Genome Mining and Predictive Functional Profiling of Acidophilic Rhizobacterium *Pseudomonas fluorescens* Pt14

Pooja Rani¹ · Nitish Kumar Mahato¹ · Anukriti Sharma¹ ·
Desiraju Lakshmi Narsimha Rao² · Komal Kamra³ · Rup Lal¹

Received: 28 January 2017 / Accepted: 5 April 2017 / Published online: 18 April 2017
© Association of Microbiologists of India 2017

Abstract *Pseudomonas fluorescens* Pt14 is a non-pathogenic and acidophilic bacterium isolated from acidic soil (pH 4.65). Genome sequencing of strain Pt14 was performed using Single Molecule Real Time (SMRT) sequencing to get insights into unique existence of this strain in acidic environment. Complete genome sequence of this strain revealed a chromosome of 5,841,722 bp having 5354 CDSs and 88 RNAs. Phylogenomic reconstruction based on 16S rRNA gene, Average Nucleotide Identity (ANI) values and marker proteins revealed that strain Pt14 shared a common clade with *P. fluorescens* strain A506 and strain SS101. ANI value of strain Pt14 in relation to strain A506 was found 99.23% demonstrating a very close sub-species association at genome level. Further, orthology determination among these three phylogenetic neighbors revealed 4726 core proteins. Functional analysis elucidated significantly higher abundance of sulphur metabolism ($>1\times$) which could be one of the reasons for the survival of strain Pt14 under acidic conditions (pH 4.65). Acidophilic bacteria have capability to oxidize sulphur into sulphuric acid which in turn can make the soil

acidic and genome-wide analysis of *P. fluorescens* Pt14 demonstrated that this strain contributes towards making the soil acidic.

Keywords *Pseudomonas fluorescens* · Genomics · Acidophilic · SMRT sequencing

Introduction

The Genus *Pseudomonas* is characterized by aerobic, rod-shaped, Gram-stain negative gammaproteobacterium having multiple polar flagella. Member of this genus are ubiquitous in water, soil, fly ash dumpsite, oil-contamination site and have predominant associations with the plant rhizosphere [1–6]. Some species of genus *Pseudomonas* are human pathogens e.g. *Pseudomonas aeruginosa* [7, 8] and some are plant pathogens e.g. *Pseudomonas syringae* [9–11]. However, *Pseudomonas fluorescens* is a non-pathogenic rhizobacterium that is known for its ability to produce a soluble fluorescent pigment pyoverdine, a type of siderophore [12]. These fluorescent siderophores are high affinity iron chelating agents helping bacteria in uptake of iron from their surroundings during iron deficiency [13]. Strains belonging to species *P. fluorescens* are beneficial in agriculture as they have been used as bio-control agents [12, 14]. These bacteria protect plants from plant pathogens by producing secondary metabolites like hydrogen cyanide (HCN), antibiotics, fluorescein and pyocyanin [15]. Additionally, due to widespread colonizing tendency, *P. fluorescens* is known to constrain the growth of fungi and other microorganisms by competing with them [14]. At the time of writing this manuscript, genome sequences of 95 strains of *P. fluorescens* were publicly

Nucleotide sequence Accession Number: The complete genome sequence of *Pseudomonas fluorescens* strain Pt14 has been deposited at NCBI under Genbank Accession Number CP017296.1.

✉ Rup Lal
ruplal@gmail.com

¹ Molecular Biology Laboratory, Department of Zoology, University of Delhi, Delhi 110007, India

² All India Network Project on Soil Biodiversity and Biofertilizers, Indian Institute of Soil Science, Bhopal 462038, India

³ Ciliate Biology Laboratory, SGTB Khalsa College, University of Delhi, Delhi 110007, India

available in the database (<http://www.ncbi.nlm.nih.gov/genome>).

In order to get pH specific functional insights modulating the genome of strain Pt14, we sequenced and annotated the complete genome of *P. fluorescens* Pt14, an environmental strain which was isolated from acidic rhizosphere soil (pH 4.65) of rice crop at Tinsukia, Assam, India.

Materials and Methods

DNA Isolation, Genome Sequencing and Assembly

High molecular weight DNA of *P. fluorescens* strain Pt14 was isolated using Cetyl trimethyl ammonium bromide (CTAB) NaCl method of DNA isolation [16]. In this method, cells are lysed and proteins are removed by proteinase K and remaining cell debris, polysaccharides and protein are removed by selective precipitation with CTAB. For purification chloroform:isoamylalcohol (24:1) mix was used. DNA was precipitated by 0.6 volume of isopropanol. The whole genome of strain Pt14 was sequenced at the McGill University and Génome Québec Innovation Centre (Montréal, QC, Canada) using the PacBio RS II DNA or Single Molecule Real Time (SMRT) sequencing. A total 1,126,041 raw sub-reads (1.6 Gb) were generated using SMRT Cells in a PacBio RSII sequencer. These subreads were assembled into a complete replicon of size of 5,841,722 bp using the Hierarchical Genome Assembly Process (HGAP).

Genome Annotation

Genome sequence was circularized by using UGENE [17]. For basic annotations, the genome was uploaded on RAST server version 4.0 [18] and the NCBI Prokaryotic Genomes Annotation Pipeline (PGAP; <http://www.ncbi.nlm.nih.gov/genomes/static/Pipeline.html>). KEGG Automatic Annotation Server (KAAS) [19] was used for functional annotations of genes present in the genome by amino acids sequence comparisons against a manually curated set of ortholog groups. Average Nucleotide Identity (ANI) analysis [20] with its neighbors was done to confirm the phylogenetic status of strain Pt14. Genome was uploaded on CRISPR finder [21] and IS finder [22] respectively to retrieve CRISPR arrays and IS elements. PHAST online server [23] was used to detect the intact phage ORFs present in the genome. To validate the genome assembly by the presence of essential genes, genome of strain Pt14 was uploaded on Amphora Net [24]. RNAmmer 1.2 server [25] and ARAGORN [26] were used to predict the number of copies of rRNAs and tRNAs present in the genome, respectively.

Phylogenomic Reconstruction and Functional Analyses

Phylogenomic reconstruction was performed at three levels including 16S rRNA genes, ANI values and bacterial marker proteins (Fig. 1) using close phylogenetic neighbors of *P. fluorescens* Pt14 on the basis of RAST result, namely *P. fluorescens* A506, *P. fluorescens* SS101, *P. fluorescens* SBW25, *P. fluorescens* WH6, *P. fluorescens* BS2, *P. fluorescens* NZ052, *P. fluorescens* Q2-87 and *P. fluorescens* Pfo-1 isolated from different habitats pear phyllosphere, wheat rhizosphere soil, leaf surface of sugar beet, wheat rhizosphere soil, mushroom cap, wheat rhizosphere and agricultural soil, respectively [27–31] (Table 1). 16S rRNA gene sequences of these strains were retrieved using RNAmmer server, aligned using CLUSTALX version 2.1 program [32] and Maximum Likelihood (ML) tree was constructed using MEGA software package version 6 [33]. To calculate the evolutionary distance matrix, Jukes and Cantor [34] model was used and a phylogenetic tree was constructed using Neighbour Joining method [35]. Based on a similarity matrix of the ANI values [20] among the *Pseudomonas fluorescens* strains, a phylogenetic tree was constructed using MeV_4_9_0 software and MEGA6 [33] software using hierarchical clustering [36]. Similarly, using 28 essential proteins derived by FragGeneScan1.20 [37] (DnaG, Pkg, NusA, RplA, RplB, RplC, RplD, RplE, RplF, RplE, RplF, RplL, RplM, RplN, RplP, RplS, PyrG, RpmA, RpoB, SmpB, RpsB, RpsC, RpsE, RpsI, RpsJ, RpsK, RpsM and Tsf) which were present in genomes of all the strains, a tree was constructed on MEGA6 [33].

In order to represent the relative abundance of pathways present in the genome of strain Pt14 and its nearest neighbor *P. fluorescens* A506 (clustered together with strain Pt14), pie charts were prepared using top 40 enriched pathways with the help of Krona [38] (Fig. 2).

Further, in order to understand the core-pan genome dynamics across *P. fluorescens* strains, orthologous proteins were estimated among three most closely related phylogenetic neighbors i.e., *P. fluorescens* Pt14, *P. fluorescens* A506 and *P. fluorescens* SS101 using OrthoVenn web server [39] and a venn diagram was constructed (Fig. 3).

Results and Discussion

General Genome Features

Complete genome sequence of *P. fluorescens* Pt14 was circularized into a chromosome of size 5,841,722 bp using UGENE with a read coverage of 100X. G + C content of the genome of *P. fluorescens* Pt14 was 60.3%, which is in

Fig. 1 Phylogenetic tree based on **a** 16S rRNA genes, **b** average nucleotide identity (ANI) and **c** essential marker proteins (DnaG, P_{gk}, NusA, PyrG, RplA, RplB, RplC, RplD, RplE, RplF, RplE, RplF, RplL, RplM, RplN, RplP, RplS, RpmA, RpoB, RpsB, RpsC, RpsE,

RpsI, RpsJ, RpsK, RpsM, SmpB and Tsf). In all the three methods, strains Pt14, A506 and SS101 shared a common clade and strain Pt14 clustered with strain A506

Table 1 Comparative genome features of *Pseudomonas fluorescens* strains

Name of strains of <i>Pseudomonas fluorescens</i>	Source of isolation	Genome size (bp)	%G + C content	% similarity with Pt14 (ANI value)
Pt14	Rhizosphere soil of rice plant	5,841,722	60.3	—*
A506	Pear phyllosphere	6,019,547	60.0	99.23
SS101	Rhizosphere soil of wheat plant	6,179,143	60.0	95.1
SBW25	Leaf surface of sugar beet plant	6,722,539	60.5	86.99
WH6	Rhizosphere soil of wheat	6,270,867	60.6	86.92
BS2	Soil	6,123,259	60.6	86.86
NZ052	Button mushroom	6,827,647	60.1	86.92
Q2-87	Rhizosphere soil of wheat plant	6,368,168	60.6	82.65
Pfo-1	Agricultural soil	6,438,405	60.5	82.61

* % similarity of strain Pt14 is not written, as we calculated % similarity of strain Pt14 with respect to other strains of the group

accordance with range of genomes belonging to the genus *Pseudomonas*. RAST server data analyses revealed the presence of 5354 CDSs and 88 RNAs. A short palindromic

repeat (CRISPR) element regularly interspersed in the genome and one insertion sequence (IS element) were retrieved from the genome. Two functionally intact

Fig. 2 Abundance of metabolic pathways in **a** *P. fluorescens* Pt14 and **b** *P. fluorescens* A506. Sulphur metabolism is present in 2× fold in Pt14, while in A506 it is 1× fold

Fig. 3 Distribution of core, accessory and strain specific orthologous gene clusters of *P. fluorescens* Pt14, *P. fluorescens* A506 and *P. fluorescens* SS101. A total of 4726 core orthologous proteins were shared by all three strains. Strain Pt14 and A506 have 292 accessory genes in common while 34 genes were shared by strain Pt14 and SS101. Strains Pt14, A506 and SS101 harbor 9, 5 and 11 unique strain specific genes, respectively

prophages of size 20.1 and 31.2 kb with G + C content 59.48 and 58.21% respectively were also detected. All 31 essential protein coding bacterial phylogenetic markers genes namely *dnaG*, *nusA*, *pyrG*, *rplA*, *rplB*, *rplC*, *rplD*, *rplE*, *rplF*, *rplK*, *rplL*, *rplM*, *rplN*, *rplP*, *rplS*, *rplT*, *pgk*, *rpmA*, *rpoB*, *rpsB*, *rpsC*, *rpsE*, *rpsI*, *rpsJ*, *rpsK*, *rpsM*, *rpsB*, *infC*, *smgB*, *frr* and *tsf* were present in the genome of strain Pt14 marking its complete status. Seven copies of 5S and six copies each of 23S and 16S rRNAs and 74 copies of tRNAs were present in the genome of strain Pt14. Whole genome based ANI topology confirmed that the closest neighbor of strain Pt14 is *P. fluorescens* A506 (complete genome sequence accession number CP003041) with maximum percentage similarity (99.23%).

Phylogenomic Reconstruction of Strain Pt14 and Functional Analyses

Phylogenetic analyses based on 16S rRNA gene, ANI and essential marker proteins showed that *P. fluorescens* Pt14, *P. fluorescens* A506 and *P. fluorescens* SS101 share same clade (Fig. 1) despite being isolated from different habitats i.e. rhizosphere soil of rice, phyllosphere and wheat rhizosphere, respectively. This reflected that phylogeny of strains does not depend upon the source of isolation. Strains Pt14 and A506 shared a monophyletic clade as their genomes showed 99.23% similarity demonstrating same species status. It is also evident from the three phylogenetic

trees, that the relative position of these strains remained same despite being constructed by using three different methods (Fig. 1). However, the percent ANI shared in all the neighbors ranged from 99.23% (*P. fluorescens* A506) to 82.61% (*P. fluorescens* Pfo-1) (Table 1).

Since, strains Pt14 and A506 shared maximum similarity as per the whole genome based phylogenetic tree, most of the pathways were present in similar incidences except sulphur metabolism which was present in higher fold ($>1\times$) in strain Pt14 (Fig. 2). Higher folds of sulphur metabolism present in strain Pt14 can be attributed to its ability to survive in acidic environments (pH 4.65) as acidophilic bacteria are mostly sulphur oxidizers [40–41]. These bacteria oxidize reduced sulphur to sulphuric acid thus making the habitat acidic. Also, genes involved in sulphur metabolism *cysH*, *soxB* and *soxD* were present in the genome. Sulphur oxidizing potential of such bacterium has been used in agriculture and biotechnology [43]. Hence, genetic engineering of strain Pt14 can develop it as a potential bioleaching agent to remove sulphur from contaminated soil.

Genome wide analysis of orthologous clusters revealed a total of 4726 proteins encompassing the core genome which is significantly higher owing to the sub-species status of the three genomes (Fig. 3). Apart from house-keeping genes, one-third of the core genome accounted for hypothetical proteins of unknown functions [44]. The orthologous proteins were further clustered into groups called singletons based on similar functions. 292 proteins encoding accessory or dispensable genes were found common between strains Pt14 and A506 owing to significantly high similarity (99.23%). Only 34 protein subgroups for dispensable genes were found common between these two strains. Accessory genes have specific functions that are related to survival in different niches and are associated with specific antibody or secondary metabolite production or colonization on host organisms or biofilm formation. Nine and five unique/strain specific protein subgroups were present in each genome of strains Pt14 and A506 whereas 11 subgroups were found in strain SS101. Strain specific genes or protein subgroups of metabolic islands present among the strains of same species are result of horizontal gene transfer during the course of evolution [44]. Here we observed, though strain Pt14 shares 99.23% similarity with strain A506, four strain specific protein subgroups are present required for its survival.

Conclusions

Recent advancements in high throughput next generation sequencing technologies have eased acquiring complete genome sequences of bacteria. Genomic insights into a

bacterial isolate reveal individual gene signatures and also allow predictive functional potential [45]. Here, genome versus genome analysis using ANI values showed that *P. fluorescens* Pt14 and *P. fluorescens* A506 are highly similar (99.23% similarity). Both strains clustered together in phylogenetic tree irrespective of their different habitat as well as showed feeble differences in abundance of metabolic pathways. These results illustrated sub-species level phylogeny between both the genomes. Strain Pt14 was isolated from rhizosphere soil having pH 4.65 from the field where cultivation of rice had been in practice for the past 50 years using local varieties without chemical fertilizers. The fact that sulphur metabolism is multifold in strain Pt14 in comparison to strain *P. fluorescens* A506 is a pointer to the possibility of strain Pt14 being responsible for acidity of the soil by oxidizing reduced sulphur into sulphuric acid from where it was isolated. Most of the above discussed predictions are solely based on genome sequence information obtained using bioinformatic tools in the absence of direct experimental validation. Further experiments are needed to gain insights into metabolic potential of this strain. Moreover, this study can provide a foundation for future studies focused on understanding of versatile metabolic potentials of the species *Pseudomonas fluorescens*.

Acknowledgements The authors acknowledge funds from Government of India under project from Department of Biotechnology (DBT), All India Network Project on Soil Biodiversity-Biofertilizers Indian Council of Agricultural Research (ICAR) and National Bureau of Agriculturally Important Microorganisms (NBAIM). PR, NKM and AS gratefully acknowledge ICAR, DBT and NBAIM respectively for providing research fellowships.

References

- Spiers AJ, Buckling A, Rainey PB (2000) The causes of *Pseudomonas* diversity. Microbiology 146:2345–2350. doi:10.1099/00221287-146-10-2345
- Prakash O, Kumari K, Lal R (2007) *Pseudomonas delhiensis* sp. nov., from a fly ash dumping site of a thermal powerplant. Int J Syst Evol Microbiol 57:527–531. doi:10.1099/ijs.0.64456-0
- Gupta SK, Kumari R, Prakash O, Lal R (2008) *Pseudomonas panipatensis* sp. nov., isolated from an oil-contaminated site. Int J Syst Evol Microbiol 58:1339–1345. doi:10.1099/ijs.0.65401-0
- Grewal S, Vakhlu J, Gupta V, Sangwan N, Kohli P, Nayyar N, Rani P, Sance SS, Lal R (2014) Draft genome sequence of *Pseudomonas* sp. strain JMM, a sediment-hosted environmental isolate. Genome Announc 2:000879–14. doi:10.1128/genomeA.00879-14
- Sharma A, Sangwan N, Negi V, Kohli P, Khurana JP, Rao DLN, Lal R (2015) Pan-genome dynamics of *Pseudomonas* gene complements enriched across hexachlorocyclohexane dumpsite. BMC Genom 16:313. doi:10.1186/s12864-015-1488-2
- Korshunova TY, Ramírez-Bahena MH, Chetverikov SP, Igual JM, Peix Á, Loginov O (2016) *Pseudomonas turukhanskensis* sp. nov., isolated from oil-contaminated soils. Int J Syst Evol Microbiol 66:4657–4664. doi:10.1099/ijsem.0.001406
- Lomholt JA, Poulsen K, Kilian M (2001) Epidemic population structure of *Pseudomonas aeruginosa*: evidence for a clone that is pathogenic to the eye and that has a distinct combination of virulence factors. Infect Immun 69:6284–6295. doi:10.1128/IAI.69.10.6284-6295.2001
- Kamath KS, Pascovici D, Penesyan A, Goel A, Venkatakrishnan V, Paulsen IT, Packer NH, Molloy MP (2016) *Pseudomonas aeruginosa* cell membrane protein expression from phenotypically diverse cystic fibrosis isolates demonstrates host-specific adaptations. J Proteome Res 15:2152–2163. doi:10.1021/acs.jproteome.6b00058
- Young JM (2010) Taxonomy of *Pseudomonas syringae*. J Plant Pathol 92:S5–S14. doi:10.4454/jpp.v92i1sup.2501
- Arrebola E, Cazorla FM, Durán VE, Rivera E, Olea F, Codina JC, Pérez-García A, Vicente A (2003) Mangotoxin: a novel antimetabolite toxin produced by *Pseudomonas syringae* inhibiting ornithine/arginine biosynthesis. Physiol Mol Plant Pathol 63:117–127. doi:10.1128/AEM.03007-12
- Nowell RW, Laue BE, Sharp PM, Green S (2016) Comparative genomics reveals genes significantly associated with woody hosts in the plant pathogen *Pseudomonas syringae*. Mol Plant Pathol 17:1409–1424. doi:10.1111/mpp.12423
- Garrido-Sanz D, Meier-Kolthoff JP, Göker M, Martín M, Rivilla R, Redondo-Nieto M (2016) Genomic and genetic diversity within the *Pseudomonas fluorescens* complex. PLoS ONE 11:e0150183. doi:10.1371/journal.pone.0150183
- Miethke M, Marahiel MA (2007) Siderophore-based iron acquisition and pathogen control. Microbiol Mol Biol Rev 71:413–451. doi:10.1128/MMBR.00012-07.20
- Barahona E, Navazo A, Martínez-Granero F, Zea-Bonilla T, Pérez-Jiménez RM, Martín M, Rivilla R (2011) *Pseudomonas fluorescens* F113 mutant with enhanced competitive colonization ability and improved biocontrol activity against fungal root pathogens. Appl Environ Microbiol 77:5412–5419. doi:10.1128/AEM.00320-11
- Vanitha S, Ramjagathesh R (2014) Biocontrol Potential of *Pseudomonas fluorescens* against coleus root rot disease. J Plant Pathol Microb 5:216. doi:10.4172/2157-7471.1000216
- Murray MG, Thompson WF (1980) Rapid isolation of high molecular weight plant DNA. Nucleic Acids Res 8:4321–4326. doi:10.1093/nar/8.19.4321
- Okonechnikov K, Golosova O, Fursov M (2012) Unipro UGENE: a unified bioinformatics toolkit. Bioinformatics 28:1166–1167. doi:10.1093/bioinformatics/bts091
- Aziz RK, Bartels D, Best AA, DeJongh M, Disz T, Edwards RA, Formsma K, Gerdes S, Glass EM, Kubal M, Meyer F, Olsen GJ, Olson R, Osterman AL, Overbeek RA, McNeil LK, Paarmann D, Paczian T, Parrello B, Pusch GD, Reich C, Stevens R, Vassieva O, Vonstein V, Wilke A, Zagnitko O (2008) The RAST Server: rapid annotations using subsystems technology. BMC Genom 9:75. doi:10.1186/1471-2164-9-75
- Moriya Y, Itoh M, Okuda S, Yoshizawa A, Kanehisa M (2007) KAAS: an automatic genome annotation and pathway reconstruction server. Nucleic Acids Res 35:W182–W185. doi:10.1093/nar/gkm321
- Konstantinidis KT, Tiedje JM (2005) Towards a genome-based taxonomy for prokaryotes. J Bacteriol 187:6258–6264. doi:10.1128/JB.187.18.6258-6264.2005
- Grissa I, Vergnaud G, Pourcel C (2007) CRISPRFinder: a web tool to identify clustered regularly interspaced short palindromic repeats. Nucleic Acids Res 35:W52–W57. doi:10.1093/nar/gkm360
- Siguier P, Perochon J, Lestrade L, Mahillon J, Chandler M (2006) ISfinder: the reference centre for bacterial insertion sequences. Nucleic Acids Res 34:D32–D36. doi:10.1093/nar/gkj014

23. Zhou Y, Liang Y, Lynch K, Dennis JJ, Wishart DS (2011) PHAST: a fast phage search tool. *Nucleic Acids Res* 39:W347–W352. doi:[10.1093/nar/gkr485](https://doi.org/10.1093/nar/gkr485)
24. Kerepesi C, Bánky D, Grolmusz V (2014) AmphoraNet: the webserver implementation of the AMPHORA2 metagenomic workflow suite. *Gene* 533:538–540. doi:[10.1016/j.gene.2013](https://doi.org/10.1016/j.gene.2013)
25. 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* 35:3100–3108. doi:[10.1093/nar/gkm160](https://doi.org/10.1093/nar/gkm160)
26. Laslett D, Canback B (2004) ARAGORN, a program for the detection of transfer RNA and transfer-messenger RNA genes in nucleotide sequences. *Nucleic Acids Res* 32:11–16. doi:[10.1093/nar/gkf459](https://doi.org/10.1093/nar/gkf459)
27. Loper JE, Hassan KA, Mavrodi DV, Davis EW II, Lim CK, Shaffer BT, Elbourne LD, Stockwell VO, Hartney SL, Breakwell K, Henkels MD, Tetu SG, Rangel LI, Kidarsa TA, Wilson NL, van de Mortel JE, Song C, Blumhagen R, Radune D, Hostetler JB, Brinkac LM, Durkin AS, Kluepfel DA, Wechter WP, Anderson AJ, Kim YC, Pierson LS, Pierson EA, Lindow SE, Kobayashi DY, Raaijmakers JM, Weller DM, Thomashow LS, Allen AE, Paulsen IT (2012) Comparative genomics of plant-associated *Pseudomonas* spp.: insights into diversity and inheritance of traits involved in multitrophic interactions. *PLoS Genet* 8:e1002784. doi:[10.1371/journal.pgen.1002784](https://doi.org/10.1371/journal.pgen.1002784)
28. Silby MW, Cerdeño-Tárraga AM, Vernikos GS, Giddens SR, Jackson RW, Preston GM, Zhang XX, Moon CD, Gehrig SM, Godfrey SA, Knight CG, Malone JG, Robinson Z, Spiers AJ, Harris S, Challis GL, Yaxley AM, Harris D, Seeger K, Murphy L, Rutter S, Squares R, Quail MA, Saunders E, Mavromatis K, Brettin TS, Bentley SD, Hotherhall J, Stephens E, Thomas CM, Parkhill J, Levy SB, Rainey PB, Thomson NR (2009) Genomic and genetic analyses of diversity and plant interactions of *Pseudomonas fluorescens*. *Genom Biol* 10:R51. doi:[10.1186/gb-2009-10-5-r51](https://doi.org/10.1186/gb-2009-10-5-r51)
29. Kimbrel JA, Givan SA, Halgren AB, Creason AL, Mills DJ, 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. doi:[10.1186/1471-2164-11-522](https://doi.org/10.1186/1471-2164-11-522)
30. Stabler RA, Negus D, Pain A, Taylor PW (2013) Draft genome sequences of *Pseudomonas fluorescens* BS2 and *Pseudomonas noertemannii* BS8, soil bacteria that cooperate to degrade the poly- γ -D-glutamic acid anthrax capsule. *Genome Announc* 1:e00057-12. doi:[10.1128/genomeA.00057-12](https://doi.org/10.1128/genomeA.00057-12)
31. Winsor GL, Griffiths EJ, Lo R, Dhillon BK, Shay JA, Brinkman FS (2016) Enhanced annotations and features for comparing thousands of *Pseudomonas* genomes in the *Pseudomonas* genome database. *Nucleic Acids Res* 44:D646–D653. doi:[10.1093/nar/gkv1227](https://doi.org/10.1093/nar/gkv1227)
32. Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, Valentin F, Wallace IM, Wilm A, Lopez R, Thompson JD, Gibson TJ, Higgins DG (2007) Clustal W and clustal X version 2.0. *Bioinformatics* 23:2947–2948. doi:[10.1093/bioinformatics/btm404](https://doi.org/10.1093/bioinformatics/btm404)
33. Tamura K, Stecher G, Peterson D, Filipski A, Kumar S (2013) MEGA6: molecular evolutionary genetics analysis version 6.0. *Mol Biol Evol* 30:2725–2729. doi:[10.1093/molbev/mst197](https://doi.org/10.1093/molbev/mst197)
34. Jukes TH, Cantor CR (1969) Evolution of protein molecules. In: Munro HN (ed) *Mammalian protein metabolism*, vol 3. Academic Press, New York, pp 21–132
35. Saitou N, Nei M (1987) The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol Biol Evol* 4:406–425. doi:[10.1093/oxfordjournals.molbev.a040454](https://doi.org/10.1093/oxfordjournals.molbev.a040454)
36. Eisen MB, Spellman PT, Brown PO, Botstein D (1998) Cluster analysis and display of genome-wide expression patterns. *Proc Natl Acad Sci USA* 95:14863–14868. doi:[10.1073/pnas.95.25.14863](https://doi.org/10.1073/pnas.95.25.14863)
37. Rho M, Tang H, Ye Y (2010) FragGeneScan: predicting genes in short and error-prone reads. *Nucleic Acids Res* 38:e191. doi:[10.1093/nar/gkq747](https://doi.org/10.1093/nar/gkq747)
38. Ondov BD, Bergman NH, Phillippy AM (2011) Interactive metagenomic visualization in a Web browser. *BMC Bioinformatics* 12:385. doi:[10.1186/1471-2105-12-385](https://doi.org/10.1186/1471-2105-12-385)
39. Wang Y, Coleman-Derr D, Chen G, Gu YQ (2015) OrthoVenn: a web server for genome wide comparison and annotation of orthologous clusters across multiple species. *Nucleic Acids Res* 43:W78–W84. doi:[10.1093/nar/gkv487](https://doi.org/10.1093/nar/gkv487)
40. Johnson DB, Hallsberg KB (2008) Carbon, iron and sulfur metabolism in acidophilic micro-organisms. *Adv Microb Physiol* 54:201–255. doi:[10.1016/S0065-2911\(08\)00003-9](https://doi.org/10.1016/S0065-2911(08)00003-9)
41. Oren A (2010) Acidophiles. In: eLS. John Wiley & Sons Ltd, Chichester. <http://www.els.net>. doi:[10.1002/9780470015902.a0000336.pub2](https://doi.org/10.1002/9780470015902.a0000336.pub2)
42. Zhang X, Liu X, Liang Y, Fan F, Zhang X, Yin H (2016) Metabolic diversity and adaptive mechanisms of iron and/or sulfur oxidizing autotrophic acidophiles in extremely acidic environments. *Environ Microbiol Rep* 8:738–751. doi:[10.1111/1758-2229.12435](https://doi.org/10.1111/1758-2229.12435)
43. Chen SY, Lin JG (2004) Bioleaching of heavy metals from livestock sludge by indigenous sulfur-oxidizing bacteria: effects of sludge solids concentration. *Chemosphere* 54:283–289. doi:[10.1016/j.chemosphere.2003.08.009](https://doi.org/10.1016/j.chemosphere.2003.08.009)
44. Tettelin H, Maignani V, Cieslewicz MJ, Donati C, Medini D, Ward NL, Angiuoli SV, Crabtree J, Jones AL, Durkin AS, DeBoy RT (2005) Genome analysis of multiple pathogenic isolates of *Streptococcus agalactiae*: implications for the microbial “pan-genome”. *Proc Natl Acad Sci USA* 102:13950–13955. doi:[10.1073/pnas.0506758102](https://doi.org/10.1073/pnas.0506758102)
45. Zhao L, Li H, Zhu Z, Wakefield MR, Fang Y, Ye Y (2017) Genomic sequencing of a strain of *Acinetobacter baumannii* and potential mechanisms to antibiotics resistance. *Infect Genet Evol* 50:20–24. doi:[10.1016/j.meegid.2017.02.001](https://doi.org/10.1016/j.meegid.2017.02.001)