Genetic determinants of virulence and extensive drug resistance in *Pseudomonas aeruginosa* PPA14 isolated from eggplant rhizosphere

Sakthivel Ambreetha[1,2,#], Govindasamy Parshatd[3], Christian Castellanos[2], Giri Narasimhan[4], Dananjeyan Balachandar[1,*], Trevor Cickovski[4,*] and Kalai Mathee[2,5,*]

[1]Department of Agricultural Microbiology, Tamil Nadu Agricultural University, Coimbatore, Tamil Nadu, India.

[2]Department of Human and Molecular Genetics, Herbert Wertheim College of Medicine, Florida International University, Miami, Florida, USA.

[#]Current affiliation – Developmental Biology and Genetics, Indian Institute of Science, Bengaluru, Karnataka, India

[3]Faculty of Applied Sciences, Asian Institute of Medicine, Science and Technology, Bedong, Kedah, Malaysia.

[4]Knight Foundation School of Computing and Information Sciences, Florida International University, Miami, Florida, USA.

[5]Biomolecular Sciences Institute, Florida International University, Miami, Florida, USA.

*Corresponding authors:

Kalai Mathee (matheelabfor65roses@gmail.com)

Trevor Cickovski (movingpictures83@gmail.com)

Dananjeyan Balachandar (dbalu@tnau.ac.in)

Running title. Genome analyses of plant-associated *Pseudomonas aeruginosa.*

Conflict of Interest. The authors declare no conflict of interest.

Abstract

Pseudomonas aeruginosa is one of the Priority Level I critical pathogens that are least sensitive to antibiotics and can cause fatal hospital-acquired infections. This bacterium is predominantly present in the agricultural ecosystem. However, there are very limited studies on health threats associated with *P. aeruginosa* strains flourishing in edible plants. Previously, we isolated and characterized 18 *P. aeruginosa* strains from vegetable plants directly harvested from the farms. In the current work, it has been hypothesized that plant-associated *P. aeruginosa* harbors genetic determinants for virulence and resistance. To test this hypothesis, *in vitro* resistome profiles of the plant-associated *P. aeruginosa* strains were assessed based on the Kirby-Bauer disk diffusion method. Hierarchical clustering analysis was done to identify the plant-associated strains that are phenotypically similar to clinical isolates. An eggplant-associated strain, PPA14, that exhibited high virulence and extensive *in vitro* resistance against eight antibiotic classes was selected for complete genome analyses. The PPA14 genome was sequenced using the Solexa-Illumina and Oxford-Nanopore platforms, assembled, and annotated. The presence of virulence-related and antibiotic resistance (ABR) genes were predicted using the ABRicate tool and validated based on standard reference databases such as VFDB, NCBI AMRFinderPlus, MEGARes, CARD, and ResFinder. IslandViewer4 tool was used to predict the genes acquired through horizontal gene transfer. Additionally, comparative analyses of all the plant-associated and environmental *P. aeruginosa* genomes characterized so far were done using the Roary tool. The PPA14 genome size was 6.72 Mbp, encoding 6315 open reading frames. The genome harbored 49 ABR genes, including those coding for multiple families of efflux pumps that collectively confer resistance against 11 antibiotic classes. In addition, we detected 225 virulence-related genes, 83 genomic islands, and 235 unique genes in the PPA14 genome. Over 4% of the PPA14 genome is devoted to conferring virulence and extensive drug resistance. Our report highlights the health threat associated with an eggplant-associated *P. aeruginosa*.

Keywords

Comparative genomics, Genomic Islands, Extensive drug resistance, Plant-associated *Pseudomonas aeruginosa,* Virulence-associated genes

Abbreviations

ABR, Antibiotic resistance; MDR, Multi-drug resistant; HGT, Horizontal gene transfer; PPA, plant-associated *Pseudomonas aeruginosa*; LB, Luria Bertani; CTAB, hexadecyl-trimethyl ammonium bromide; Prokka, rapid prokaryotic genome annotation; RAST, Rapid annotation using subsystem technology; NCBI, National Center for Biotechnology Information; CARD, Comprehensive antibiotic resistance database; CGE, Center for Genomic Epidemiology; VFDB, Virulence factor database; GI, genomic island; CGView, Circular genome viewer; ORF, Open reading frames; RAxML, Randomized axelerated maximum likelihood; ANOVA, Analysis of variance; Mbp, Million base pairs; RND, Resistance-nodulation-division; MFS, Major facilitator superfamily; MATE, Multidrug and toxic compound extrusion; SMR, Small multidrug resistance; CDS, protein-coding sequences

Data Summary

All sequence data generated in this study were deposited in NCBI GenBank (Accession no. MT734694 to MT734711).

Introduction

Antibiotic-resistant (ABR) pathogens are emerging threats to the global health system. The World Health Organization has recorded 7,00,000 annual deaths due to ABR pathogens (1). *Pseudomonas aeruginosa* is one of the dreadful ESCAPE (*Enterococcus faecium, Staphylococcus aureus, Klebsiella pneumoniae, Acinetobacter baumannii, Pseudomonas aeruginosa, and Enterobacter species*) pathogens that cannot be eliminated by common antibiotics (2). In specific, carbapenem-resistant *P. aeruginosa* is a Priority Level-I Critical Pathogen that has been declared a serious threat (1, 3). *P. aeruginosa* has become a leading cause of life-threatening hospital-acquired infections due to its multi-drug resistant (MDR) phenotype. The MDR *P. aeruginosa*-associated infections can increase the fatality rate by 18 to 61% (4-6). Unfortunately, the MDR *P. aeruginosa* has been detected in diverse ecosystems. Improper disposal of hospital wastes potentially discharges several MDR *P. aeruginosa* into water bodies (7). Numerous MDR *P. aeruginosa* strains have been identified in wastewater effluents, polluted water bodies, hydrocarbon-contaminated soil, and compost (8-13). Contaminated water channels and farm animals have been identified as the huge reservoirs of the MDR *P. aeruginosa* strains (14, 15). These strains could be easily transmitted to agricultural fields and plants.

In our previous studies, we isolated and characterized 18 *P. aeruginosa* strains from the rhizospheric and endophytic niches of edible plants (cucumber, tomato, eggplant, and chili) harvested directly from the farms (16, 17). DNA fingerprinting and 16S rDNA-based phylogenetic analyses showed that these strains are evolutionarily related to the clinical *P. aeruginosa* isolates (16). Nearly 50-80% of these plant-associated *P. aeruginosa* strains exhibited multiple virulence traits including lytic activities (hemolysis, proteolysis, and lipolysis), biofilm formation, swarming motility, and production of rhamnolipids, pyocyanin, and siderophores (16, 17). Also, 10-40% of death occurred in the *C. elegans* model when the *P. aeruginosa* strains from the agricultural system were given as the food source (17). These reports indicate the presence of virulent and pathogenic *P. aeruginosa* strains in edible plants.

Furthermore, *P. aeruginosa* contamination has been detected in fresh vegetables at more than 50% of the supermarkets (18). Contaminated agricultural produces could be a major source of human transmission of *P. aeruginosa* (19, 20). Since *P. aeruginosa* is transmittable across the plant-animal-human interface (19-26), it is important to analyze the genetic determinants of virulence and pathogenicity in the non-clinical strains. However, the genotypic characteristics of plant-associated *P. aeruginosa* strains have not been studied as extensively as that of clinical isolates. Only six out of 214 *P. aeruginosa* complete genomes

available in the *Pseudomonas* database belong to the plant-associated strains (https://pseudomonas.com/). There is a clear gap in the genomic characterization of the agricultural *P. aeruginosa* strains.

The omnipresence of *P. aeruginosa* is often supported by a large and complex genome (5.5 to 7.4 Mbp; (27)). *P. aeruginosa* flourishes in diverse habitats by frequently acquiring new genetic elements or genomic islands through horizontal gene transfer (HGT) (28-31). Nearly 3000 plasmids potentially acquired through HGT have been identified in various *P. aeruginosa* strains (31). As a result, this bacterium has evolved with a highly versatile genome. Only 6.6% of its genome comprises core essential genes while the rest could be variable among the strains (32). Such variable regions in the *P. aeruginosa* genome contribute to the strain-level metabolic and functional diversity and are termed the regions of genomic plasticity (28, 33, 34). Importantly, 5 to 12% of the variable regions harbor virulence-related and drug-resistance genes (31). Previous reports suggest that the pathogenicity of *P. aeruginosa* is often related to genomic flexibility (35, 36). Henceforth, complete genome analyses of plant-associated *P. aeruginosa* strains might help to identify if they harbor genetic elements that confer virulence, pathogenicity, and ABR.

In the current study, we looked into the *in vitro* ABR profile of 18 plant-associated *P. aeruginosa* strains. An eggplant rhizospheric strain, PPA14 that exhibited extensive resistance was selected for complete genome analyses. It was hypothesized that plant-associated *P. aeruginosa* (PPA) acquires virulence-related and drug-resistance genes from the environment. To test this hypothesis, we examined the genomic islands, virulence-related, and ABR genes in the PPA14 genome. In addition, the PPA14 genome was compared with other plant-associated and environmental *P. aeruginosa* strains to determine the unique and shared genes among the non-clinical isolates.

Methods

Bacterial strains used in this study

Eighteen plant-associated *P. aeruginosa* strains (PPA01-18) used in this study (Table 1) were previously isolated and characterized from vegetable plants such as cucumber, tomato, eggplant, and chili (16, 17). Clinical strains of *P. aeruginosa*, PAO1, ATCC10145, and ATCC9027 were used as controls for phenotypic assays (37-39). The *P. aeruginosa* PA14 genome was used as the reference for genome assembly (40). The complete genomes of six plant-associated strains and 16 environmental strains of *P. aeruginosa* that were previously sequenced and deposited in the *Pseudomonas* database (https://pseudomonas.com/) were used

for comparative genomic analyses along with the standard reference genomes, PAO1 and PA14 (40, 41).

Antibiotic resistance assay

The ABR profiles of the *P. aeruginosa* strains were tested using the Kirby-Bauer disk diffusion method (42, 43). The three clinical strains of *P. aeruginosa*, ATCC10145, ATCC9027, and PAO1 were used as controls. Antibiotic susceptibility test was done using octa- and dodeca discs (Himedia) with 18 different antibiotics from eight classes namely aminoglycosides (gentamycin, streptomycin, and kanamycin), fluoroquinolone (ciprofloxacin), cephalosporins of first (cefadroxil), second (cefuroxime, and cefaclor), and third (cefotaxime, and cefoperazone) generations, penicillins (ampicillin, augmentin, and penicillin-G), macrolides (erythromycin, azithromycin, and clarithromycin), nitrofuran (nitrofurantoin), polymyxin (colistin), sulphonamide (co-trimoxazole), and tetracycline. Briefly, the strains were grown overnight at 37°C in Luria Bertani (LB) broth. After 24 h, 10% of the OD_{600}~1.0 adjusted overnight cultures were added into 15 ml of LB medium at a lukewarm temperature and poured into sterile Petri plates. After solidification, the antibiotic discs were placed over the inoculated media, and the plates were incubated overnight at 37°C. The occurrence of the inhibition zone around a particular disc was indicative of susceptibility toward the corresponding antibiotic. The percentage of antibiotic resistance exhibited by each of the tested strains and the percentage of strains resistant to each of the tested antibiotics were computed.

Hierarchical clustering of *P. aeruginosa* strains

The phenotypic and metabolic traits of the 18 PPA strains and the three clinical strains, ATCC10145, ATCC9027, and PAO1 were characterized previously (16, 17). In the current study, all these *P. aeruginosa* strains were grouped based on their phenotypic relatedness using the NCSS 2020 Statistical Software (https://www.ncss.com/software/ncss/). A hierarchical-clustered heat map with a double dendrogram was generated based on the Manhattan distance metric and the similarity between the clusters was defined using Ward's minimum variance (44). The heat map and the dendrogram were used to identify the most virulent plant-associated strains that are closely related to the clinical strains.

Extraction and quantification of *P. aeruginosa* PPA14 genomic DNA

The genomic DNA of *P. aeruginosa* PPA14 was extracted using the hexadecyl-trimethyl ammonium bromide (CTAB) method (45) and quantified using the NanoDrop (Thermo Scientific, Nanodrop 2000c) spectrophotometer. The absorbance relation of the DNA at 260/280 was indicative of its purity.

Whole-genome sequencing

The complete genomic DNA of the *P. aeruginosa* PPA14 strain was sequenced using the Solexa-Illumina, and Oxford-Nanopore technologies platforms, for short and long reads, respectively (46-48) by the Genotypic Technology, Bangalore, India (www.genotypic.co.in).

Genome assembly and scaffolding

The *P. aeruginosa* PPA14 genome reads were assembled using eleven different software tools (Fig. 1; (49)). The short reads generated by the Solexa-Illumina sequencing were assembled using SPAdes v3.13.0 (50), IDBA v1.1.3 (51), and Megahit v1.1.3 (52), SKESA v2.3.039 (53), and Unicycler v0.4.711 (54). The long reads generated by the Oxford-Nanopore technologies were assembled using Canu 1.8 (55), Flye 2.3.7 (56), and Unicycler v0.4.7 (54). Unicycler v0.4.7 (54), SPAdes v3.13.0 (57), and IDBA-hyb v1.1.1 (51) were used to assemble the hybrid reads formed by the combination of short and long reads. The contig level assemblies generated by each of these tools were further assembled into scaffolds using MeDuSa v1.6 (58) with *P. aeruginosa* PA14 as the reference genome.

Comparison and evaluation of genomic assemblies

All 11 genome assemblies were compared using four different tools, QUAST 5.0.1 (59), BUSCO (60), Circlator (61), and Qualimap v2.2.2 (62); Fig. 2). The assemblies were evaluated based on their genome and annotation features such as number of scaffolds, genome size, length of the shortest contig that covers 50% of the genome size (N50), orthologous gene score, number of repeat regions, and number of rRNAs and tRNAs (49). The best-assembled genome was selected based on these metrics and used for further downstream analyses.

Gene annotation

The assembled genome of PPA14 was annotated using the rapid prokaryotic genome annotation (Prokka) and Rapid annotation using subsystem technology (RAST) computational tools to predict its functional motifs (63-65). Furthermore, the predicted gene

functions were manually curated using the standard *P. aeruginosa* PAO1, PA14, PAK, and PA7 genomes as the references.

Identification of ABR and virulence genes

The ABRicate tool (https://github.com/tseemann/abricate) was used to identify the genetic determinants of *P. aeruginosa* PPA14 drug resistance and virulence. The AMR genes were detected based on the AMRFinderPlus, National Center for Biotechnology Information (NCBI), comprehensive antibiotic resistance database (CARD; (66)), and ResFinder, Center for Genomic Epidemiology (CGE; (67)). The virulence genes were identified based on the virulence factor database (VFDB; http://www.mgc.ac.cn/VFs/; (68)).

Identification of genomic islands

The Island Viewer 4 tool was used to identify the genomic islands (GI) harbored by the *P. aeruginosa* PPA14 genome (69). The Island Viewer 4 tool predicted the GIs in the PPA14 genome using the ensemble method by integrating three different tools, IslandPath-DIMOB, SIGI-HMM, and IslandPick (69). IslandPath-DIMOB detects the GIs based on the presence of at least one mobile genetic element coupled with dinucleotide bias in more than eight consecutive genes (70). SIGI-HMM predicts the GIs based on the hidden Markov model depending on biased codon usage (71, 72). IslandPick uses a comparative genomics approach on monophyletic strains to identify the GIs (73).

Genome map

The genome map of *P. aeruginosa* PPA14 was created and visualized using the circular genome viewer (CGView) server ((74); http://cgview.ca/). The open reading frames (ORFs), functional motifs, GC content, GC bias, predicted GIs and ABR genes were integrated and visualized in the PPA14 genome map.

Comparative genomic analyses of agricultural, and environmental *P. aeruginosa* strains

Complete genomes of *P. aeruginosa* strains that were previously isolated from plants (CR1, YB01, F9676, M18, CFSAN084950, and PPA14) and environments (N17-1, RD1-3, YL84, IMP66, IMP67, IMP68, PA1RG, MTB-1, KRP1, B10W, JB2, Ocean-1175, Ocean-1155, HS9, FA-HZ1, and DN1) were downloaded from the *Pseudomonas* database for the comparative genomic analyses (75). Two reference *P. aeruginosa* genomes, PAO1 and

PA14, were included as controls (Table 1). Computational tools such as Prokka (65), Roary, the pan-genome pipeline (76), and FastTree (77) were used for comparative genomic analyses of the 25 *P. aeruginosa* strains. The pan-genome profile of these strains produced based on the 'gene presence-absence matrix' produced by the Roary tool was visualized using Phandango (https://jameshadfield.github.io/phandango/#/; (78)). The unique genes harbored by the PPA14 genome were identified from the pan-genome profile and their functions were predicted by BLAST analyses against the reference genomes in the NCBI database. The maximum composite likelihood method-based phylogenetic analysis was performed using the complete genome of the 25 *P. aeruginosa* strains analyzed in this study. The phylogenetic tree was constructed using the randomized axelerated maximum likelihood (RAxML) computation tool (79).

Statistical analyses

All the phenotypic experiments were performed in triplicates. All data were subjected to a one-way analysis of variance (ANOVA) with a *p*-value of 0.05, and Duncan's multiple range test (DMRT) was performed between individual means to reveal any significant difference (XLSTAT, version 2010.5.05 add-in with Windows Excel). The hierarchical-clustered heat map with a double dendrogram was created using NCSS 2020 statistical software (NCSS, Kaysville, USA) to cluster the *P. aeruginosa* strains based on their phenotypic characteristics. Data analysis and scientific graphing were done in OriginPro version 8.5 (OriginLab®, USA).

Results

Phenotypic drug-resistance profile of *P. aeruginosa* strains

In vitro ABR profile of the plant-associated and clinical *P. aeruginosa* strains was assessed using 19 antibiotics belonging to eight different classes including aminoglycosides, beta-lactams, tetracyclines, fluoroquinolones, macrolides, nitrofurans, polymyxin, and sulfonamides (list of antibiotics is given in the "Materials and Methods").

The efficacy of different antibiotics in curtailing *P. aeruginosa* growth was estimated based on the percentage of strains resistant to a particular antibiotic (Fig. 3A). All of the plant-associated *P. aeruginosa* strains were inhibited by a fluoroquinolone (10 µg/ml ciprofloxacin). Cefotaxime (30 µg/ml), colistin (10 µg/ml), and gentamycin (10 µg/ml) belonging to third-generation cephalosporin, polymyxin, and aminoglycoside classes, respectively were effective against 84-95% of the PPA strains. A macrolide (15 µg/ml

erythromycin), and a second-generation cephalosporin (30 µg/ml cefaclor), prevented the growth of 77 to 88% of the PPA strains. Only 10% of the tested strains were sensitive to first-generation cephalosporins (30 µg/ml cefadroxil), nitrofuran (300 µg/ml nitrofurantoin), and sulphonamide (25 µg/ml co-trimaxazole). The rest of the antibiotics exhibited an average resistance and inhibited about 30 to 60% of the PPA strains tested in this study (Fig. 3A).

In vitro resistance profile of each *P. aeruginosa* strain was recorded individually (Fig. 3B). The three clinical isolates (ATCC10145, ATCC9027, and PAO1) were resistant against 47 to 68% of the antibiotics. None of the plant-associated *P. aeruginosa* strains were susceptible to all the 19 antibiotics tested in this work. Importantly, 12 out of 18 PPA strains tested in this study had higher drug resistance than the *P. aeruginosa* PAO1. The overall resistance percentage of the plant-associated strains varied between 26 to 84%. The strains, PPA12, and PPA14 isolated from the eggplant rhizosphere exhibited the lowest and highest level of drug resistance, respectively. Seven out of the nine endophytic (PPA03/cucumber; PPA07, PPA08/tomato; PPA16, PPA17, PPA18/chili) and two out of the nine rhizospheric (PPA13, PPA14/eggplant) strains resisted more than 50% of the tested antibiotics. The rest of the strains (PPA01, PPA02, PPA04/cucumber; PPA05, PPA06, PPA09, PPA10/tomato; PPA11, PPA12/eggplant; PPA15/chili) exhibited relatively low antibiotic resistance.

Phenotypic relatedness of agricultural and clinical *P. aeruginosa* strains

The plant-beneficial traits, virulence factors, and pathogenicity levels of plant-associated (PPA01-PPA18) and clinical (ATCC10145, ATCC9027, and PAO1) *P. aeruginosa* strains used in this study were previously determined (16, 17). This includes production of pyocyanin, rhamnolipid, siderophores, ammonia, and indole-3 acetic acid, formation of biofilm, swarming motility, solubilization of complex soil minerals, hemolytic, lipolytic, and proteolytic activities, biocontrol of phytopathogenic bacteria (*Xanthomonas oryzae*) and fungi (*Pythium aphanidermatum*, *Rhizactonia solani*, and *Fusarium oxysporum*), and pathogenicity against the *Caenorhabditis elegans* model (16, 17).

In the current work, a hierarchically-clustered heat map with a double dendrogram was created based on these phenotypic traits and drug-resistance profiles to identify the plant-associated *P. aeruginosa* strains that are closely related to the clinical isolates (Fig. 4). The cells in the heat map were color-coded from blue (low) to red (high) based on the relative abundance of the traits across the tested strains. All the tested strains clustered into three distinct groups (top dendrogram). Cluster A was occupied by seven strains (PPA01/cucumber; PPA05, PPA06, PPA07, PPA09/tomato; PPA11, PPA12/eggplant).

Two cucumber-associated strains (PPA02 and PPA04) were found in cluster B. Cluster C had three clinical isolates (ATCC10145, AT9027, and PAO1) and nine plant-associated strains (PPA03/cucumber; PPA08, PPA10/tomato; PPA13, PPA14/eggplant; PPA15, PPA16, PPA17, PPA18/chili). These nine strains exhibited shared virulence, pathogenicity, and drug resistance with the tested clinical isolates. Among these plant-associated strains, PPA14 isolated from an eggplant rhizosphere caused 35% mortality in the *C. elegans* model (17) and exhibited extensive resistance against eight antibiotic classes (aminoglycosides, fluoroquinolone, three-generation cephalosporins, penicillins, macrolides, nitrofuran, sulphonamide, and tetracycline) used in the current work (Fig. 3). Since the *P. aeruginosa* PPA14 expressed relatively higher pathogenicity and drug-resistance among the 18 tested plant-associated strains and also got clustered with the clinical *P. aeruginosa* isolates (Fig. 4), this strain was selected for the complete genome analyses.

P. aeruginosa PPA14 genome sequencing and assembly

The whole genome of *P. aeruginosa* PPA14 was sequenced using two platforms, Solexa-Illumina and Oxford-Nanopore technologies (46-48). The Illumina and Nanopore sequencing provided short and long raw reads, respectively. These reads were assembled individually and also combined to generate hybrid assemblies (refer to 'Materials and Methods' for the list of assembly tools). MeDuSa v1.6 (58) was used for scaffolding the assemblies based on *P. aeruginosa* PA14 as the reference genome. In total, 11 draft genome assemblies of PPA14 were generated which included five short-read, three long-read, and three hybrid-read assemblies. The 11 draft genome assemblies of the PPA14 were compared using QUAST 5.0.1 (59), BUSCO (60), Circlator (61), and Qualimap v2.2.2 (62). Each of these assemblies was evaluated based on their contiguity and completeness (49). The number of scaffolds generated by each assembly tool was considered the primary metric for contiguity. The short-assemblies had a higher number of scaffolds compared to long- and hybrid-assemblies (Fig. 5a). Only Hybrid-SPAdes, Hybrid-IDBA, Long-Canu, and Long-Unicycler could produce a single circular scaffold that reflected their contiguity. The short-assemblies also had low performance based on their i) draft genome length and ii) length of the shortest contig that covers 50% of the genome size (N50; Fig. 5b). The Long-Canu and Short-IDBA assemblies had the largest (6.74 Mbp) and shortest (6.59 Mbp) draft genome, respectively. The contiguity of an assembly is determined based on its N50 (49). In this study, the draft genome length was equal to the N50 only in the assemblies

produced by the Hybrid-SPAdes, Hybrid-IDBA, Long-Canu, and Long-Unicycler. An orthologous gene score was generated for each of the assemblies based on the presence of conserved genes. The assemblies with more than 99.5% orthologous gene scores were considered to be the best (60). Three short-assemblies (Megahit, SKESA, and Unicycler) and two hybrid assemblies (IDBA, and Unicycler) had a 99.9% orthologous gene score (Fig. 6a). All three long assemblies had poor scores based on their orthologous genes (75 to 95%). The Short-Megahit (203) and Short-IDBA (39) had the highest and lowest number of repeat regions, respectively. Seven out of eleven assemblies (Short-SPAdes and SKESA; Long-Canu, Flye, and Unicycler; Hybrid-SPAdes and Unicycler) had nearly 100 repeat regions. The *P. aeruginosa* genomes harbor a minimum of 12 ribosomal RNAs (rRNAs) and 64 transfer RNAs (tRNAs). On comparing all the PPA14 genome assemblies generated in this study, the tRNA numbers ranged from 36 to 64 while the rRNA numbers ranged from 2 to 12. Only three out of eleven assembly tools, Long-Flye, Hybrid-SPAdes, and Hybrid-Unicycler generated genomes with the optimal number of RNAs (Fig. 6b). Particularly, the rRNA count was too low (>8) in the rest of the assemblies.

Overall, the Hybrid-Unicycler assembly had the best scores based on all the tested parameters and was selected for downstream analysis.

Prediction of gene functions in the PPA14 genome

The size of the PPA14 genome was 6.72 million base pairs (Mbp) carrying 6315 ORFs. The predicted ORFs were individually analyzed to determine their functional motifs. In addition to the Prokka and RAST-based gene annotations, each ORF was individually evaluated based on previous records to avoid misrepresenting the gene products. Manual curation aims to offer high-quality genome annotations by correcting the inaccuracies in the software outputs. Missing genes, start codon misalignments, and disrupted genes were manually evaluated and fixed. Gene functions were predicted based on the homologous sequences in *P. aeruginosa* genomes that were previously characterized and deposited in reference databases. Out of 6315 open reading frames in the PPA14 genome, functions were predicted for 5042 genes. The remaining 20% of the genome was scored hypothetical as those genes did not have identifiable functions. The PPA14 genome had 64 tRNA, and 12 rRNA genes that comprised of four sets of 23S, 5S, and 16S subunits. It was also observed that the origin of replication or chromosomal replication initiator protein DnaA was located at 483 bp in the PPA14 genome. This gene was followed by DNA polymerase III beta subunit, DNA recombination, and repair protein RecF, DNA

gyrase subunit B, and Putative helicase. The order of these genes was similar to the reference *P. aeruginosa* genomes such as PAO1, PA14, PAK, and PA7 (https://pseudomonas.com/). This further validates the gene annotation process carried out in the current study.

Genomic islands harbored by the PPA14 genome

GI refers to the cluster of genes that have been potentially acquired by a bacterium through HGT. Such gene clusters play a significant role in bacterial genome evolution. In the current study, IslandPath-DIMOB, SIGI-HMM, and IslandPick tools collectively identified 83 GIs in the PPA14 genome. The largest GI detected was 48 kilobase pairs (Kbp) long constituting 54 ORFs. It was observed that 34 out of 83 GIs were longer than 10 Kbp. In total, over 982 genes of horizontal origins were detected in these islands. Linear maps representing the functional motifs of the genes harbored in each GI were created for all the 83 predicted islands. Our results show that nearly 15% of the PPA14 genome carries genes that were acquired through HGT.

Genetic determinants of the PPA14 antibiotic resistance

A list of 49 ABR genes was identified in the PPA14 genome based on the reference databases, NCBI AMRFinderPlus, MEGARes, CARD, and ResFinder (Table 2). The PPA14 genome carries a set of *aph* genes that code for aminoglycoside O-phosphotransferase conferring resistance against kanamycin, gentamycin, and streptomycin. The genes (*blaPDC-374* and *blaOXA-50*) that code for class C and D beta-lactamases, respectively conferring resistance against cephalosporins and carbapenems were also identified. The PPA14 genome also had a fosfomycin resistance gene *foxA* coding for FosA family glutathione transferase, and a chloramphenicol resistance gene, *catB7*, coding for chloramphenicol O-acetyltransferase. Additionally, *crpP* coding for a ciprofloxacin resistance protein was also found. Apart from these ABR genes that confer direct resistance to a specific antibiotic, 35 genes that code for multi-drug efflux pumps were detected in the PPA14 genome. This includes the resistance-nodulation-division (RND) superfamily efflux pumps, *mex, mux, tri, and opr* that could eliminate a wide spectrum of antibiotics. Moreover, the major facilitator superfamily (MFS), the multidrug and toxic compound extrusion (MATE), and the small multidrug resistance (SMR) family pumps that could resist aminoglycosides, bicyclomycin, and fluoroquinolones were also found. Additionally, the position of 49 ABR genes and 83 genomic islands were located in the PPA14 genome map and visualized using the CGView

(Fig. 7). It was found that two sets of *aph(6')-Id*, *aph(3")-Ib* genes, *mexXY,* and *crpP* were present within three genomic islands.

Genetic determinants of the PPA14 virulence

The virulence-related genes in the PPA14 genome were identified based on the VFDB database (68). More than 200 genes that could confer virulence to the *P. aeruginosa* PPA14 were detected (Table 3). The potential products of these genes included the major virulence factors such as phenazines, rhamnolipids, siderophores (pyochelin, pyoverdine), alginate, lipopolysaccharides, flagellar, pili-related, and secretion system proteins, and exotoxins (Table 3). In addition, the PPA14 genome had a list of genes that could produce lytic enzymes such as protease, lipase, elastase, alkaline and serine proteinase, hemolytic phospholipase, and peptidoglycan hydrolase. Chemotaxis, motility-associated, and ABC transporter genes were also detected in the PPA14 genome (Table 3).

Genomic plasticity in plant-associated and environmental *P. aeruginosa* strains

The complete genomes of *P. aeruginosa* strains that were previously isolated and sequenced from plants and the environment (Table 1) were compared with the PPA14 genome. The *P. aeruginosa* reference strains, PAO1 and PA14, were used as outliers. These reference strains had 6.2-6.5 million base pairs (Mbp) genome size (Fig. 8). The genomes of the plant-associated strains (CR1, YB01, F9676, M18, CFSAN084950, and PPA14) ranged from 6.1 to 6.7 Mbp while the environmental strains (N17-1, RD1-3, YL84, IMP66, IMP67, IMP68, PA1RG, MTB-1, KRP1, B10W, JB2, Ocean-1175, Ocean-1155, HS9, FA-HZ1, and DN1) had the genome sizes of 6.3 to 6.9 Mbp. Overall, the genomes of three environmental strains, ocean-1155, ocean-1175, and DN1 were the largest (6.9 Mbp). Among the plant-associated *P. aeruginosa* strains, PPA14, isolated in the current study had the highest genome size (6.7 Mbp).

The total number of protein-coding sequences (CDS), CDS that code for proteins with assigned functions, and non-hypothetical proteins were compared between the plant-associated (CR1, YB01, F9676, M18, CFSAN084950, and PPA14) and environmental strains (N17-1, RD1-3, YL84, IMP66, IMP67, IMP68, PA1RG, MTB-1, KRP1, B10W, JB2, Ocean-1175, Ocean-1155, HS9, FA-HZ1, and DN1) along with the PAO1 and PA14 (Fig. 9). The PAO1 and PA14 genomes had 5.8K and 6.1K CDS, respectively. The number of CDS harbored by the plant-associated *P. aeruginosa* strains varied from 5.7K to 6.2K. The *P. aeruginosa* PPA14 had the highest number of CDS among the plant-associated strains.

However, eight out of 16 environmental strains (KRP1, B10W, JB2, Ocean-1175, Ocean-1155, HS9, FA-HZ1, and DN1) had more than 6.2K CDS. The environmental strains with the smallest (N17-1) and largest (DN1) genome sizes had 5.9K and 6.6K CDS, respectively. Nearly 81-83% of the genes in the reference genomes (PAO1, and PA14) coded for proteins with well-known functions. In all of the tested strains, nearly 75-80% of the genes coded for known proteins while the rest were hypothetical.

Pan-genome profile of plant-associated, and environmental *P. aeruginosa* strains

The number of genes shared across the 23 *P. aeruginosa* strains isolated from the plant (CR1, YB01, F9676, M18, CFSAN084950, and PPA14), and environmental (N17-1, RD1-3, YL84, IMP66, IMP67, IMP68, PA1RG, MTB-1, KRP1, B10W, JB2, Ocean-1175, Ocean-1155, HS9, FA-HZ1, and DN1) ecosystems (Table 1) were estimated using the Roary tool (76). The *P. aeruginosa* reference genomes, PAO1 and PA14, were used as controls (Table 1). The pan-genome is comprised of four components including hardcore, softcore, shell, and cloud genes that were shared by 100%, 95-99%, 15-94%, and less than 15% of the tested genomes, respectively (76). The *P. aeruginosa* genomes compared in this study had 3324 hardcore, 1185 softcore, 1379 shell, and 6736 cloud genes (Fig. 10).

The *P. aeruginosa* pan-genome profiles were visualized using the Phandango interface ((78); Fig. 11). The matrix represented the core and accessory genes identified among the plant-associated, environmental, and reference strains. The hard and softcore were the components of the core genome while the shell and cloud were the components of the accessory genome. Out of 13294 genes identified among the 25 *P. aeruginosa* strains, only 39% occupied the core genome. The remaining 61% were the accessory genes that were absent at least in 22 out of the 25 tested genomes. The number of genes that were unique to each of these *P. aeruginosa* genomes was identified based on the Roary output. The reference strains, PAO1 and PA14, had 53 and 28 unique genes, respectively (Fig. 12). Nearly 65 to 2021 genes were exclusively found in the plant-associated *P. aeruginosa* strains (CR1, YB01, F9676, M18, CFSAN084950, and PPA14). The number of unique genes in the environmental strains (N17-1, RD1-3, YL84, IMP66, IMP67, IMP68, PA1RG, MTB-1, KRP1, B10W, JB2, Ocean-1175, Ocean-1155, HS9, FA-HZ1, and DN1) ranged from 12 to 590. An agricultural *P. aeruginosa*, CR1, isolated from the chili rhizosphere had the highest number of unique genes (2021). An environmental *P. aeruginosa* strain, IMP68 isolated from crude oil had the lowest number (12). The PPA14 genome characterized in the current study had 235 unique genes.

Evolutionary relatedness of plant-associated, environmental, and clinical *P. aeruginosa*

The evolutionary distance between the *P. aeruginosa* strains compared in this study was estimated using the Maximum Composite Likelihood method (80). The phylogenetic analysis clustered the strains based on their pan-genome profile (Fig. 12). The Chili rhizospheric isolate (CR1) and three environmental isolated (PA1RG/hospital sewage; YL84/ compost; RD1-3/ landfill) were the outliers distinct from all the other *P. aeruginosa* strains (Fig. 12). An arugula leaf isolate. CFSAN084950, clustered with N17-1 isolated from farm soil. The rest of the *P. aeruginosa* strains formed two major clades, one with PAO1 and the other with PA14. In the PAO1 clade, three plant-associated (M18/melon rhizosphere; YB01/tobacco stem; F9676/diseased rice sample) and six environmental (IMP66, IMP67, IMP68/crude oil; HS9/soil; FA-HZ1/wastewater; KRP1/methanogenic sludge) *P. aeruginosa* strains were found. The PA14 clade had two plan-associated (PPA14/eggplant rhizosphere; L10/halobiotic reed) and six environmental (B10W/wastewater; JB2/soil; MTB-1/hexachlorocyclohexane contaminated soil; DN1/China soil; Ocean-1155, Ocean-1175/open ocean) *P. aeruginosa* strains.

Discussion

P. aeruginosa is one of the major pathogens involved in hospital-acquired infections and fatalities. Unfortunately, agricultural plants and soil harbor virulent *P. aeruginosa* strains (16, 17) which could be potentially disseminated to humans and animals. In such a case, it is crucial to assess the level of risks associated with the plant-associated *P. aeruginosa*. However, there aren't many studies on genomic characterization of virulence and pathogenicity in plant-associated *P. aeruginosa*. In the current work, we hypothesized that virulence and drug-resistance genes could be found in the *P. aeruginosa* flourishing in the agricultural ecosystem. To test this hypothesis, we screened the drug-resistance profiles of 18 plant-associated *P. aeruginosa* strains that were previously isolated from cucumber, tomato, eggplant, and chili (16). Among these, an extensively resistant strain PPA14 that shared its virulence factors with the clinical *P. aeruginosa* isolates was selected for genomic analyses. Complete genome analyses of the plant-associated *P. aeruginosa* PPA14 to identify the genetic determinants of its virulence and ABR were the major focus of this paper. Additionally, we have also identified the unique genes in PPA14 and the GIs that were acquired through HGT.

P. aeruginosa constantly evolves resistance against multiple antibiotics making itself a Priority Level-I critical pathogen (1). The plant-associated *P. aeruginosa* strains tested in this study exhibited i*n vitro* resistance against cephalosporins, aminoglycosides, macrolides, nitrofurans, tetracyclines, and sulfonamides (Fig. 3a). Nine out of the 18 plant-associated strains (PPA03/cucumber; PPA07, PPA08, PPA10/tomato; PPA13, PPA14/eggplant; PPA16, PPA17, PPA18/chili) had higher resistance than the *P. aeruginosa* PAO1. Drug-resistant *P. aeruginosa* strains have been previously detected in sewage water, effluents, wastewater treatment plants, hydrocarbon-polluted soil, compost, and other contaminated soil and water bodies (8-13, 81). Furthermore, it has been predicted that ABR pathogens get disseminated into the agricultural ecosystem through livestock excreta, manures, and contaminated irrigation systems (82). In specific, the excessive use of antibiotics for the treatment of plant- and animal diseases could lead to the prevalence of ABR pathogens in the agricultural system (83-85). In the present work, we found extensive drug resistance in 50% of the tested plant-associated *P. aeruginosa* (Fig. 3a). Colistin and gentamycin are some of the last resort antibiotics recommended for treating extensively drug-resistant pathogens (86, 87). Even those antibiotics did not control 100% of the agricultural *P. aeruginosa* strains tested in our study (Fig. 3b). Particularly, an eggplant rhizosphere strain, PPA14, was resistant against 85% of the tested antibiotics (Fig. 3a). A detailed phenotypic comparison was done between 18 plant-associated *P. aeruginosa* strains and three clinical isolates (16, 17). Among these, the strains isolated from a wound (PAO1), outer ear infection (ATCC9027), clinical setting (ATCC10145), cucumber endophyte (PPA03), tomato endophytes (PPA08; PPA10), eggplant rhizosphere (PPA13; PPA14), chili rhizosphere (PPA15), and chili endophytes (PPA16; PPA17; PPA18) exhibited shared phenotypic characteristics (Fig. 4). These results highlight the shared virulence and antibiotic resistance between agricultural and clinical *P. aeruginosa* strains. Among the plant-associated *P. aeruginosa* strains that clustered with the clinical isolates, PPA14 was the extensively resistant one and hence was selected for the complete genome analyses.

In the current study, we have detected 49 ABR and 225 virulence-associated genes in the *P. aeruginosa* strain that was isolated from an eggplant rhizosphere (Tables 2 and 3). The environmental occurrence of such ABR pathogens is a globally emerging health concern. Some of the previous studies argue that environmental *P. aeruginosa* strains are highly sensitive to most antibiotics unlike the clinical isolates (88, 89). However, virulent and ABR *P. aeruginosa* strains have been reported in several natural environments like oil-contaminated soil, freshwater springs, domestic sewage, hospital wastewaters, and

wastewater treatment lagoons (7, 11, 12, 90-92). Intensive use of antibiotics in poultry farming subsequently leads to the dissemination of ABR *P. aeruginosa* in nearby soil (93). Nearly 75% of the *P. aeruginosa* strains previously isolated from different freshwater resources carried virulence factor genes and exhibited resistance against carbapenems, β-lactams, piperacillin, ceftazidime, and ciprofloxacin antibiotics (94). Also, there have been many reports on plant surfaces and vegetables carrying numerous strains of *P. aeruginosa* (16, 95). More than 200 *P. aeruginosa* strains have been isolated from vegetables collected from different farms and supermarkets (96). However, there were very limited attempts to look into the virulence and ABR determinants in *P. aeruginosa* strains flourishing in farm vegetables. The eggplant-associated strain, PPA14 characterized in the current study harbored genes that could confer resistance against 11 antibiotic classes including penicillins, tetracyclines, cephalosporins, fluoroquinolone, macrolides, sulfonamides, glycopeptides, aminoglycosides, carbapenems, fosfomycin, and monobactam (Table 2; Fig. 7). This strain also possessed 83 GIs that constitutes 15% of its total genetic material. In previously reported *P. aeruginosa* genomes, the genes acquired through HGT have maximumly constituted only to 10% of its genetic material (97-99). This indicates the presence highly versatile and hypermutable genome in the *P. aeruginosa* PPA14 that has acquired a high number of new genetic elements. Moreover, an arsenal of virulence determinants including the phenazines, rhamnolipids, siderophores (pyochelin, pyoverdine), alginate, lipopolysaccharides, flagellar, pili-related, and secretion system proteins, exotoxins, protease, lipase, elastase, alkaline, and serine proteinase, hemolytic phospholipase, peptidoglycan hydrolase, chemotaxis, motility-associated, and ABC transporter genes were also found in the PPA14 genome (Table 3). The incidence of multiple virulence, and ABR genes in the microbes flourishing in agricultural systems could confer alarming threats to human health (100-103).

Conclusion

The presence of multiple virulence and ABR genes in an eggplant-associated strain reveals the health threats associated with non-clinical *P. aeruginosa*. Furthermore, 12 out of 18 plant-associated *P. aeruginosa* strains characterized in this work exhibited higher *in vitro* resistance than the clinical isolate PAO1. These reports indicate the rising risk of ABR pathogens in agricultural produces. Consumption of raw vegetables (salads) could potentially disseminate ABR pathogens into humans and animals. As it is highly challenging to treat the infections caused by antibiotic-insensitive pathogens the current scenario might escalate the ABR-associated mortality rates in the future. The WHO has already predicted 10 million annual

deaths by 2050 due to ABR infections (1). The incidence of ABR pathogens in agricultural produces could be reduced in the future by preventing the (i) excessive use of antibiotics on farm animals (ii) dumping of pesticides on farmlands and (iii) use of sewage-contaminated water for irrigation.

Funding information

Author Contributions

The genomic experiments were conceived and designed by SA and KM. Phenotypic experiments were designed by SA and DB. All the experiments in this paper were performed by SA unless specified otherwise. Whole genome assembly, manual curation, and identification of genomic islands were done by TC, CC, and GP, respectively. Critical analyses of the data and manuscript preparation were done by SA and KM. Finally, all the authors were involved in the critical review of the paper.

Acknowledgment

Conflicts of interest

The authors declare no conflict of interest.

Ethical statement

The work did not involve humans or experimental animals.

Table 1. *Pseudomonas aeruginosa* **strains used in this study**

Ecosystem	Strain	Isolation Source	References
Plant	PPA01	Cucumber Rhizosphere	(16)
	PPA02	Cucumber Rhizosphere	(16)
	PPA03	Cucumber Endophyte	(16)
	PPA04	Cucumber Rhizosphere	(16)
	PPA05	Tomato Endophyte	(16)
	PPA06	Tomato Rhizosphere	(16)
	PPA07	Tomato Endophyte	(16)
	PPA08	Tomato Endophyte	(16)
	PPA09	Tomato Rhizosphere	(16)
	PPA10	Tomato Endophyte	(16)
	PPA11	Eggplant Endophyte	(16)
	PPA12	Eggplant Rhizosphere	(16)
	PPA13	Eggplant Rhizosphere	(16)
	PPA14	Eggplant Rhizosphere	(16)
	PPA15	Chili Rhizosphere	(16)
	PPA16	Chili Endophyte	(16)
	PPA17	Chili Endophyte	(16)
	PPA18	Chili Endophyte	(16)
	CR1	Chili rhizosphere	(104)
	M18	Melon rhizosphere	(105)
	YB01	Tobacco Stem	(106)
	F9676	Diseased rice sample	(107)
	CFSAN084950	Arugula Leaf	(108)
	L10	Halobiotic reed	(109)
Environment	N17-1	Farm Soil	(110)
	RD1-3	Landfill	(111)
	YL84	Compost	(112)
	PA1RG	Hospital sewage	(113)
	IMP66	Crude oil	(114)
	IMP67	Crude oil	(114)
	IMP68	Crude oil	(114)
	FA-HZ1	Wastewater	(115)
	MTB-1	T-HCH contaminated soil	(116)
	KRP1	Methanogenic sludge	(117)

	B10W	Wastewater	(118)
	HS9	Soil	(119)
	JB2	Soil	(120)
	Ocean-1175	Open ocean	(121)
	Ocean-1155	Open ocean	(121)
	DN1	China Soil	(122)
Clinical	PAO1	Wound	(38)
	PA14	Burn ward	(22)
	ATCC9027	Otitis externa	(37)
	ATCC10145	Unknown	(39)

Table 2. List of antibiotic resistance predicted in the PPA14 genome

Efflux Pumps		
Family	*Genes*	*Confer Antibiotic Resistance*
	triABC	Triclosan
RND	*mexAB; mexCD-oprJ; mexEF-opmB; mexMN; mexGHI-opmD; mexJKL; mexPQ-opmE;mexWY; opmH;*	Aminocoumarin; Aminoglycoside; Carbapenem; Cephalosporin; Cephamycin; Diaminopyrimidine; Fluoroquinolone; Macrolide; Monobactam; Penam; Peptide; Phenicol; Sulfonamide; Tetracycline; Triclosan
	muxABC-oprN	Aminocoumarin; Macrolide; Monobactam; Tetracycline
MATE	*pmpM*	Aminoglycoside; Benzalkonium Chloride; Fluoroquinolone
SMR	*emrE*	Aminoglycoside
MFS	*bcr-1*	Bicyclomycin
Antibiotic Inactivation		
Genes		*Confer Antibiotic Resistance*
aph(3'')-Ib; aph(3')-Iib; aph(6)-Id		Aminoglycoside
arnA; basS		Polymyxin
blaOXA-846; blaOXA-50;blaPDC-374		Carbapenem;Cephalosporin; Penam; Monobactam
catB7		Phenicol
crpP		Fluoroquinolone
fosA		Fosfomycin

Table 3. Virulence-related genes in plant-associated *Pseudomonas aeruginosa* PPA14

Virulence Factors	Genes
Alginate	*algGI8EAKXDJ44LF, algZPRQ, algC, algB, algW, algU; mucABCD, mucE, mucP*
Flagellar proteins	*flgBCDEFGHIJK; motAB, motCD, motY; fleN fliEFGHIJ, fleQRS; fliAKLMNOPQR; flhABF;*
Pyoverdine	*pvdA, pvdF, pvdG, pvdH, pvdL, pvdMNO, pvdP, pvdQ, pvdS; ptxR; mbtH*
Pyochelin	*pchABCD, pchEFGHI, pchR; fptA*
Phenazine	*phzA1B1C1D1E1F1G1*
Pyocyanin	*phzM, phzS*
Rhamnolipids	*rhlABRI, rhlC*
Protease and elastase	*lasABI*
Alkaline proteinase	*aprA*
Exotoxins	*exoU, exoT, exoY*
Haemolytic phospholipase	*plcH*
Phosphatase	*pppAB*
Flagellin	*flgANM*
Pili-related proteins	*pilB, pilRS, pilF; fimV*
Paerucumarin	*pvcABCD*
Exotoxin	*toxA*
Lipopolysaccharide	*waaACFGP*
Twitching motility	*pilGHIJKTU*
Pepditoglycan hydrolase	*tse1, tse2, tse3*
Chemotaxis	*chpABCDE*
Serine protein kinase	*ppkA; tagR*
ABC transporter protein	*tagST, tagQ*
Lipoprotein	*lip1*
Type II secretion system	*xcpPQRSTUVWXYZ*
Type III secretion system	*pscBCDEFGHIJKL; pscNOPQRSTU; exsABCDE; popBDN; pcr1234DGHRV;*
Type VI secretion system	*fha1; hsiA1B1C1E1F1G1H1J1; hcp1; clpV1; vgrG1aG1b; tssM1; dotU1*

References

1.   **Talebi Bezmin Abadi A, Rizvanov AA, Haertlé T, Blatt NL**. World Health Organization report: Current crisis of antibiotic resistance. *BioNanoScience*. 2019;9(4):778-88. doi: 10.1007/s12668-019-00658-4.

2.   **Rice LB**. Federal funding for the study of antimicrobial resistance in nosocomial pathogens: no ESKAPE. *J Infect Dis*. 2008;197(8):1079-81. 10.1086/533452.

3.   AR Threats Report. Antibiotic resistance threats in the United States, 2019. *U.S. Department of Health and Human Services Centers for Disease Control and Prevention.* https://www.cdc.gov/drugresistance/biggest-threats.html 2019 [cited 2022 December 04].

4.   **Emerson J, Rosenfeld M, McNamara S, Ramsey B, Gibson RL**. *Pseudomonas aeruginosa* and other predictors of mortality and morbidity in young children with cystic fibrosis. *Pediatr Pulmonol*. 2002;34(2):91-100. https://doi.org/10.1002/ppul.10127.

5.   **Araújo D, Shteinberg M, Aliberti S, Goeminne PC, Hill AT, Fardon TC, et al.** The independent contribution of *Pseudomonas aeruginosa* infection to long-term clinical outcomes in bronchiectasis. *European Respiratory Journal*. 2018;51(2):1701953. 10.1183/13993003.01953-2017.

6.   **Kunz Coyne AJ, El Ghali A, Holger D, Rebold N, Rybak MJ**. Therapeutic strategies for emerging multidrug-resistant *Pseudomonas aeruginosa*. *Infectious Diseases and Therapy*. 2022;11(2):661-82. 10.1007/s40121-022-00591-2.

7.   **Fuentefria DB, Ferreira AE, Corção G**. Antibiotic-resistant *Pseudomonas aeruginosa* from hospital wastewater and superficial water: Are they genetically related? *Journal of Environmental Management*. 2011;92(1):250-5. https://doi.org/10.1016/j.jenvman.2010.09.001.

8.   **Kaszab E, Szoboszlay S, Dobolyi C, Háhn J, Pék N, Kriszt B**. Antibiotic resistance profiles and virulence markers of *Pseudomonas aeruginosa* strains isolated from composts. *Bioresource Technol*. 2011;102(2):1543-8. 10.1016/j.biortech.2010.08.027.

9.   **Kaszab E, Radó J, Kriszt B, Pászti J, Lesinszki V, Szabó Á, et al.** Groundwater, soil and compost, as possible sources of virulent and antibiotic-resistant *Pseudomonas aeruginosa*. *International Journal of Environmental Health Research*. 2021;31(7):848-60. 10.1080/09603123.2019.1691719.

10.   **Pirnay J-P, Matthijs S, Colak H, Chablain P, Bilocq F, Van Eldere J, et al.** Global *Pseudomonas aeruginosa* biodiversity as reflected in a Belgian river. *Environ Microbiol*. 2005;7(7):969-80. 10.1111/j.1462-2920.2005.00776.x.

11.   **Wheater DWF, Mara DD, Jawad L, Oragui J**. *Pseudomonas aeruginosa* and *Escherichia coli* in sewage and fresh water. *Water Res*. 1980;14(7):713-21. 10.1016/0043-1354(80)90245-6.

12.   **Ullah A, Durrani R, Ali G, Ahmed S**. Prevalence of antimicrobial resistant *Pseudomonas aeruginosa* in fresh water spring contaminated with domestic sewage. *Journal of Biological and Food Science Research*. 2012;1(2):19-22.

13.   **Hosu MC, Vasaikar S, Okuthe GE, Apalata T**. Molecular detection of antibiotic-resistant genes in *Pseudomonas aeruginosa* from nonclinical environment: Public health implications in mthatha, eastern cape province, south africa. *Int J Microbiol*. 2021;2021:8861074. 10.1155/2021/8861074.

14.   **Haenni M, Bour M, Châtre P, Madec J-Y, Plésiat P, Jeannot K**. Resistance of animal strains of *Pseudomonas aeruginosa* to carbapenems. *Front Microbiol*. 2017;8. 10.3389/fmicb.2017.01847.

15.   **Gharieb R, Saad M, Khedr M, El Gohary A, Ibrahim H**. Occurrence, virulence, carbapenem resistance, susceptibility to disinfectants and public health hazard of *Pseudomonas aeruginosa* isolated from animals, humans and environment in intensive farms. *J Appl Microbiol*. 2022;132(1):256-67. 10.1111/jam.15191.

16.   **Ambreetha S, Marimuthu P, Mathee K, Balachandar D**. Rhizospheric and endophytic *Pseudomonas aeruginosa* in edible vegetable plants share molecular and metabolic traits with clinical isolates. *J Appl Microbiol*. 2021;132(4):3226– 48. doi: 10.1111/jam.15317.

17.   **Ambreetha S, Marimuthu P, Mathee K, Balachandar D**. Plant-associated *Pseudomonas aeruginosa* strains harbor multiple virulence traits critical for human infection. *J Med Microbiol*. 2022;71(8):1-18. 10.1099/jmm.0.001493.

18. **Allydice-Francis K, Brown PD**. Diversity of antimicrobial resistance and virulence determinants in *Pseudomonas aeruginosa* associated with fresh vegetables. *Int J Microbiol*. 2012;2012:426241. 10.1155/2012/426241.

19. **Kominos SD, Copeland CE, Grosiak B, Postic B**. Introduction of *Pseudomonas aeruginosa* into a hospital via vegetables. *Appl Microbiol*. 1972;24(4):567-70. 10.1128/am.24.4.567-570.1972.

20. **Correa CMC, Tibana A, Filho PPG**. Vegetables as a source of infection with *Pseudomonas aeruginosa* in a University and Oncology Hospital of Rio de Janeiro. *J Hosp Infect*. 1991;18(4):301-6. 10.1016/0195-6701(91)90187-D.

21. **Schroth M, Cho J, Green S, Kominos S**. Epidemiology of *Pseudomonas aeruginosa* in agricultural areas In: Young V, editor. *Pseudomonas aeruginosa:* Ecological aspects and patient colonization: New York: Raven Press; 1977. p. 1-29.

22. **Schroth MN, Cho JJ, Green SK, Kominos SD**. Epidemiology of *Pseudomonas aeruginosa* in agricultural areas. *J Med Microbiol*. 2018;67(8):1191-201. 10.1099/jmm.0.000758.

23. **Plotnikova JM, Rahme LG, Ausubel FM**. Pathogenesis of the human opportunistic pathogen *Pseudomonas aeruginosa* PA14 in *Arabidopsis. Plant Physiol*. 2000;124(4):1766-74. 10.1104/pp.124.4.1766.

24. **Walker TS, Bais HP, Déziel E, Schweizer HP, Rahme LG, Fall R, et al.** *Pseudomonas aeruginosa*-plant root interactions. Pathogenicity, biofilm formation, and root exudation. *Plant Physiol*. 2004;134(1):320-31. 10.1104/pp.103.027888.

25. **Fernandes MR, Sellera FP, Moura Q, Carvalho MPN, Rosato PN, Cerdeira L, et al.** Zooanthroponotic transmission of drug-resistant *Pseudomonas aeruginosa*, Brazil. *Emerg Infect Dis*. 2018;24(6):1160-2. 10.3201/eid2406.180335.

26. **Foti M, Giacopello C, Fisichella V, Latella G**. Multidrug-resistant *Pseudomonas aeruginosa* isolates from captive reptiles. *Journal of Exotic Pet Medicine*. 2013;22(3):270-4. 10.1053/j.jepm.2013.08.007.

27. **Klockgether J, Cramer N, Wiehlmann L, Davenport C, Tümmler B**. *Pseudomonas aeruginosa* genomic structure and diversity. *Front Microbiol*. 2011;2(150). 10.3389/fmicb.2011.00150.

28. **Mathee K, Narasimhan G, Valdes C, Qiu X, Matewish JM, Koehrsen M, et al.** Dynamics of *Pseudomonas aeruginosa* genome evolution. *PNAS*. 2008;105(8):3100-5. 10.1073/pnas.0711982105.

29. **Qiu X, Kulasekara BR, Lory S**. Role of horizontal gene transfer in the evolution of *Pseudomonas aeruginosa* virulence. In: de Reuse H, Bereswill S, editors. Microbial Pathogenomics. 62009. p. 126–39. doi: 10.1159/000235767.

30. **San Millan A, Toll-Riera M, Qi Q, MacLean RC**. Interactions between horizontally acquired genes create a fitness cost in *Pseudomonas aeruginosa*. *Nat Commun*. 2015;6(1):6845. 10.1038/ncomms7845.

31. **Freschi L, Vincent AT, Jeukens J, Emond-Rheault J-G, Kukavica-Ibrulj I, Dupont M-J, et al.** The *Pseudomonas aeruginosa* pan-genome provides new insights on its population structure, horizontal gene transfer, and pathogenicity. *Genome Biol Evol*. 2019;11(1):109-20. 10.1093/gbe/evy259.

32. **Poulsen BE, Yang R, Clatworthy AE, White T, Osmulski SJ, Li L, et al.** Defining the core essential genome of *Pseudomonas aeruginosa*. *PNAS*. 2019;116(20):10072-80. 10.1073/pnas.1900570116.

33. **Römling U, Schmidt KD, Tümmler B**. Large genome rearrangements discovered by the detailed analysis of 21 *Pseudomonas aeruginosa* clone C isolates found in environment and disease habitats. *Journal of Molecular Biology*. 1997;271(3):386-404. https://doi.org/10.1006/jmbi.1997.1186.

34. **Heuer T, Bürger C, Tümmler B**. Smith/Birnstiel mapping of genome rearrangements in *Pseudomonas aeruginosa*. *ELECTROPHORESIS*. 1998;19(4):495-9. https://doi.org/10.1002/elps.1150190407.

35. **Woods PW, Haynes ZM, Mina EG, Marques CNH**. Maintenance of *S. aureus* in co-culture with *P. aeruginosa* while growing as biofilms. *Front Microbiol*. 2019;9(3291). 10.3389/fmicb.2018.03291.

36.   **Harrison JJ, Almblad H, Irie Y, Wolter DJ, Eggleston HC, Randall TE, et al.** Elevated exopolysaccharide levels in *Pseudomonas aeruginosa* flagellar mutants have implications for biofilm growth and chronic infections. *PLOS Genetics*. 2020;16(6):e1008848. doi: 10.1371/journal.pgen.1008848.

37.   **Haynes WC**. *Pseudomonas aeruginosa*---its characterization and identification. *Microbiol*. 1951;5(5):939-50. 10.1099/00221287-5-5-939.

38.   **Holloway BW**. Genetic recombination in *Pseudomonas aeruginosa*. *J Gen Microbiol*. 1955;13(3):572-81. 10.1099/00221287-13-3-572.

39.   **Picard B, Denamur E, Barakat A, Elion J, Goullet P**. Genetic heterogeneity of *Pseudomonas aeruginosa* clinical isolates revealed by esterase electrophoretic polymorphism and restriction fragment length polymorphism of the ribosomal RNA gene region. *J Med Microbiol*. 1994;40(5):313-22. 10.1099/00222615-40-5-313.

40.   **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. 10.1186/gb-2006-7-10-r90.

41.   **Stover CK, Pham XQ, Erwin AL, Mizoguchi SD, Warrener P, Hickey MJ, et al.** Complete genome sequence of *Pseudomonas aeruginosa* PAO1, an opportunistic pathogen. *Nature*. 2000;406(6799):959-64. 10.1038/35023079.

42.   **Bauer AW, Perry DM, Kirby WMM**. Single-disk antibiotic-sensitivity testing of staphylococci: An analysis of technique and results. *AMA Archives of Internal Medicine*. 1959;104(2):208-16. 10.1001/archinte.1959.00270080034004.

43.   **Hudzicki J**. Kirby-Bauer disk diffusion susceptibility test protocol. *American society for microbiology*. 2009;15:55-63.

44.   **Ward JH**. Hierarchical grouping to optimize an objective function. *Journal of the American Statistical Association*. 1963;58(301):236-44. 10.2307/2282967.

45.   **Melody SC**. Plant Molecular Biology - A laboratory manual. New York, USA: Springer-Verlag Berlin Heidelberg; 1997. 529 p.

46.   **Sanger F, Nicklen S, Coulson AR**. DNA sequencing with chain-terminating inhibitors. *PNAS*. 1977;74(12):5463-7. 10.1073/pnas.74.12.5463.

47.   **Jain M, Olsen HE, Paten B, Akeson M**. The Oxford Nanopore MinION: delivery of nanopore sequencing to the genomics community. *Genome Biol*. 2016;17(239):1-11. 10.1186/s13059-016-1103-0.

48.   **Shendure J, Ji H**. Next-generation DNA sequencing. *Nature Biotechnol*. 2008;26(10):1135-45. 10.1038/nbt1486.

49.   **Molina-Mora JA, Chinchilla-Montero D, Chavarría-Azofeifa M, Ulloa-Morales AJ, Campos-Sánchez R, Mora-Rodríguez R, et al.** Transcriptomic determinants of the response of ST-111 Pseudomonas aeruginosa AG1 to ciprofloxacin identified by a top-down systems biology approach. *Sci Rep*. 2020;10(1):13717. 10.1038/s41598-020-70581-2.

50.   **Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, et al.** SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J Comput Biol*. 2012;19(5):455-77. 10.1089/cmb.2012.0021.

51.   **Peng Y, Leung HCM, Yiu SM, Chin FYL**, editors. IDBA – a practical iterative de bruijn graph de novo assembler2010; Berlin, Heidelberg: Springer Berlin Heidelberg.

52.   **Li D, Liu C-M, Luo R, Sadakane K, Lam T-W**. MEGAHIT: an ultra-fast single-node solution for large and complex metagenomics assembly via succinct de Bruijn graph. *Bioinformatics (Oxford, England)*. 2015;31(10):1674-6. 10.1093/bioinformatics/btv033.

53.   **Souvorov A, Agarwala R, Lipman DJ**. SKESA: strategic k-mer extension for scrupulous assemblies. *Genome Biol*. 2018;19(1):153. 10.1186/s13059-018-1540-z.

54.   **Wick RR, Judd LM, Gorrie CL, Holt KE**. Unicycler: Resolving bacterial genome assemblies from short and long sequencing reads. *PLOS Computational Biology*. 2017;13(6):e1005595. 10.1371/journal.pcbi.1005595.

55.   **Koren S, Walenz BP, Berlin K, Miller JR, Bergman NH, Phillippy AM**. Canu: scalable and accurate long-read assembly via adaptive k-mer weighting and repeat separation. *Genome research*. 2017;27(5):722-36. 10.1101/gr.215087.116.

56. **Kolmogorov M, Yuan J, Lin Y, Pevzner PA**. Assembly of long, error-prone reads using repeat graphs. *Nature Biotechnol*. 2019;37(5):540-6. 10.1038/s41587-019-0072-8.

57. **Antipov D, Hartwick N, Shen M, Raiko M, Lapidus A, Pevzner PA**. plasmidSPAdes: assembling plasmids from whole genome sequencing data. *Bioinformatics (Oxford, England)*. 2016;32(22):3380-7. 10.1093/bioinformatics/btw493.

58. **Bosi E, Donati B, Galardini M, Brunetti S, Sagot MF, Lió P, et al.** MeDuSa: a multi-draft based scaffolder. *Bioinformatics (Oxford, England)*. 2015;31(15):2443-51. 10.1093/bioinformatics/btv171.

59. **Gurevich A, Saveliev V, Vyahhi N, Tesler G**. QUAST: quality assessment tool for genome assemblies. *Bioinformatics (Oxford, England)*. 2013;29(8):1072-5. 10.1093/bioinformatics/btt086.

60. **Simão FA, Waterhouse RM, Ioannidis P, Kriventseva EV, Zdobnov EM**. BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics (Oxford, England)*. 2015;31(19):3210-2. 10.1093/bioinformatics/btv351.

61. **Liao Y-C, Cheng H-W, Wu H-C, Kuo S-C, Lauderdale T-LY, Chen F-J**. Completing circular bacterial genomes with assembly complexity by using a sampling strategy from a single MinION run with barcoding. *Front Microbiol*. 2019;10(2068). 10.3389/fmicb.2019.02068.

62. **Okonechnikov K, Conesa A, García-Alcalde F**. Qualimap 2: advanced multi-sample quality control for high-throughput sequencing data. *Bioinformatics (Oxford, England)*. 2016;32(2):292-4. 10.1093/bioinformatics/btv566.

63. **Aziz RK, Bartels D, Best AA, DeJongh M, Disz T, Edwards RA, et al.** The RAST Server: Rapid annotations using subsystems technology. *BMC Genomics*. 2008;9(1):75. 10.1186/1471-2164-9-75.

64. **Brettin T, Davis JJ, Disz T, Edwards RA, Gerdes S, Olsen GJ, et al.** RASTtk: A modular and extensible implementation of the RAST algorithm for building custom annotation pipelines and annotating batches of genomes. *Sci Rep*. 2015;5(1):8365. 10.1038/srep08365.

65. **Seemann T**. Prokka: rapid prokaryotic genome annotation. *Bioinformatics (Oxford, England)*. 2014;30(14):2068-9. 10.1093/bioinformatics/btu153.

66. **Alcock BP, Raphenya AR, Lau TTY, Tsang KK, Bouchard M, Edalatmand A, et al.** CARD 2020: antibiotic resistome surveillance with the comprehensive antibiotic resistance database. *Nucleic Acids Res*. 2020;48(D1):D517-d25. 10.1093/nar/gkz935.

67. **Bortolaia V, Kaas RS, Ruppe E, Roberts MC, Schwarz S, Cattoir V, et al.** ResFinder 4.0 for predictions of phenotypes from genotypes. *The Journal of antimicrobial chemotherapy*. 2020;75(12):3491-500. 10.1093/jac/dkaa345.

68. **Chen L, Yang J, Yu J, Yao Z, Sun L, Shen Y, et al.** VFDB: a reference database for bacterial virulence factors. *Nucleic Acids Res*. 2005;33(Database issue):D325-8. 10.1093/nar/gki008.

69. **Bertelli C, Laird MR, Williams KP, Simon Fraser University Research Computing Group, Lau BY, Hoad G, et al.** IslandViewer 4: expanded prediction of genomic islands for larger-scale datasets. *Nucleic Acids Res*. 2017;45(W1):W30-W5. 10.1093/nar/gkx343.

70. **Hsiao WWL, Ung K, Aeschliman D, Bryan J, Finlay BB, Brinkman FSL**. Evidence of a large novel gene pool associated with prokaryotic genomic islands. *PLOS Genetics*. 2005;1(5):e62. 10.1371/journal.pgen.0010062.

71. **Waack S, Keller O, Asper R, Brodag T, Damm C, Fricke WF, et al.** Score-based prediction of genomic islands in prokaryotic genomes using hidden Markov models. *BMC Bioinformatics*. 2006;7(1):142. 10.1186/1471-2105-7-142.

72. **Merkl R**. SIGI: score-based identification of genomic islands. *BMC Bioinformatics*. 2004;5(1):22. 10.1186/1471-2105-5-22.

73. **Langille MGI, Hsiao WWL, Brinkman FSL**. Evaluation of genomic island predictors using a comparative genomics approach. *BMC Bioinformatics*. 2008;9(1):329. 10.1186/1471-2105-9-329.

74. **Grant JR, Stothard P**. The CGView Server: a comparative genomics tool for circular genomes. *Nucleic Acids Res*. 2008;36(suppl_2):W181-W4. 10.1093/nar/gkn179.

75. **Winsor GL, Griffiths EJ, Lo R, Dhillon BK, Shay JA, Brinkman Fiona SL**. Enhanced annotations and features for comparing thousands of Pseudomonas genomes in the Pseudomonas genome database. *Nucleic Acids Res*. 2015;44(D1):D646-D53. 10.1093/nar/gkv1227.

76. **Page AJ, Cummins CA, Hunt M, Wong VK, Reuter S, Holden MT, et al.** Roary: rapid large-scale prokaryote pan genome analysis. *Bioinformatics (Oxford, England)*. 2015;31(22):3691-3. 10.1093/bioinformatics/btv421.

77. **Price MN, Dehal PS, Arkin AP**. FastTree 2 – approximately Maximum-Likelihood trees for large alignments. *PLoS One*. 2010;5(3):e9490. 10.1371/journal.pone.0009490.

78. **Hadfield J, Megill C, Bell SM, Huddleston J, Potter B, Callender C, et al.** Nextstrain: real-time tracking of pathogen evolution. *Bioinformatics (Oxford, England)*. 2018;34(23):4121-3. 10.1093/bioinformatics/bty407.

79. **Stamatakis A**. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics (Oxford, England)*. 2014;30(9):1312-3. 10.1093/bioinformatics/btu033.

80. **Tamura K, Nei M, Kumar S**. Prospects for inferring very large phylogenies by using the neighbor-joining method. *PNAS*. 2004;101(30):11030-5. doi: 10.1073/pnas.0404206101.

81. **Pitondo-Silva A, Gonçalves GB, Stehling EG**. Heavy metal resistance and virulence profile in *Pseudomonas aeruginosa* isolated from Brazilian soils. *J Biol Chem*. 2016;124(8):681-8. 10.1074/jbc.M312080200.

82. **Thanner S, Drissner D, Walsh F**. Antimicrobial resistance in agriculture. *mBio*. 2016;7(2):10.1128/mbio.02227-15. doi:10.1128/mbio.02227-15.

83. **Falkiner FR**. The consequences of antibiotic use in horticulture. *The Journal of antimicrobial chemotherapy*. 1998;41(4):429-31.

84. **McManus PS, Stockwell VO, Sundin GW, Jones AL**. Antibiotic use in plant agriculture. *Annu Rev Phytopathol*. 2002;40(1):443-65. 10.1146/annurev.phyto.40.120301.093927.

85. **Schnabel EL, Jones AL**. Distribution of Tetracycline Resistance Genes and Transposons among Phylloplane Bacteria in Michigan Apple Orchards. *Appl Environ Microbiol*. 1999;65(11):4898-907. doi:10.1128/AEM.65.11.4898-4907.1999.

86. **Karaiskos I, Friberg LE, Galani L, Ioannidis K, Katsouda E, Athanassa Z, et al.** Challenge for higher colistin dosage in critically ill patients receiving continuous venovenous haemodiafiltration. *International Journal of Antimicrobial Agents*. 2016;48(3):337-41. https://doi.org/10.1016/j.ijantimicag.2016.06.008.

87. **Gonzalez-Padilla M, Torre-Cisneros J, Rivera-Espinar F, Pontes-Moreno A, López-Cerero L, Pascual A, et al.** Gentamicin therapy for sepsis due to carbapenem-resistant and colistin-resistant Klebsiella pneumoniae. *Journal of Antimicrobial Chemotherapy*. 2014;70(3):905-13. 10.1093/jac/dku432.

88. **Gholami S, Tabatabaei M**. The prevalence of virulence factors in human and environmental isolates of *Pseudomonas aeruginosa*. *Avicenna J Clin Microbiol Infect*. 2019;6(1):9-14. 10.34172/ajcmi.2019.03.

89. **Ramsay KA, Wardell SJT, Patrick WM, Brockway B, Reid DW, Winstanley C, et al.** Genomic and phenotypic comparison of environmental and patient-derived isolates of *Pseudomonas aeruginosa* suggest that antimicrobial resistance is rare within the environment. *J Med Microbiol*. 2019;68(11):1591-5. 10.1099/jmm.0.001085.

90. **Alonso A, Rojo F, Martínez JL**. Environmental and clinical isolates of *Pseudomonas aeruginosa* show pathogenic and biodegradative properties irrespective of their origin. *Environ Microbiol*. 1999;1(5):421-30. 10.1046/j.1462-2920.1999.00052.x.

91. **Aumeran C, Paillard C, Robin F, Kanold J, Baud O, Bonnet R, et al.** *Pseudomonas aeruginosa* and *Pseudomonas putida* outbreak associated with contaminated water outlets in an oncohaematology paediatric unit. *J Hosp Infect*. 2007;65(1):47-53. 10.1016/j.jhin.2006.08.009.

92. **Petit SMC, Lavenir R, Colinon-Dupuich C, Boukerb AM, Cholley P, Bertrand X, et al.** Lagooning of wastewaters favors dissemination of clinically relevant *Pseudomonas aeruginosa*. *Research in Microbiology*. 2013;164(8):856-66. 10.1016/j.resmic.2013.06.007.

93. **Kousar S, Rehman N, Javed A**. Intensive poultry farming practices influence antibiotic resistance profiles in *Pseudomonas aeruginosa* inhabiting nearby soils. *Infection and drug resistance*. 2021;29(14):4511-6. 10.2147/idr.s324055.

94. **Liew SM, Rajasekaram G, Puthucheary SDA, Chua KH**. Antimicrobial susceptibility and virulence genes of clinical and environmental isolates of *Pseudomonas aeruginosa*. *PeerJ*. 2019;7:e6217. 10.7717/peerj.6217.

95.   **Green SK, Schroth MN, Cho JJ, Kominos SD, Vitanza-Jack VB**. Agricultural plants and soil as a reservoir for *Pseudomonas aeruginosa*. *Appl Microbiol*. 1974;28(6):987-91. 10.1128/am.28.6.987-991.1974.

96.   **Schwaiger K, Helmke K, Hölzel CS, Bauer J**. Antibiotic resistance in bacteria isolated from vegetables with regards to the marketing stage (farm vs. supermarket). *Int J Food Microbiol*. 2011;148(3):191-6. 10.1016/j.ijfoodmicro.2011.06.001.

97.   **Spencer DH, Kas A, Smith EE, Raymond CK, Sims EH, Hastings M, et al.** Whole-genome sequence variation among multiple isolates of *Pseudomonas aeruginosa*. *J Bacteriol*. 2003;185(4):1316-25. 10.1128/JB.185.4.1316-1325.2003.

98.   **Shen K, Sayeed S, Antalis P, Gladitz J, Ahmed A, Dice B, et al.** Extensive genomic plasticity in *Pseudomonas aeruginosa* revealed by identification and distribution studies of novel genes among clinical isolates. *Infect Immun*. 2006;74(9):5272-83. doi: 10.1128/iai.00546-06.

99.   **Battle SE, Rello J, Hauser AR**. Genomic islands of Pseudomonas aeruginosa. *FEMS Microbiol Lett*. 2009;290(1):70-8. 10.1111/j.1574-6968.2008.01406.x.

100.   **Durso LM, Cook KL**. Impacts of antibiotic use in agriculture: what are the benefits and risks? *Curr Opin Microbiol*. 2014;19:37-44. 10.1016/j.mib.2014.05.019.

101.   **Allen HK, Donato J, Wang HH, Cloud-Hansen KA, Davies J, Handelsman J**. Call of the wild: antibiotic resistance genes in natural environments. *Nature Reviews Microbiology*. 2010;8(4):251-9. 10.1038/nrmicro2312.

102.   **Lau CH-F, Engelen Kv, Gordon S, Renaud J, Topp E**. Novel antibiotic resistance determinants from agricultural soil exposed to antibiotics widely used in human medicine and animal farming. *Appl Environ Microbiol*. 2017;83(16):e00989-17. doi:10.1128/AEM.00989-17.

103.   **Browne S, Bhatia S, Sarkar N, Kaushik M**. Chapter 11 - Antibiotic-resistant bacteria and antibiotic-resistant genes in agriculture: a rising alarm for future. In: Singh P, Sillanpää M, editors. Degradation of Antibiotics and Antibiotic-Resistant Bacteria from Various Sources: Academic Press; 2023. p. 247-74. https://doi.org/10.1016/B978-0-323-99866-6.00017-9.

104.   **Sood U, Hira P, Kumar R, Bajaj A, Rao DLN, Lal R, et al.** Comparative genomic analyses reveal core-genome-wide genes under positive selection and major regulatory hubs in outlier strains of *Pseudomonas aeruginosa*. *Front Microbiol*. 2019;10(53). 10.3389/fmicb.2019.00053.

105.   **Wu D-Q, Ye J, Ou H-Y, Wei X, Huang X, He Y-W, et al.** Genomic analysis and temperature-dependent transcriptome profiles of the rhizosphere originating strain *Pseudomonas aeruginosa* M18. *BMC Genomics*. 2011;12(1):438. 10.1186/1471-2164-12-438.

106.   **Kubba Y, Auda J**. Cloning and expression of a protease gene from *Pseudomonas aeruginosa* into *E. Coli*. *Plant Archives*. 2021;21(1):1741-5.

107.   **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;211:77-8. 10.1016/j.jbiotec.2015.07.015.

108.   **Chen Z, Erickson DL, Meng J**. Benchmarking long-read assemblers for genomic analyses of bacterial pathogens using oxford nanopore sequencing. *Int J Mol Sci*. 2020;21(23):9161.

109.   **Wu T, Xu J, Xie W, Yao Z, Yang H, Sun C, et al.** *Pseudomonas aeruginosa* L10: A hydrocarbon-degrading, biosurfactant-producing, and plant-growth-promoting endophytic bacterium isolated from a reed (*Phragmites australis*). *Front Microbiol*. 2018;9(1087):1-12. 10.3389/fmicb.2018.01087.

110.   **Sangare L, Zhao Y, Folly YME, Chang J, Li J, Selvaraj JN, et al.** Aflatoxin B1 degradation by a Pseudomonas strain. *Toxins*. 2014;6(10):3028-40.

111.   **Hou L, Xi J, Liu J, Wang P, Xu T, Liu T, et al.** Biodegradability of polyethylene mulching film by two *Pseudomonas* bacteria and their potential degradation mechanism. *Chemosphere*. 2022;286:131758. https://doi.org/10.1016/j.chemosphere.2021.131758.

112.   **Chan K-G, Yin W-F, Lim YL**. Complete genome sequence of *Pseudomonas aeruginosa* strain YL84, a quorum-sensing strain isolated from compost. *Genome announcements*. 2014;2(2):e00246-14.

113.   **Li G, Lu S, Shen M, Le S, Tan Y, Li M, et al.** Complete genome sequence of *Pseudomonas aeruginosa* phage-resistant variant PA1RG. *Genome announcements*. 2016;4(1):e01761-15.

114.   **Xu A, Wang D, Ding Y, Zheng Y, Wang B, Wei Q, et al.** Integrated Comparative Genomic Analysis and Phenotypic Profiling of *Pseudomonas aeruginosa* Isolates From Crude Oil. *Front Microbiol*. 2020;11(519):1-13. 10.3389/fmicb.2020.00519.

115.   **Ali F, Hu H, Xu P, Tang H**. Complete genome sequence of Pseudomonas aeruginosa FA-HZ1, an efficient dibenzofuran-degrading bacterium. *Genome announcements*. 2017;5(7):e01634-16.

116.   **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 announcements*. 2014;2(1):e01130-13.

117.   **Berger C, Rückert C, Blom J, Rabaey K, Kalinowski J, Rosenbaum MA**. Estimation of pathogenic potential of an environmental *Pseudomonas aeruginosa* isolate using comparative genomics. *Sci Rep*. 2021;11(1):1370. 10.1038/s41598-020-80592-8.

118.   **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 announcements*. 2017;5(3):e01489-16.

119.   **Huang L, Wang W, Zanaroli G, Xu P, Tang H**. Hexabromocyclododecanes are dehalogenated by CYP168A1 from a Pseudomonas strain HS9. *Appl Environ Microbiol*. 2021:AEM. 00826-21.

120.   **Hickey WJ, Focht DD**. Degradation of mono-, di-, and trihalogenated benzoic acids by *Pseudomonas aeruginosa* JB2. *Appl Environ Microbiol*. 1990;56(12):3842-50. doi:10.1128/aem.56.12.3842-3850.1990.

121.   **Kumagai Y, Yoshizawa S, Nakamura K, Ogura Y, Hayashi T, Kogure K**. Complete and draft genome sequences of eight oceanic *Pseudomonas aeruginosa* strains. *Genome Announcements*. 2017;5(44):e01255-17. doi:10.1128/genomeA.01255-17.

122.   **Dong W, He C, Li Y, Huang C, Chen F, Ma Y**. Complete genome sequence of a versatile hydrocarbon degrader, *Pseudomonas aeruginosa* DN1 isolated from petroleum-contaminated soil. *Gene Reports*. 2017;7:123-6. https://doi.org/10.1016/j.genrep.2017.04.001.
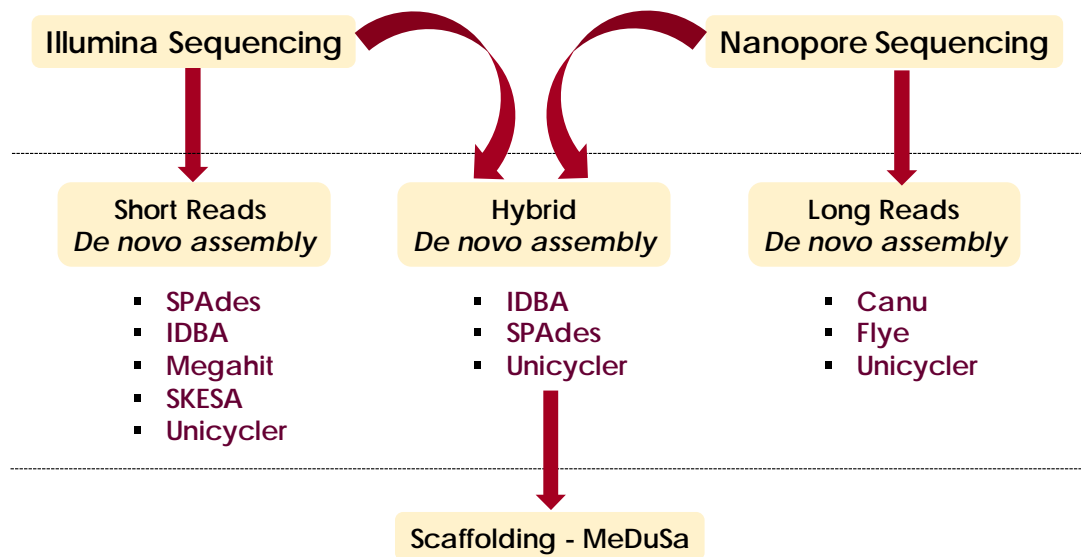
Fig. 1. *Pseudomonas aeruginosa* **PPA14 genome sequencing and assembly.** The flow chart represents the different sequencing platforms and assembly tools used for the sequencing and assembly of the PPA14 genome (Method adapted from Molina-Mora et al. 2020). SPAdes v3.13.0, IDBA v1.1.3. and Megahit v1.1.3, SKESA v2.3.039, and Unicycler v0.4.711, Canu 1.8, and Flye 2.3.7 were the software versions used.
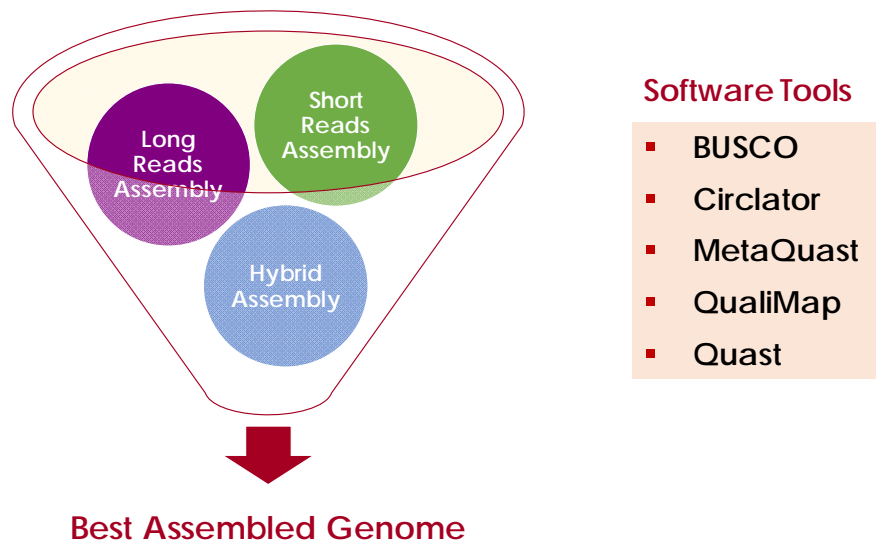
Fig. 2. Comparison and evaluation of *P. aeruginosa* PPA14 genome assemblies. The figure represents the list of tools used to select the best-assembled draft genome by screening all the short, long, and hybrid assemblies of the PPA14 genome (Method adapted from Molina-Mora et al. 2020).
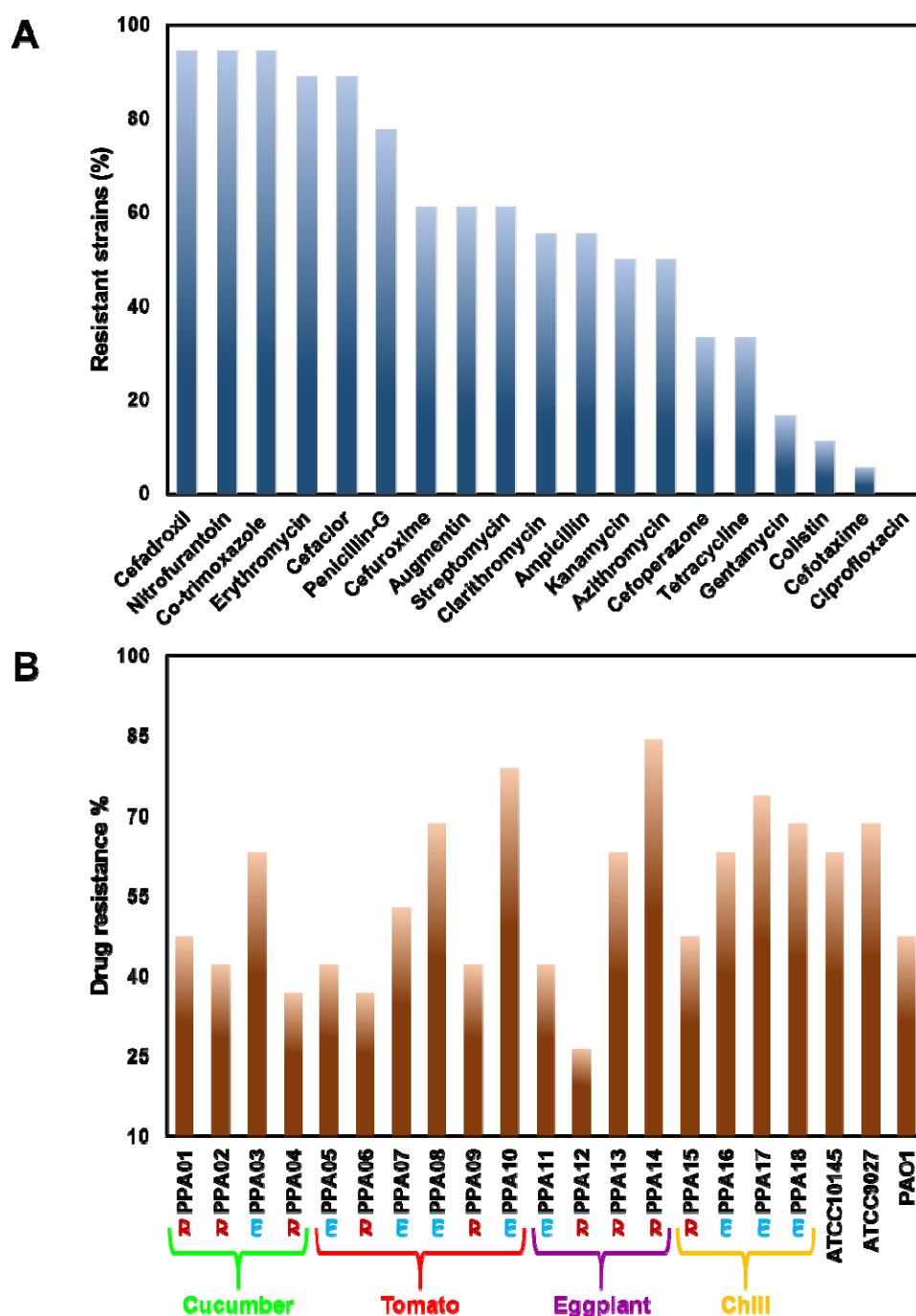
Fig 3. Antibiotic resistance profile of *P. aeruginosa* strains. (A) The efficacy of different antibiotics in inhibiting plant-associated *P. aeruginosa* strains. (B) The percentage of drug resistance exhibited by each of the tested plant-associated and clinical strains of *P. aeruginosa.* The strains are color-coded based on their source of isolation. Green – cucumber; red – tomato; purple – eggplant; yellow – chili; black – clinical isolates.
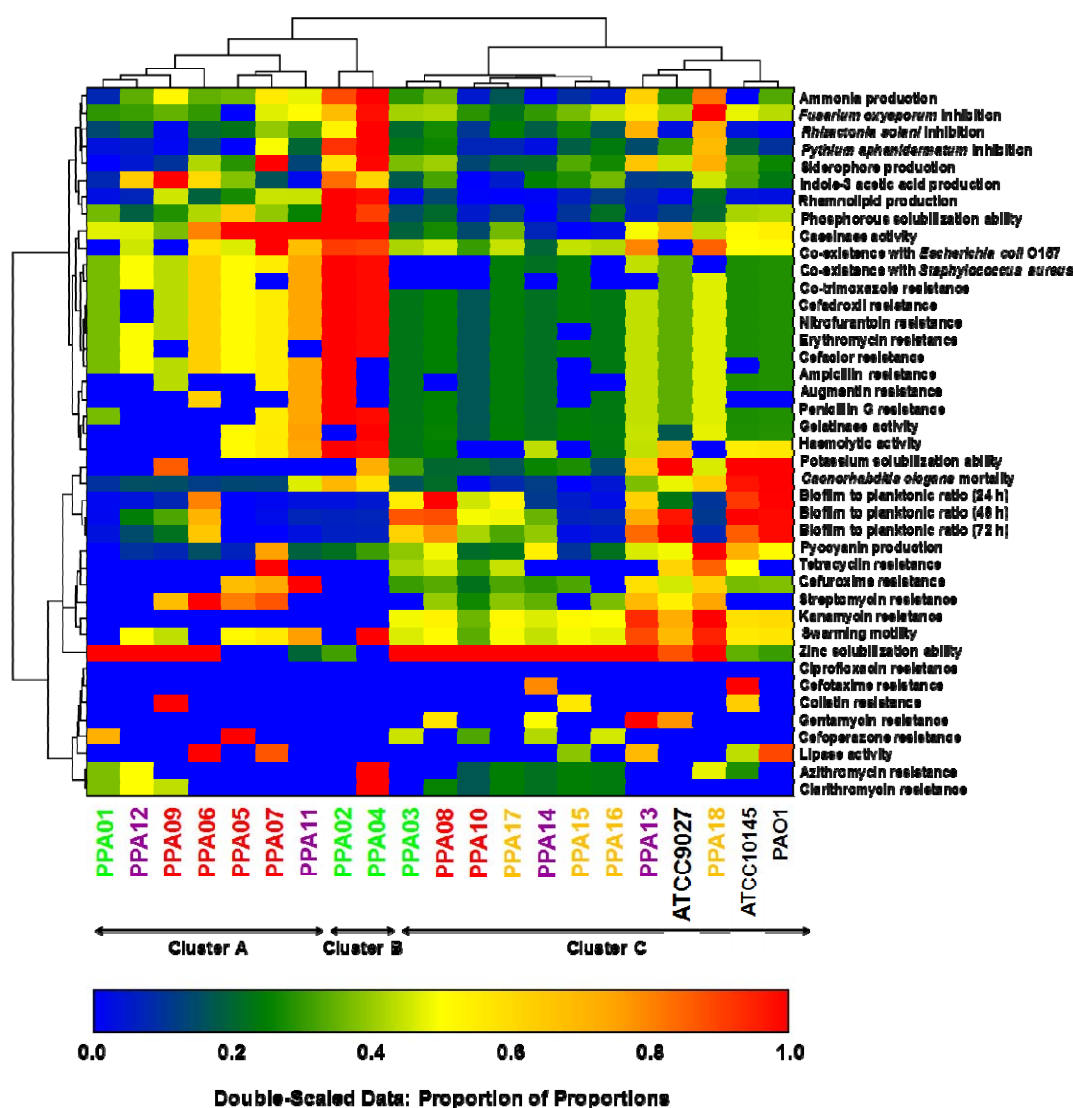
Fig. 4. Hierarchical clustering analysis of *P. aeruginosa* strains. Double dendrograms (top and left) depict the relatedness between strains and between traits, respectively. Hierarchical clustering was done based on Ward minimum variance with Manhattan distance as the distance metric. The proportion method was used for variable scaling. Cell colors correspond to the relative abundance of the traits after mean-centering and unit-scaling, ranging from blue (low value) to red (high value). The strain names are color-coded based on their isolation source, green – cucumber, red – tomato, purple – eggplant, yellow – chili, black – clinical origin.
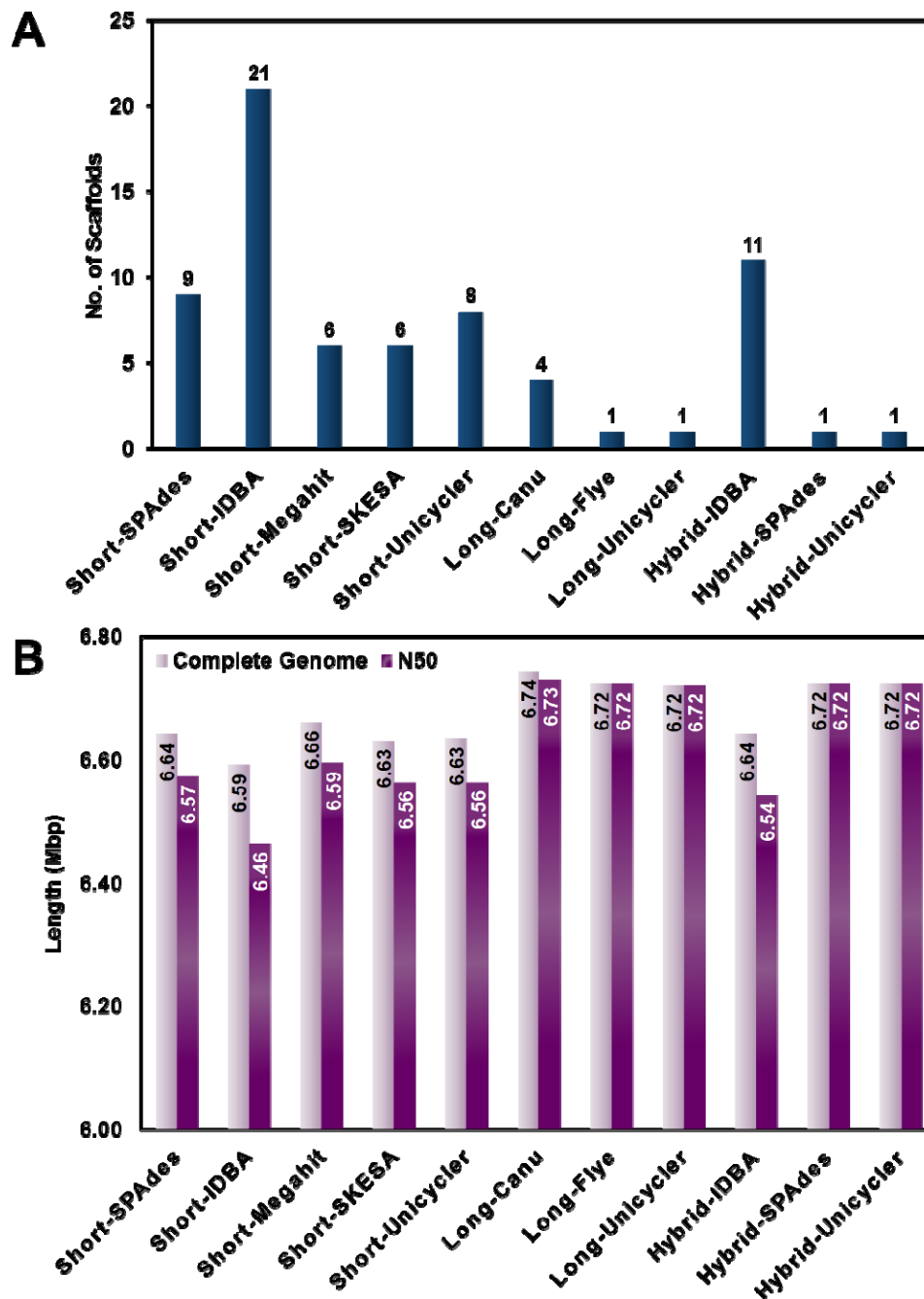
Fig. 5. Variations in PPA14 genome contiguity when assembled using different assembly tools. (a) The number of scaffolds generated during the PPA14 genome assembly by different tools (b) The genome size and N50 value of the PPA14 genome assemblies produced by each tool (Molina-Mora et al. 2020). N50 - length of the shortest contig that covers 50% of the genome size.
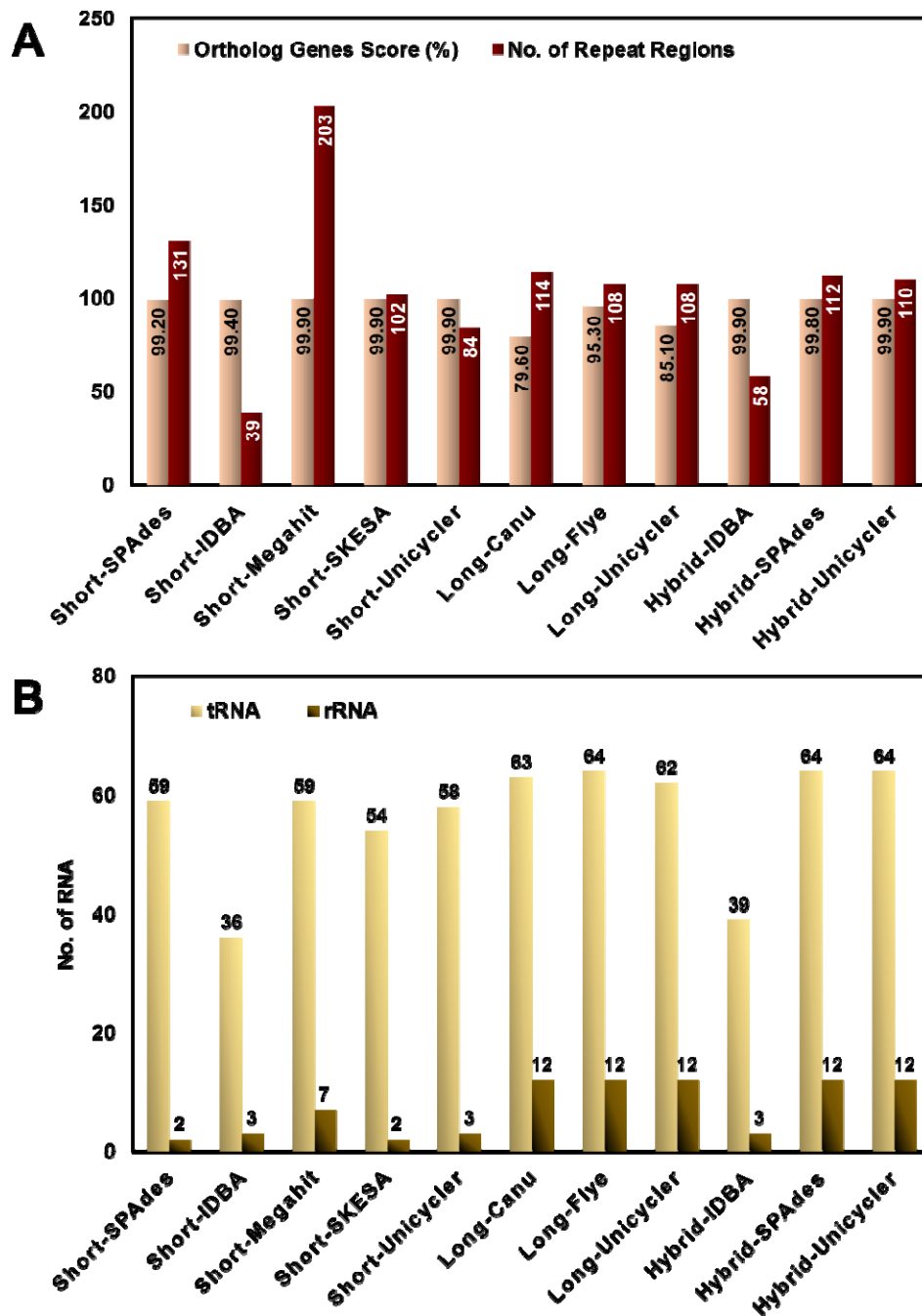
Fig. 6. Variations in PPA14 genome completeness when assembled using different tools. (a) Orthologous genes score and the number of repeat regions in different genome assemblies. (b) The number of tRNAs and rRNAs identified in different genome assemblies (Molina-Mora et al. 2020).
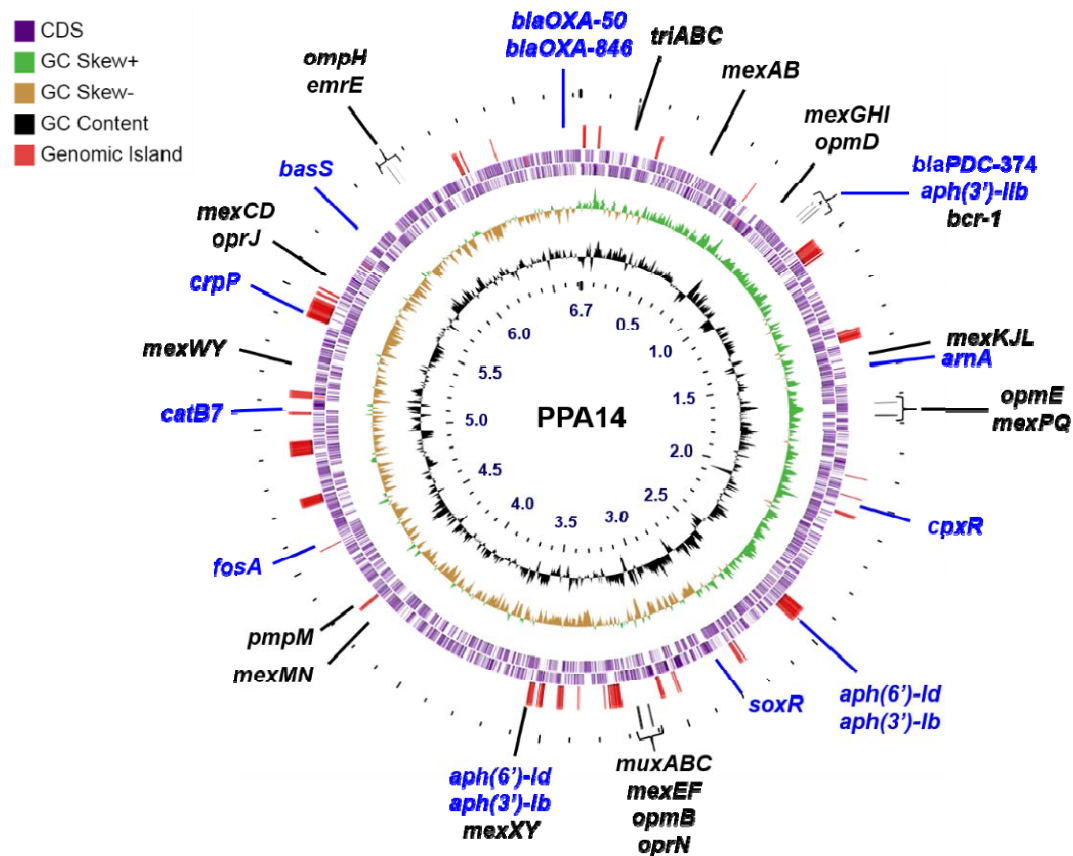
Fig. 7. Genome map. The PPA14 genome map representing the AMR genes and genomic islands. The genes in blue font code for proteins that confer direct resistance against specific antibiotics while the genes in black font code for proteins that act as efflux pumps against a wide spectrum of drugs.
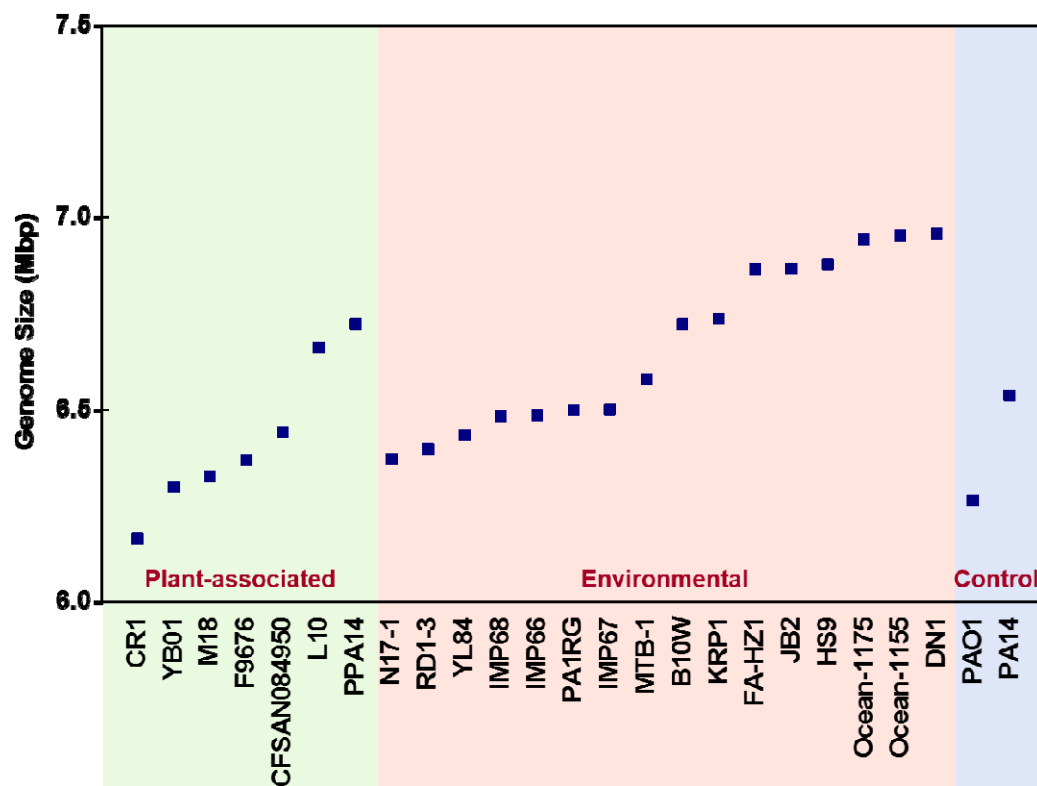
Fig. 8. Genomic size of plant-associated, and environmental *Pseudomonas aeruginosa* strains. The graph depicts the strain level variations in the *P. aeruginosa* genome size. The strains are colored based on their source of origin. The plant-associated, environmental, and control strains of *P. aeruginosa* are given green, orange, and blue background colors, respectively.
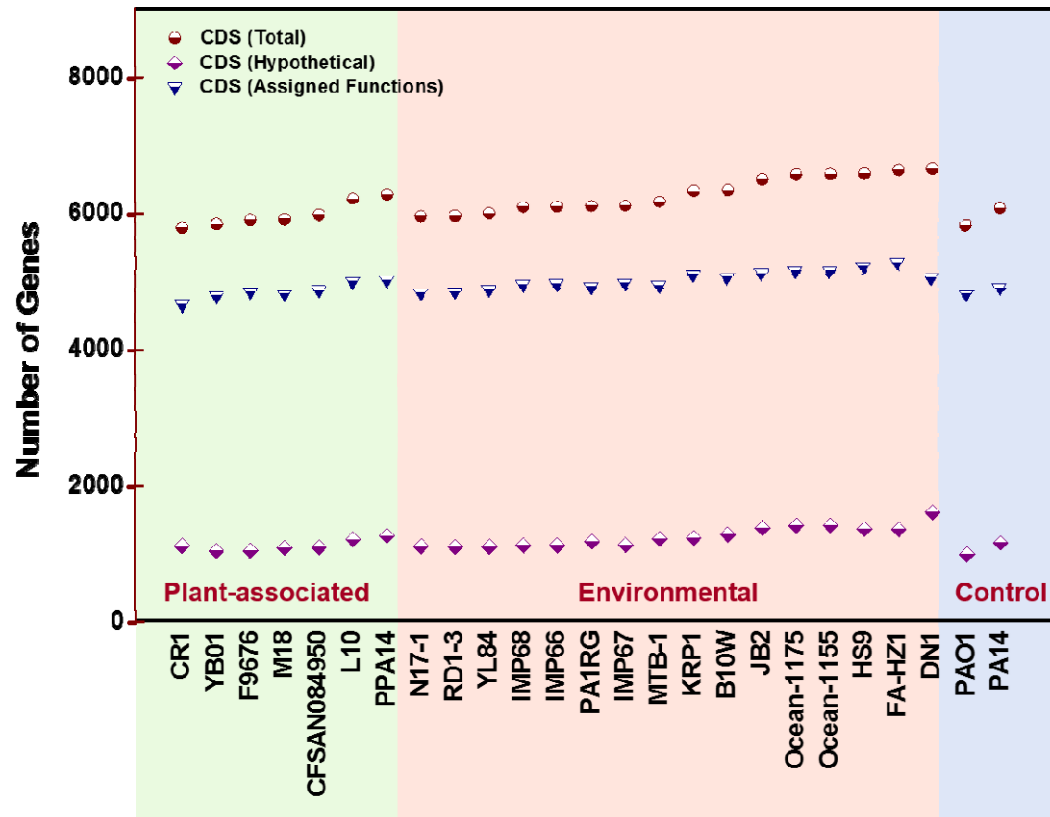
Fig. 9. Genomic features of plant-associated, and environmental strains of *Pseudomonas aeruginosa.* The graph represents CDS present in each strain (brown circle), the CDS that codes for proteins with functional assignments (blue triangle), and hypothetical proteins (purple diamond). The plant-associated, environmental, and control strains of *P. aeruginosa* are given green, orange, and blue background colors, respectively.
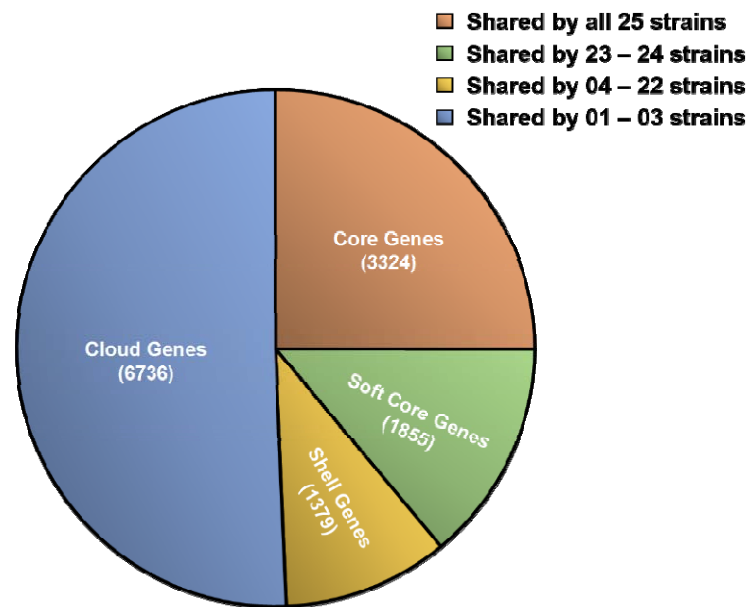
Fig. 10. Pan-genome components of *P. aeruginosa* strains. The pie chart represents the number of CDS identified as the hardcore, softcore, shell, and cloud genes while comparing the complete genomes of *P. aeruginosa* strains isolated from plants and environmental ecosystems.
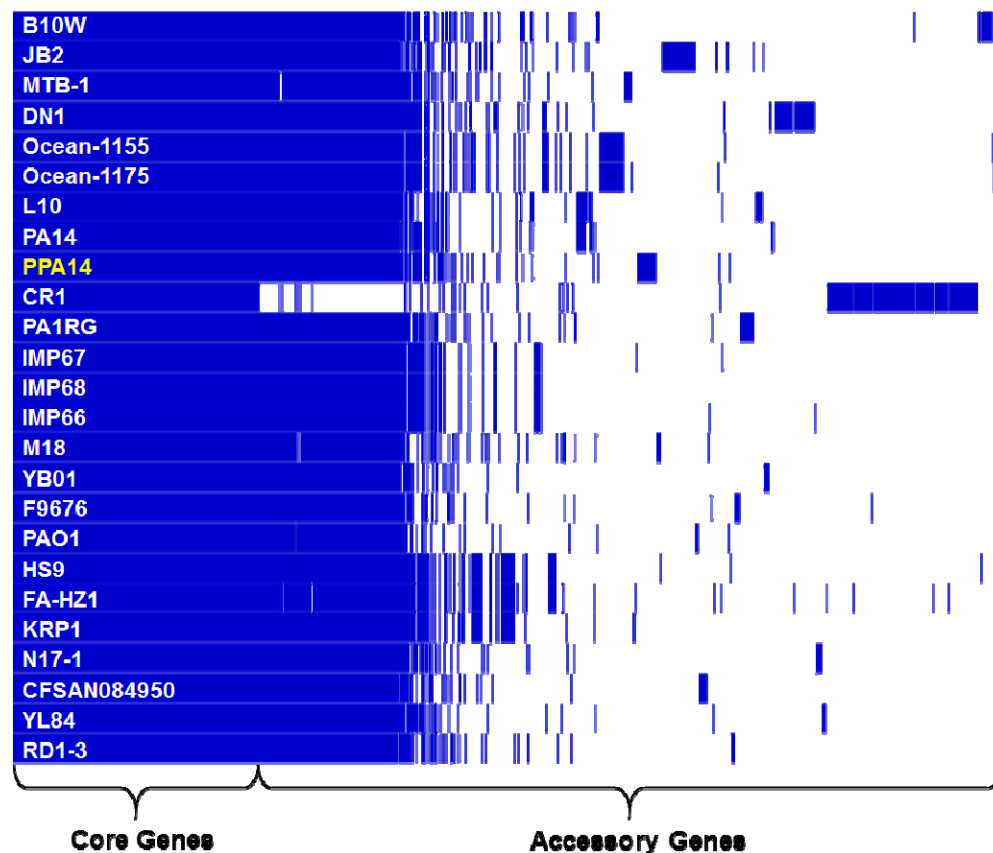
Fig. 11. Pan-genome profiles of plant-associated and environmental *Pseudomonas aeruginosa* strains. Visualization of the pan-genome comprising of the core and accessory genes. The *P. aeruginosa*, PPA14, isolated in the current study is highlighted in yellow.
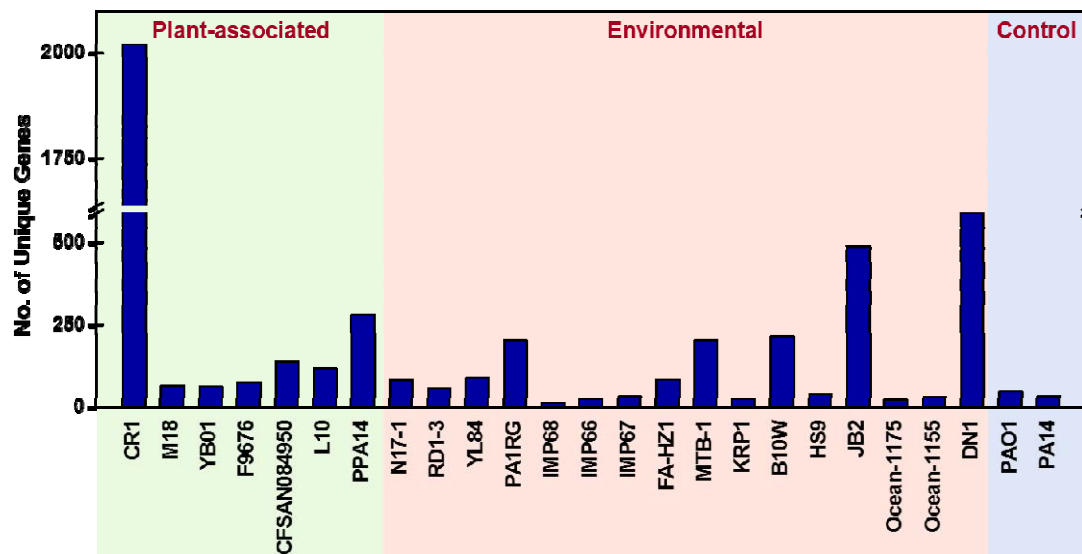
Fig. 12. Unique genes harbored by *Pseudomonas aeruginosa* genomes. The graph represents the strain-level variations in the number of unique genes detected in the *P. aeruginosa* genome. The plant-associated, environmental, and control strains of *P. aeruginosa* are given green, orange, and blue background colors, respectively.
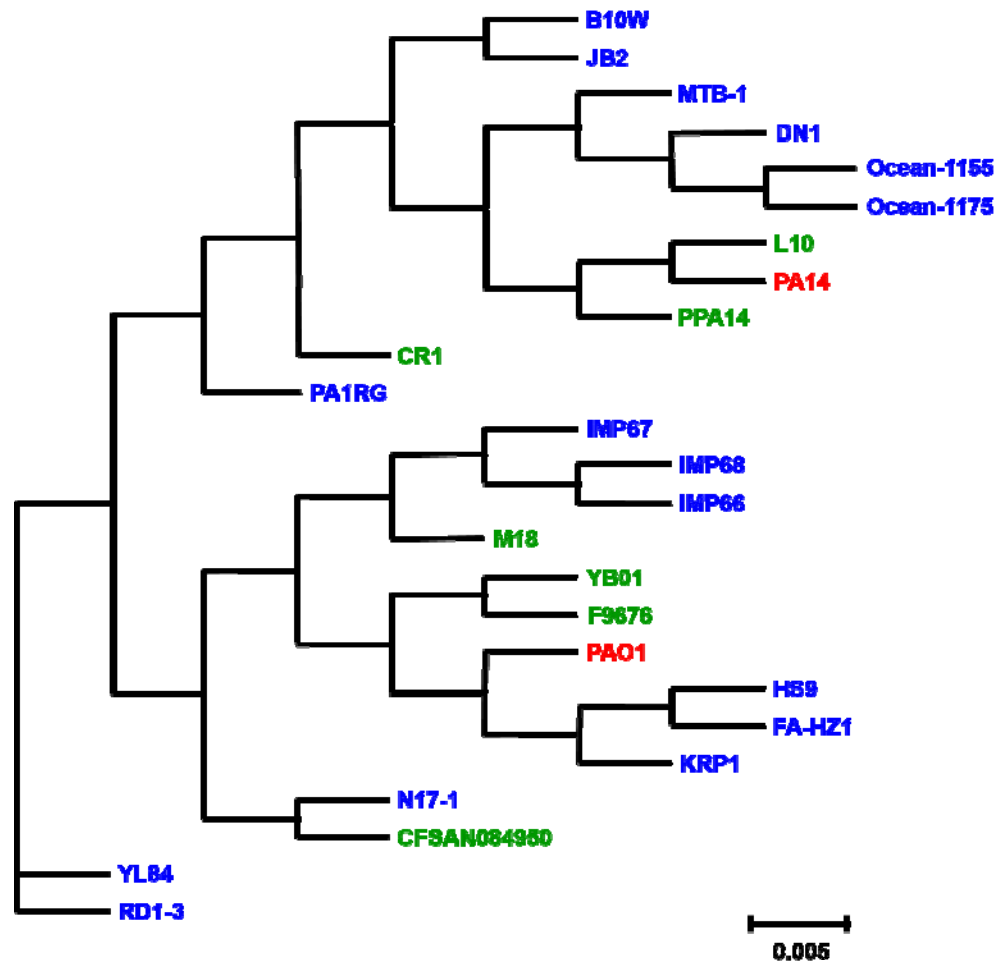
Fig. 13. Phylogenetic analyses of *Pseudomonas aeruginosa* strains. The phylogenetic tree was constructed based on the Maximum Composite Likelihood method. The scale value represents the evolutionary distance. The agricultural, environmental, and clinical strains are in green, blue, and red fonts, respectively