Draft Genome Sequences of 10 Environmental *Pseudomonas aeruginosa* Strains Isolated from Soils, Sediments, and Waters

Shailab D. Shrestha,[a] David S. Guttman,[b,c] Gabriel G. Perron[a]

Reem-Kayden Center for Science and Computation, Department of Biology, Bard College, Annandale-on-Hudson, New York, USA[a]; Department of Cell and Systems Biology, University of Toronto, Toronto, Ontario, Canada[b]; Centre for the Analysis of Genome Evolution and Function, University of Toronto, Toronto, Ontario, Canada[c]

ABSTRACT *Pseudomonas aeruginosa* is an important opportunistic pathogen that has the ability to grow in a range of environmental niches. Here, we report the draft genome sequences of 10 environmental strains of the bacterium isolated from soils, sediments, and waters in various locations in North America and South Africa.

Address correspondence to David S. Guttman, guttman@utoronto.ca, or Gabriel G. Perron, gperron@bard.edu.

D.S.G. and G.G.P. contributed equally to this article.

*P*seudomonas aeruginosa is an important opportunistic pathogen often associated with multiple-drug resistance in hospital environments (1). The bacterium is especially problematic in immunocompromised individuals, such as patients with cystic fibrosis, in whom it causes severe health problems (2). In addition to human hosts, the bacterium has the ability to grow in a range of ecological niches, including plant and animal tissues of all kinds, as well as soils and aquatic environments (3). Even though *P. aeruginosa* demonstrates extensive genetic and phenotypic diversity in ecological niches (4, 5), the majority of whole-genome sequencing (WGS) studies considering the organism focus on medically important strains (6–9).

To investigate the genomic diversity of *P. aeruginosa* in the environment, we obtained the whole-genome sequences of 10 *P. aeruginosa* strains isolated from different locations across North America and South Africa. Strains Pae85, Pae100, Pae102, and Pae110 were isolated from soils and sediments at various sites around the greater Toronto area in Ontario, Canada, while Pae111, Pae112, and Pae113 were isolated from farm soil samples taken outside Maysville, KY. Strain Pae160 was isolated from a freshwater pond on a golf course in South Africa. Strains PaEB1 and PaEB6 were isolated from water sediment collected at two sites along the Saw Kill, a tributary of the Hudson River in Dutchess County, NY. The two strains were extracted from positive Pseudalert test kits (Idexx Laboratories, Westbrook, MA).

Isolated bacterial cultures were grown in Luria-Bertani broth overnight at 37°C in a shaking incubator. Cells were harvested and genomic DNA was extracted using the DNeasy UltraClean microbial kit (Qiagen, Carlsbad, CA). For the Pae strains, WGS was performed on a MiSeq platform at the University of Toronto Centre for the Analysis of Genome Evolution and Function (Toronto, Canada), according to the manufacturer's protocol, to generate 250-bp paired-end reads. For the PaEB strains, WGS was performed on a HiSeq platform by Wright Labs at Juniata College (Huntingdon, PA), according to the manufacturer's protocol, to generate 150-bp paired-end reads. For every strain, adapters were trimmed using Cutadapt version 1.14 (10) and quality filtered using Trimmomatic version 0.36 (11). Draft assemblies were conducted using SPAdes version 3.10.0 (12), with k-mer sizes of 21, 33, 55, 77, 99, and 127. Contigs shorter than 500 bp or with fewer than four reads were removed from the assembly. Finally, assembly improvement was undertaken as described by Page et al. (13), where

TABLE 1 Summary of the draft genome sequences for *Pseudomonas aeruginosa* environmental strains

Isolate	Origin	Location	No. of contigs	Genome size (bp)	G+C content (%)	N_{50} (bp)	Median read depth (×)	GenBank accession no.
Pae85	Soil	King City, Ontario, Canada	100	6,360,006	65.95	140,050	29	NKYD00000000
Pae100	Soil	King City, Ontario, Canada	74	6,436,962	66.42	196,319	26	NKYC00000000
Pae102	Soil	King City, Ontario, Canada	90	6,437,018	66.42	121,945	20	NKYB00000000
Pae110	Sediment	Toronto, Ontario, Canada	75	6,438,971	66.42	234,265	29	NKYA00000000
Pae111	Soil	Maysville, KY	71	6,385,725	66.47	206,050	22	NKXZ00000000
Pae112	Soil	Maysville, KY	68	6,444,483	66.39	231,479	40	NKXY00000000
Pae113	Sediment	Maysville, KY	65	6,393,508	66.47	245,535	45	NKXX00000000
Pae160	Water	South Africa	98	6,836,824	66.01	188,795	30	NKXW00000000
PaEB1	Water	Annandale, NY	65	6,336,070	66.50	735,709	337	NKXV00000000
PaEB6	Water	Annandale, NY	93	6,467,958	66.41	489,845	274	NKXU00000000

contigs were scaffolded using SSPACE (14) and sequence gaps were filled using GapFiller (15). The assembly with the largest N_{50} value was used for further analysis.

The average number of contigs per genome was 80, with a standard deviation of 13.9. The draft genomes ranged in size from 6,336,070 bp to 6,836,824 bp, with an average G+C content of 66.3% (Table 1). The N_{50} of the draft genomes ranged from 121,945 bp to 735,709 bp, with an average of medians of 278,999.2 bp and an average median coverage of 85.2× (Table 1).

Accession number(s). The whole-genome shotgun projects have been deposited at DDBJ/ENA/GenBank (see Table 1). The versions described in this paper are NKYD01000000 (Pae85), NKYC01000000 (Pae100), NKYB01000000 (Pae102), NKYA01000000 (Pae110), NKXZ01000000 (Pae111), NKXY01000000 (Pae112), NKXX01000000 (Pae113), NKXW01000000 (Pae160), NKXV01000000 (PaEB1), and NKXU01000000 (PaEB6).

REFERENCES

1. Livermore DM. 2002. Multiple mechanisms of antimicrobial resistance in *Pseudomonas aeruginosa*: our worst nightmare? Clin Infect Dis 34:634–640. https://doi.org/10.1086/338782.
2. Winstanley C, O'Brien S, Brockhurst MA. 2016. *Pseudomonas aeruginosa* evolutionary adaptation and diversification in cystic fibrosis chronic lung infections. Trends Microbiol 24:327–337. https://doi.org/10.1016/j.tim.2016.01.008.
3. Pirnay JP, Bilocq F, Pot B, Cornelis P, Zizi M, Van Eldere J, Deschaght P, Vaneechoutte M, Jennes S, Pitt T, De Vos D. 2009. *Pseudomonas aeruginosa* population structure revisited. PLoS One 4:e7740. https://doi.org/10.1371/journal.pone.0007740.
4. Pirnay JP, Matthijs S, Colak H, Chablain P, Bilocq F, Van Eldere J, De Vos D, Zizi M, Triest L, Cornelis P. 2005. Global *Pseudomonas aeruginosa* biodiversity as reflected in a Belgian river. Environ Microbiol 7:969–980. https://doi.org/10.1111/j.1462-2920.2005.00776.x.
5. Selezska K, Kazmierczak M, Müsken M, Garbe J, Schobert M, Häussler S, Wiehlmann L, Rohde C, Sikorski J. 2012. *Pseudomonas aeruginosa* population structure revisited under environmental focus: impact of water quality and phage pressure. Environ Microbiol 14:1952–1967. https://doi.org/10.1111/j.1462-2920.2012.02719.x.
6. Dettman JR, Rodrigue N, Aaron SD, Kassen R. 2013. Evolutionary genomics of epidemic and nonepidemic strains of *Pseudomonas aeruginosa*. Proc Natl Acad Sci U S A 110:21065–21070. https://doi.org/10.1073/pnas.1307862110.
7. Dettman JR, Rodrigue N, Kassen R. 2014. Genome-wide patterns of recombination in the opportunistic human pathogen *Pseudomonas aeruginosa*. Genome Biol Evol 7:18–34. https://doi.org/10.1093/gbe/evu260.
8. Freschi L, Jeukens J, Kukavica-Ibrulj I, Boyle B, Dupont MJ, Laroche J, Larose S, Maaroufi H, Fothergill JL, Moore M, Winsor GL, Aaron SD, Barbeau J, Bell SC, Burns JL, Camara M, Cantin A, Charette SJ, Dewar K, Déziel É, Grimwood K, Hancock REW, Harrison JJ, Heeb S, Jelsbak L, Jia

B, Kenna DT, Kidd TJ, Klockgether J, Lam JS, Lamont IL, Lewenza S, Loman N, Malouin F, Manos J, McArthur AG, McKeown J, Milot J, Naghra H, Nguyen D, Pereira SK, Perron GG, Pirnay J-P, Rainey PB, Rousseau S, Santos PM, Stephenson A, Taylor V, Turton JF, Waglechner N, et al. 2015. Clinical utilization of genomics data produced by the international *Pseudomonas aeruginosa* consortium. Front Microbiol 6:1036. https://doi.org/10.3389/fmicb.2015.01036.
9. Stewart L, Ford A, Sangal V, Jeukens J, Boyle B, Kukavica-Ibrulj I, Caim S, Crossman L, Hoskisson PA, Levesque R, Tucker NP. 2014. Draft genomes of 12 host-adapted and environmental isolates of *Pseudomonas aeruginosa* and their positions in the core genome phylogeny. Pathog Dis 71:20–25. https://doi.org/10.1111/2049-632X.12107.
10. Martin M. 2011. Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet J 17:10–12. https://doi.org/10.14806/ej.17.1.200.
11. Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 30:2114–2120. https://doi.org/10.1093/bioinformatics/btu170.
12. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Prjibelski AD, Pyshkin AV, Sirotkin AV, Vyahhi N, Tesler G, Alekseyev MA, Pevzner PA. 2012. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. J Comput Biol 19:455–477. https://doi.org/10.1089/cmb.2012.0021.
13. Page AJ, De Silva N, Hunt M, Quail MA, Parkhill J, Harris SR, Otto TD, Keane JA. 2016. Robust high-throughput prokaryote de novo assembly and improvement pipeline for Illumina data. Microb Genom 2:e000083. https://doi.org/10.1099/mgen.0.000083.
14. Boetzer M, Henkel CV, Jansen HJ, Butler D, Pirovano W. 2011. Scaffolding pre-assembled contigs using SSPACE. Bioinformatics 27:578–579. https://doi.org/10.1093/bioinformatics/btq683.
15. Boetzer M, Pirovano W. 2012. Toward almost closed genomes with Gap-Filler. Genome Biol 13:R56. https://doi.org/10.1186/gb-2012-13-6-r56.