novo assembler. Giga Sci 1:18

Mulholland V, Hinton JCD, Sidebotham J, Toth IK, Hyman LJ, Perombelon MCM, Reeves PJ, Salmond GPC (1993) A pleiotropic reduced virulence (rvi-) mutant of *Erwinia carotovora* subspecies *atroseptica* is defective in flagella assembly proteins that are conserved in plant and animal bacterial pathogens. Mol Microbiol 10:1154–1154

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

Parkhill J, Dougan G, James KD, Thomson NR, Pickard D, Wain J, Churcher C, Mungall KL, Bentley SD, Holden MT et al (2001a) Complete genome sequence of a multiple drug resistant *Salmonella enterica* serovar Typhi CT18. Nature 413:848–852

Parkhill J, Wren BW, Thomson NR, Titball RW, Holden MT, Prentice MB, Sebaihia M, James KD, Churcher C, Mungall KL et al (2001b) Genome sequence of *Yersinia pestis*, the causative agent of plague. Nature 413:523–527

Perna NT, Plunkett G 3rd, Burland V, Mau B, Glasner JD, Rose DJ, Mayhew GF, Evans PS, Gregor J, Kirkpatrick HA et al (2001) Genome sequence of enterohaemorrhagic *Escherichia coli* O157:H7. Nature 409:529–533

Pitman AR, Wright PJ, Galbraith MD, Harrow SA (2008) Biochemical and genetic diversity of *Pectolytic enterobacteria* causing soft rot disease of potatoes in New Zealand. Australas. Plant Pathol 37:559–568

Reeves PJ, Whitcombe D, Wharam S, Gibson M, Allison G, Bunce N, Barallon R, Douglas P, Mulholland V, Stevens S et al (1993) Molecular cloning and characterization of 13 out genes from *Erwinia carotovora* subspecies *carotovora*: genes encoding members of a general secretion pathway (GSP) widespread in gram-negative bacteria. Mol Microbiol 8:443–456

Sandkvist M (2001) Type II secretion and pathogenesis. Infect Immun 69:3523–3535

Shikata M, Hayashi N, Fujimoto A, Nakamura T, Matsui N, Ishiyama A, Maekawa Y, Gotoh N (2016) The pilT gene contributes to type III ExoS effector injection into epithelial cells in *Pseudomonas aeruginosa*. J Infect Chemother 22:216–220

Silverman JM, Austin LS, Hsu F, Hicks KG, Hood RD, Mougous JD (2011) Separate inputs modulate phosphorylation-dependent and -independent type VI secretion activation. Mol Microbiol 82:1277–1290

Takagi H, Kakuta Y, Okada T, Yao M, Tanaka I, Kimura M (2005) Crystal structure of archaeal toxin-antitoxin RelE-RelB complex with implications for toxin activity and antitoxin effects. Nat Struct Mol Biol 12:327–331

Tatusov RL, Koonin EV, Lipman DJ (1997) A genomic perspective on protein families. Science 278:631–637

Toth IK, Bell KS, Holeva MC, Birch PRJ (2003) Soft rot erwiniae: from genes to genomes. Mol Plant Pathol 4:17–30

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

Tseng TT, Tyler BM, Setubal JC (2009) Protein secretion systems in bacterial-host associations, and their description in the Gene Ontology. BMC Microbiol 9:S2

Unterholzner SJ, Hailer B, Poppenberger B, Rozhon W (2013) Characterisation of the stbD/E toxin-antitoxin system of pEP36, a plasmid of the plant pathogen *Erwinia pyrifoliae*. Plasmid 70:216–225

van der Wolf JM, Nijhuis EH, Kowalewska MJ, Saddler GS, Parkinson N, Elphinstone JG, Pritchard L, Toth IK, Lojkowska E, Potrykus M et al (2014) *Dickeya solani* sp. nov., a pectinolytic plant-pathogenic bacterium isolated from potato (*Solanum tuberosum*). Int J Syst Evol Microbiol 64:768–774

Welch RA, Burland V, Plunkett G, 3rd, Redford P, Roesch P, Rasko D, Buckles EL, Liou SR, Boutin A, Hackett J et al (2002) Extensive mosaic structure revealed by the complete genome sequence of uropathogenic *Escherichia coli*. Proc Natl Acad Sci USA 99:17020–17024

Yaganza ES, Tweddell RJ, Arul J (2014) Postharvest application of organic and inorganic salts to control potato (*Solanum tuberosum L.*) storage soft rot: plant tissue-salt physicochemical interactions. J Agric Food Chem 62:9223–9231

Yang J, Tian B, Liang L, Zhang KQ (2007) Extracellular enzymes and the pathogenesis of nematophagous fungi. Appl Microbiol Biot 75:21–31