Journal of Plant Pathology
https://doi.org/10.1007/s42161-019-00328-0

ORIGINAL ARTICLE

Comparative genomic analysis of subspecies of *Pantoea stewartii* reveals distinct variations

Peng Li[1] · Ying Zhang[1] · Yuqingqing Sun[1] · Xingxing Wu[2,3] · Zhiyuan Wang[4] · Jianuan Zhou[5] · Xiaofan Zhou[5]

Received: 15 November 2018 / Accepted: 17 May 2019

Abstract

Pantoea stewartii subsp. *stewartii* (*Pnss*) and *P. stewartii* subsp. *indologenes* (*Pnsi*) are closely related plant pathogens that differ in their host specificities. *Pnss* is the causal agent of Stewart's wilt of corn, whereas *Pnsi* causes disease on millets but not corn. Comparative genomics is a valuable method for characterizing the differences between genomes, but there are few studies on this important quarantine pathogen. Here, we compared publicly available genomes of seven strains of *Pnss* and three strains of *Pnsi*. Pan- and core-genome analyses showed that strains isolated from close geographical regions are more similar in their genome structures. Gene content and collinearity analyses further revealed numerous strain-specific genes. In particular, the *Pnss* type strain DC283 contained over 1200 additional genes compared with other strains. Importantly, we also identified eleven genes that are only present in *Pnsi* genomes and thus may be useful to distinguish between *Pnss* and *Pnsi* strains. Overall, this study characterized the common and distinct genomic features of *Pnss* and *Pnsi*, which lay the foundation for future development of molecular methods to detect the Stewart's wilt pathogen in maize for quarantine regulations and distinguish its two subspecies.

Keywords Quarantine pathogen · Maize · Virulence variation · Genome evolution

Introduction

The bacterial pathogen *Pantoea stewartii* consists of two closely related subgroups, namely *Pantoea stewartii* subsp. *stewartii* (*Pnss*) and *P. stewartii* subsp. *indologenes* (*Pnsi*). The *Pnss* type strain causes the Stewart's bacterial wilt and leaf blight of maize and sweat corn (Braun 1982), resulting in huge economic losses (Roper 2011) and thus has been classified as an important quarantine organism in many countries (Pataky and Ikin 2003a). On the other hand, *Pnsi* is avirulent to maize and corn, but instead causes leaf spots on foxtail millet (*Setaria italica*), pearl millet (*Pennisetum americanum*) (Mergaert et al. 1993), and onion (Stumpf et al. 2018). Because the two subgroups are highly similar in their phenotypes, to date, a number of molecular detection methods have been described, such as SNP analysis of house-keeping genes (Gehring et al. 2014), MALDI-TOF analysis (Wensing et al. 2010), mini-primer PCR assay (Xu et al. 2010), proteomic and genetic analysis (Wu et al. 2007), specific PCR primers from virulence genes (Coplin et al. 2002), and enzyme-linked immunosorbent assay (Lamka et al. 1991). However, a sensitive method that can robustly distinguish one subgroup from the other is still lacking.

Numerous experiments have shown that comparative genomic analysis is really a high-throughput and valuable method to study the variations between the similar species. For example, by using the comparative genomic analysis, the host adapted genes and virulence variation mechanisms were

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s42161-019-00328-0) contains supplementary material, which is available to authorized users.

✉ Peng Li
dayuleepong@163.com

✉ Xiaofan Zhou
xiaofan_zhou@scau.edu.cn

[1] Ministry of Education Key Laboratory for Ecology of Tropical Islands, Hainan Provincial Key Laboratory for Tropical Plant and Animal Ecology, College of Life Sciences, Hainan Normal University, Haikou 571158, China

[2] Southwest Forestry University, Kunming 650224, China

[3] Yunnan Forestry Technological College, Kunming 670224, China

[4] Dehong Teachers' College, Mangshi 678400, China

[5] Integrative Microbiology Research Centre, Guangdong Province Key Laboratory of Microbial Signals and Disease Control, College of Agriculture, South China Agricultural University, Guangzhou 510642, China

Published online: 24 May 2019

characterized fully in another notorious bacterial wilt disease pathogen *Ralstonia solanacearum* (Li et al. 2016; Peeters et al. 2013; Prior et al. 2016).

In this study, we compared the publicly available genomes of seven *Pnss* strains and three *Pnsi* strains. Pan- and core-genome analyses were designed to assess the entire core-genome characterizations of *Pnss* and *Pnsi* type strains, as well to address the variations among these strains. The comparative analysis results of seven *Pnss* and three *Pnsi* type strains with available genome sequences revealed functional strategies of their core-genome, provided detailed genome characterizations between the two closely related subspecies, and served as useful clues for distinguishing *Pnss* from other bacterium. It will be of significant interest to find appropriate molecular markers for accurate detection of virulent *Pnss* and avirulent *Pnsi* type strains, and help to explain the genome evolution mechanisms among these *P. stewartii* species.

Materials and methods

Genomic comparisons

The genome sequences of seven *Pnss* and three *Pnsi* type strains (Table 1) were downloaded from the NCBI database. Dubious regions were manually curated in CLC genomics workbench browser. Also, a separate *de novo* assembly was generated using raw reads downloaded from NCBI. The GC content calculation and gene annotation were performed using $CL_{GENOMICS}$ (http://www.chunlab.com). The evolutionary genealogy of genes: non-supervised orthologous groups (EggNOG) (Powell et al. 2012) analysis was performed to generate functional annotations for coding sequences.

Pan- and core-genome analysis

After constructing pan-genome from multiple genomes, all CDSs were clustered into pan-genome orthologous groups (POGs) (Vernikos et al. 2015). POG contains at least one CDS (called singleton POGs) and highly conserved POGs are found in all genomes, which comprises the core-genome (cut off = 100%). Pan-genome's gene content information is converted into binary matrix. Presence of a gene is coded as 1 and absence as 0, respectively. Then, similarity values can be calculated either between two genomes or between two POGs. We use Jaccard coefficient ($J(A,B) = \frac{|A \cap B|}{|A \cup B|}$; A: Genes in genome A; B: Genes in genome B) to calculate similarities which is widely used for presence/absence type data. In this study, two different calculations are provided: (1) complete calculation: analysis based on all POGs; (2) differential calculation: analysis based on POGs excluding those present in all genomes (core-genome) and singletons (present in single genome). Gene content (presence/absence) data together with clustering can be used to generate heat maps through UPGMA clustering (Comandatore et al. 2013), which is used to analyze the gene presence and absence matrix of all CDSs. The above analyses were performed using EzBioCloud's Comparative Genomics Database. The pan/core genome curves were then calculated from POGs (Zakham et al. 2012).

Collinearity analysis

Pairwise whole-genome alignments among these strains were constructed and visualized using Mauve progressive alignment Mauve software (Darling et al. 2004). Prior to the Mauve analysis, we use the MUMmer (Kurtz et al. 2004) software to rearrange the contigs of draft genomes referencing

Table 1 General characterization of *Pantoea stewartii* strains

Subgroup	Strain	Genome size (bp)	GC%	Contigs	tRNA	rRNA	ORFs	Country
Pnss	DC283	5,314,090	53.78	Complete	72	21	4883	America
	A206	4,654,611	53.80	19	78	22	4301	Canada
	S301	4,486,406	53.93	27	76	14	4128	Canada
	RSA36	4,790,394	53.73	76	64	11	4426	India
	NS381	4,693,001	53.85	51	69	5	4319	India
	RSA30	4,760,375	53.68	84	65	5	4393	India
	RSA13	4,766,500	53.67	48	67	6	4408	India
Pnsi	LMG2632	4,681,235	53.70	35	64	9	4474	South Africa
	M073a	4,817,607	53.92	39	68	5	4465	Malaysia
	M009	4,821,705	53.90	56	70	5	4476	Malaysia

to the complete genome sequences of strain DC283, and the figure was generated by the Mauve rearrangement viewer. OrthoMCL (Chen et al. 2006) cluster analyses were performed to identify the set of genes unique to strain DC283 with following parameters: p value cut-off = 1×10^{-5}, identity cut-off = 100%, percent match cut-off = 100%.

Results

General genomic features of *P. stewartii* strains

In this study, we examined genomes of the ten *P. stewartii* strains currently available in the NCBI database (Table 1), including seven *Pnss* strains (A206, DC283, NS381, RSA13, RSA30, RSA36, S301) and three *Pnsi* strains (M073a, LMG2632, M009). The genome of the *Pnss* type strain DC283 is the only complete one, consisting of one chromosome and eleven plasmids, whereas all the other nine genomes are in the draft status and are comprised of from 19 to 84 contigs. The strain DC283 has a total genome size of 5.3 Mbp and 4883 protein-coding genes, the vast majority of which is attributed to the chromosome (4.5 Mbp); the eleven plasmids vary greatly in their sizes (4.3–304.6 Kbp) and numbers of genes (5–302). In comparison, the other nine genomes have highly similar G + C content as DC283 (53.67–53.93%), but are smaller in size (4.4–4.8 Mbp) and contain fewer protein-coding genes (4128-4476), probably because their assemblies are still incomplete. Interestingly, among the nine incomplete genomes, both the largest two (M009 and M073a) and the smallest one (LMG2632) belong to the *Pnsi* subgroup.

Pan- and core-genome analysis

To better understand the genomic characteristics of the ten *Pnss* and *Pnsi* strains, we conducted pan- and core-genome analyses which are commonly used approaches to identify conserved and divergent components of bacterial genomes. In result, 44,831 coding sequences (CDSs) in the ten genomes were clustered into 6796 pan-genome orthologous group (POGs), of which 3346 POGs are present in all ten species (referred to as core-POGs hereafter) (Fig. 1). We then performed functional annotation of the core-POGs, which were successfully classified into 19 eggNOG functional categories; 22 POGs (0.66%) were involved in two or more eggNOG groups, whereas 69 POGs (2.05%) were failed to be classified into any group (Table 2). Except for the core gene group predicted to function unknown (863, 25.62%), the largest functional group was carbohydrate transport and metabolism which contained 313 genes (9.29%), followed by amino acid transport and metabolism (288 genes, 8.55%), transcription (246 genes, 7.30%), inorganic ion transport and metabolism (223 genes, 6.62%), cell wall/membrane/envelope biogenesis (200 genes, 5.94%), translation, ribosomal structure and biogenesis (171 genes, 2.11%), energy production and conversion (165 genes, 4.90%), replication, recombination and repair (133 genes, 3.95%), signal transduction mechanisms (131 genes, 3.89%), coenzyme transport and metabolism (110 genes, 3.27%), as well as posttranslational modification and protein turnover chaperones (109 genes, 3.24%), and so on.

In addition, we also performed pan-genome analyses separately for the five strains (DC283, M073a, LMG2632, S301 and NS381) from different continents and the other five strains (M009, NS381, RSA13, RSA30 and RSA36) from Asia. Results showed that 3363 core-POGs (cut off = 100%) were found among the five strains from different continents (Fig. 2a), accounting for 51.82% of their pan-genome (6490 genes), and nearly equal to the number of core-POGs (3346) shared by all the ten strains. Among the five strains from Asia, 3850 core-POGs were found (Fig. 2b), accounting for 75.85% of their pan-genome (5076 genes).

The four India strains (RSA13, RSA30, RSA36 and NS381) of *Pnss* group were selected to compare the core-gene families within the same subgroup. Results showed that 3979 core-gene families were found, accounting for 84.30% of their pan-genome (4720 genes) (Fig. 2c); among the three strains of *Pnsi* group, 3961 core-gene families were found (Fig. 2d), accounting for 82.32% of their pan-genome (4812 genes). Thus, besides the 3346 core genes present in all ten genomes, there were 633 additional genes shared by the four India *Pnss* strains and 615 additional genes common to the three *Pnsi* strains, respectively, accounting for 15.9% and 15.5% of their core-genomes.

Gene presence and absence analysis

Gene content (presence/absence) data together with clustering were used to generate heat map by clustering the genome sequences. Substantial gene deletions and insertions were found among the ten *P. stewartii* strains (Fig. 3). The two strains (A206, S301) from Canada with the similar gene distribution, the four strains from India belong to one group, while the strain NS381 with more variants than the other three strains (RSA13, 30, 36); the three *Pnsi* type strains (LMG2632, M073a, M009) formed a single clade and were more divergent from the *Pnss* type strains, but distinct gene rearrangements were also found between the South Africa strain (LMG2632) and the two Malaysia strains (M073a,

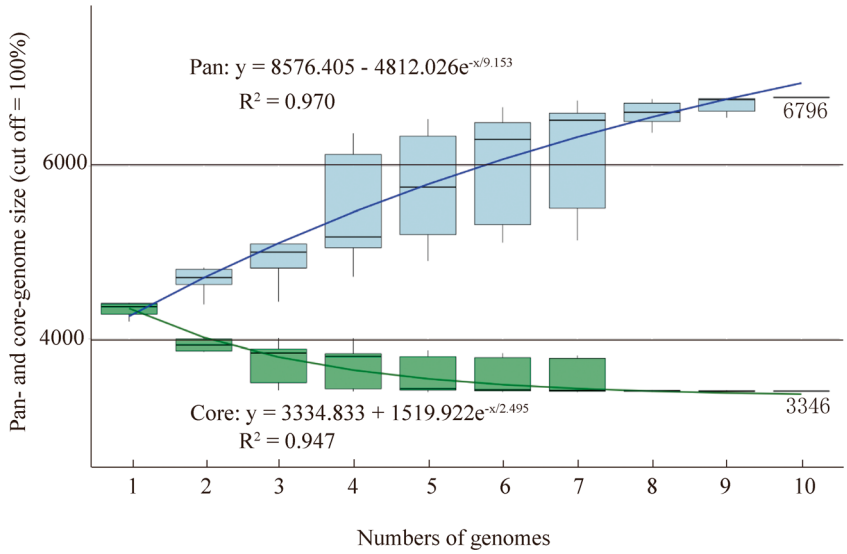

Fig. 1 Core- and pan-genome curve of ten *P. stewartii* strains

M009). Notably, there were ~1200 extra genes in *Pnss* type strain DC283 (Fig. 3). Taking a thorough analysis of these additional genes, results demonstrated that there were 645 genes of unknown function existed, among of which, at least 270 genes were from plasmid, 127 genes were transposon elements, 34 genes involved in virulence, and 29 genes involved in protein transport were also found (Supplemental information Table 1, SI Table 1).

Interestingly, we also found eleven genes that are only present in the three *Pnsi* type strains (Table 3). These genes

Table 2 Functional categories based on eggNOG in core-POGs

Description	Number of POGs	(%)
Function unknown	863	25.62
Carbohydrate transport and metabolism	313	9.29
Amino acid transport and metabolism	288	8.55
Transcription	246	7.30
Inorganic ion transport and metabolism	223	6.62
Cell wall/ membrane/ envelope biogenesis	200	5.94
Translation, ribosomal structure and biogenesis	171	2.11
Energy production and conversion	165	4.90
Replication, recombination and repair	133	3.95
Signal transduction mechanisms	131	3.89
Coenzyme transport and metabolism	110	3.27
Posttranslational modification, protein turnover, chaperones	109	3.24
Nucleotide transport and metabolism	80	2.38
Lipid transport and metabolism	75	2.23
Intracellular trafficking, secretion, and vesicular transport	46	1.37
Defense mechanisms	38	1.13
Cell motility	38	1.13
Secondary metabolites biosynthesis, transport and catabolism	36	1.07
Cell cycle control, cell division, chromosome partitioning	34	1.01
Not in eggNOG	69	2.05
	3368	100

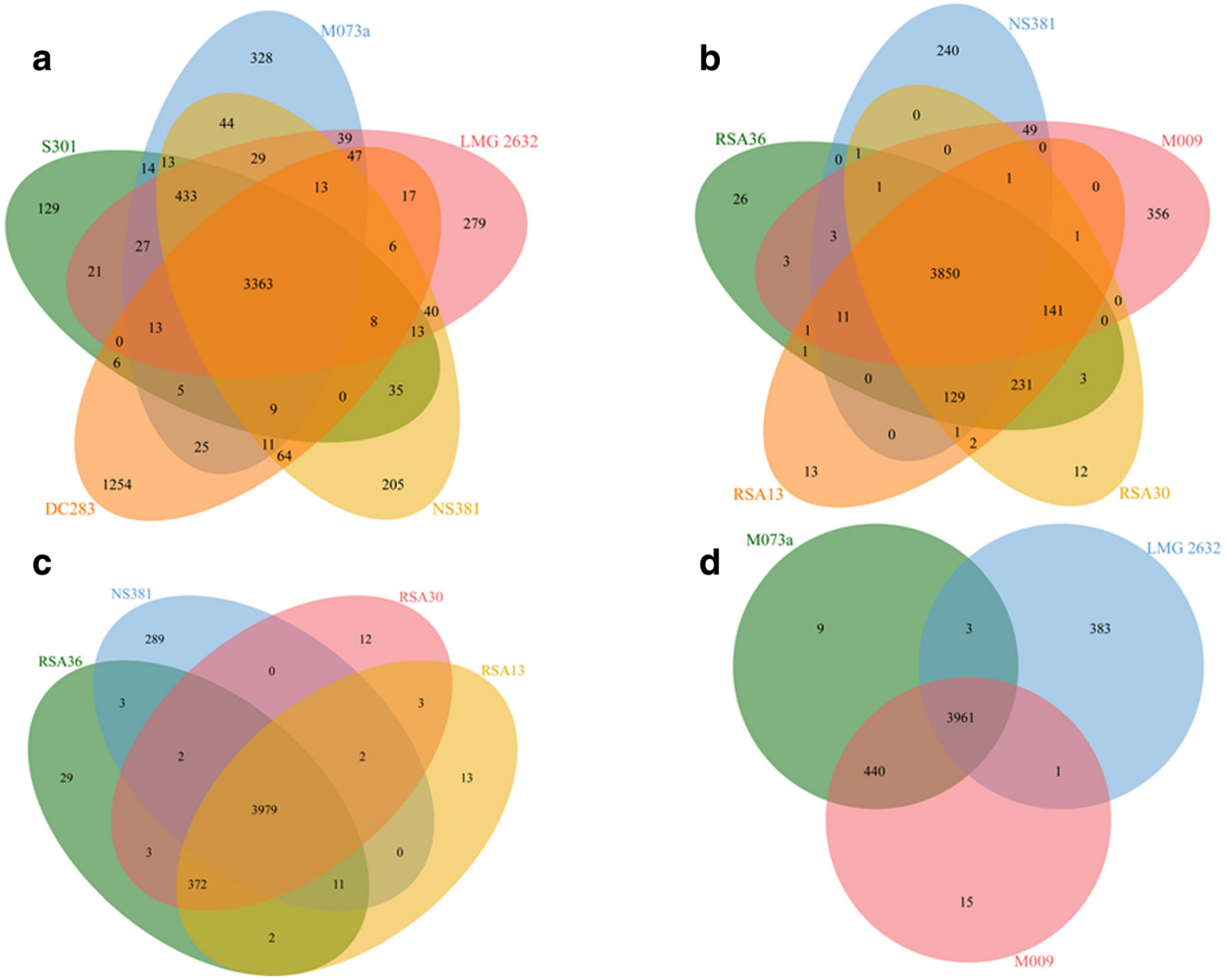

Fig. 2 Venn diagrams for deduced proteins of five *P. stewartii* strains from different regions (**a**), five *P. stewartii* strains from Asia (**b**), four *Pnss* strains from India (**c**), and three *Pnsi* strains (**d**). The overlapping sections indicate shared numbers of deduced proteins. Values were calculated by OrthoMCL cluster analyses with the parameters: *p* value Cut-off = 1×10^{-5}, identity cut-off = 100%, percent match cut-off = 100%. The overlapping sections indicate shares numbers of gene families, the numbers in brackets mean the genes of the corresponding gene families

are highly conserved at sequence level among *Pnsi* strains, but no homologous sequences were found in the *Pnss* strains even at a relatively low cut off value (50%). Five of the eleven *Pnsi* specific genes encode for hypothetical proteins, while the other six genes encode for acetyltransferase, type III effector, multidrug MFS transporter, glutamine ABC transporter permease, glutamine ABC transporter substrate-binding protein, and DNA-binding protein, respectively. All of these eleven genes were highly sequence conserved among the *Pnsi* type strains (coverage 100%, identity 100%).

Collinearity analyses

To evaluate the genome variation of these *P. stewartii* strains, the complete genome sequence of strain DC283 was aligned with the other four strains (A206, LMG2632, M073a, RSA13). Results showed that numerous gene deletion, insertion, and inversions distributed across the genome between strain DC283 and the other four strains (SI Fig. 1). Furthermore, contigs homologous to the plasmids of DC283 were found in the strains M073a and RSA13 (SI Fig. 1A, B), but not in the strains A206 and LMG2632 (SI Fig. 1C, D).

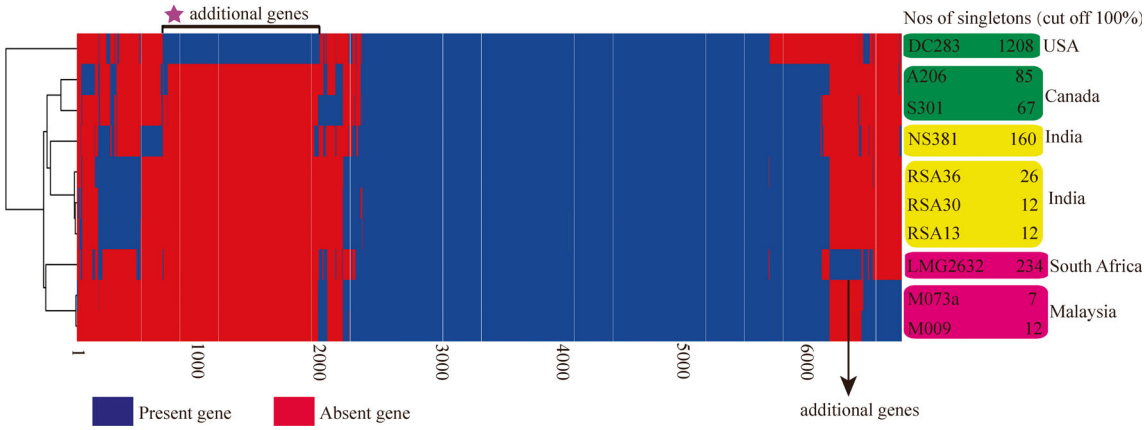

Fig. 3 Heat map based on gene content (presenece/absence) by complete calculation. UPGMA clustering is carried out for both genomes and POGs (genes). Genomes are clustered using presence/absence of genes and genes are clustered using presence/absence of genomes. Here, the numbers of singletons of each strain is also listed

Discussion

More than 60 countries from North America, Europe and Asia have placed strict quarantine regulations on the pathogens of *Pnss* (Pataky and Ikin 2003b), for its destructive damage on maize. However, there still lacks an effective and reliable approach to distinguish between *Pnss* and *Pnsi* strains due to their highly similar phenotypes. Therefore, it becomes increasingly important to discover genomic characteristics that can be used to distinguish *Pnss* strains from members of *Pnsi* or other closely related *Pantoea* species, such as *P. ananatis* (Coplin et al. 2002; Coutinho and Venter 2009; Wensing et al. 2010).

By analyzing the general genome features of *Pnss* and *Pnsi* strains, we found considerable variation in their genome sizes, which might be partly attributed to the fact the most genome are still in the draft status. Notably, the *Pnss* type strain DC283 is the only complete genome and is the largest one in size (Fig. 3); it harbors more than 1200 genes that are not found in other, including many genes distributed on plasmids or belong to transposable elements (e.g., transposons, bacteriophages). Previous studies have demonstrated that transposable elements can be a double-edged sword in the evolution of plant pathogens. For instance, different bacteriophages could either enhance or repress the virulence of *R. solanacearum* species and thus affect the outcome of pathogen-host interactions in completely opposite ways (Addy et al. 2012a, b). In that regards, these plasmid genes and transposable elements are likely to have important roles during the evolution of *P. stewartii*, particularly in shaping the variation of virulence and host range among strains.

Core-genome analyses have shown that strains isolated from the same regions share more core-gene families, indicating a correlation between geographic distribution and evolutionary relationship. Evidences have showed that the core-genomes might be relatively free of combination in other bacterial groups (Daubin et al. 2003; Lerat et al. 2005), based our data about 16% core genome of the *P. stewartii* species were involved in their core genome combination, the similar result was also reported in the *Streptococcus* species (Lefebure and Stanhope 2007). In addition, geographic distribution for core genome evolution may also related to their genome evolution, for a higher number of core genome genes (3850) existed among the five Asia strains, though they belonged to different subgroups.

The results of our core-genome analysis have also highlighted a number of genes required for infection in the genomes of both *Pnsi* and *Pnss* subgroups, including 46 genes involved in secretion, 38 genes involved in defense mechanism, and 38 genes involved in motility. These infection-related genes might have formed the molecular basis for the role of *Psni* and *Pnss* as plant pathogens.

PCR assays targeting strain specific sequence variations can be used to differentiate closely related strains (Fegan and Prior 2005; Gehring and Geider 2012; Genin 2010). One major finding of our comparative genomic study is the identification of eleven *Pnsi*-specific genes which can be highly useful for the characterization of these phenotypically highly similar *Pnss* and *Pnsi* strains to decide about possible contamination of maize imports with the Stewart's wilt pathogen.

In conclusion, ten publicly available *P. stewartii* genomes were compared and their shared and unique genome characteristics were analyzed. Some typical additional or unique

Table 3 Loci and function information of eleven unique genes in *Pnsi* type strains

Strain Loci	M073a	LMG 2632	M009	Function
1	JSXF01000012, 251533-251631	JPKO01000035, 104051-10414	JRWI01000011, 244349-244447	Hypothetical protein
2	SXF01000012, 65145-65558	JPKO01000035, 290075-290488	JRWI01000011, 57961-58374	Acetyltransferase
3	JSXF01000012, 68614-69600	JPKO01000035, 286034-287019	JRWI01000011, 61430-62416	Type III effector
4	JSXF01000012, 252973-254031	JPKO01000035, 101651-102709	JRWI01000011, 245789-246847	Hypothetical protein
5	JSXF01000012, 66282-67592	JPKO01000035, 288041-289351	JRWI01000011, 59098-60408	Hypothetical protein
6	JSXF01000014, 582990-584147	PKO01000019, 10389-11546	JRWI01000011, 346096-347253	Multidrug MFS transporter
7	JSXF01000014, 210736-212064	JPKO01000026, 9915-11243	JRWI01000005, 194712-196040	Hypothetical protein
8	JSXF01000014, 210185-210739	JPKO01000026, 9364-9918	JRWI01000005, 196037-196591	Hypothetical protein
9	JSXF01000014, 208600-209256	JPKO01000026, 7781-8435	JRWI01000005, 197520-198176	Glutamine ABC transporter permease
10	JSXF01000014, 209307-210059	JPKO01000026, 8486-9238	RWI01000005, 196717-197469	Glutamine ABC transporter substrate-binding protein
11	JSXF01000014, 212194-213051	JPKO01000026, 11373-12230	JRWI01000005, 193725-194582	DNA-binding protein

genes regions existed in them, especially in the strain DC283 with more than 1200 additional genes, which made it different from other strains significantly. Moreover, we also identified eleven genes that are only present in *Pnsi* genomes and thus may be useful to distinguish between *Pnss* and *Pnsi* strains. Overall, this study characterized the common and distinct genomic features of *Pnss* and *Pnsi*, which lay the foundation for future development of molecular methods to detect the Stewart's wilt pathogen in maize for quarantine regulations and distinguish its two subspecies.

Acknowledgments This work was supported by the State Key Laboratory for Conservation and Utilization of Subtropical Agro–bioresources (No. SKLCUSA–b201715). National Key Project for Basic Research of China (973 Program, No. 2015CB150600).

Author contributions PL and XZ designed the experiments, PL, YZ, YS, XY, ZY, and XZ analyzed the data and wrote the paper, XW, ZW, JZ and XZ revised the manuscript.

Compliance with ethical standards The authors declare no potential conflicts of interest, and the research involving no human participants and/or animals.

References

Addy HS, Askora A, Kawasaki T, Fujie M, Yamada T (2012a) The filamentous phage φRSS1 enhances virulence of Phytopathogenic *Ralstonia solanacearum* on tomato. Phytopathology 102:244–251

Addy HS, Askora A, Kawasaki T, Fujie M, Yamada T (2012b) Loss of virulence of the Phytopathogen *Ralstonia solanacearum* through infection by φRSM filamentous phages. Phytopathology 102:469–477

Braun EJ (1982) Ultrastructural investigation of resistant and susceptible maize inbreds infected with *Erwinia stewartii*. Phytopathology 72: 159–166

Chen F, Mackey AJ, Stoeckert CJ Jr, Roos DS (2006) OrthoMCL-DB: querying a comprehensive multi-species collection of ortholog groups. Nucleic Acids Res 34:D363–D368

Comandatore F, Sassera D, Montagna M, Kumar S, Koutsovoulos G, Thomas G, Repton C, Babayan SA, Gray N, Cordaux R, Darby A, Makepeace B, Blaxter M (2013) Phylogenomics and analysis of shared genes suggest a single transition to mutualism in *Wolbachia* of nematodes. Genome Biol Evol 5:1668–1674

Coplin DL, Majerczak DR, Zhang YX, Kim WS, Jock S, Geider K (2002) Identification of *Pantoea stewartii* subsp. *stewartii* by PCR and strain differentiation by PFGE. Plant Dis 86:304–311

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

Darling ACE, Mau B, Blattner FR, Perna NT (2004) Mauve: multiple alignment of conserved genomic sequence with rearrangements. Genome Res 14:1394–1403

Daubin V, Moran NA, Ochman H (2003) Phylogenetics and the cohesion of bacterial genomes. Science 301:829–832

Fegan M, Prior P (2005) How complex is the *Ralstonia solanacearum* species complex. In: Allen C, Prior P, Hayward A (eds) Bacterial wilt disease and the Ralstonia solanacearum species complex. American Phytopathological Society, Madison, pp 449–461

Gehring I, Geider K (2012) Identification of *Erwinia* species isolated from apples and pears by differential PCR. J Microbiol Methods 89:57–62

Gehring I, Wensing A, Gernold M, Wiedemann W, Coplin DL, Geider K (2014) Molecular differentiation of *Pantoea stewartii* subsp *indologenes* from subspecies *stewartii* and identification of new isolates from maize seeds. J Appl Microbiol 116:1553–1562

Genin S (2010) Molecular traits controlling host range and adaptation to plants in *Ralstonia solanacearum*. New Phytol 187:920–928

Kurtz S, Phillippy A, Delcher AL, Smoot M, Shumway M, Antonescu C, Salzberg SL (2004) Versatile and open software for comparing large genomes. Genome Biol 5:R12–R12

Lamka GL, Hill JH, Mcgee DC, Braun EJ (1991) Development of an immunosorbent assay for seed-borne *Erwinia stewartii* in corn seed. Phytopathology 81:839–846

Lefebure T, Stanhope MJ (2007) Evolution of the core and pan-genome of *Streptococcus*: positive selection, recombination, and genome composition. Genome Biol 8:R71–R71

Lerat E, Daubin V, Ochman H, Moran NA (2005) Evolutionary origins of genomic repertoires in bacteria. PLoS Biol 3:e130–e130

Li P, Wang DC, Yan JL, Zhou JN, Yue DY, Jiang ZD, Cao BH, He ZF, Zhang L-H (2016) Genomic analysis of phylotype I strain EP1 reveals substantial divergence from other strains in the *Ralstonia solanacearum* species complex. Front Microbiol 7:01719

Mergaert J, Verdonck L, Kersters K (1993) Transfer of *Erwinia ananas* (synonym, *Erwinia uredovora*) and *Erwinia stewartii* to the genus *Pantoea* emend as *Pantoea ananas* (serrano 1928) comb. nov. and *Pantoea stewartii* (smith 1898) comb. nov., respectively, and description of *Pantoea stewartii* subsp *indologenes* subsp. nov. Int J Syst Bacteriol 43:162–173

Pataky J, Ikin R (2003a) Pest risk analysis : the risk of introducing Erwinia stewartii in maize seed. The Innternational Seed Federation, Nyon

Pataky J, Ikin R (2003b) Pest risk analysis: the risk of introducing *Erwinia stewartii* in maize seed. The International Seed Federation, Nyon

Peeters N, Guidot A, Vailleau F, Valls M (2013) *Ralstonia solanacearum*, a widespread bacterial plant pathogen in the post-genomic era. Mol Plant Pathol 14:651–662

Powell S, Szklarczyk D, Trachana K, Roth A, Kuhn M, Muller J, Arnold R, Rattei T, Letunic I, Doerks T, Jensen LJ, von Mering C, Bork P (2012) eggNOG v3.0: orthologous groups covering 1133 organisms at 41 different taxonomic ranges. Nucleic Acids Res 40:D284–D289

Prior P, Ailloud F, Dalsing BL, Remenant B, Sanchez B, Allen C (2016) Genomic and proteomic evidence supporting the division of the plant pathogen *Ralstonia solanacearum* into three species. BMC Genomics 17:90

Roper MC (2011) *Pantoea stewartii* subsp. *stewartii*: lessons learned from a xylem-dwelling pathogen of sweet corn. Mol Plant Pathol 12:628–637

Stumpf S, Kvitko B, Gitaitis R, Dutta B (2018) Isolation and characterization of novel *Pantoea stewartii* subsp. *indologenes* strains exhibiting center rot in onion. Plant Dis 102:727–733

Vernikos G, Medini D, Riley DR, Tettelin H (2015) Ten years of pan-genome analyses. Curr Opin Microbiol 23:148–154

Wensing A, Zimmermann S, Geider K (2010) Identification of the corn pathogen *Pantoea stewartii* by mass spectrometry of whole-cell extracts and its detection with novel PCR primers. Appl Environ Microbiol 76:6248–6256

Wu Q, Jiang Z, Liao J, Chen Z, Li H, Mei M, Zhang L-H (2007) Identification of genetic markers to distinguish the virulent and avirulent subspecies of *Pantoea stewartii* by comparative proteomics and genetic analysis. Appl Microbiol Biotechnol 74:186–193

Xu R, Chen Q, Robleh Djama Z, Tambong JT (2010) Miniprimer PCR assay targeting multiple genes: a new rapid and reliable tool for genotyping *Pantoea stewartii* subsp. *stewartii*. Lett Appl Microbiol 50:216–222

Zakham F, Aouane O, Ussery D, Benjouad A, Ennaji MM (2012) Computational genomics-proteomics and phylogeny analysis of twenty one mycobacterial genomes (tuberculosis & non tuberculosis strains). Microb Info Exp 2:7