

RESEARCH ARTICLE

# Next generation sequencing reveals the antibiotic resistant variants in the genome of *Pseudomonas aeruginosa*

Babu Ramanathan<sup>1\*</sup>, Hassan Mahmood Jindal<sup>2</sup>, Cheng Foh Le<sup>3</sup>, Ranganath Gudimella<sup>4</sup>, Arif Anwar<sup>4</sup>, Rozaimi Razali<sup>4</sup>, Johan Poole-Johnson<sup>4</sup>, Rishya Manikam<sup>5</sup>, Shamala Devi Sekaran<sup>2\*</sup>

**1** Department of Biological Sciences, School of Science and Technology, Sunway University, Kuala Lumpur, Malaysia, **2** Department of Medical Microbiology, Faculty of Medicine, University of Malaya, Kuala Lumpur, Malaysia, **3** School of Pharmacy, Faculty of Science, University of Nottingham Malaysia Campus, Jalan Broga, Selangor, Malaysia, **4** Sengenics, High Impact Research (HIR), University of Malaya, Kuala Lumpur, Malaysia, **5** Department of Trauma and Emergency, University Malaya Medical Centre, Kuala Lumpur, Malaysia

\* [babur@sunway.edu.my](mailto:babur@sunway.edu.my) (BR); [shamalamy@yahoo.com](mailto:shamalamy@yahoo.com) (SDS)



**OPEN ACCESS**

**Citation:** Ramanathan B, Jindal HM, Le CF, Gudimella R, Anwar A, Razali R, et al. (2017) Next generation sequencing reveals the antibiotic resistant variants in the genome of *Pseudomonas aeruginosa*. PLoS ONE 12(8): e0182524. <https://doi.org/10.1371/journal.pone.0182524>

**Editor:** Chiyu Zhang, Institut Pasteur of Shanghai Chinese Academy of Sciences, CHINA

**Received:** February 16, 2017

**Accepted:** July 19, 2017

**Published:** August 10, 2017

**Copyright:** © 2017 Ramanathan et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data Availability Statement:** All relevant data are within the paper and its Supporting Information files.

**Funding:** This study was supported by University of Malaya High Impact Research Grant (reference number: UM.C/HIR/MOHE/MED/40, account number: H-848 20001-E000079) and University of Malaya Research Grant (UMRG Project no. RP020C-14AFR and RP001C-13ICT). Support from internal research grant (INTS-2017-SST-RCBS-01)

## Abstract

Rapid progress in next generation sequencing and allied computational tools have aided in identification of single nucleotide variants in genomes of several organisms. In the present study, we have investigated single nucleotide polymorphism (SNP) in ten multi-antibiotic resistant *Pseudomonas aeruginosa* clinical isolates. All the draft genomes were submitted to Rapid Annotations using Subsystems Technology (RAST) web server and the predicted protein sequences were used for comparison. Non-synonymous single nucleotide polymorphism (nsSNP) found in the clinical isolates compared to the reference genome (PAO1), and the comparison of nsSNPs between antibiotic resistant and susceptible clinical isolates revealed insights into the genome variation. These nsSNPs identified in the multi-drug resistant clinical isolates were found to be altering a single amino acid in several antibiotic resistant genes. We found mutations in genes encoding efflux pump systems, cell wall, DNA replication and genes involved in repair mechanism. In addition, nucleotide deletions in the genome and mutations leading to generation of stop codons were also observed in the antibiotic resistant clinical isolates. Next generation sequencing is a powerful tool to compare the whole genomes and analyse the single base pair variations found within the antibiotic resistant genes. We identified specific mutations within antibiotic resistant genes compared to the susceptible strain of the same bacterial species and these findings may provide insights to understand the role of single nucleotide variants in antibiotic resistance.

## Background

*Pseudomonas aeruginosa* is a major opportunistic pathogen causing acute and chronic infections in human community. It is a fascinating, ubiquitous, gram-negative bacterium that can

and a publication support grant from Sunway University is gratefully acknowledged.

**Competing interests:** The authors have read the journal's policy and the authors of this manuscript have the following competing interests: SDS is an academic editor for PLOS ONE and BR is a reviewer in the peer review panel. The authors confirm that this does not alter their adherence to all PLOS ONE policies on sharing data and materials.

thrive at low densities within the range of 4°C to 42°C and involves in a range of interactions with eukaryotic hosts [1, 2]. The infection leads to most serious manifestations such as bacteremia, pneumonia, urinary tract infections and wound infections [3]. *P. aeruginosa* is intrinsically resistant to many antibiotics and the challenge is selection of the most appropriate antibiotic as the organism can develop resistance during treatment [4, 5].

The infection caused by multi-drug resistant (MDR) *P. aeruginosa* often leads to morbidity, mortality, chronic care and an increase towards the overall cost of treating the infection [6, 7]. Mutational changes found within the resistant genes on plasmids or mutations that alter the function of genes encoded by chromosomes may collectively contribute to the antibiotic resistant mechanisms [8]. The antibiotic resistant mechanisms include the acquisition of extended spectrum  $\beta$ -lactamases, carbapenemases, aminoglycoside-modifying enzymes and 16S ribosomal ribonucleic acid (rRNA) methylases. Mutational changes causing the up regulation of multidrug efflux pumps, depression of ampC, modification of antimicrobial targets and changes in the outer membrane permeability barrier have also been reported [8].

Epidemiological typing of *P. aeruginosa* requires the identification of stable distinguishing characteristics. Current molecular epidemiological typing techniques in practice such as variable-number tandem repeat (VNTR) and pulsed-field gel electrophoresis (PFGE) have helped to reveal the outbreaks of *P. aeruginosa* in hospital settings [9, 10]. However, limited resolution, cost and complex workflow of these systems make the typing in reference laboratories be performed in batch mode that affects the real time sequencing efforts. In addition, the traditional typing methods lacks in providing insights into the evolutionary significance of pathogens [11]. Multilocus sequence typing (MLST) to identify specific genes in clinical isolates has also been reported to be an effective genotyping method [12].

Application of next-generation sequencing (NGS) technology for whole genome, whole proteome and whole transcriptome sequencing is becoming popular. NGS is quick and high throughput technique that follows a single unified workflow. Using this technique, it is possible to identify the single-base-pair mutations within same bacterial species that can replace the traditional molecular typing methods for bacterial pathogens [13]. This approach provides greater sequence resolution than traditional methods by delivering a definitive catalog of genetic polymorphisms, particularly single-nucleotide polymorphisms (SNPs). WGS also associates epidemiology to genome evolution, genome structure, pathogen biology and gene content; which provides insights to biological markers, such as antibiotic resistance and virulence factors [14].

Identification of SNPs in bacterial genomes play a key role to determine the relationship between antibiotic resistant clinical isolates and tracing their evolutionary counterpart [15, 16]. Several studies have reported the genetic variants in cystic fibrosis disease conditions where the pathogen has emerged because of mutation and long-term colonization [17, 18]. In addition, genome sequencing facilitates high throughput screening for specific mutations within antibiotic resistant genes compared to the susceptible and/or reference strain of the same bacterial species. Thus, in the present study, we examined single nucleotide variations within the antibiotic resistant and susceptible *P. aeruginosa* clinical isolates collected from patients admitted at the University Malaya Medical Centre, Kuala Lumpur, Malaysia. These nsSNPs found within the antibiotic resistant isolates could potentially alter the amino acid sequence and may affect the stability and/or the function of the resulting protein expression.

## Materials and methods

### Bacterial strains and culture condition

*P. aeruginosa* isolates were collected from different hospital wards (Table 1) at the University Malaya Medical Centre, Kuala Lumpur, Malaysia over a period of 1 year from April 2009 to

**Table 1. Selected *P. aeruginosa* clinical isolates included in the study from University Malaya Medical Centre, Kuala Lumpur, Malaysia over a period of 1 year from April 2009 to March 2010.**

Isolate	Ward	Specimen
PAS1	Surgery	Wound
PAS2	Orthopedic	Wound
PAS3	Surgery	Urine
PAS4	Surgery	Urine
PAS5	Surgery	Wound
PAS6	Medical	Urine
PAS7	Surgery	Urine
PAS8	Pediatric	Blood
PAS9	Pediatric	Wound
PAS10	Medical	Blood

<https://doi.org/10.1371/journal.pone.0182524.t001>

March 2010. The clinical strains were isolated from urine, wound, blood and indwelling medical devices. Isolates were frozen with Luria Bertani (LB) broth containing 30% V/V glycerol and stored at -80°C until used. The protocol for this study was approved by the Medical Ethics Committee, University Malaya Medical Centre, Kuala Lumpur, Malaysia.

### Antimicrobial susceptibility testing

The antimicrobial susceptibility of isolates was tested using the disc diffusion method (E-test®) against following nine different antibiotics; piperacillin/tazobactam (PPT), ceftazidime (CAZ), aztreonam (AZT), amikacin (AK), gentamicin (GN), ciprofloxacin (CIP), imipenem (IMP), meropenem (MPM) and colistin (CL). The reference strain *P. aeruginosa* ATCC # 27853 was used to validate the technique. The calibrated carrier strip and the intersection of the strip to the edge of the inhibitory zone indicates the minimum inhibitory concentration (MIC) [19]. The clinical isolates were chosen for whole genome sequencing based on the antimicrobial susceptibility levels screened against different antibiotics.

### High-throughput whole genome sequencing

Genomic DNA was isolated from *P. aeruginosa* clinical isolates using the DNeasy DNA extraction kit (Qiagen, USA) according to the manufacturer's instructions. The samples were then quantified using Qubit. All genomic DNA were fragmented using Covaris S2 at the temperature of 5.5 to 6°C for 40 seconds. The fragmented DNA was end repaired, added with dA base and ligated with Illumina indexed adapters. Size selections of the samples were performed using Invitrogen 2% agarose E-gels. The selected DNA fragments with adapter molecules on both ends underwent 10 cycles of PCR amplification and was sequenced using whole genome shotgun Illumina Hiseq 2000 flow cell V3 with paired-end libraries (~200 base-pair insert size).

### Read pre-processing

Adapter sequences and low quality reads were removed with quality score filter of  $\geq 30$  using PRINSEQ version 0.20.3 [20]. The following types of reads were removed: reads having 'N' in more than 10% of the total bases of that read, reads with Phred quality score less than 20 and reads shorter than 50 bp.

## Genome assembly

The sequenced reads were assembled using SPAdes Assembler version 3.8.1 [21]. Assembler was run using iterative kmer lengths ranging from 27 to 77. The assemblies were mapped against *P. aeruginosa* reference genome POA1 to evaluate the core genome average identities and completeness. The removal of probable contaminants was performed by BLAST against the assembly sequences of mitochondrial and primates'/rodents chromosome databases. The finalized draft genomes were submitted to Rapid Automated using Subsystem Technology (RAST) [22] server for gene predictions and annotations. Prokka Version 1.11 (Prokaryotic annotation) was used to perform the gene prediction [23]. The prediction is relied on the existing annotation resources such proteins and coding DNA sequences (CDS).

## Analysis of single nucleotide polymorphism (SNP)

To identify genetic differences at genome level, the sequencing reads of each strain were mapped to its corresponding reference genome *P. aeruginosa* POA1 using Bowtie v0.12.0 software [24]. High-confidence SNP variants data sets were created by applying series of filters. The Variants were identified and extracted using Samtools with the following parameters: each variant must be supported by at least 10 reads (-d = 10), only the variants that were supported by maximum of 10,000 reads were considered for downstream analysis (-D = 10,000) and the reads must be supported by mapping quality of equal or larger than 25 (MQ = 25). Variants harbored by more than 90% of reads were included for further analysis and the depth of coverage was 25x. The effects of nsSNPs were annotated using SnpEFF.

## Genome comparison

Protein Basic Local Alignment Search Tool (BLASTp) was used to match the sequence similarities between genomes of clinical isolates and the *P. aeruginosa* reference genome PAO1 [25]. Genome visualization of ten clinical isolates and the similarity between *P. aeruginosa* reference genome (NC\_002516.2) was performed using BLAST Ring Image Generator (BRIG) [26].

## Heat map

Heat maps were generated using Complex Heat map package from Bioconductor in R [27]. Clusters are predicted using Euclidean distance method.

## Results

### Antibiotic-resistance profiles of *P. aeruginosa* clinical isolates

A total of ten *P. aeruginosa* clinical isolates with resistant to three or more of the nine antibiotics tested were chosen for the study. The MIC values of nine antibiotics tested against the bacterial isolates are shown in Table 2. Isolates PAS2 and PAS10 were highly resistant to all the antibiotics tested. In contrary, isolates PAS4 and PAS8 were susceptible to at least 3 antibiotics (Imipenem, Meropenem and Collistin). PAS1 and PAS7 were only susceptible to Amikacin and Piperacillin/Tazobactam respectively. All ten isolates were resistant to Cefazidime, Aztreonam, Gentamicin and Ciprofloxacin.

### Genome sequencing and assembly

The sequencing consisted of 1 lane 100 bp paired-end reads, yielded approximately 0.8 Gbp to 2.8 Gbp for each clinical isolates. The draft genome assemblies for the ten *Pseudomonas* clinical isolates have been submitted to NCBI bioproject under the project accession number

**Table 2. Minimum inhibitory concentration (MIC) of the ten *P. aeruginosa* clinical isolates used in the current study.**

Isolate	IMP (≥32)	MPM (≥32)	CAZ (≥48)	AZT (≥16)	PPT (≥24)	GN (≥32)	AK (≥24)	CIP (≥4)	CL (≥3)
PAS1	32	32	256	32	24	32	16	32	4
PAS2	32	32	256	32	32	48	24	32	6
PAS3	2	32	256	256	256	384	24	32	2
PAS4	2	3	256	256	256	1024	48	32	2
PAS5	1.5	32	256	24	128	1024	192	4	2
PAS6	1.5	3	256	64	256	1024	32	32	3
PAS7	32	32	256	24	16	48	32	32	3
PAS8	1.5	1	256	16	256	1024	98	12	1.5
PAS9	3	1.5	256	16	128	1024	128	16	6
PAS10	32	32	48	16	256	1024	256	32	4

Abbreviations: IMP, imipenem; MPM, meropenem; CAZ, ceftazidime; AZT, aztreonam; PPT, piperacillin/tazobactam; GN, gentamicin; AK, amikacin; CIP, ciprofloxacin; CL, colistin.

The bracketed values indicate the breakpoints for the classification of resistance for the respective antibiotics.

<https://doi.org/10.1371/journal.pone.0182524.t002>

PRJNA388357 (<http://www.ncbi.nlm.nih.gov/bioproject/388357>). More than 80% of the reads are above Phred quality score of 30 indicating high quality sequencing data. The sequenced data were subjected to contamination screening and the genome sizes of final assemblies were ranging from 6.6 Mbp to 7.0 Mbp. The de novo assembly of genome sequence data revealed that the number of contigs (>200 bp) varied from 185 to 492 for each genome. These contigs were aligned to the reference genome of *P. aeruginosa* PAO1, and manual genome finishing was performed. The maximum contig size among the genomes was 485,560 bp aligned to PAS4. Genomes with the mean identity ranged from 98.77 for PAS 7 to 98.83 for PAS 9. The GC content ranged from 65.79 for PAS 1 to 66.19 for PAS 2. A summary of the genome sequence data and assembly are shown in Table 3.

## Genome architecture

Fig 1 contains the circular map of the ten newly sequenced clinical isolates compared to the reference genome PAO1. The inner most circle represents the reference genome of *P. aeruginosa* (NC\_002516.2) and outer most circle with labels represent the protein-coding regions (CDS) with annotation on the genome. The shared identity of each isolate with the reference genome is represented in different colors, which denotes the BLASTn matches between 70% to 100% nucleotide identities. The blank spaces in the rings represent the matches less than 70% or no BLAST matches to the reference genome. The nsSNPs harbored on the genes compared to the reference genome are represented as individual gene names.

## Genome polymorphism

We focused our analysis on identifying nsSNPs that results in amino acid changes within the functional gene sequences. In addition, we also screened for point mutations that leads to premature termination of codons or nucleotide deletions that create frame-shift mutations. We initially screened the nsSNPs in clinical isolates compared to the *P. aeruginosa* reference genome, PAO1. Fig 1 shows the genome organization of all the clinical isolates. The CDS regions (in dark) and the genes harboring these regions containing nsSNPs were illustrated in Fig 1. As high as 18,876 nsSNPs were identified in Amikacin resistant isolates when compared to the PAO1 genome; whereas, only 81 nsSNPs were identified in Piperacillin resistant isolates

**Table 3. Illumina sequencing results for *P. aeruginosa* isolates.**

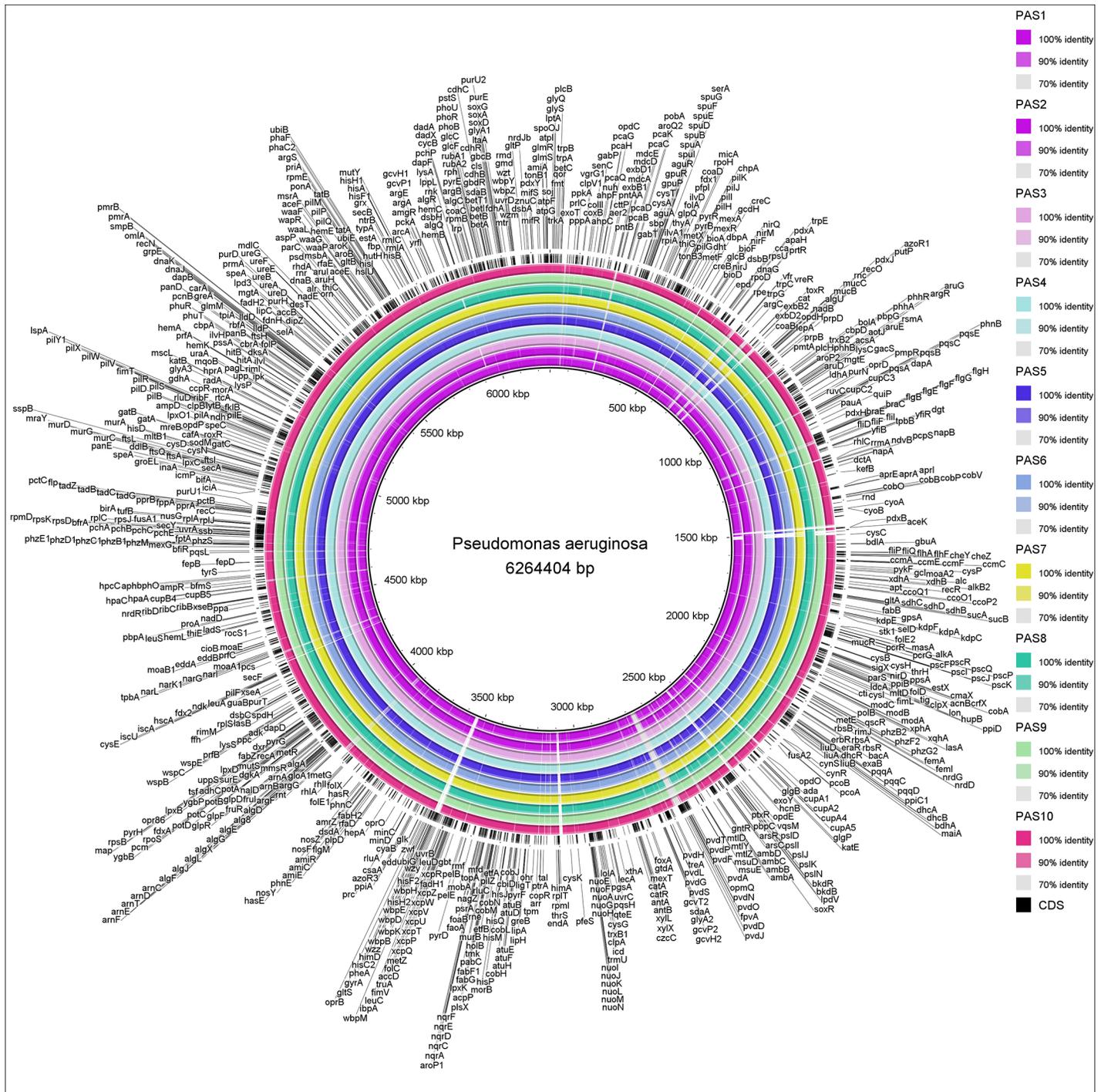
Strain	Number of contig	Genome Size (bp)	Reads per pair	Minimum Contig Size (bp)	Maximum Contig Size (bp)	N50 Size (bp)	% of $\geq$ Q30 bases	GC Content (%)	Completeness (%)	Average Identity (%)
PAS1	325	6,909,968	28282644	200	404,308	102,152	84.02	65.79	80.34	98.82
PAS2	243	6,627,283	25135336	200	379,250	119,180	84.25	66.19	79.06	98.79
PAS3	322	6,731,789	20991654	200	327,510	102,161	91.29	66.12	79.68	98.80
PAS4	185	6,753,415	8898394	205	485,560	118,034	85.17	66.03	79.80	98.80
PAS5	230	6,954,778	9978228	200	404,098	105,908	84.81	65.93	81.22	98.81
PAS6	189	7,036,089	7154400	200	377,351	113,718	84.42	65.99	81.21	98.82
PAS7	492	6,851,779	10743970	200	319,361	91,918	84.64	65.90	80.25	98.77
PAS8	293	6,762,669	8407562	205	238,159	90,287	84.73	66.00	79.61	98.81
PAS9	279	7,050,477	8326856	200	363,660	102,446	83.79	65.85	81.65	98.83
PAS10	239	6,956,100	27086434	200	377,138	93,007	83.63	65.93	81.21	98.82

<https://doi.org/10.1371/journal.pone.0182524.t003>

(Fig 2). Majority of the nsSNPs were present within the PA2018, PA2019 and PA5471 loci that code for aminoglycoside resistant genes *mexAB-oprM* and *mexXY-oprM* responsible for the regulation of multi-drug efflux pump (Table 4).

Then we compared the resistant clinical isolates with the susceptible strains to extract specific nsSNPs found only in resistant isolates. The antibiotic profiles of Ceftazidime, Aztreonam, Gentamicin and Ciprofloxacin were not included as all 10 isolates were resistant to these antibiotics. The isolates classified as resistant for 5 different antibiotics were screened further to identify the nsSNPs. For example, PAS 1 was susceptible to Amikacin and 9 isolates (PAS 2–10) were resistant to the antibiotic. Hence we compared the nsSNPs present within resistant isolates. As shown in S1 Table, the amikacin resistant isolates possess two unique single nucleotide deletions that results in a frameshift mutation that truncates the protein products. Deletion at nucleotide position 1275766 results in a frameshift mutation and truncation of protein NapA (periplasmic nitrate reductase protein). Similarly, deletion at nucleotide position 1402975 results in truncation of 3-mercaptopyruvate sulfurtransferase. The detailed list of unique nsSNPs identified in the resistant clinical isolates compared to susceptible isolates for 5 different antibiotics is shown in the S1 Table, S2 Table, S3 Table, S4 Table and S5 Table. Differential mutation screening between the isolates were performed to identify specific nsSNPs associated with the particular antibiotic resistances.

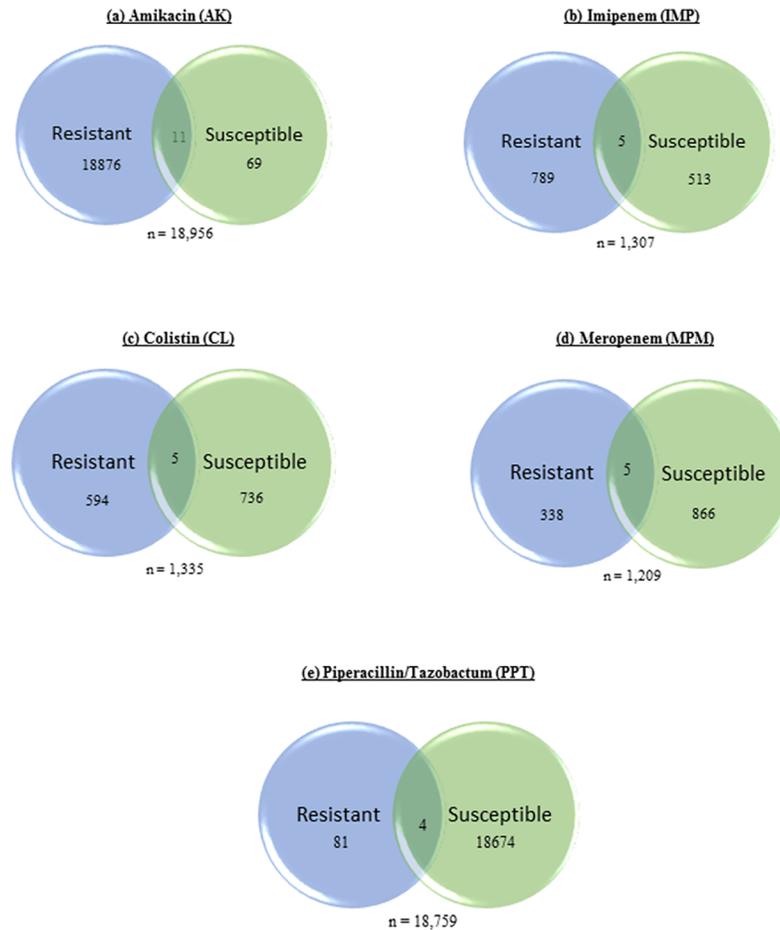
In order to investigate nsSNPs in genes annotated to be associated in antibiotic resistance [8] or virulence factors [28], we specifically extracted the nsSNPs presented in various antibiotic resistant genes of resistant and susceptible clinical isolates. The number of nsSNPs shared within a gene across all clinical isolates is shown in Fig 3. For example, 6 nsSNPs were found within the colistin resistant isolate (PAS1 and 2) when compared to colistin susceptible isolates (PAS 3 and 8) at the gene PA3292 that codes for a hypothetical protein. Similarly, 6 nsSNPs were also seen in all the isolates except PAS 8 at the PA1923 gene that



**Fig 1. Genome map of clinical isolates compared to *P. aeruginosa* reference genome (NC\_002516.2).** All ten isolates were separated in rings. The inner most circle represents the reference genome and outer most circle with labels represent the CDS (in dark). The percentage similarity between each genome are represented in different colors. Map and underlying analysis were performed with the BLAST Ring Image Generator (BRIG) (<http://sourceforge.net/projects/brig/>).

<https://doi.org/10.1371/journal.pone.0182524.g001>

codes for cobaltochelate subunit CobN. Many of the polymorphic genes among these clinical isolates were mainly associated with cell wall, multi-drug efflux pump, protein



**Fig 2. Venn diagram showing the non-synonymous single nucleotide polymorphism (nsSNP's) in antibiotic resistant and susceptible clinical isolates aligned to the reference genome PAO1.** The unique nsSNP's identified in (a) Amikacin, (b) Imipenem, (c) Colistin, (d) Meropenem and (e) Peperacillin/Tazobactam resistant and susceptible isolates were depicted in the figure. The profiles of antibiotics Ceftazidime, Aztreonam, Gentamicin and Ciprofloxacin were not included in the comparison since all ten clinical isolates were resistant to these antibiotics.

<https://doi.org/10.1371/journal.pone.0182524.g002>

secretion, DNA replication and genes involved in repair mechanism. In addition, nsSNPs were found within genes responsible for resistance to  $\beta$ -lactams; transporter (PA4218), transcriptional regulators (PA1184, PA0032, PA5293 and PA2383) and a hypothetical protein (PA3693). These were all collectively reported to play a role in decreased antibiotic uptake and affecting cell wall permeability [29]. Notably, a SNP at position 2213177 in all antibiotic resistant isolates changed tryptophan into a stop codon. Furthermore, two frame shift mutations at the *napA* and *PA1292* genes of Amikacin resistant genes and a mutation resulting in stop codon at amino acid position Q65 in the Meropenem resistant isolates might have implications in the drug resistant mechanism. These findings indicate that the nsSNPs identified in the clinical isolates may contribute towards antibiotic resistance between the resistant clinical isolates examined in the study. Further targeted evidence using a large cohort of clinical isolates and site directed mutational analysis of these antibiotic resistant genes will provide further insights in the antibiotic resistant mechanism of the *P. aeruginosa*.

**Table 4. Key nsSNPs identified in the antibiotic resistant genes of *P. aeruginosa* multi-drug resistant isolates compared to PAO1 reference genome.**

SNP in isolate	SNP in PAO1	Position in PAO1	PAO1 locus	PAO1 amino acid	SNP amino acid	Amino acid position	Protein
G	T	2208200	PA2018	N	T	1036	Resistance–Nodulation–Cell division (RND) multidrug efflux transporter
C	G	2208789	PA2018	Q	E	840	Resistance–Nodulation–Cell division (RND) multidrug efflux transporter
T	C	2209400	PA2018	S	N	636	Resistance–Nodulation–Cell division (RND) multidrug efflux transporter
G	C	2209541	PA2018	G	A	589	Resistance–Nodulation–Cell division (RND) multidrug efflux transporter
C	T	2209680	PA2018	T	A	543	Resistance–Nodulation–Cell division (RND) multidrug efflux transporter
C	T	2209701	PA2018	I	V	536	Resistance–Nodulation–Cell division (RND) multidrug efflux transporter
G	A	2211441	PA2019	W	R	358	Resistance–Nodulation–Cell division (RND) multidrug efflux membrane fusion protein
C	A	2211522	PA2019	L	V	331	Resistance–Nodulation–Cell division (RND) multidrug efflux membrane fusion protein
G	T	2211528	PA2019	K	Q	329	Resistance–Nodulation–Cell division (RND) multidrug efflux membrane fusion protein
C	T	6159991	PA5471	I	V	237	ArmZ
T	C	6160365	PA5471	S	N	112	ArmZ
G	A	6160582	PA5471	C	R	40	ArmZ
A	G	4722060	PA4218	L	F	267	AmpP
C	T	34646	PA0032	L	P	8	Probable Transcriptional regulator
T	C	1286548	PA1184	A	T	81	Probable Transcriptional regulator
T	G	5959101	PA5293	D	E	75	Probable Transcriptional regulator
A	C	4135835	PA3693	A	S	47	Macro domain-containing protein
G	A	2635770	PA2383	S	P	68	Probable Transcriptional regulator
G	A	6159972	PA5471	V	A	243	ArmZ
G	A	2211441	PA2019	W	R	358	Resistance–Nodulation–Cell division (RND) multidrug efflux membrane fusion protein
A	G	2213177	PA2020	W	*	167	MexZ
C	T	5959144	PA5293	Q	R	61	Probable Transcriptional regulator

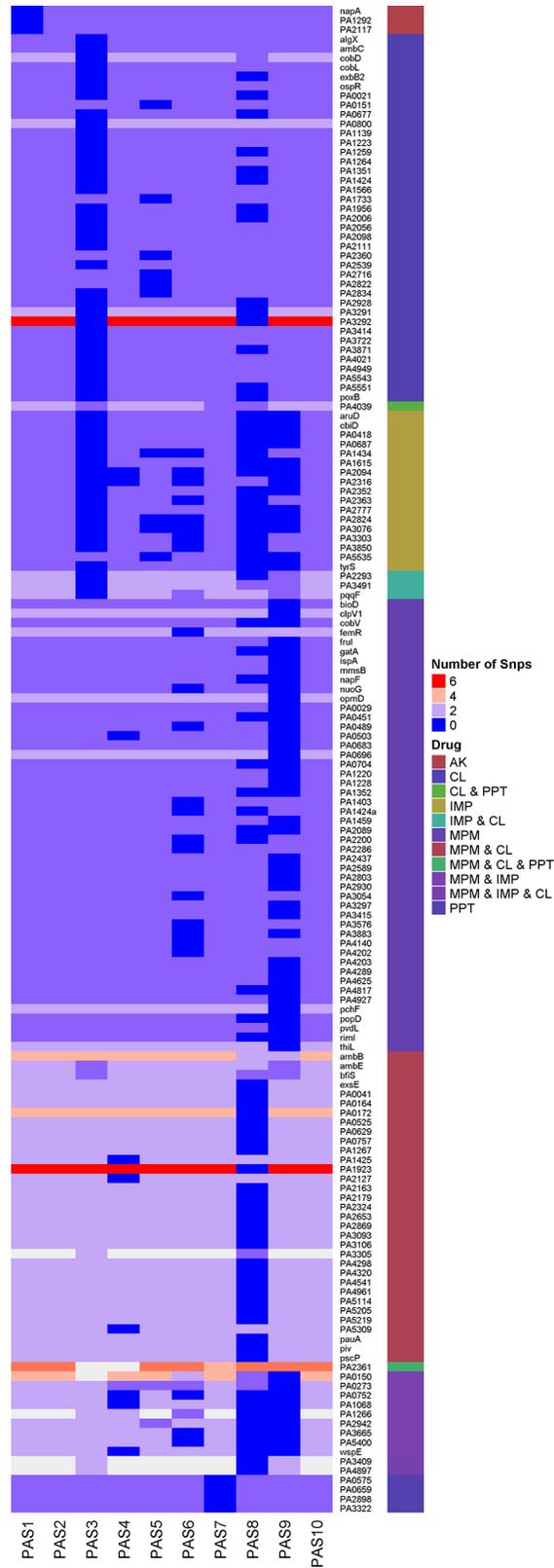
\* Indicates mutation that leads to a stop codon

<https://doi.org/10.1371/journal.pone.0182524.t004>

## Discussion

Whole genome sequencing promises high resolution, high-throughput genome sequencing with insights to the repertoire of genetic polymorphism [3, 11]. This approach can be used to identify the single nucleotide polymorphism that either alters a single amino acid or leads to a stop codon, and frame-shift mutations that alters a gene sequence itself resulting in truncated proteins. Here, we have investigated the genetic variation in *P. aeruginosa* hospital isolates by high-throughput NGS technology and comparative genomics.

The clinical isolates included in this study were selected based on their multidrug resistant profiles against a total of nine antibiotics. To study the genomic variation among the isolates with varying antibiotic resistance characteristics, we analyzed the genomic sequences of the isolates using WGS technique and compared the data with *P. aeruginosa* reference genome PAO1.



**Fig 3. Heat-map representing the number of non-synonymous SNPs found within different antibiotic-resistant clinical isolates of *P. aeruginosa*.** The number of SNP's identified in each clinical isolate with

respect to their antibiotic resistant profile is depicted in the heat-map. The red colour indicates there are 6 nsSNPs harbouring in that gene among different isolates and a blue colour represents there are no nsSNPs within the gene among isolates.

<https://doi.org/10.1371/journal.pone.0182524.g003>

Several non-synonymous SNPs that may play an important role in antibiotic resistance have been identified, with insights to enhance our current understanding of the factors that influence the antibiotic resistance regime of hospital isolates. Though we have studied the nsSNPs of entire genome of all hospital isolates, we specifically extracted nsSNPs that were in genes reported to play a major implication in antibiotic resistance [28] within the antibiotic resistant and susceptible clinical isolates. Several mechanisms of antibiotic resistance in *P. aeruginosa* have been reported: by altering porins or outer membrane properties, resistance through enzymatic modification of drugs, and resistance through efflux systems and chromosomal mutations within regulatory genes [30–33].

We have identified specific nsSNPs within the *mexXY-oprM* and *mexAB-oprM* genes, which are known to play a major role in antibiotic resistance through efflux pumps and both these genes part the same efflux porin OprM [34, 35]. In addition, the efflux system MexXY-OprM is also recognized to be contributing in aminoglycoside resistance in *P. aeruginosa*, and MexAB-OprM is reported to confer resistance to  $\beta$ -lactams, aminoglycosides and fluoroquinolones resistance [31, 36]. Also a polymorphism resulting in a stop codon at locus PA2020 corresponding to a *mexZ* transcriptional regulator may have implications in drug resistance [37]. Collectively, *P. aeruginosa* may use a combination of all these mechanisms to achieve multi-drug antibiotic resistance.

The unique nsSNPs identified within the genes coding for aminoglycoside modifying enzymes (AME's) and 16S rRNA methylase genes, mutations leading to colistin resistance by altering the bacterial outer membrane, and resistant to  $\beta$  lactam by hyper production of  $\beta$ -lactamase AmpC may together play a crucial role in multi-drug resistance. Our results confirm the assumption of other authors that the drug resistance may be the result of a pool of pathogenicity-related genes that interact in various combinations [29, 38, 39]. In addition to the unique mutations, there is high degree of sequence conservation between resistant and susceptible clinical isolates. This conservation within the virulence genes of different *P. aeruginosa* clinical isolates may be a factor that is challenging to devise strategies for the treatment of infections. In addition, the extensive conservation of virulence genes in the genomes regardless of drug resistance may be due to the fact that the disease-causing ability of this opportunistic pathogen relies on a set of highly conserved pathogenic mechanisms [29]. Moreover, the resistance of one isolate for several antibiotics may be due to the pleiotropic effects of resistance causing gene [39].

Drug resistance to several antibiotics may be due to a combination of mutations resulting in overexpression of several multi-drug efflux pumps, alteration of expression of enzymes and structural components involved in peptidoglycan outer membrane stability, mutations affecting gyrase activity and mutations taking effect in aminoglycoside phosphotransferases [29]. In our study, the analysis of the genomes of drug resistant clinical isolates possessed non-synonymous mutations in at least one gene known to be involved in antibiotic resistance (for example, *oprD*, *gyrA* or *B*, *mex*-type efflux systems, *parC*, *rpoB*, *baeS*). The genes carrying these mutations are coding for proteins involved in outer membrane permeability, multi- drug efflux pumps, gyrases and drug modifying enzymes (aminoglycoside modification enzymes and/ or  $\beta$ -lactamases) [39]. However, a limitation of this study is that further targeted evidence with site directed mutational analysis on the reported nsSNPs harboring the resistant genes is

required. Such analysis will provide further insights to the overall phenotype alteration and synergistic effect of these nsSNPs in the antibiotic resistant mechanism.

In conclusion, this study demonstrates the potential of next generation high-throughput whole genome sequencing to compare the genome polymorphism between clinical isolates. We have identified a number of key mutations that may play a key role in altering antibiotic resistant genes leading to genetic polymorphism. This should form the basis of our future research to study the prevalence of such multidrug resistance *P. aeruginosa* in the Asia Pacific region. This technology can be utilised in conjunction with current epidemiological studies, diagnostic assays and antimicrobial susceptibility tests to reveal the genetic variation and pathogen biology of “high-risk” pathogens including *P. aeruginosa* [11].

## Supporting information

**S1 Table. Non-synonymous SNP’s identified in Amikacin resistant isolates.** The Amikacin susceptible clinical isolate PAS1 was compared against all other resistant isolates. (DOCX)

**S2 Table. Non-synonymous SNP’s in Meropenem resistant isolates.** The Meropenem susceptible clinical isolates PAS 4, 6, 8 and 9 were compared against the rest of the resistant isolates. (DOCX)

**S3 Table. Non-synonymous SNP’s in Imipenem resistant isolates.** The Imipenem resistant isolates PAS 1, 2, 7 and 10 were compared against the rest of the susceptible isolates. (DOCX)

**S4 Table. Non-synonymous SNP’s in Colistin resistant isolates.** The Colistin susceptible isolates PAS 3, 4, 5 and 8 were compared against the rest of the resistant isolates. (DOCX)

**S5 Table. Non-synonymous SNP’s in Piperacillin/Tazobactam resistant isolates.** The Tazobactam susceptible isolate PAS 7 was compared against the rest of the resistant isolates. (DOCX)

## Author Contributions

**Conceptualization:** SDS CFL.

**Data curation:** BR CFL.

**Formal analysis:** BR HMJ RG JPJ RR.

**Funding acquisition:** RM SDS.

**Software:** AA.

**Writing – original draft:** BR SDS.

**Writing – review & editing:** BR SDS CFL.

## References

1. Fujitani S, Sun H-Y, Yu VL, Weingarten JA. Pneumonia due to pseudomonas aeruginosa: Part i: epidemiology, clinical diagnosis, and source. Chest. 2011; 139(4):909–19. <https://doi.org/10.1378/chest.10-0166> PMID: 21467058

2. Grisaru-Soen L L-G, N K, H B, JH P, A B. *Pseudomonas aeruginosa* bacteremia in children: analysis of trends in prevalence, antibiotic resistance and prognostic factors. *The Pediatric Infectious Disease Journal*. 2000; 19(10):959–63. PMID: [11055596](#)
3. Snyder LA, Loman NJ, Faraj LA, Levi K, Weinstock G, Boswell TC, et al. Epidemiological investigation of *Pseudomonas aeruginosa* isolates from a six-year-long hospital outbreak using high-throughput whole genome sequencing. *Euro surveillance: bulletin Europeen sur les maladies transmissibles = European communicable disease bulletin*. 2013; 18(42). Epub 2013/11/02. PMID: [24176582](#).
4. Nathwani D, Raman G, Sulham K, Gavaghan M, Menon V. Clinical and economic consequences of hospital-acquired resistant and multidrug-resistant *Pseudomonas aeruginosa* infections: a systematic review and meta-analysis. *Antimicrob Resist Infect Control*. 2014; 3(1):1–16. <https://doi.org/10.1186/2047-2994-3-32> PMID: [25371812](#)
5. Septimus EJ, Kuper KM. Clinical Challenges in Addressing Resistance to Antimicrobial Drugs in the Twenty-First Century. *Clin Pharmacol Ther*. 2009; 86(3):336–9. <https://doi.org/10.1038/clpt.2009.122> PMID: [19571803](#)
6. Aloush V, Navon-Venezia S, Seigman-Igra Y, Cabili S, Carmeli Y. Multidrug-resistant *Pseudomonas aeruginosa*: risk factors and clinical impact. *Antimicrobial agents and chemotherapy*. 2006; 50(1):43–8. Epub 2005/12/27. <https://doi.org/10.1128/AAC.50.1.43-48.2006> PMID: [16377665](#); PubMed Central PMCID: [PMCPmc1346794](#).
7. Lye DC, Earnest A, Ling ML, Lee TE, Yong HC, Fisher DA, et al. The impact of multidrug resistance in healthcare-associated and nosocomial Gram-negative bacteraemia on mortality and length of stay: cohort study. *Clinical microbiology and infection: the official publication of the European Society of Clinical Microbiology and Infectious Diseases*. 2012; 18(5):502–8. Epub 2011/08/20. <https://doi.org/10.1111/j.1469-0691.2011.03606.x> PMID: [21851482](#).
8. Lister PD, Wolter DJ, Hanson ND. Antibacterial-resistant *Pseudomonas aeruginosa*: clinical impact and complex regulation of chromosomally encoded resistance mechanisms. *Clinical microbiology reviews*. 2009; 22(4):582–610. Epub 2009/10/14. <https://doi.org/10.1128/CMR.00040-09> PMID: [19822890](#); PubMed Central PMCID: [PMCPmc2772362](#).
9. Turton JF, Turton SE, Yearwood L, Yarde S, Kaufmann ME, Pitt TL. Evaluation of a nine-locus variable-number tandem-repeat scheme for typing of *Pseudomonas aeruginosa*. *Clinical microbiology and infection: the official publication of the European Society of Clinical Microbiology and Infectious Diseases*. 2010; 16(8):1111–6. Epub 2009/09/08. <https://doi.org/10.1111/j.1469-0691.2009.03049.x> PMID: [19732093](#).
10. Martin K, Baddal B, Mustafa N, Perry C, Underwood A, Constantidou C, et al. Clusters of genetically similar isolates of *Pseudomonas aeruginosa* from multiple hospitals in the UK. *Journal of medical microbiology*. 2013; 62(Pt 7):988–1000. Epub 2013/04/06. <https://doi.org/10.1099/jmm.0.054841-0> PMID: [23558134](#).
11. Robinson ER, Walker TM, Pallen MJ. Genomics and outbreak investigation: from sequence to consequence. *Genome medicine*. 2013; 5(4):36. Epub 2013/05/16. <https://doi.org/10.1186/gm440> PMID: [23673226](#); PubMed Central PMCID: [PMCPmc3706975](#).
12. Aguilar-Rodea P, Zuniga G, Rodriguez-Espino BA, Olivares Cervantes AL, Gamino Arroyo AE, Moreno-Espinosa S, et al. Identification of extensive drug resistant *Pseudomonas aeruginosa* strains: New clone ST1725 and high-risk clone ST233. *PloS one*. 2017; 12(3):e0172882. Epub 2017/03/03. <https://doi.org/10.1371/journal.pone.0172882> PMID: [28253282](#); PubMed Central PMCID: [PMCPMC5333833](#).
13. Metzker ML. Sequencing technologies—the next generation. *Nature reviews Genetics*. 2010; 11(1):31–46. Epub 2009/12/10. <https://doi.org/10.1038/nrg2626> PMID: [19997069](#).
14. Sabat AJ, Budimir A, Nashev D, Sa-Leao R, van Dijk J, Laurent F, et al. Overview of molecular typing methods for outbreak detection and epidemiological surveillance. *Euro surveillance: bulletin Europeen sur les maladies transmissibles = European communicable disease bulletin*. 2013; 18(4):20380. Epub 2013/02/02. PMID: [23369389](#).
15. Lewis T, Loman NJ, Bingle L, Jumaa P, Weinstock GM, Mortiboy D, et al. High-throughput whole-genome sequencing to dissect the epidemiology of *Acinetobacter baumannii* isolates from a hospital outbreak. *The Journal of hospital infection*. 2010; 75(1):37–41. Epub 2010/03/20. <https://doi.org/10.1016/j.jhin.2010.01.012> PMID: [20299126](#).
16. Octavia S, Lan R. Single-nucleotide-polymorphism typing and genetic relationships of *Salmonella enterica* serovar Typhi isolates. *Journal of clinical microbiology*. 2007; 45(11):3795–801. Epub 2007/08/31. <https://doi.org/10.1128/JCM.00720-07> PMID: [17728466](#); PubMed Central PMCID: [PMCPmc2168493](#).
17. Bragonzi A, Paroni M, Nonis A, Cramer N, Montanari S, Rejman J, et al. *Pseudomonas aeruginosa* microevolution during cystic fibrosis lung infection establishes clones with adapted virulence. *American*

- journal of respiratory and critical care medicine. 2009; 180(2):138–45. Epub 2009/05/09. <https://doi.org/10.1164/rccm.200812-1943OC> PMID: 19423715.
18. Jelsbak L, Johansen HK, Frost AL, Thogersen R, Thomsen LE, Ciofu O, et al. Molecular epidemiology and dynamics of *Pseudomonas aeruginosa* populations in lungs of cystic fibrosis patients. *Infection and immunity*. 2007; 75(5):2214–24. Epub 2007/01/31. <https://doi.org/10.1128/IAI.01282-06> PMID: 17261614; PubMed Central PMCID: PMC1865789.
  19. Laliitha MK, Manayani DJ, Priya L, Jesudason MV, Thomas K, Steinhoff MC. E test as an alternative to conventional MIC determination for surveillance of drug resistant *Streptococcus pneumoniae*. *The Indian journal of medical research*. 1997; 106:500–3. Epub 1998/01/24. PMID: 9439095.
  20. Schmieder R, Edwards R. Quality control and preprocessing of metagenomic datasets. *Bioinformatics (Oxford, England)*. 2011; 27(6):863–4. Epub 2011/02/01. <https://doi.org/10.1093/bioinformatics/btr026> PMID: 21278185; PubMed Central PMCID: PMC3051327.
  21. 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. *Journal of computational biology: a journal of computational molecular cell biology*. 2012; 19(5):455–77. Epub 2012/04/18. <https://doi.org/10.1089/cmb.2012.0021> PMID: 22506599; PubMed Central PMCID: PMC3342519.
  22. 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; 9:75. Epub 2008/02/12. <https://doi.org/10.1186/1471-2164-9-75> PMID: 18261238; PubMed Central PMCID: PMC2265698.
  23. Seemann T. Prokka: rapid prokaryotic genome annotation. *Bioinformatics (Oxford, England)*. 2014; 30(14):2068–9. Epub 2014/03/20. <https://doi.org/10.1093/bioinformatics/btu153> PMID: 24642063.
  24. Langmead B, Trapnell C, Pop M, Salzberg SL. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome biology*. 2009; 10(3):R25. Epub 2009/03/06. <https://doi.org/10.1186/gb-2009-10-3-r25> PMID: 19261174; PubMed Central PMCID: PMC2690996.
  25. Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, et al. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic acids research*. 1997; 25(17):3389–402. Epub 1997/09/01. PMID: 9254694; PubMed Central PMCID: PMC146917.
  26. Alikhan NF, Petty NK, Ben Zakour NL, Beatson SA. BLAST Ring Image Generator (BRIG): simple prokaryote genome comparisons. *BMC genomics*. 2011; 12:402. Epub 2011/08/10. <https://doi.org/10.1186/1471-2164-12-402> PMID: 21824423; PubMed Central PMCID: PMC3163573.
  27. Gu Z, Eils R, Schlesner M. Complex heatmaps reveal patterns and correlations in multidimensional genomic data. *Bioinformatics (Oxford, England)*. 2016. <https://doi.org/10.1093/bioinformatics/btw313> PMID: 27207943
  28. Chen L, Xiong Z, Sun L, Yang J, Jin Q. VFDB 2012 update: toward the genetic diversity and molecular evolution of bacterial virulence factors. *Nucleic acids research*. 2012; 40(Database issue):D641–5. Epub 2011/11/10. <https://doi.org/10.1093/nar/gkr989> PMID: 22067448; PubMed Central PMCID: PMC3245122.
  29. Jansen G, Mahrt N, Tueffers L, Barbosa C, Harjes M, Adolph G, et al. Association between clinical antibiotic resistance and susceptibility of *Pseudomonas* in the cystic fibrosis lung. *Evolution, Medicine, and Public Health*. 2016. <https://doi.org/10.1093/emph/eow016> PMID: 27193199
  30. Gooderham WJ, Hancock RE. Regulation of virulence and antibiotic resistance by two-component regulatory systems in *Pseudomonas aeruginosa*. *FEMS microbiology reviews*. 2009; 33(2):279–94. Epub 2009/02/27. <https://doi.org/10.1111/j.1574-6976.2008.00135.x> PMID: 19243444.
  31. Alekshun MN, Levy SB. Molecular mechanisms of antibacterial multidrug resistance. *Cell*. 2007; 128(6):1037–50. Epub 2007/03/27. <https://doi.org/10.1016/j.cell.2007.03.004> PMID: 17382878.
  32. Poole K. Efflux-mediated antimicrobial resistance. *The Journal of antimicrobial chemotherapy*. 2005; 56(1):20–51. Epub 2005/05/26. <https://doi.org/10.1093/jac/dki171> PMID: 15914491.
  33. Poole K. Outer membranes and efflux: the path to multidrug resistance in Gram-negative bacteria. *Current pharmaceutical biotechnology*. 2002; 3(2):77–98. Epub 2002/05/23. PMID: 12022261.
  34. Hocquet D, Vogne C, El Garch F, Vejux A, Gotoh N, Lee A, et al. MexXY-OprM efflux pump is necessary for a adaptive resistance of *Pseudomonas aeruginosa* to aminoglycosides. *Antimicrobial agents and chemotherapy*. 2003; 47(4):1371–5. Epub 2003/03/26. PubMed Central PMCID: PMC152483. <https://doi.org/10.1128/AAC.47.4.1371-1375.2003> PMID: 12654672
  35. Vogne C, Aires JR, Bailly C, Hocquet D, Plesiat P. Role of the multidrug efflux system MexXY in the emergence of moderate resistance to aminoglycosides among *Pseudomonas aeruginosa* isolates from patients with cystic fibrosis. *Antimicrobial agents and chemotherapy*. 2004; 48(5):1676–80. Epub 2004/04/24. PubMed Central PMCID: PMC400545. <https://doi.org/10.1128/AAC.48.5.1676-1680.2004> PMID: 15105120

36. Li XZ, Poole K, Nikaïdo H. Contributions of MexAB-OprM and an EmrE homolog to intrinsic resistance of *Pseudomonas aeruginosa* to aminoglycosides and dyes. *Antimicrobial agents and chemotherapy*. 2003; 47(1):27–33. Epub 2002/12/25. PubMed Central PMCID: PMCPmc149025. <https://doi.org/10.1128/AAC.47.1.27-33.2003> PMID: 12499164
37. Wei Q, Tarighi S, Dotsch A, Haussler S, Musken M, Wright VJ, et al. Phenotypic and genome-wide analysis of an antibiotic-resistant small colony variant (SCV) of *Pseudomonas aeruginosa*. *PloS one*. 2011; 6(12):e29276. Epub 2011/12/24. <https://doi.org/10.1371/journal.pone.0029276> PMID: 22195037; PubMed Central PMCID: PMCPmc3240657.
38. Sanchez D, Gomila M, Bennasar A, Lallucat J, Garcia-Valdes E. Genome analysis of environmental and clinical *P. aeruginosa* isolates from sequence type-1146. *PloS one*. 2014; 9(10):e107754. Epub 2014/10/21. <https://doi.org/10.1371/journal.pone.0107754> PMID: 25329302; PubMed Central PMCID: PMCPMC4198096.
39. 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. Epub 2014/04/30. <https://doi.org/10.1186/1471-2164-15-318> PMID: 24773920; PubMed Central PMCID: PMCPMC4234422.