



# Complete Genome Sequence of *Pseudomonas aeruginosa* Reference Strain PAK

 Amy K. Cain,<sup>a,b</sup>  Laura M. Nolan,<sup>c</sup>  Geraldine J. Sullivan,<sup>b</sup>  Cynthia B. Whitchurch,<sup>d</sup>  Alain Filloux,<sup>c</sup>  Julian Parkhill<sup>a,e</sup>

<sup>a</sup>Wellcome Sanger Institute, Wellcome Trust Genome Campus, Hinxton, Cambridge, United Kingdom

<sup>b</sup>Department of Molecular Sciences, Macquarie University, North Ryde, Australia

<sup>c</sup>Imperial College London, MRC Centre for Molecular Bacteriology and Infection, South Kensington, London, United Kingdom

<sup>d</sup>The ithree institute, Faculty of Science, University of Technology Sydney, Ultimo, Australia

<sup>e</sup>Department of Veterinary Medicine, University of Cambridge, Cambridge, United Kingdom

**ABSTRACT** We report the complete genome of *Pseudomonas aeruginosa* strain PAK, a strain which has been instrumental in the study of a range of *P. aeruginosa* virulence and pathogenesis factors and has been used for over 50 years as a laboratory reference strain.

*Pseudomonas aeruginosa* is one of the deadly and highly drug-resistant, hospital-associated ESKAPE pathogens (*Enterococcus faecium*, *Staphylococcus aureus*, *Klebsiella pneumoniae*, *Acinetobacter baumannii*, *Pseudomonas aeruginosa*, and *Enterobacter* species), and it has been flagged by the WHO as a species of “greatest public health concern.” *P. aeruginosa* strain K (PAK) is one of the most well-studied *P. aeruginosa* laboratory strains, first described for its Pf1 phage sensitivity in 1966 (1) and its hyperpiliation in 1972 (2), and it is still in use today. This virulent exemplar strain highly expresses pili and flagella and contains glycosylation and pathogenicity islands (3, 4). Although over 500 publications involving the PAK strain exist to date (PubMed database), a long-read, closed, and high-quality genome annotation sequence is lacking.

Whole-genomic PAK DNA was sourced from the Filloux laboratory collection at Imperial College London, which originally was kindly gifted by Stephen Lory at Harvard Medical School. The strain was revived from storage in glycerol at  $-80^{\circ}\text{C}$  by streaking the bacteria onto an LB agar plate and growing it overnight. A single colony was then selected for overnight growth at  $37^{\circ}\text{C}$  in LB broth, with continuous shaking. The DNA was prepared using a phenol-chloroform method (5), and  $2\ \mu\text{g}$  was sequenced using the long-read Pacific Biosciences RS II sequencing platform (PacBio, USA). Genomic DNA (gDNA) was needle sheared to 26,522 bp (4 complete passes using a 26-gauge 2-in. blunt-ended needle), and the library was sequenced using the manufacturer’s protocol with the P4-C2 chemistry kit on 5 single-molecule real-time (SMRT) cells, producing 126,553 total reads and an average read length of 3,494 bp. *De novo* assembly of these reads was performed using the Hierarchical Genome Assembly Process 3 (HGAP3), on the SMRT Analysis pipeline 2.2.0, into a single contig with  $67.2\times$  coverage, which was linearized at *dnaA* using Circlator (6).

The PAK chromosome consists of 6,395,872 bp and contains 5,757 coding DNA sequences (CDSs) and 111 RNAs, with an average GC content of 66.44%. Default parameters were used for all software unless otherwise specified. Using the *Pseudomonas* multilocus sequence type (MLST) database (7), PAK was categorized as sequence type 693 (ST693). *De novo* annotation was performed using PROKKA 1.12 (8), followed by a secondary annotation transfer from reference strain PAO1 with RATT (9) using a cutoff of 90% similarity. Comparisons of PAK with reference strain PAO1 revealed 44,127 unique single nucleotide polymorphisms (SNPs), determined using SNP-calling

**Citation** Cain AK, Nolan LM, Sullivan GJ, Whitchurch CB, Filloux A, Parkhill J. 2019. Complete genome sequence of *Pseudomonas aeruginosa* reference strain PAK. *Microbiol Resour Announc* 8:e00865-19. <https://doi.org/10.1128/MRA.00865-19>.

**Editor** David A. Baltrus, University of Arizona

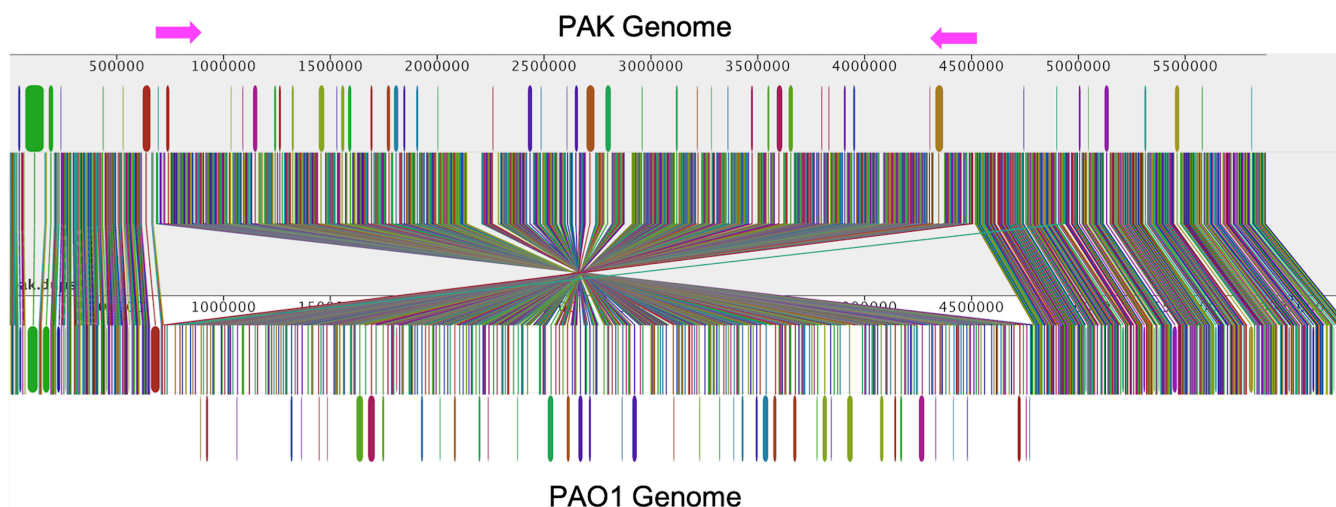
**Copyright** © 2019 Cain et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

Address correspondence to Alain Filloux, [a.filloux@imperial.ac.uk](mailto:a.filloux@imperial.ac.uk).

**Received** 29 July 2019

**Accepted** 9 September 2019

**Published** 10 October 2019



**FIG 1** Comparison of the *Pseudomonas aeruginosa* strain PAK genome and reference strain PAO1 genome. Sequences were linearized at position 1 of PAK and visualized using Mauve (13). Orthologous regions between the genomes are indicated by the locally colinear blocks matched by vertical lines. Inverted magenta arrows mark the potential rRNA inversion point (arrows not to scale).

methods as previously described (10), and identified 5,537 orthologous genes, as defined by BLASTN (11). We identified a large inversion of 4,185,263 bp between the strains using ACT (12), and visualized it using Mauve (13) (Fig. 1). Two sets of complementary rRNA genes of 5,752 bp with 99.63% sequence identity flank the inversion site.

Further manual annotation was performed. Five acquired resistance genes were identified using ResFinder 2.0 (14), all of which are also in PAO1, as follows: *aph(3')-IIb* (aminoglycoside), *ampC*, *bla<sub>OXA-50</sub>* (beta-lactamase), *catB7* (chloramphenicol), and *fosA* (fosfomycin). PHASTER (15) was used to identify the bacteriophage Pf1. Interestingly, Pf1 has long been known to be quite specific for PAK, yet Pf1 does not always integrate into the PAK chromosome (16). IslandViewer 4 (17) found 8 further genomic islands. The type IV pilus genes *pilA* to *pilD*, *pilF*, *pilG* to *pilK*, *chpA* to *chpC*, *fimL*, *fimS*, *algR*, *fimT*, *fimU*, *pilV* to *pilX*, *pilY1*, *pilY2*, *pilE*, and *pilM* to *pilQ* and the 3 type VI secretion system (T6SS) clusters, H1, H2, and H3 were identified. The H1 cluster spanned from *tagQ1* to PA0101, the H2 cluster from *tssA2* to *stk1*, and the H3 cluster from *sfa3* to PA2375.

This high-quality genome sequence will advance future understanding of this pathogen and help integrate almost 50 years of research on PAK into a modern genomic context.

**Data availability.** This whole-genome sequence has been deposited in GenBank under accession number [LR657304](https://doi.org/10.1093/genbank/RRY001), with the raw PacBio reads accessible under European Nucleotide Archive (ENA) sample accession number [ERS484051](https://doi.org/10.1093/ena/ERS484051).

## ACKNOWLEDGMENTS

This work was supported by the Wellcome Trust (grant number WT098051). A.K.C. was supported by an Australian Research Council (ARC) Discovery Early Career Researcher Award (DECRA) fellowship (DE180100929); L.M.N. was supported by MRC grant number MR/N023250/1 and a Marie Curie fellowship (PIIF-GA-2013-625318). A.F. was supported by MRC grant numbers MR/K001930/1 and MR/N023250/1 and Biotechnology and Biological Sciences Research Council (BBSRC) grant number BB/N002539/1.

We thank Thomas Otto at the Wellcome Sanger Institute for helpful discussions of annotation transfer.

## REFERENCES

1. Takeya K, Amako K. 1966. A rod-shaped *Pseudomonas* phage. *Virology* 28:163–165. [https://doi.org/10.1016/0042-6822\(66\)90317-5](https://doi.org/10.1016/0042-6822(66)90317-5).
2. Bradley DE. 1972. A study of pili on *Pseudomonas aeruginosa*. *Genet Res* 19:39–51. <https://doi.org/10.1017/S0016672300014257>.

3. Arora SK, Bangera M, Lory S, Ramphal R. 2001. A genomic island in *Pseudomonas aeruginosa* carries the determinants of flagellin glycosylation. *Proc Natl Acad Sci U S A* 98:9342–9347. <https://doi.org/10.1073/pnas.161249198>.
4. Vallet I, Olson JW, Lory S, Lazdunski A, Filloux A. 2001. The chaperone/usher pathways of *Pseudomonas aeruginosa*: identification of fimbrial gene clusters (cup) and their involvement in biofilm formation. *Proc Natl Acad Sci U S A* 98:6911–6916. <https://doi.org/10.1073/pnas.111551898>.
5. Sambrook J, Fritsch EF, Maniatis T. 1989. *Molecular cloning: a laboratory manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
6. Hunt M, De Silva N, Otto TD, Parkhill J, Keane JA, Harris SR. 2015. Circlator: automated circularization of genome assemblies using long sequencing reads. *Genome Biol* 16:294. <https://doi.org/10.1186/s13059-015-0849-0>.
7. Curran B, Jonas D, Grundmann H, Pitt T, Dowson CG. 2004. Development of a multilocus sequence typing scheme for the opportunistic pathogen *Pseudomonas aeruginosa*. *J Clin Microbiol* 42:5644–5649. <https://doi.org/10.1128/JCM.42.12.5644-5649.2004>.
8. Seemann T. 2014. Prokka: rapid prokaryotic genome annotation. *Bioinformatics* 30:2068–2069. <https://doi.org/10.1093/bioinformatics/btu153>.
9. Otto TD, Dillon GP, Degraeve WS, Berriman M. 2011. RATT: rapid annotation transfer tool. *Nucleic Acids Res* 39:e57. <https://doi.org/10.1093/nar/gkq1268>.
10. Harris SR, Cartwright EJP, Török ME, Holden MTG, Brown NM, Ogilvy-Stuart AL, Ellington MJ, Quail MA, Bentley SD, Parkhill J, Peacock SJ. 2013. Whole-genome sequencing for analysis of an outbreak of methicillin-resistant *Staphylococcus aureus*: a descriptive study. *Lancet Infect Dis* 13:130–136. [https://doi.org/10.1016/S1473-3099\(12\)70268-2](https://doi.org/10.1016/S1473-3099(12)70268-2).
11. Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, Madden TL. 2009. BLAST+: architecture and applications. *BMC Bioinformatics* 10:421. <https://doi.org/10.1186/1471-2105-10-421>.
12. Carver TJ, Rutherford KM, Berriman M, Rajandream M-A, Barrell BG, Parkhill J. 2005. ACT: the Artemis comparison tool. *Bioinformatics* 21:3422–3423. <https://doi.org/10.1093/bioinformatics/bti553>.
13. Darling AC, Mau B, Blattner FR, Perna NT. 2004. Mauve: multiple alignment of conserved genomic sequence with rearrangements. *Genome Res* 14:1394–1403. <https://doi.org/10.1101/gr.2289704>.
14. Zankari E, Hasman H, Cosentino S, Vestergaard M, Rasmussen S, Lund O, Aarestrup FM, Larsen MV. 2012. Identification of acquired antimicrobial resistance genes. *J Antimicrob Chemother* 67:2640–2644. <https://doi.org/10.1093/jac/dks261>.
15. Arndt D, Grant JR, Marcu A, Sajed T, Pon A, Liang Y, Wishart DS. 2016. PHASTER: a better, faster version of the PHAST phage search tool. *Nucleic Acids Res* 44:W16–W21. <https://doi.org/10.1093/nar/gkw387>.
16. Secor PR, Sweere JM, Michaels LA, Malkovskiy AV, Lazzareschi D, Katznelson E, Rajadas J, Birnbaum ME, Arrigoni A, Braun KR, Evanko SP, Stevens DA, Kaminsky W, Singh PK, Parks WC, Bollyky PL. 2015. Filamentous bacteriophage [*sic*] promote biofilm assembly and function. *Cell Host Microbe* 18:549–559. <https://doi.org/10.1016/j.chom.2015.10.013>.
17. Bertelli C, Laird MR, Williams KP, Simon Fraser University Research Computing Group, Lau BY, Hoard G, Winsor GL, Brinkman FSL. 2017. IslandViewer 4: expanded prediction of genomic islands for larger-scale datasets. *Nucleic Acids Res* 45:W30–W35. <https://doi.org/10.1093/nar/gkx343>.