Virulence analysis of 81 of Pseudomonas aeruginosa genomes available in public sequence databases

Marcin Brzozowski (marcin.brzozowski@pum.edu.pl)
Pomorski Uniwersytet Medyczny w Szczecinie https://orcid.org/0000-0002-1376-7812
Joanna Barbara Jursa-Kulesza
Department of Medical Microbiology, Pomeranian Medical University in Szczecin
Danuta Kosik-Bogacka
Department of Biology and Medical Parasitology, Pomeranian Medical University in Szczecin

Research article

Keywords: Pseudomonas aeruginosa, Virulence factors, Core genome, Accessory genome, Comparative genomics

Posted Date: June 6th, 2019

DOI: https://doi.org/10.21203/rs.2.10146/v1

Abstract

Background Pseudomonas aeruginosa is a pathogen capable of causing a wide range of severe opportunistic infections. Its genome contains numerous virulence genes encoding secretion systems of different types, structures responsible for adhesion and motility, toxins, proteases, siderophores, and others. The aim of this study is to analyse virulence, population structure, and distribution of highly divergent genes among 81 P. aeruginosa strains available in whole genome sequence databases. Results For this purpose, 260 virulence genes were searched in 81 different P. aeruginosa whole genomes that were available in databases. We identified most of the virulence genes as core and softcore genes. The most of the highly divergent sequences encoding pyoverdines, flagella and pilA were acknowledged as accessory, because of the differences in sequence among different alleles of those genes. Phylogenetic tree revealed the existence of three genetic groups of P. aeruginosa. Strains of the first clade were characterised as ExoS positive, whiles genomes of the second clade were ExoU positive. The member of third clade, PA7 strain was the only strain deprived of all T3SS genes. The analysis of pyoverdine locus facilitated finding a new pyoverdine type similar to pyoverdine type III. This newly described variant was present in 7 different strains. It contained a gene that was probably created by the fusion of pilD and pilI genes. In order to determine the coexistence of genes encoding exoenzymes, flagella and pyoverdines, Pearson correlation coefficients were calculated. There were significant correlations between genes encoding ExoS/ExoU-type strains and genes encoding type-A/type-B flagella. The correlation also occurred between Conclusion This study facilitates describing genetic differences of various P. aeruginosa strains based on Pseudomonas aeruginosa whole genome information from online databases. We conclude that most P. aeruginosa virulence genes are present in more than 95% of available genomes of the species. There are correlations of occurrence of different P. aeruginosa accessory virulence genes.

Introduction

Pseudomonas aeruginosa is a gram-negative, rod-shaped, opportunistic pathogen capable of causing many life-threatening or chronic infections. It is one of the main etiological factors of nosocomial infections, especially for immunocompromised patients, with extended burns, aids, and cystic fibrosis.

Possession of five different secretion systems by *P. aeruginosa* has been described up to date (T1SS, T2SS, T3SS, T5SS, T6SS). Different secretion systems have different effector proteins. T1SS is responsible for the secretion of alkaline protease *aprA*, *lasA*, *lasB*, and exotoxin *toxA* are secreted by T2SS. T3SS is a major virulence determinant of *P. aeruginosa.* Exoenzymes *exoS*, *exoU*, *exoT*. *exoY* are secreted by this system. The presence of *exoS* and *exoU* is mutually exclusive among *P. aeruginosa* isolates *exoS* and *exoT* sequences are highly similar, sequence identity between those genes is 76% [1]. Interestingly, there are significant differences between strains harbouring different exotoxins. Based on recent phylogenetical analyses published by Stewart et al [2], whether the *P. aeruginosa* is *exoS* or *exoU* positive or both negative, it belongs to different genetic clade.

Other important *P. aeruginosa* virulence factors include structures responsible for adhesion (pili, flagella), motility (flagella), slime production (alginate), siderophores biosynthesis (pyoverdine) or quorum sensing mechanism [3-8].

Some of the *P. aeruginosa* virulence genes can be categorized into various types with different amino acid sequences [9-12]. Genes encoding pyoverdine, flagella, *pilA* fimbrial protein and lipopolysccharide (LPS) can be highlighted as the ones with divergent sequences. The highest known sequence diversity in *P. aeruginosa* is observed for *pilA* gene. The pairwise differences between *pilA* variants are up to 71.3% [13]. The long chain of polysaccharide (O-antigen) is a component of bacterial LPS. Region encoding B-band of *P. aeruginosa* O-antigen is also highly divergent between different strains. Eleven conservative sets of genes were recognised in this region [11]. Each group of genes is highly divergent from one another. Pyoverdine is a major *P. aeruginosa* siderophore. Based on the differences in sequence of *fvpA*, *pvdE*, *pvdD*, *pvdJ* and *pvdI* genes in the pyoverdine locus, three different pyoverdine types were identified. According to the different gene set, bacteria produce structurally distinct pyoverdines [12]. Sequence diversity is exhibited in sequences of flagella encoding genes as well. Flagella of type A or B are produced by different *P. aeruginosa* isolates. The production of each type of flagella is facilitated by the possession of different gene set in flagellin biosynthesis locus. Amino acid sequences of different chains building flagella are identical between types in 63-65% [14].

Over the last decade, rapid development of new sequencing methods increased effectiveness of sequencing while reducing the cost. These recent advantages help researchers to sequence hundreds of thousands of full genomes of different organisms. This development has a great impact on microbiology. Whole genome sequences have been utilised to highlight genetic variations within species or to assess the size of the pan, core genome of different microorganisms. Core genes are described as the genes that are present in all strains of one microorganism. Those sequences are usually responsible for encoding essential factors for bacteria. Term ‘accessory genes’ refers to the genes that can be found in a subset of strains of a species. Those genes are accountable for interspecies variability. They contribute to individual features of different strains, for example: the ability to colonise different host organisms [15]. Together with the increasing number of whole sequenced genomes, it is now possible to precisely assess the size of the core and accessory genome. In recent studies, core genome of *P. aeruginosa* was assessed based on over a dozen complete genomes [16-18]. Genomic data was also used recently to analyse bacterialvirulence, antimicrobial resistance, metabolism and population structure [19, 20].

In this work, 81 *P. aeruginosa* whole genomes of different strains from environmental and clinical sources were used. We analysed the virulence and population structure of those strains. In this analysis, we redefined core, softcore, and accessory virulence genome of *P. aeruginosa*. Distribution of different variants of highly divergent genes (HDGs) encoding flagella and pyoverdines in downloaded genomes was also investigated. We decided to examine if there is correlation of occurrence of different accessory genes encoding exoenzymes, flagella and pyoverdines.

Results

Sequences of five genes including *algC*, *pilA*, *pscP*, *vgrg1b* were not detected with gene finding software. Among all investigated genes, 109 were acknowledged as core genes, 101 as softcore genes, and 50 as accessory genes (Supplementary table S1).

In a group of genes associated with T3SS, 40 of them are softcore genes and 4 are accessory genes. Genes encoding exoenzymes *exoU*, *exoS*, *exoY*, *exsE*, and *pscP* were accessory. *ExoU* was found in 23, *exoS* in 57, *exoY* in 70 and *exsE* in 75 genomes out of 81. Effector proteins of T3SS, *exoS* and *exoU* are mutually exclusive in examined genomes. The conducted phylogenetic analysis of *P. aeruginosa* strains showed the existence of two large genetic clads. Strains of the first clade were characterised as ExoS-positive, whiles genomes of the second clade were ExoU-positive. We have not found any genome containing both *exoS* and *exoU* genes. Dendrogram is shown in figure 1. PA7 was the only strain deprived of all T3SS genes. Strain PA7 is genetically distant from other *P. aeruginosa* groups and is was not included in the tree.

Figure 1. Evolutionary relationship of various *Pseudomonas aeruginosa* strains based on five housekeeping genes. The tree was constructed using the neighbour-joining (NJ) method with 1,000 bootstraps. The presence of genes encoding *exoU*, *exoS*, pyoverdine and flagellar types of different strains is shown to the right of the tree.

In a group of genes encoding Type IV pili 16 sequences were present in all genomes, and 5 were identified as soft-core genes. Sequence of *pilA* gene and sequences in locus *fimT-pilE* were identified as accessory. Gene *pilA* was searched manually as its sequence was not recognised by gene finding software. We detected 76 genomes containing different variants of *pilA* gene. All variants had similar conservative sites and were located in locus characteristic of this gene. *PilA* sequences were highly divergent. Sequence similarity network (SSN) has been constructed. Our SSN divided all pilA sequences into nine different groups. Separation of the groups is demonstrated in figure 2.

Figure 2. Sequence similarity network of various *pilA* genes of investigated *P. aeruginosa* strains. Sequences in each group have at least 80% sequence identity over 80% query coverage. Based on applied conditions *pilA* genes were divided into nine different groups.

Among genes associated with alginate production, only *mucA* is identified as accessory. The genes encoding effector toxins of the T2SS are core and softcore. *ToxA* and *plcH* sequences are found in 80 genomes, whereas *lasA* and *lasB* in all the analysed genomes. Alkaline protease *aprA* - effector protein of T1SS is found in 80 strains. Rhamnolipids encoding genes *rhlA* and *rhlB* are core sequences. Genes encoding T6SS are qualified as core and softcore genes except the sequence of *fha1* gene which is accessory and is found in 76 genomes. All sequences encoding phenazines except *phzH* are qualified as core genes. In the most strains, more than one copy of each gene encoding phenazines has been detected. Distribution of all genes in core, softcore and accessory genome is illustrated in figure 4.

P. aerugiosa have flagella of type A or type B, depending on the set of genes that strain possesses [10]. Among all the genes involved in flagella biosynthesis, 34 of 49 are found in core and softcore genome. Genetic variants of genes *flgK*, *flgL*, *fliC*, *fliD*, *flel/flag*, *fliS*, *fliS'* and *fliT* encoding A-type flagella are present in 50-51 (61,7%-63%) analysed genomes. Variants of those genes encoding B-type flagella are found in 30 (37%) of strains.

Different genes involved in pyoverdine biosynthesis were core, softcore, and accessory. All accessory genes were found in pyoverdine divergent locus. In this locus, three known sets of genes can be found. Alleles of these different sets are mutually exclusive. First set, known as type I is present in 27 of 81 analysed genomes. Alleles of pyoverdine type II are the most abundant. Genes of these type are found in 37 of 81 strains. Pyoverdine type III is found in 16 strains. The analysis of pyoverdine divergent locus in analysed genomes revealed the existence of previously unknown set of genes similar to the genes of pyoverdine type III. We refer to this set of genes as type IIIb. This type contains *pvdE*, *fpvA* and siderophore-interacting protein *(sip)* gene alleles of pyoverdine type III. However, type IIIb is devoid of *pvdI*, *pvdJ* or *pvdD* genes. Instead there is a long coding sequence which is a 5' side highly similar to *pvdI* of type III, and at 3' side almost identical as *pvdD* gene of type III. We referred to this sequence as *pvdID*. There is no similarity between *pvdID* sequence and *pvdJ* gene. Figure 3 demonstrates the locus structure of pyoverdine type IIIa and IIIb. Type IIIb occurred in strains isolated in Brazil.

As we have analysed the distribution of HDGs between different *P. aeruginosa* strains, we have also decided to determine the correlation of occurrence of different accessory genes encoding exoenzymes, flagella and pyoverdines. Symmetric similarity matrix was formed after all r-values were calculated (additional files 2 and 3). There were significant correlations between genes encoding: (1) exoenzymes exoS/exoU and type-A/type-B flagella [r = -0.32/0.32, p <0.01], (2) type-II pyoverdine and flagella type-A/type-B [r = 0,37/-037; p < 0.01]. There were also significant correlations between occurrence of *exoS* and *exoU* genes and between occurrence of genes encoding different flagella types. However those correlations here were obvious as those genes are mutually exclusive. It is possible for both *exoU* and *exoS* strains to possess different types of flagella or pyoverdine. The only exception are genes of pyoverdine type III. There is no *exoU* positive genome with this type of pyoverdine.

Figure 3. Structure of pyoverdine type IIIb in comparison with type IIIa.

Figure 4. Distribution of genes associated with different virulence factors in core, softcore, and accessory genome of *P. aeruginosa*.

Discussion

The purpose of this research was to characterise virulence of 81 *P. aeruginosa* strains isolated worldwide, based on their genomic data. We regrouped virulence genes according to whether they are a part of core, softcore or accessory genome and analyse the distribution of HDGs between different strains that were isolated worldwide. This research facilitated demonstrating possible genetic differences and similarities in virulence between various *P. aeruginosa* strains.

P. aeruginosa has multiple virulence factors including toxins, proteases, pyoverdines, pili, flagella, secretion systems, or quorum sensing mechanism. In a group of T3SS encoding genes, no core gene was found. Lack of T3SS core genes is caused by the fact that one of the strains PA7 is deprived of T3SS. PA7 also does not have effector exoenzymes *exoS*, *exoU*, *exoY* and *exoT* of this system. According to this research and recently published articles, PA7 strain is genetically distant to other *P. aeruginosa* strains. There are known cases of isolation of other *exoS* and *exoU* negative PA7-like strains related to PA7 [2, 62, 63]. Research published by Stewart et al. 2014, demonstrates the existence of three genetic clades within *P. aeruginosa* species [2]. The first group was characterised as exoS positive, second as *exoU* positive. The third group was devoid of both *exoU* and *exoS* genes. In this group, PA7 strain and other PA7-like strains were found.

The phylogenetic analysis conducted in the research has similar structure to that published by Stewart et. al. [2]. Similarly, *exoS* and *exoU* positive strains are separated into two clades, and PA7 is the third branch of the tree. According to Kulasekara et al. 2006, both *exoS* and *exoU* genes together with whole T3SS apparatus were acquired by horizontal gene transfer from other microorganisms [64]. It explains the existence of the third PA7 group. In this clade, there are *P. aeruginosa* strains that have never acquired T3SS system with effector proteins. More frequent isolation of T3SS positive strains can be explained by potential evolutionary benefits caused by this system, and in result there is a genetic advantage of T3SS positive strains over T3SS negative strains.

Genes encoding pyoverdines, flagella and LPS and *pilA* have divergent sequences between different strains [9, 11, 12]. The pairwise differences between *pilA* variants are up to 71,3% [13].

PilA gene cannot be determined as core gene using standard cut-off values as this gene has highly divergent sequences between different genomes [10]. Although *pilA* sequence is not core or softcore, its function is still probably very important as different variants of this gene were found in 76 genomes. The presence of *pilA* in genomes was confirmed using manual search of genomes, as sequence of this gene were not detected by gene finding software.

Large number of genes encoding pyoverdines were acknowledged as accessory genes, because their sequence similarities between alleles of those genes exhibited more than cut-off values used for separating gene families. However, we did not find any genome lacking different gene sets of pyoverdines. Different variants of *fpvA*, *pvdE*, *pvdY* genes were found in all investigated genomes, and all found variants were located in the pyoverdine locus. Based on these results, disregarding that pyoverdine genes are accessory, it is seen that pyoverdine production and pyoverdine receptors are crucial for *P. aeruginosa* and therefore those accessory genes have core functions.

Similarly to *exoS* and *exoU* genes it is suggested that different types of pyoverdine genes were acquired by horizontal gene transfer. History of horizontal gene transfer is suggested by unusual codon usage and oligonucleotide usage in pyoverdine region [12]. Nevertheless, *exoS* and *exoU* distribution in various genomes is consistent with genetic clades of phylogenetic tree representing evolutionary relationships of *P. aeruginosa* strains. This is not observed for pyoverdine types. It is possible to find both pyoverdine

types of type I and II in *exoU* and ExoS clade. To explain these differences, we suggest that sets of genes encoding pyoverdines were transferred to *P. aeruginosa* strains of different genetic clades or they were transferred between different *P. aeruginosa* strains. Another question arises there. Why genes of different pyoverdine variants are mutually exclusive. We suppose that the possible benefits for strains harbouring two pyoverdines gene sets may be insufficient in comparison to the disadvantages, e.g. additional energy and resource consumption. In our study, we did not observe genomes with mixtures of alleles of different types, but the existence of strains with type I and type II pvd genes was described previously [12]. Those isolates may stand as strains that after acquisition of two types of pvd genes, recombined them and as a result the mixture strain containing type I and type II pvd genes was created. Throughout the analysis of pyoverdine locus in 7 strains, we found new undescribed coding sequence. This gene at 5' side is nearly identical with a part of *pvdD*(3) gene of pyoverdine type III, and at 3' site is similar to *pvdI*(3). This gene is presumably a fusion gene that was created after deletion of the locus between *pvdD* and pvdI genes. This gene has been found in 7 investigated genomes.

Genes encoding different flagella types A and B are distributed between ExoU and ExoS clades. Similarly to pyoverdine genes, in this case horizontal gene transfer is also presumably responsible for distribution of different sets of genes between ExoU and ExoS clades. We also noticed that there is significant correlation between ExoS/ExoU clades and type of flagella. Most of strains of ExoU clade had also of type-A flagella. This correlation is difficult to explain as there is no direct linkage of function between those genes. However both exoenzymes and flagella are acknowledged as virulence factors, and therefore we can suppose that there is some kind of beneficial effect for *exoU* strains to also harbour flagella of type II gene sets. There was also a significant correlation between occurrence of genes encoding pyoverdines of type II with different flagella type. In this situation beneficial effect could also explain this correlation. Similarly to pyoverdines, most flagella encoding genes were acknowledged as accessory. However it is seen that most of those genes have core functions, because through the analysis, we did not find a genome deprived of flagellar gene set.

Conclusion

We analysed 81 whole genome sequences of *P. aeruginosa* strain available in NCBI Reference Sequence Database. This study facilitates describing genetic differences between various *P. aeruginosa* strains. We conclude that the most *P. aeruginosa* virulence genes are present in more than 95% of available genomes of the species. There is significant correlation in distribution of different variants of flagella and pyoverdine sets of genes with ExoU and ExoS clades. The analysis of pyoverdine locus facilitated finding a new pyoverdine type similar to pyoverdine type III. This newly described variant was present in 7 different strains. It contained a gene that was probably created by the fusion of *pilD* and *pilI* genes.

Methods

Bacterial isolates and their source of genomic data

We used 81 whole genome sequences of *P. aeruginosa*. The nucleotide sequenceswere downloaded from NCBI Reference Sequence Database (https://www.ncbi.nlm.nih.gov/refseq/)[24]. A description of the isolates is shown in table 1. The subject sequences originated from clinical and environmental sources. Genomes of chosen isolates had different sizes ranging from 6.04 to 7.5 Mb.

Table 1. Description of the whole genome sequences of *Pseudomonas aeruginosa*.

Virulence genes

P. aeruginosa reference virulence geneswere downloaded from Virulence Factors of Pathogenic Bacteria database (https://www.mgc.ac.cn/VFs/) and UniProt Knowledgebase (https://www.uniprot.org/uniprot/) [52, 53]. In the research, we used 260 genes associated with virulence, encoding various secretion systems, type IV pili, flagella, alginate, lipopolysaccharide (LPS), proteins responsible for pigmentation, proteases, toxins, and components of quorum sensing mechanism (Additional file 1).

Analysis of core and accessory virulence genes

In the first stage of our analysis, the genes were recognised in the genomes with bioinformatical tool Prodigal [54]. We evaluated the capacity of Prodigal to correctly predict the presence of reference virulence genes in reference genomes PAO1 and UCBPA-14. Sequences of correctly predicted genes were used as a database in Diamond Blast [55]. HDG and genes that were not identified by Prodigal were searched manually in the annotated genomes available on RefSeq database and Pseudomonas Genome Database website [24, 56]. To identify gene homologues, we performed reference genes against all genes similarity search using Diamond Blast [55]. The genes have been clustered with the removal of similarities <80% sequence identity and <80% query and subject coverage. Sequences were recognised as core genes when homologues of reference genes were detected in all genomes. Soft core genes were identified as genes present in 95% of genomes. Below this value, all genes were qualified as accessory genes.

Variants of highly divergent genes in different genomes

The presence of different variants of HDG encoding *pilA*, flagella, and pyoverdines between different genomes was verified. The variants of those HDGs were searched in annotated genomes available in RefSeq database and Pseudomonas Genome Database using inbuilt BLAST software and known reference gene sequences from UniProt database [53, 56]. Coding sequences that were found exhibiting at least 30% sequence identity, 30% query coverage with the refence genes were identified as possible variants of HDG. In the next step, genomic location of each identified possible variant was determined. When the location of the sequence was the same as for reference gene, we referred to the sequence as a

genetic variant of HDG. The presence and location of the genes in specific locus were determined with the use of simple bioinformatic tools Blast and NCBI's Genome Workbench graphical viewer [57, 58].

Phylogenetic analysis

Sequences of several core genes *aroE*, *carA*, *gyrB*, *rpoB*, *rpoD* were extracted from the genomes from Pseudomonas Genome Database [56]. Those core genes were used in phylogenetic analyses of different *P. aeruginosa* strains [59]. Genes were concatenated with SequenceMatrix software v.1.8 and compared with each other with the use of bioinformatic tool MEGA v.7.0.26 [60, 61]. The phylogenetic tree was constructed using the neighbour-joining (NJ) method with 1,000 bootstraps.

Cooccurrence of different virulence genes in different genomes.

Correlation in distribution of accessory virulence genes in analysed genomes was additionally investigated. We focused on occurrence of non-divergent genes *exoS* and *exoU* with different variants of flagella and different sets of pyoverdine locus genes in analysed genomes.

Abbreviations

HDG(s): Highly divergent gene(s), LPS: Lipopolysaccharide, NJ: Neighbour-joining, PVD: Pyoverdine, SSN: Sequence similarity network, T1SS: Type I secretion system, T2SS: Type II secretion system, T3SS: Type III secretion system, T5SS: Type V secretion system, T6SS: Type VI secretion system

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Availability of data and materials

All data generated or analysed during this study are included in this published article [and its supplementary information files].

Competing interests

The authors declare that they have no competing interests

Funding

The publication of this article is supported by the Department of Microbiology, Immunology and

Laboratory Medicine, Pomeranian Medical University.

Authors' contributions

MB designed the study and performed the bioinformatic analyses. All authors contributed with writing, reviewing and editing. All authors read and approved the final version of the manuscript.

Acknowledgments

Agata Filipowska (Department of Information Systems, University of Economics and Busines), Bartosz Wojciuk (Department of Immunological Diagnostics, Pomeranian Medical University) for assistance with statistical analysis.

References

1. Filloux A. Protein Secretion Systems in Pseudomonas aeruginosa: An Essay on Diversity, Evolution, and Function. Front Microbiol. 2011;2:155. https://doi.org/10.3389/fmicb.2011.00155.
2. Stewart L, Ford A, Sangal V, Jeukens J, Boyle B, Kukavica-Ibrulj I, et al. Draft genomes of 12 host-adapted and environmental isolates of Pseudomonas aeruginosa and their positions in the core genome phylogeny. Pathog Dis. 2014;71:20-25. https://doi.org/10.1111/2049-632X.12107.
3. Doig P, Todd T, Sastry PA, Lee KK, Hodges RS, Paranchych W, et. al. Role of pili in adhesion of Pseudomonas aeruginosa to human respiratory epithelial cells. Infect Immun. 1988;56:1641-1646.
4. Feldman M, Bryan R, Rajan S, Scheffler L, Brunnert S, Tang H, et al. Role of flagella in pathogenesis of Pseudomonas aeruginosa pulmonary infection. Infect Immun 1998;66:43-51.
5. Gacesa P. Bacterial alginate biosynthesis–recent progress and future prospects. Microbiology. 1998;144 (Pt 5):1133-1143.
6. Meyer JM, Stintzi A, De Vos D, Cornelis P, Tappe R, Taraz K, et. al. Use of siderophores to type pseudomonads: the three Pseudomonas aeruginosa pyoverdine systems. Microbiology. 1997;143(Pt1):35-43. https://doi.org/10.1099/00221287-143-1-35.
7. Rocchetta HL, Burrows LL, Lam JS. Genetics of O-antigen biosynthesis in Pseudomonas aeruginosa. Microbiol Mol Biol Rev. 1999;63:523-553.
8. Rumbaugh KP, Griswold JA, Hamood AN. The role of quorum sensing in the in vivo virulence of Pseudomonas aeruginosa. Microbes Infect. 2000;2:1721-1731.
9. Arora SK, Dasgupta N, Lory S, Ramphal R. Identification of two distinct types of flagellar cap proteins, FliD, in Pseudomonas aeruginosa. Infect Immun. 2000;68:1474-1479.
10. Klockgether J, Wurdemann D, Wiehlmann L, Binnwies T, Ussery D, Tummler B. Genome Diversity of Pseudomonas aeruginosa. In: Cornelis P, editor. Pseudomonas : genomics and molecular biology, Norfolk: Caister Acad Press; 2008. P. 24-25.

11. Lam JS, Taylor VL, Islam ST, Hao Y, Kocincova D. Genetic and Functional Diversity of Pseudomonas aeruginosa Lipopolysaccharide. Front Microbiol. 2011;2:118. https://doi.org/10.3389/fmicb.2011.00118.
12. Smith EE, Sims EH, Spencer DH, Kaul R, Olson MV. Evidence for diversifying selection at the pyoverdine locus of Pseudomonas aeruginosa. J Bacteriol. 2005;187:2138-2147. https://doi.org/10.1128/JB.187.6.2138-2147.2005.
13. Kiewitz C, Tummler B. Sequence diversity of Pseudomonas aeruginosa: impact on population structure and genome evolution. J Bacteriol. 2000;182:3125-3135.
14. Spangenberg C, Heuer T, Burger C, Tummler B. Genetic diversity of flagellins of Pseudomonas aeruginosa. FEBS Lett. 1996;396:213-217.
15. Vernikos G, Medini D, Riley DR, Tettelin H. Ten years of pan-genome analyses. Curr Opin Microbiol. 2015;23:148-154. https://doi.org/10.1016/j.mib.2014.11.016.
16. Grosso-Becerra MV, Santos-Medellin C, Gonzalez-Valdez A, Mendez JL, Delgado G, Morales-Espinosa R, et . al. Pseudomonas aeruginosa clinical and environmental isolates constitute a single population with high phenotypic diversity. BMC Genomics. 2014;15:318. https://doi.org/10.1186/1471-2164-15-318.
17. Ozer EA, Allen JP, Hauser AR. Characterization of the core and accessory genomes of Pseudomonas aeruginosa using bioinformatic tools Spine and AGEnt. BMC Genomics. 2014;15:737. https://doi.org/10.1186/1471-2164-15-737.
18. Valot B, Guyeux C, Rolland JY, Mazouzi K, Bertrand X, Hocquet D. What It Takes to Be a Pseudomonas aeruginosa? The Core Genome of the Opportunistic Pathogen Updated. PLoS One. 2015;10:e0126468. https://doi.org/10.1371/journal.pone.0126468.
19. Holt KE, Wertheim H, Zadoks RN, Baker S, Whitehouse CA, Dance D, et al. Genomic analysis of diversity, population structure, virulence, and antimicrobial resistance in Klebsiella pneumoniae, an urgent threat to public health. Proc Natl Acad Sci U S A. 2015;112(27):E3574-81. https://doi.org/10.1073/pnas.1501049112.
20. Lòpez-Fernàndez S, Sonego P, Moretto M, Pancher M, Engelen K, Pertot, et al. Whole-genome comparative analysis of virulence genes unveils similarities and differences between endophytes and other symbiotic bacteria. Frontiers in microbiology. 2015;6,419. https://doi.org/10.3389/fmicb.2015.00419.
21. Ichise YK, Kosuge T, Uwate M, Nakae T, Maseda H. Complete Genome Sequence of Pseudomonas aeruginosa Strain 8380, Isolated from the Human Gut. Genome Announc. 2015;3(3):e00520-15. https://doi.org/10.1128/genomeA.00520-15.
22. Karna SL, Chen T, Chen P, Peacock TJ, Abercrombie JJ, Leung KP. Genome Sequence of a Virulent Pseudomonas aeruginosa Strain, 12-4-4(59), Isolated from the Blood Culture of a Burn Patient. Genome Announc. 2016;4(2):e00079-16. https://doi.org/10.1128/genomeA.00079-16.
23. Boyle B, Fernandez L, Laroche J, Kukavica-Ibrulj I, Mendes CM, Hancock RW, et al. Complete genome sequences of three Pseudomonas aeruginosa isolates with phenotypes of polymyxin B adaptation

and inducible resistance. J Bacteriol. 2012;194:529-530. https://doi.org/10.1128/JB.06246-11.

24. O'Leary NA, Wright MW, Brister JR, Ciufo S, Haddad D, McVeigh R, et al. Reference sequence (RefSeq) database at NCBI: current status, taxonomic expansion, and functional annotation. Nucleic Acids Res. 2016;4;44(D1):D733-45. https://doi.org/10.1093/nar/gkv1189.
25. Zhong C, Nelson M, Cao G, Sadowsky MJ, Yan T. Complete Genome Sequence of the Triclosan- and Multidrug-Resistant Pseudomonas aeruginosa Strain B10W Isolated from Municipal Wastewater. Genome Announc. 2017;5(3):e01489-16. https://doi.org/10.1128/genomeA.01489-16.
26. Sanjar F, Karna SL, Chen T, Chen P, Abercrombie JJ, Leung KP. Whole-Genome Sequence of Multidrug-Resistant Pseudomonas aeruginosa Strain BAMCPA07-48, Isolated from a Combat Injury Wound. Genome Announc. 2016;4(4):e00547-16. https://doi.org/10.1128/genomeA.00547-16.
27. Yin Y, Withers TR, Johnson SL, Yu HD. Draft Genome Sequence of a Mucoid Isolate of Pseudomonas aeruginosa Strain C7447m from a Patient with Cystic Fibrosis. Genome Announc. 2013;1(5):e00837-13. https://doi.org/10.1128/genomeA.00837-13.
28. Valot B, Rohmer L, Jacobs MA, Miller SI, Bertrand X, Hocquet D. Comparative Genomic Analysis of Two Multidrug-Resistant Clinical Isolates of ST395 Epidemic Strain of Pseudomonas aeruginosa Obtained 12 Years Apart. Genome Announc. 2014;2(3):e00515-14. https://doi.org/10.1128/genomeA.00515-14.
29. Rau MH, Marvig RL, Ehrlich GD, Molin S, Jelsbak L. Deletion and acquisition of genomic content during early stage adaptation of Pseudomonas aeruginosa to a human host environment. Environ Microbiol. 2012;14:2200-2211. https://doi.org/10.1111/j.1462-2920.2012.02795.x.
30. Shi Z, Ren D, Hu S, Hu X, Wu L, Lin H, et. al. Whole genome sequence of Pseudomonas aeruginosa F9676, an antagonistic bacterium isolated from rice seed. J Biotechnol 2015;211:77-78. https://doi.org/10.1016/j.jbiotec.2015.07.015.
31. Silo-Suh LA, Suh SJ, Ohman DE, Wozniak DJ, Pridgeon JW. Complete Genome Sequence of Pseudomonas aeruginosa Mucoid Strain FRD1, Isolated from a Cystic Fibrosis Patient. Genome Announc. 2015;3(2):e00153-15. https://doi.org/10.1128/genomeA.00153-15.
32. Winstanley C, Langille MG, Fothergill JL, Kukavica-Ibrulj I, Paradis-Bleau C, Sanschagrin F, et al. Newly introduced genomic prophage islands are critical determinants of in vivo competitiveness in the Liverpool Epidemic Strain of Pseudomonas aeruginosa. Genome Res. 2009;19:12-23. https://doi.org/10.1101%2Fgr.086082.108.
33. Wu DQ, Ye J, Ou HY, Wei X, Huang X, He YW, et. al. Genomic analysis and temperature-dependent transcriptome profiles of the rhizosphere originating strain Pseudomonas aeruginosa M18. BMC Genomics. 2011;12:438. https://doi.org/10.1186/1471-2164-12-438.
34. Ohtsubo Y, Sato T, Kishida K, Tabata M, Ogura Y, Hayashi T, et. al. Complete Genome Sequence of Pseudomonas aeruginosa MTB-1, Isolated from a Microbial Community Enriched by the Technical Formulation of Hexachlorocyclohexane. Genome Announc 2014;2(1):e01130-13. https://doi.org/10.1128/genomeA.01130-13.

35. Sangare L, Zhao Y, Folly YM, Chang J, Li J, Selvaraj JN, et. al. Aflatoxin B(1) degradation by a Pseudomonas strain. Toxins. 2014;6:3028-3040. https://doi.org/10.3390/toxins6103028.
36. Tada T, Miyoshi-Akiyama T, Shimada K, Shiroma A, Nakano K, Teruya K, et al. A Carbapenem-Resistant Pseudomonas aeruginosa Isolate Harboring Two Copies of blaIMP-34 Encoding a Metallo-beta-Lactamase. PLoS One. 2016;11:e0149385. https://doi.org/10.1371/journal.pone.0149385.
37. Miyoshi-Akiyama T, Kuwahara T, Tada T, Kitao T, Kirikae T. Complete genome sequence of highly multidrug-resistant Pseudomonas aeruginosa NCGM2.S1, a representative strain of a cluster endemic to Japan. J Bacteriol. 2011;193:7010. https://doi.org/10.1128/JB.06312-11.
38. Wang K, Chen YQ, Salido MM, Kohli GS, Kong JL, Liang HJ, et al. The rapid in vivo evolution of Pseudomonas aeruginosa in ventilator-associated pneumonia patients leads to attenuated virulence. Open Biol. 2017;7(9):170029. https://doi.org/10.1098/rsob.170029.
39. Lu S, Le S, Li G, Shen M, Tan Y, Zhao X, et al. Complete Genome Sequence of Pseudomonas aeruginosa PA1, Isolated from a Patient with a Respiratory Tract Infection. Genome Announc. 2015;3(6):e01453–15. https://doi.org/10.1128/genomeA.01453-15.
40. Nascimento AP, Ortiz MF, Martins WM, Morais GL, Fehlberg LC, Almeida LG, et. al. Intraclonal Genome Stability of the Metallo-beta-lactamase SPM-1-producing Pseudomonas aeruginosa ST277, an Endemic Clone Disseminated in Brazilian Hospitals. Front Microbiol. 2016;7:1946. https://doi.org/10.3389/fmicb.2016.01946.
41. Feng Y, Jonker MJ, Moustakas I, Brul S, Ter Kuile BH. Dynamics of Mutations during Development of Resistance by Pseudomonas aeruginosa against Five Antibiotics. Antimicrob Agents Chemother. 2016;60:4229-4236. https://doi.org/10.1128/AAC.00434-16.
42. Liu J, Li L, Peters BM, Li B, Chen D, Xu Z, Shirtliff ME. Complete genomic analysis of multidrug-resistance Pseudomonas aeruginosa Guangzhou-Pae617, the host of megaplasmid pBM413. Microb Pathog. 2018;117:265-269. https://doi.org/10.1016/j.micpath.2018.02.049.
43. Deraspe M, Alexander DC, Xiong J, Ma JH, Low DE, Jamieson FB, et al. Genomic analysis of Pseudomonas aeruginosa PA96, the host of carbapenem resistance plasmid pOZ176. FEMS Microbiol Lett. 2014;356:212-216. https://doi.org/10.1111/1574-6968.12435.
44. Qiu D, Eisinger VM, Head NE, Pier GB, Yu HD. ClpXP proteases positively regulate alginate overexpression and mucoid conversion in Pseudomonas aeruginosa. Microbiology. 2008;154:2119-2130. https://doi.org/10.1099/mic.0.2008/017368-0.
45. Vergnaud G, Midoux C, Blouin Y, Bourkaltseva M, Krylov V, Pourcel C. Transposition Behavior Revealed by High-Resolution Description of Pseudomonas Aeruginosa Saltovirus Integration Sites. Viruses. 2018;10(5):245. https://doi.org/10.3390/v10050245.
46. Jeukens J, Boyle B, Bianconi I, Kukavica-Ibrulj I, Tummler B, Bragonzi A, et. al. Complete Genome Sequence of Persistent Cystic Fibrosis Isolate Pseudomonas aeruginosa Strain RP73. Genome Announc. 2013;1(4):e00568-13. https://doi.org/10.1128/genomeA.00568-13.
47. Eckweiler D, Bunk B, Sproer C, Overmann J, Haussler S. Complete Genome Sequence of Highly Adherent Pseudomonas aeruginosa Small-Colony Variant SCV20265. Genome Announc.

2014;2(1):e01232-13. https://doi.org/10.1128/genomeA.01232-13 .

48. Lee DG, Urbach JM, Wu G, Liberati NT, Feinbaum RL, Miyata S, et al. Genomic analysis reveals that Pseudomonas aeruginosa virulence is combinatorial. Genome Biol. 2006;7(10):R90. https://doi.org/10.1186/gb-2006-7-10-r90.
49. Miller CL, Chen T, Chen P, Leung KP. Genome Sequence of Highly Virulent Pseudomonas aeruginosa Strain VA-134, Isolated from a Burn Patient. Genome Announc. 2016;4(1):e01662-15. https://doi.org/10.1128/genomeA.01662-15.
50. Murugan N, Malathi J, Umashankar V, Madhavan HN. Unraveling genomic and phenotypic nature of multidrug-resistant (MDR) Pseudomonas aeruginosa VRFPA04 isolated from keratitis patient. Microbiol Res. 2016;193:140-149. https://doi.org/10.1016/j.micres.2016.10.002.
51. Chan KG, Yin WF, Lim YL. Complete Genome Sequence of Pseudomonas aeruginosa Strain YL84, a Quorum-Sensing Strain Isolated from Compost. Genome Announc. 2014;2(2):e00246-14. https://doi.org/10.1128/genomeA.00246-14.
52. Chen L. Zheng D. Liu B. Yang J. Jin Q. VFDB 2016: hierarchical and refined dataset for big data analysis–10 years on. Nucleic Acids Res. 2016;44:D694–D697. https://doi.org/10.1093/nar/gkv1239.
53. UniProt Consortium. The universal protein resource (UniProt). Nucleic Acids Res. 2007;36(Database issue):D190-5. https://doi.org/10.1093/nar/gkm895.
54. Hyatt D, Chen GL, Locascio PF, Land ML, Larimer FW, Hauser LJ. Prodigal: prokaryotic gene recognition and translation initiation site identification. BMC Bioinformatics. 2010;11:119. https://doi.org/10.1186/1471-2105-11-119.
55. Buchfink B, Xie C, Huson DH. Fast and sensitive protein alignment using DIAMOND. Nat Methods. 2015; 12:59-60. https://doi.org/10.1038/nmeth.3176.
56. Winsor GL, Lam DK, Fleming L, Lo R, Whiteside MD, Yu, N. Y, et al. Pseudomonas Genome Database: improved comparative analysis and population genomics capability for Pseudomonas genomes. Nucleic acids research, 2010;39(Database issue),D596-600. https://doi.org/10.1093/nar/gkq869.
57. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. J Mol Biol. 1990;Oct5;215(3):403-10.
58. NCBI's Sequence Viewer. [https://www.ncbi.nlm.nih.gov/projects/sviewer/].
59. Gomila M, Pena A, Mulet M, Lalucat J, Garcia-Valdes E. Phylogenomics and systematics in Pseudomonas. Front Microbiol. 2015;6:214. https://doi.org/10.3389/fmicb.2015.00214.
60. Vaidya G, Lohman DJ, Meier R. SequenceMatrix: concatenation software for the fast assembly of multi-gene datasets with character set and codon information. Cladistics. 2011;27:171-180. https://doi.org/10.1111/j.1096-0031.2010.00329.x.
61. Tamura K, Dudley J, Nei M, Kumar S. MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. Mol Biol Evol. 2007;24:1596-1599. https://doi.org/10.1093/molbev/msm092.

62. Hilker R, Munder A, Klockgether J, Losada PM, Chouvarine P, Cramer N, et al. Interclonal gradient of virulence in the Pseudomonas aeruginosa pangenome from disease and environment. Environ Microbiol. 2015;17:29-46. https://doi.org/10.1111/1462-2920.12606.
63. Klockgether J, Tummler B. Recent advances in understanding Pseudomonas aeruginosa as a pathogen. F1000Res. 2017;6:1261. https://doi.org/10.12688/f1000research.10506.1.
64. Kulasekara BR, Kulasekara HD, Wolfgang MC, Stevens L, Frank DW, Lory S. Acquisition and evolution of the exoU locus in Pseudomonas aeruginosa. J Bacteriol. 2006;188:4037-4050. https://doi.org/10.1128/JB.02000-05.

Table

Table 1. Description of the whole genome sequences of *Pseudomonas aeruginosa*.

Strain name	Origin	Genome size (Mb)	GC content (%)	Refseq accession number	Reference
PAO1	Clinical	6.26	66.8%	NC_002516.2	Stover et al. (2000)
FA-HZ1	Environmental	6.87	66.2%	NZ_CP017353.1	Ali et al. (2017)
W45909	Clinical	6.78	66.2%	NZ_CP008871.2	Yan et al. (unpublished)
NHmuc	Unknown	6.21	66.6%	NZ_CP013479.1	Irvine et al. (unpublished)
SCVJan	Animal	6.21	66.6%	NZ_CP013478.1	Irvine et al. (unpublished)
SCVFeb	Animal	6.21	66.6%	NZ_CP013477.1	Irvine et al. (unpublished)
8380	Clinical	6.61	66.2%	NZ_AP014839.1	Ichise et al. (2015)
NCTC 10332	Unknown	6.32	66.5%	NZ_LN831024.1	Unknown
DHS01	Clinical	7.06	65.8%	NZ_CP013993.1	Valot et al. (2014)
PA1	Clinical	6.50	66.4%	NC_022808.2	Lu et al. (2015)
H27930	Clinical	6.57	66.2%	NZ_CP008860.2	Yan et al. (unpublished)
T52373	Clinical	6.32	66.5%	NZ_CP008867.1	Yan et al. (unpublished)
F22031	Clinical	6.60	66.2%	NZ_CP007399.1	Yan and Xavier (unpublished)
DK2	Clinical	6.40	66.3%	NC_018080.1	Rau et al. (2012)
PA154197	Unknown	6.44	66.4%	NZ_CP014866.1	Cao et al. (unpublished)
PA121617	Clinical	6.43	66.4%	NZ_CP016214.1	Liu et al. (2018)
ATCC 27853	Clinical	6.82	66.1%	NZ_CP015117.1	Feng et al. (2016)
F9670	Clinical	6.79	66.1%	NZ_CP008873.1	Yan and Xavier (unpublished)
H5708	Clinical	6.33	66.5%	NZ_CP008859.2	Yan and Xavier (unpublished)
T38079	Clinical	6.79	66.1%	NZ_CP008866.2	Yan and Xavier (unpublished)
S86968	Clinical	6.93	66.0%	NZ_CP008865.2	Yan et. al. (unpublished)
F63912	Clinical	6.62	66.2%	NZ_CP008858.2	Yan and Xavier (unpublished)
W36662	Clinical	6.79	66.2%	NZ_CP008870.2	Yan et. al. (unpublished)
T63266	Clinical	6.46	66.3%	NZ_CP008868.1	Yan and Xavier (unpublished)
12-4-4(59)	Clinical	6.43	66.3%	NZ_CP013696.1	Karna et al. (2016)
RP73	Clinical	6.34	66.5%	NC_021577.1	Jeukens et al. (2013)
PA1088	Clinical	6.72	66.1%	NZ_CP015001.1	Nascimento et al. (2016)
PA7790	Clinical	7.02	66.0%	NZ_CP014999.1	Nascimento et al. (2016)
PA8281	Clinical	6.93	66.0%	NZ_CP015002.1	Nascimento et al. (2016)
PA11803	Clinical	7.01	66.0%	NZ_CP015003.1	Nascimento et al. (2016)
USMARC-41639	Animal	6.36	66.4%	NZ_CP013989.1	Harhay et al. (unpublished)
19BR	Clinical	6.74	66.1%	NZ_AFXJ01000001.1	Boyle et al. (201(2)
213BR	Clinical	6.72	66.1%	NZ_AFXK01000001.1	Boyle et al. (201(2)
PA_D21	Clinical	6.64	66.2%	NZ_CP012582.1	Wang et al. (2017)
PA_D5	Clinical	6.69	66.2%	NZ_CP012579.1	Wang et al. (2017)
PA_D25	Clinical	6.68	66.2%	NZ_CP012584.1	Wang et al. (2017)
PA_D22	Clinical	6.68	66.2%	NZ_CP012583.1	Wang et al. (2017)
PA_D16	Clinical	6.68	66.2%	NZ_CP012581.1	Wang et al. (2017)
PA_D9	Clinical	6.65	66.2%	NZ_CP012580.1	Wang et al. (2017)
PA_D2	Clinical	6.64	66.2%	NZ_CP012578.1	Wang et al. (2017)
PA_D1	Clinical	6.64	66.2%	NZ_CP012585.1	Wang et al. (2017)
VA-134	Clinical	6.40	66.4%	NZ_CP013245.1	Miller et al. (2016)
W60856	Clinical	6.90	66.2%	NZ_CP008864.2	Yan et al. (unpublished)
VRFPA04	Clinical	6.81	66.5%	NZ_CP008739.1	N et al. (2016)
X78812	Clinical	6.37	66.4%	NZ_CP008872.2	Yan and Xavier (unpublished)

W16407	Clinical	6.81	65.9%	NZ_CP008869.2	Yan et al. (unpublished)
PA96	Clinical	6.44	66.3%	NZ_CP007224.1	Deraspe et al. (2014)
YL84	Environmental	6.43	66.4%	NZ_CP007147.1	Chan et al. (2014)
N17-1	Environmental	6.37	66.4%	NZ_CP014948.1	Sangare et al. (2014)
IOMTU 133	Clinical	6.90	66.0%	NZ_AP017302.1	Tada et al. (unpublished)
H47921	Clinical	6.84	66.1%	NZ_CP008861.1	Yan and Xavier (unpublished)
F23197	Clinical	6.54	66.2%	NZ_CP008856.2	Yan et al. (unpublished)
Carb01 63	Clinical	7.50	65.6%	NZ_CP011317.1	Kraak and van der Zee (unpublished)
FRD1	Clinical	6.71	66.1%	NZ_CP010555.1	Silo-Suh et al. (2015)
C7447m	Clinical	6.26	66.4%	NC_022360.1	Yin et al. (2013a)
BAMC 07-48	Clinical	7.02	66.0%	NZ_CP015377.1	Sanjar et al. (2016)
PAER4_119	Unknown	6.50	66.4%	NZ_CP013113.1	Walter (unpublished)
F30658	Clinical	7.27	65.8%	NZ_CP008857.1	Yan and Xavier (unpublished)
MTB-1	Environmental	6.58	66.2%	NC_023019.1	Ohtsubo et al. (2014)
PACS2	Clinical	6.49	66.3%	NZ_AAQW01000001.1	Smith et al. (unpublished)
B10W	Environmental	6.72	66.2%	NZ_CP017969.1	Zhong et al. (2017)
PcyII-10	Clinical	6.29	66.5%	NZ_LT673656.1	Vergnaud et al. (2018)
F9676	Environmental	6.37	66.5%	NZ_CP012066.1	Shi et al. (2015)
LESlike1	Clinical	6.51	66.4%	NZ_CP006984.1	Jeukens et al. (2014)
LESB65	Clinical	6.53	66.4%	NZ_CP006983.1	Jeukens et al. (2014)
LES431	Clinical	6.55	66.3%	NC_023066.1	Jeukens et al. (2014)
LESB58	Clinical	6.60	66.3%	NC_011770.1	Winstanley et al. (2009)
M18	Environmental	6.33	66.5%	NC_017548.1	Wu et al. (2011)
SCV20265	Clinical	6.73	66.3%	NC_023149.1	Eckweiler et al. (2014)
S0490	Clinical	7.10	66.0%	NZ_CP011369.1	Kraak and Van der Zee (unpublished)
DN1	Environmental	6.64	66.3%	NZ_CP017099.1	Ma (unpublished)
NCGM1900	Clinical	6.81	66.0%	NZ_AP014622.1	Tada et al. (unpublished)
NCGM1984	Clinical	6.85	66.0%	NZ_AP014646.1	Tada et al. (2016)
B136-33	Clinical	6.42	66.4%	NC_020912.1	Lo et al. (unpublished)
NCGM2.S1	Clinical	6.76	66.1%	NC_017549.1	Miyoshi-Akiyama et al. (2011)
M37351	Clinical	6.90	66.0%	NZ_CP008863.1	Yan and Xavier (unpublished)
M1608	Clinical	6.48	66.0%	NZ_CP008862.2	Yan and Xavier (unpublished)
UCBPP-PA14	Clinical	6.54	66.3%	NC_008463.1	Lee et al. (2006)
NCGM257	Clinical	7.09	65.9%	NZ_AP014651.1	Akiyama et al. (unpublished)
PA7	Clinical	6.59	66.4%	NC_009656.1	Roy et al. (2010)

Supplemental File Legend

Additional file 1: List of the *P. aeruginosa* virulence genes used in the study.

Additional file 2: Symmetric similarity matrix with calculated r and p values.

Additional file 3: Correlogram display of correlation matix for *exoS*, *exoU*, genes and pyoverdine and flagella gene sets in analysed genomes.

Figures

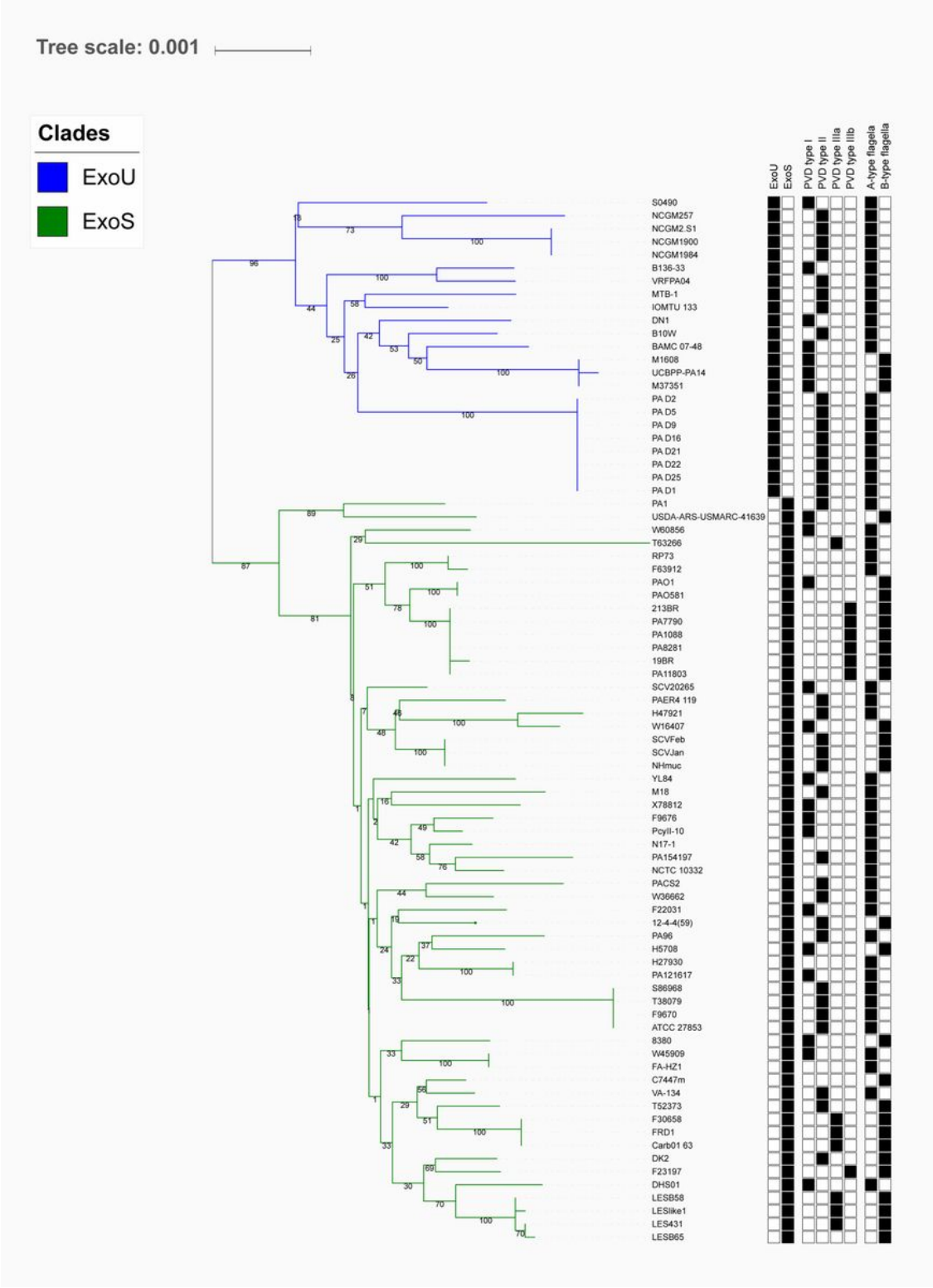

Figure 1

Evolutionary relationship of various Pseudomonas aeruginosa strains based on five housekeeping genes. The tree was constructed using the neighbour-joining (NJ) method with 1,000 bootstraps. The presence of genes encoding exoU, exoS, pyoverdine and flagellar types of different strains is shown to the right of the tree.

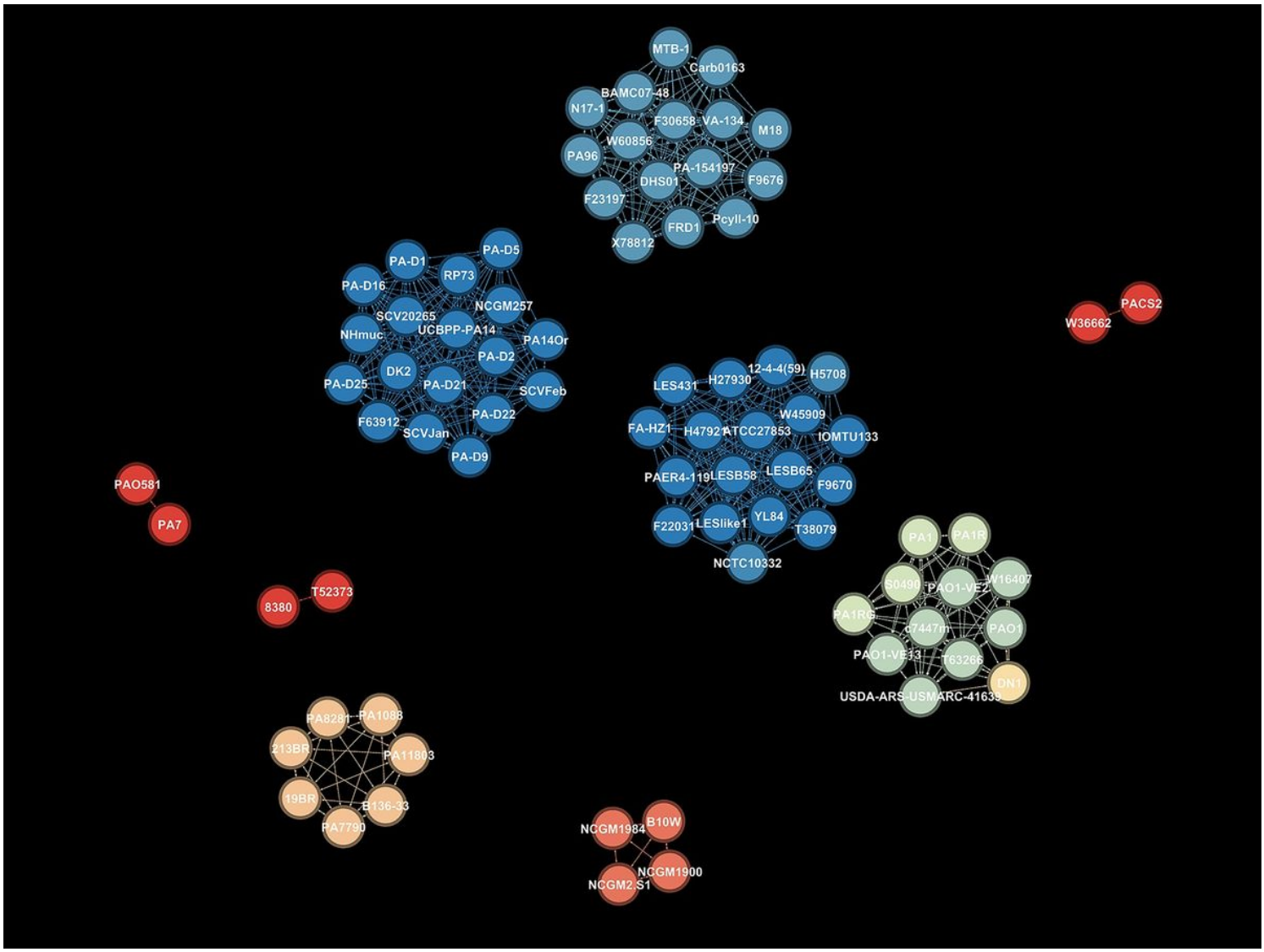

Figure 2

Sequence similarity network of various pilA genes of investigated P. aeruginosa strains. Sequences in each group have at least 80% sequence identity over 80% query coverage. Based on applied conditions pilA genes were divided into nine different groups.

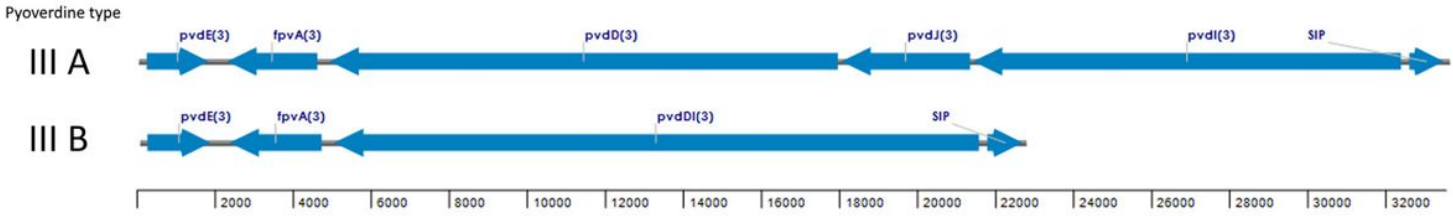

Figure 3

Structure of pyoverdine type IIIb in comparison with type IIIa.

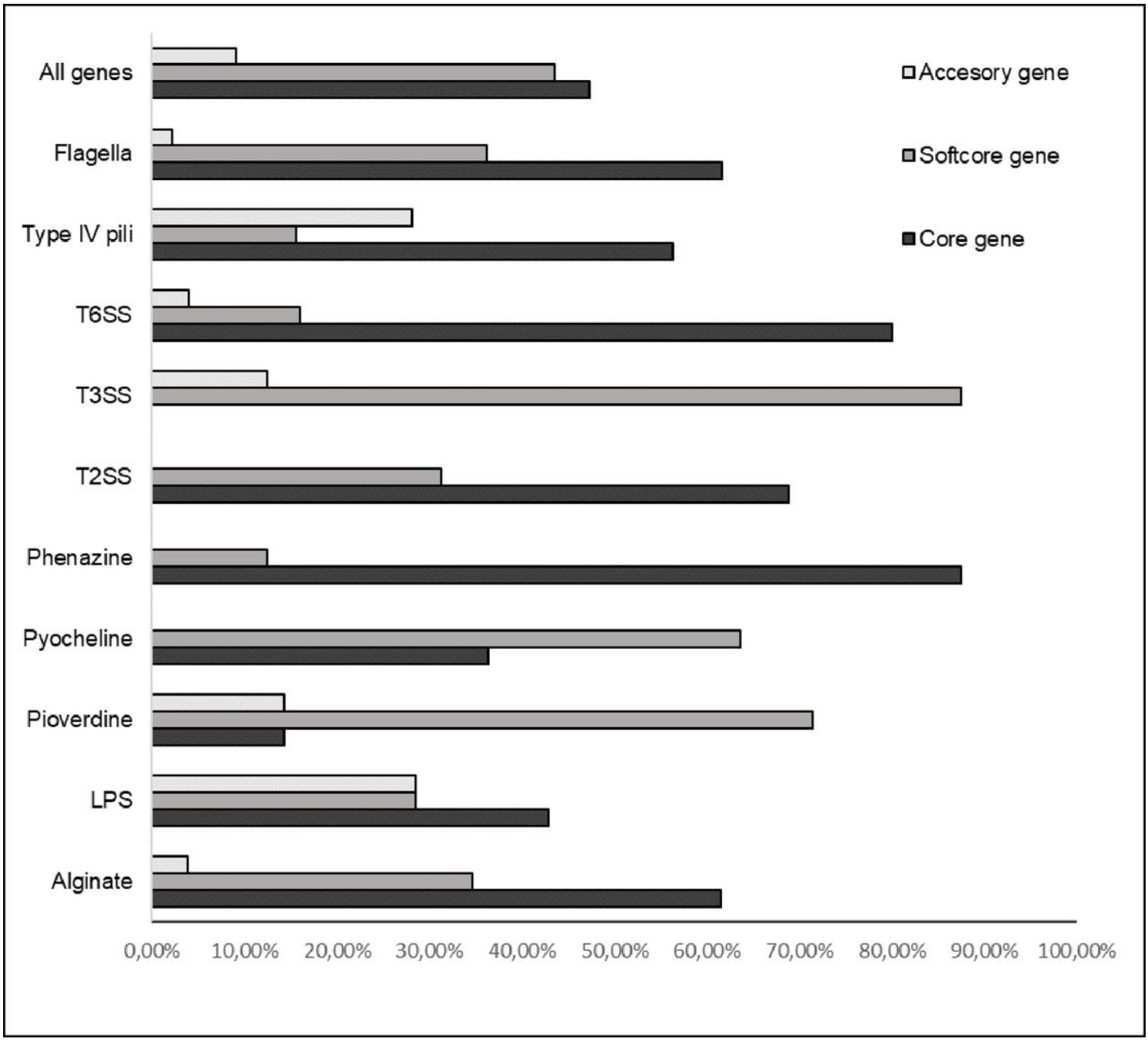

Figure 4

Distribution of genes associated with different virulence factors in core, softcore, and accessory genome of P. aeruginosa.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Additionalfile3.tif
- Additionalfile1.docx
- Virulenceanalysiswithtablesandfigures.docx
- Additionalfile2.xlsx