Insights into the Draft Genome Sequence of the Kiwifruit-Associated Pathogenic Isolate *Pseudomonas fluorescens* AHK-1

Lixin Zhang¹ · Zhiran Wu¹ · Xia Wang¹ · Genjia Tan¹ · Jianghua Song²

Received: 15 December 2018 / Accepted: 19 February 2019 / Published online: 1 March 2019 © Springer Science+Business Media, LLC, part of Springer Nature 2019

Abstract

Pseudomonas fluorescens is a physiologically diverse species of bacteria present in many habitats, which possesses multifunctional traits that provide it with the capability to exhibit biological control activities, promote plant health or cause plant disease. Here, we present the draft genome sequence of the kiwifruit-associated pathogenic isolate AHK-1 of *P. fluorescens*, which was isolated from the diseased leaves of kiwifruit plants. The genome size of AHK-1 was found to be 7,035,786 bp, with a G+C content of 60.88%. It is predicted to contain a total of 6327 genes, of which 3998 were homologous to genes in the other two sequenced *P. fluorescens* isolates (SBW25 and GcM5-1A) and 946 were unique to AHK-1 based on comparative genomic analysis. Furthermore, we identified several candidate virulence factors in the genome of AHK-1, including the *fliA* gene encoding flagellar biosynthetic protein for biosynthesis, and the genes for components of type VI, III, and IV secretion systems. This genomic resource will serve as a reference for better understanding the genetics of pathogenic and non-pathogenic strains, and will help to elucidate the pathogenic mechanisms of *P. fluorescens* associated with plant disease.

Introduction

Pseudomonas fluorescens is a physiologically diverse species of bacteria that are found in soil, rhizosphere, living plants and water, as well as in contaminated human blood products and respiratory samples. Some *P. fluorescens* isolates benefit plants by their capacity to colonize plant surfaces and producing antibiotics toxic to target pathogens [1]. In contrast, some isolates are recognized as an opportunistic pathogen to humans [2]. Another isolates were documented with the ability to negatively affect growth of plants [3] or be pathogenic to plants [4, 5].

To date, the genome sequences were generated for several plant-associated isolates of *P. fluorescens*, most of them

Lixin Zhang and Zhiran Wu have contributed equally to this work.

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s00284-019-01655-2) contains supplementary material, which is available to authorized users.

☐ Jianghua Song jianghua_80@126.com

¹ Department of Plant Pathology, College of Plant Protection, Anhui Agricultural University, Hefei 230036, China

² College of Horticulture, Anhui Agricultural University, Hefei 230036, China were found to be functional as biological control agents to suppress plant disease [1]. Although the genome of a strain GcM5-1A of *P. fluorescens* associated with pine wilt disease was recently reported [6], the genome sequences associated with plant-pathogenic isolates still remain limited, and the virulence factors and evolutionary relationships of *P. fluorescens* are poorly understood. In our previous study, *P. fluorescens* strains were frequently isolated from the infected leaves and twigs of symptomatic kiwifruit plants [7], which resembled symptoms of bacterial canker disease mainly caused by *P. syringae* pv. *actinidiae* [8, 9]. When inoculated on the detached stems and twigs of potted seedlings of kiwifruit, *P. fluorescens* AHK-1 caused necrosis around the inoculated sites, and this strain was confirmed to be pathogenic on kiwifruit [7].

In order to acquire abundant molecular information to explore in greater detail the genetic characteristics of pathogenic isolates of *P. fluorescens*, the whole genome of *P. fluorescens* AHK-1 was sequenced. In the present study, we report the draft genome sequences of *P. fluorescens* AHK-1, compare the genome of AHK-1 with those of the other two sequenced *P. fluorescens* isolates, and identify putatively relevant virulence factors based on sequence annotation. The data presented in this work may lead to better understanding of the molecular basis of pathogenic and non-pathogenic isolates within *P. fluorescens* group.



Materials and Methods

Strains and DNA Extraction

The cultures of *P. fluorescens* AHK-1 were stored in 30% (v/v) glycerol at -80 °C in this study, and deposited in China Center for Type Culture Collection (CCTCC) with accession number CCTCC AB 2018073. The strain AHK-1 was shake-cultured in Luria–Bertani (LB) medium at 28 °C for 24 h, and bacterial cells were harvested by centrifugation at 5000×g for 10 min at 4 °C. Genomic DNA was extracted from the bacterial pellet of *P. fluorescens* AHK-1 using the NEB Bacterial DNA kit according to manufacturer's instructions. DNA purity was examined by 0.8% agarose gel electrophoresis, and the concentration was measured using a Qubit® 2.0 Fluorometer.

Genome Sequencing and Assembly

The whole genome sequencing of P. fluorescens AHK-1 was performed using the Illumina HiSeq 2500 platform (Illumina, USA) at the location of a sequencing service provider (Shanghai biotechnology corporation, Shanghai, China). The program FastQC was used to assess sequencing quality. The Q value of sequencing quality was used for evaluation, the relationship of Q value and sequencing error *E*-value is $Q = -10\log_{10}E$. We used Trimmomatic software [10] to filter raw reads, remove the adapter and low-quality sequences, which included the reads with ambiguous nucleotides (Q value ≤ 20) and short (≤ 45 bp) reads. The trimmed reads were de novo assembled with the SPAdes-3.5.0 software [11]. All of generated contigs were submitted to BLAST against the nr database using BLASTX [12]. Then the obtained contigs are aligned to the available genome sequence of P. fluorescens SBW25 with Mummer software [13] to be ordered.

Genome Annotation

The annotation file in Swiss-Prot library (http://www.expas y.ch/sprot/ and http://www.ebi.ac.uk/swissprot/) and the software Prokka [14, 15] were used to predict the function of genes. Then the predicted genes were annotated by the Cluster of Orthologous Groups (COG) database (http:// www.ncbi.nlm.nih.gov/COG) with an E-value cut-off of 1.0E-5. The predicted genes were classified by Blast2GO algorithm [16], to obtain the number of cataloged groups in Gene Ontology (GO) based on molecular function, cellular component and biological processes. Furthermore, pathway assignments were performed based on the online Kyoto Encyclopedia of Genes and Genomes (KEGG) Automatic Annotation Server (http://www.genome.jp/tools/kaas/) [17]. The tRNA genes were predicted using tRNAscan-SE 1.23 [18], rRNA genes using RNAmmer 1.2 [19] and sRNA using Infernal 1.1 [20]. Moreover, the candidate virulence factors were searched among all predicted genes in AHK-1 according to the methods as described previously [6].

Comparative and Phylogenetic Analysis

To determine the variation in genome content and organization between AHK-1 and other sequenced isolates of P. fluorescens, the genome of AHK-1 was compared to the two representative genomes of P. fluorescens GcM5-1A [21, 22] and SBW25 [23] by using a multiway BLASTp analysis. Genes present exclusively in an individual strain and those shared between two or three strains were counted by using Mauve software, and represented in Venn diagrams generated by Venn Diagram in R-platform [24]. The phylogenetic relationship between P. fluorescens AHK-1 and other representative isolates of P. fluorescens published previously was analyzed with Molecular Evolutionary Genetics Analysis (MEGA) software, based on (1) 16S rRNA and (2) concatenated sequences of nine highly conserved housekeeping genes including acsA, aroE, guaA, gyrB, mutL, ppsA, pyrC, recA, and rpoB [1, 2]. The Escherichia coli strain K-12 was used as outgroup. The information on the reference strains selected are shown in Table S1.

Nucleotide Sequence Accession Numbers

The Whole Genome Shotgun sequence project of *P. fluorescens* AHK-1 has been deposited at GenBank under the accession number QRBA00000000. The BioProject designation for this project is PRJNA473300.

Results and Discussion

Genome Assembly and Annotation

A total of 7,110,161 raw reads and 7,000,101 clean reads with total bases 1988 Mb were generated by Illumina pairedend sequencing. The draft genome of *P. fluorescens* AHK-1 was based on an assembly of 50 scaffolds amounting to 7,035,786 bp, with a G+C content of 60.88%. The largest scaffold was 626.42 kb and the N50 size was 331.26 kb. A total of 6327 protein-coding genes were predicted in the genome of AHK-1, and four rRNA operons, 58 tRNA loci were detected in the genome of this strain (Table 1). Of the 6327 predicted genes, 5024 (79.41%) were assigned to COGs, and 2418 genes (38.22%) in KEGG databases, respectively.

Table 1 Genomic features of P. fluorescens AHK-1

Features	AHK-1
Genome size (Mb)	7.04
No. of scaffold	50
Max scaffold length (kb)	626.42
% G+C content	60.88
Genes	6327
rRNA	4
tRNA	58
Scaffold N50 (kb)	331.26

As to the COG functional categories (Fig. S1), 13.18% (662) of the total genes were involved in "amino acid transport and metabolism" with higher proportion, and followed by 12.96% (651), 11.21% (563), and 8.92% (448) of the total genes, which are associated with "general function", "transcription", and "function unknown", respectively. The metabolic pathway analysis using KEGG orthology revealed 85 metabolic pathways. Using Blast2Go analysis, 1157, 3759, and 3482 genes participated in cellular components, molecular functions and biological processes, respectively (Fig. S2). Within the molecular functions, "catalytic activity", "transporter activity", and "binding" were highly

represented, "cellular process", "metabolic process", and "single-organism process" were the most represented GO categories within the biological process. A high percentage of genes were classified as "cell" and "cell part" under the cellular components category.

Comparative Genomic Analysis

The presence of the P. fluorescens AHK-1 orthologous coding sequences in the two genomic sequences from P. fluorescens SBW25 and GcM5-1A were assessed (Fig. 1). Within these isolates, 3998 genes formed the core and accounted for 64.87%, 76.62%, and 68.64% of the total genes in the AHK-1, GcM5-1A, and SBW25 genomes, respectively. On the other hand, 946 strain-specific genes for P. fluorescens AHK-1 was observed, which may contribute to species-specific features of this bacterium. Among them, 80.13% (758) of the genes are classified into 20 COG functional categories accounting for "general function prediction only" (11.95%), "transcription" (11.10%), "amino acid transport and metabolism" (10.36%), "signal transduction mechanisms" (8.46%), and other functions with smaller proportion. The remaining 188 unique genes (19.87%) are not classified into any COG categories (Table 2).



5	5	Ę
J	J	-

Table 2 The number of strain-specific genes of <i>P. fluorescens</i> AHK-1 associated with the COG functional categories COG functional categories	Code	Value	% of total ^a	Description
	R	113	11.95	General function prediction only
	Κ	105	11.10	Transcription
	Е	98	10.36	Amino acid transport and metabolism
	Т	80	8.46	Signal transduction mechanisms
	Р	69	7.29	Inorganic ion transport and metabolism
	S	64	6.77	Function unknown
	G	58	6.13	Carbohydrate transport and metabolism
	М	54	5.71	Cell wall/membrane/envelope biogenesis
	С	45	4.76	Energy production and conversion
	Ι	35	3.70	Lipid transport and metabolism
	U	30	3.17	Intracellular trafficking, secretion, and vesicular transport
	Q	27	2.85	Secondary metabolites biosynthesis, transport and catabolism
	Ν	27	2.85	Cell motility
	L	24	2.54	Replication, recombination and repair
	Н	22	2.33	Coenzyme transport and metabolism
	0	19	2.01	Posttranslational modification, protein turnover, chaperones
	V	11	1.16	Defense mechanisms
	J	10	1.06	Translation, ribosomal structure and biogenesis
	F	5	0.53	Nucleotide transport and metabolism
	D	3	0.32	Cell cycle control, cell division, chromosome partitioning
	_	188	19.87	Not in COGs

^aThe total is based on the unique number of protein-coding genes in AHK-1 when compared to SBW25 and GcM5-1A

Moreover, each genome analyzed in this study has 690 to nearly 1000 unique genes when compared to each other, suggesting a high variation in diversity of genome content and heterogeneity in genome organization. This is consistent with the results of the comparison of four other P. fluorescens genomes (WH6, SBW25, Pf0-1 and Pf-5) [3]. There are more predicted genes in AHK-1 than in any of the two sequenced P. fluorescens isolates. In addition, the percentage of shared orthologous groups between AHK-1 and SBW25 was 80%, between AHK-1 and GcM5-1A was 68%, and between SBW25 and GcM5-1A was 71%. This lower percentage of shared orthologous groups is consistent with those previously observed between the isolates of P. fluorescens [6, 25].

Phylogenetic Analysis

The phylogenetic tree was generated using the maximum likelihood (ML) algorithm in MEGA based on multilocus sequence analysis (Fig. 2). It is apparent that the 18 previously sequenced isolates of P. fluorescens and strain AHK-1 fall into a single large clade composed of two Sub-clades. Totally 17 P. fluorescens isolates fall into one clade, with isolate AHK-1 more distantly from others. The second clade is composed of P. fluorescens Pf0-1 and P. fluorescens R124. These results are reasonably consistent with a maximum likelihood phylogeny based on 16S rRNA (Fig. S3). Moreover, based on the observation of the phylogeny in this study, it is consistent with those results obtained from recent phylogenetic studies in which Pf0-1 represents a distinct clade clearly distinguished from SBW25 in P. fluorescens group [1, 6].

Virulence Factors

To provide a preliminary view of the genes involved in pathogenesis, we identified 72 predicted genes as candidate virulence factors of P. fluorescens AHK-1 through BLAST against the VFDB [26] (Table S2). Among these factors, the *fliA* gene encoding flagellar biosynthetic protein participate in flagellar motility, which is critical for the colonization of respective hosts by many bacterial pathogens. The sigma factor fliA is found to repress the quorum-sensing controlled transcriptional regulator HapR and allow increased expression of virulence factors in Vibrio cholera [27]. The inactivation of *fliC* gene, encoding flagellin biosynthesis, can result in increased activity of the toxin in cell culture supernatants of *Clostridium difficile* [28]. The *fliC* gene was observed in the genome of P. fluorescens GcM5-1 and considered as a crucial pathogenic factor in this strain [6]. However, BLAST searches failed to detect homology to fliC against VFDB with all of the predicted genes in AHK-1.

Fig. 2 Phylogenetic tree showing the relationship of AHK-1 with other representative isolates of P. fluorescens. The tree was generated by MEGA5 using the maximum likelihood method based on concatenated sequences of nine core housekeeping genes including acsA, aroE, guaA, gyrB, mutL, ppsA, pyrC, recA, and rpoB. Escherichia coli strain K-12 was used as outgroup. Bootstrap support for nodes (r = 1000)were shown above 50. The scale bar indicates the number of the nucleotide acid substitutions per site



We did not find a highly homologous fragment of *fliC* in the genome of AHK-1. Furthermore, several secretion systems in AHK-1 were also identified, including type VI, III, and IV secretion systems. The type III secretion system (T3SS) was considered to be necessary for full virulence of pathogenic bacteria [29, 30]. The genes for a complete and functional T3SS system were recently identified in the genomes of P. fluorescens GcM5-1A and WH6, respectively [3, 6], suggesting that T3SS may act as an important role in subverting and colonizing their hosts in the bacterium. In contrast, P. fluorescens AHK-1 was found to solely possess partial orthologs of genes encoding components of T3SS when compared with P. fluorescens GcM5-1A and P. syringae pv. tomato DC3000. One possible explanation for that is several T3SS-encoding CDSs were missing in the genome of AHK-1 because of the draft assembly described in this work. Furthermore, the type VI secretion system (T6SS) is a recently discovered virulence mechanism utilized by Gram-negative bacteria [31], and the VipA/VipB has been shown to play key roles in virulence of clinically important pathogens including V. cholera and P. aeruginosa [32]. The essential proteins for T6SS, VipA/VipB, were detected in the genome of AHK-1 in this work, suggesting which may function to be an important role in virulence mechanism of the strain.

In conclusion, we characterized the genome of *P. fluorescens* AHK-1 isolated from infected leaves of kiwifruit, which can be pathogenic and involved in kiwifruit bacterial disease. Compared to SBW25 and GcM5-1A, AHK-1 contained more strain-specific genes involved in transcription, amino acid transport and metabolism, and signal transduction mechanisms. The detected candidate virulence factors in AHK-1 provide valuably informative clues for addressing the interacting with its host. Additionally, the draft genome sequence will serve as a reference for the analysis of *P. fluorescens* isolates associated with plant disease.

Acknowledgements This work was supported by the Natural Science Foundation of Anhui Province (1608085MC56), the Key Project of Science and Technology in Anhui Province (18030701211), the National Natural Science Foundation of China (31872108), the Hefei Special Project of Science and Technology Development (ZR201711290010), the Anhui Province Program of Academic/Technology Research leader (2017H137), and the Postgraduate Innovation Fund of Anhui Agricultural University.

Compliance with Ethical Standards

Conflict of interest The authors declare no conflict of interest.

References

- Loper JE, Hassan KA, Mavrodi DV, Davis EW, Lim CK, Shaffer BT, Elbourne LDH, Stockwell VO, Hartney SL, Breakwell K et al (2012) Comparative genomics of plant-associated *Pseudomonas* spp.: insights into diversity and inheritance of traits involved in multitrophic interactions. PLoS Genet 8(7):e1002784
- Scales BS, Dickson RP, LiPuma JJ, Huffnagle GB (2014) Microbiology, genomics, and clinical significance of the *Pseudomonas fluorescens* species complex, an unappreciated colonizer of humans. Clin Microbiol Rev 27(4):927–948
- Kimbrel JA, Givan SA, Halgren AB, Creason AL, Mills DI, Banowetz GM, Armstrong DJ, Chang JH (2010) An improved, high-quality draft genome sequence of the germination-arrest

factor-producing *Pseudomonas fluorescens* WH6. BMC Genom 11:522

- 4. Li B, Wang GL, Wu ZY, Qiu W, Tang QM, Xie GL (2009) First report of bacterial head rot of broccoli caused by *Pseudomonas fluorescens* in China. Plant Dis 93(11):1219–1219
- Aiello D, Vitale A, La Ruota AD, Polizzi G, Cirvilleri G (2017) Synergistic interactions between *Pseudomonas* spp. and *Xan-thomonas perforans* in enhancing tomato pith necrosis symptoms. J Plant Pathol 99(3):731–740
- Feng K, Li RG, Chen YN, Zhao BG, Yin TM (2015) Sequencing and analysis of the *Pseudomonas fluorescens* GcM5-1A genome: a pathogen living in the surface coat of *Bursaphelenchus xylophilus*. PLoS ONE 10(10):e0141515
- Zhang LX, Li SS, He T, Tan GJ (2014) Identification and pathogenicity of *Pseudomonas fluorescens* associated with canker disease of kiwifruit in central China. J Hortic Sci Biotech 89(2):130–135
- Renzi M, Copini P, Taddei AR, Rossetti A, Gallipoli L, Mazzaglia A, Balestra GM (2012) Bacterial canker on kiwifruit in Italy: anatomical changes in the wood and in the primary infection sites. Phytopathology 102(9):827–840
- Froud KJ, Beresford RM, Cogger NC (2018) Impact of kiwifruit bacterial canker on productivity of cv. Hayward kiwifruit using observational data and multivariable analysis. Plant Pathol 67(3):671–681
- Bolger AM, Lohse M, Usadel B (2014) Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 30(15):2114–2120
- Xu XL, Cheng TY, Yang H, Yan F, Yang Y (2015) De novo sequencing, assembly and analysis of salivary gland transcriptome of *Haemaphysalis flava* and identification of sialoprotein genes. Infect Genet Evol 32:135–142
- Altschul SF, Madden TL, Schaffer AA, Zhang JH, Zhang Z, Miller W, Lipman DJ (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res 25(17):3389–3402
- Delcher AL, Salzberg SL, Phillippy AM (2003) Using MUMmer to identify similar regions in large sequence sets. Curr Protoc Bioinform https://doi.org/10.1002/0471250953.bi1003s00
- 14. Seemann T (2014) Prokka: rapid prokaryotic genome annotation. Bioinformatics 30(14):2068–2069
- Bairoch A, Apweiler R (2000) The SWISS-PROT protein sequence database and its supplement TrEMBL in 2000. Nucleic Acids Res 28(1):45–48
- Conesa A, Gotz S, Garcia-Gomez JM, Terol J, Talon M, Robles M (2005) Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. Bioinformatics 21(18):3674–3676
- Moriya Y, Itoh M, Okuda S, Yoshizawa AC, Kanehisa M (2007) KAAS: an automatic genome annotation and pathway reconstruction server. Nucleic Acids Res 35:W182–W185
- Lowe TM, Chan PP (2016) tRNAscan-SE On-line: integrating search and context for analysis of transfer RNA genes. Nucleic Acids Res 44(W1):W54–W57
- Lagesen K, Hallin P, Rodland EA, Staerfeldt HH, Rognes T, Ussery DW (2007) RNAmmer: consistent and rapid annotation of ribosomal RNA genes. Nucleic Acids Res 35(9):3100–3108
- Nawrocki EP, Kolbe DL, Eddy SR (2009) Infernal 1.0: inference of RNA alignments. Bioinformatics 25(10):1335–1337

- Guo DS, Zhao BG, Li RG, Kulinich OA, Ryss A (2008) Purification of flagellin of *Pseudomonas fluorescens* GcM5-1A carried by the pine wood nematode, *Bursaphelenchus xylophilus*, and its in vitro toxicity to a suspension of cells of *Pinus thunbergii*. Russ J Nematol 16(2):151–157
- 22. Shinya R, Morisaka H, Takeuchi Y, Ueda M, Futai K (2010) Comparison of the surface coat proteins of the pine wood nematode appeared during host pine infection and in vitro culture by a proteomic approach. Phytopathology 100(12):1289–1297
- Trippe K, McPhail K, Armstrong D, Azevedo M, Banowetz G (2013) *Pseudomonas fluorescens* SBW25 produces furanomycin, a non-proteinogenic amino acid with selective antimicrobial properties. BMC Microbiol 13:111
- Chen H, Boutros PC (2011) VennDiagram: a package for the generation of highly-customizable Venn and Euler diagrams in R. BMC Bioinform 12(1):35
- 25. Silby MW, Cerdeno-Tarraga AM, Vernikos GS, Giddens SR, Jackson RW, Preston GM, Zhang XX, Moon CD, Gehrig SM, Godfrey SA et al (2009) Genomic and genetic analyses of diversity and plant interactions of *Pseudomonas fluorescens*. Genome Biol 10:R51
- Chen LH, Zheng DD, Liu B, Yang J, Jin Q (2016) VFDB 2016: hierarchical and refined dataset for big data analysis-10 years on. Nucleic Acids Res 44(D1):D694–D697
- Tsou AM, Frey EM, Hsiao A, Liu Z, Zhu J (2008) Coordinated regulation of virulence by quorum sensing and motility pathways during the initial stages of *Vibrio cholerae* infection. Commun Integr Biol 1(1):42–44
- Aubry A, Hussack G, Chen W, KuoLee R, Twine SM, Fulton KM, Foote S, Carrillo CD, Tanha J, Logan SM (2012) Modulation of toxin production by the flagellar regulon in clostridium difficile. Infect Immun 80(10):3521–3532
- 29. Alfano JR, Charkowski AO, Deng WL, Badel JL, Petnicki-Ocwieja T, van Dijk K, Collmer A (2000) The *Pseudomonas syringae* Hrp pathogenicity island has a tripartite mosaic structure composed of a cluster of type III secretion genes bounded by exchangeable effector and conserved effector loci that contribute to parasitic fitness and pathogenicity in plants. Proc Natl Acad Sci USA 97(9):4856–4861
- Preston GM, Bertrand N, Rainey PB (2001) Type III secretion in plant growth-promoting *Pseudomonas fluorescens* SBW25. Mol Microbiol 41(5):999–1014
- Broms JE, Ishikawa T, Wai SN, Sjostedt A (2013) A functional VipA-VipB interaction is required for the type VI secretion system activity of *Vibrio cholerae* O1 strain A1552. BMC Microbiol 13(1):96–96
- 32. Sun K, Broms J, Lavander M, Gurram BK, Enquist PA, Andersson CD, Elofsson M, Sjostedt A (2014) Screening for inhibition of *Vibrio cholerae* VipA-VipB interaction identifies small-molecular compounds active against type VI secretion. Antimicrob Agents Ch 58(7):4123–4130

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.