AMERICAN SOCIETY FOR MICROBIOLOGY | genomeAnnouncements

Draft Genome Sequences of Five *Pseudomonas aeruginosa* Clinical Strains Isolated from Sputum Samples from Cystic Fibrosis Patients

M. B. Couger, Anna Wright, Erika I. Lutter, Noha Youssef

Department of Microbiology and Molecular Genetics, Oklahoma State University, Stillwater, Oklahoma, USA

We report here the draft genome sequences of five *Pseudomonas aeruginosa* isolates obtained from sputum samples from two cystic fibrosis patients with chronic colonization. These closely related strains harbor 225 to 493 genes absent from the *P. aeruginosa* POA1 genome and contain 178 to 179 virulence factors and 29 to 31 antibiotic resistance genes.

Received 4 November 2015 **Accepted** 9 November 2015 **Published** 28 January 2016

Citation Couger MB, Wright A, Lutter EI, Youssef N. 2016. Draft genome sequences of five *Pseudomonas aeruginosa* clinical strains isolated from sputum samples from cystic fibrosis patients. Genome Announc 4(1):e01528-15. doi:10.1128/genomeA.01528-15.

Copyright © 2016 Couger et al. This is an open-access article distributed under the terms of the Creative Commons Attribution 3.0 Unported license.

Address correspondence to Noha Youssef, noha@okstate.edu.

The genomes of five *Pseudomonas aeruginosa* isolates (strains 14649, 14650, 14651, 14672, and 14673), obtained at the Adult Cystic Fibrosis Clinic in Calgary, Alberta, Canada, from sputum samples from two cystic fibrosis patients with chronic colonization, were sequenced using the Illumina MiSeq platform with 300 × 2 paired-end chemistry and an average library insert size of 700 bp. Standard Illumina quality filtering settings were used for base calling. The sequence data from each clinical isolate were assembled using Velvet (1, 2), using a *k*-mer value of 101 bp and a minimum coverage of 7×. The gene models for each isolate were created using the prokaryotic gene calling software package Prodigal (3). All protein sequences from each isolate genome were functionally annotated using a combination of NCBI Blast C++ homology search (4) and HMMER 3.0 hmmscan (5, 6) against the Pfam 26.0 database (7).

Assembly N_{50} contig values of 489,659, 417,833, 308,331, 427,030, and 519,426 bp; genome sizes of 6.23, 6.37, 6.67, 6.36, and 6.36 Mbp; and G+C contents of 66.4%, 66.5%, 66.1%, 66.5%, and 66.5% were obtained for strains 14649, 14650, 14651, 14672, and 14673, respectively. A comparison of the genomes to the reference sequence of *P. aeruginosa* strain PAO1 (8) revealed strain-specific differences. The genomes of strains 14649, 14650, 14651, 14672, and 14673 contained 225, 229, 493, 230, and 227 genes, respectively, that were absent from the PAO1 genome. Conversely, 134, 94, 92, 102, and 107 genes were present in the reference genome PAO1 compared to the genomes of 14649, 14650, 14651, 14672, and 14673, respectively (see Table 1).

Protein models important to pathogenesis were identified using comparisons to the Virulence Factor Database (VFDB) (9), and antibiotic resistance genes were identified using the ARDB database (10). The genomes of strains 14649, 14650, 14651, 14672, and 14673 contained 178, 179, 179, 179, and 179 virulence factors and 31, 31, 29, 31, and 26 antibiotic resistance genes (see Table 1). These genes included those for β-lactamase, multidrug resistance efflux pumps (11), bacitracin resistance protein (12), and the efflux transporter system OprM (13).

Nucleotide sequence accession numbers. The GenBank accession numbers for the genomes of 14649, 14650, 14651, 14672, and 14673 are listed in Table 1.

TABLE 1 General genomic features of the five *P. aeruginosa* genomes sequenced in this study

	Results of genome assemblies			No. of genes[a]:		Results from comparisons to VFDB and ARDB		
P. aeruginosa strain	Genome size (Mb)	N_{50} contig value (bp)	G+C content (%)	Present in isolate genome, absent in reference genome	Absent in isolate genome, present in reference genome	No. of virulence factor genes	No. of antibiotic resistance genes	GenBank accession no.
14649	6.23	489,659	66.4	225	134	178	31	LKPS00000000
14650	6.37	471,833	66.5	229	94	179	31	LKPT00000000
14651	6.67	308,331	66.1	493	92	179	29	LKPU00000000
14672	6.36	427,030	66.5	230	102	179	31	LKPV00000000
14673	6.36	519,426	66.5	227	107	179	26	LKPW00000000

[a] Results of BLASTP against PAO1 reference genome.

FUNDING INFORMATION

Oklahoma Center for Respiratory and Infectious Diseases provided funding to Erika I. Lutter and Noha Youssef.

REFERENCES

1. **Compeau PEC, Pevzner PA, Tesler G**. 2011. How to apply de Bruijn graphs to genome assembly. Nat Biotechnol **29:**987–991. http://dx.doi.org/10.1038/nbt.2023.
2. **Zerbino DR, Birney E**. 2008. Velvet: algorithms for *de novo* short read assembly using de Bruijn graphs. Genome Res **18:**821–829. http://dx.doi.org/10.1101/gr.074492.107.
3. **Hyatt D, Chen G, Locascio PF, Land ML, Larimer FW, Hauser LJ**. 2010. Prodigal: prokaryotic gene recognition and translation initiation site identification. BMC Bioinformatics **11:**119. http://dx.doi.org/10.1186/1471-2105-11-119.
4. **Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, Madden TL**. 2009. BLAST+: architecture and applications. BMC Bioinformatics **10:**421. http://dx.doi.org/10.1186/1471-2105-10-421.
5. **Eddy SR**. 2011. Accelerated profile HMM Searches. PLoS Comput Biol **7:**e1002195. http://dx.doi.org/10.1371/journal.pcbi.1002195.
6. **Johnson LS, Eddy SR, Portugaly E**. 2010. Hidden Markov model speed heuristic and iterative HMM search procedure. BMC Bioinformatics **11:**431. http://dx.doi.org/10.1186/1471-2105-11-431.
7. **Finn RD, Bateman A, Clements J, Coggill P, Eberhardt RY, Eddy SR, Heger A, Hetherington K, Holm L, Mistry J, Sonnhammer ELL, Tate J, Punta M**. 2014. Pfam: the protein families database. Nucleic Acids Res **42:**D222–D230. http://dx.doi.org/10.1093/nar/gkt1223.
8. **Stover CK, Pham XQ, Erwin AL, Mizoguchi SD, Warrener P, Hickey MJ, Brinkman FS, Hufnagle WO, Kowalik DJ, Lagrou M, Garber RL, Goltry L, Tolentino E, Westbrock-Wadman S, Yuan Y, Brody LL, Coulter SN, Folger KR, Kas A, Larbig K, Lim R, Smith K, Spencer D, Wong GK, Wu Z, Paulsen IT, Reizer J, Saier MH, Hancock RE, Lory S, Olson MV**. 2000. Complete genome sequence of *Pseudomonas aeruginosa* PAO1, an opportunistic pathogen. Nature **406:**959–964. http://dx.doi.org/10.1038/35023079.
9. **Chen L, Xiong Z, Sun L, Yang J, Jin Q**. 2012. VFDB 2012 update: toward the genetic diversity and molecular evolution of bacterial virulence factors. Nucleic Acids Res **40:**D641–D645. http://dx.doi.org/10.1093/nar/gkr989.
10. **Liu B, Pop M**. 2009. ARDB—Antibiotic Resistance Genes Database. Nucleic Acids Res **37:**D443–D447. http://dx.doi.org/10.1093/nar/gkn656.
11. **Aeschlimann JR**. 2003. The role of multidrug efflux pumps in the antibiotic resistance of *Pseudomonas aeruginosa* and other Gram-negative bacteria. Insights from the Society of Infectious Diseases Pharmacists. Pharmacotherapy **23:**916–924.
12. **Morita Y, Tomida J, Kawamura Y**. 2012. Primary mechanisms mediating aminoglycoside resistance in the multidrug-resistant *Pseudomonas aeruginosa* clinical isolate PA7. Microbiology **158:**1071–1083. http://dx.doi.org/10.1099/mic.0.054320-0.
13. **Llanes C, Hocquet D, Vogne C, Benali-Baitich D, Neuwirth C, Plesiat P**. 2004. Clinical strains of *Pseudomonas aeruginosa* overproducing MexAB-OprM and MexXY efflux pumps simultaneously. Antimicrob Agents Chemother **48:**1797–1802. http://dx.doi.org/10.1128/AAC.48.5.1797-1802.2004.