Genome Sequences for Two *Acinetobacter baumannii* Strains Obtained Using the Unicycler Hybrid Assembly Pipeline

John M. Farrow III,[a] 🔟 Everett C. Pesci,[a] 🔟 Daniel J. Slade[b]

[a]Department of Microbiology and Immunology, The Brody School of Medicine at East Carolina University, Greenville, North Carolina, USA

[b]Department of Biochemistry, Virginia Polytechnic Institute and State University, Blacksburg, Virginia, USA

ABSTRACT Here, we report a complete genome sequence for *Acinetobacter baumannii* strain ATCC 17961, with plasmid sequences, and a high-quality (>98% complete) build for *A. baumannii* strain AB09-003. These genome sequences were generated by combining short-read Illumina and long-read Oxford Nanopore MinION sequencing data using the Unicycler hybrid assembly pipeline.

*A*cinetobacter baumannii is a Gram-negative, opportunistic bacterial pathogen that primarily causes infections in health care settings (1). This organism has become a top priority for the development of new treatments due to extremely high rates of antibiotic resistance worldwide (2). Comparing genome sequence data from both historic and current *A. baumannii* isolates may provide new insights into the emergence and increased antibiotic resistance of this pathogen. Here, we report the genome sequences of *A. baumannii* strain ATCC 17961 (CDC7788), a clinical isolate collected prior to 1964 (3, 4), and *A. baumannii* strain AB09-003, a military isolate collected from the cerebrospinal fluid of a wounded soldier and received at the University of Miami (Miami, FL) in 2009.

For DNA isolation, bacteria were grown in lysogeny broth (LB, Lennox formulation) for approximately 16 h at 37°C. DNA for short-read sequencing was purified through extractions with Tris-buffered phenol, 1:1 phenol-chloroform, and cetyltrimethylammonium bromide (CTAB) (5). Paired-end sequencing libraries were prepared with a TruSeq DNA library prep kit (Illumina) and were sequenced using the Illumina MiSeq system with v3 chemistry (2 × 300 bp) by the Genomic Sciences Laboratory at North Carolina State University, Raleigh, NC. Adapter sequences were removed using Trimmomatic v0.32 (6), and the reads were quality filtered and trimmed using the following settings: LEADING:5, SLIDINGWINDOW:4:15, CROP:295, MINLEN:40. The resulting Illumina data sets contained 1,609,266 read pairs with an average length of 293 bp (approximately 943 Mbp) for strain ATCC 17961 and 1,665,932 read pairs with an average length of 293 bp (approximately 976 Mbp) for strain AB09-003. DNA was isolated for MinION sequencing using a Wizard genomic DNA isolation kit (Promega). Sequencing libraries were prepared using a rapid sequencing kit (SQK-RAD004; Oxford Nanopore Technologies) and sequenced using a SpotON flow cell (R9.4) and a MinION sequencing device. Data acquisition was performed using MinKNOW v3.1.8 software, base calling was done using the Oxford Nanopore base caller Guppy v2.1.3 with default settings, and adapter sequences were removed using Porechop v0.2.4 with default settings (https://github.com/rrwick/Porechop). The final long-read data sets contained 74,740 reads at an average length of 6,799 bp for strain ATCC 17961 and 19,496 reads at an average length of 6,720 bp for strain AB09-003. Genome sequences for both strains were generated by combining data from both the Illumina and MinION data sets using Unicycler v0.4.3 (7), as described previously (8). After genome assembly, the contaminating sequences for Phi X174, a control for Illumina sequencing (9), were

collapsed into a single contig that was removed from the final assembly for each strain. The completed genome sequences were annotated using the NCBI Prokaryotic Genome Annotation Pipeline v5.0 (10).

For *A. baumannii* strain ATCC 17961, these procedures resulted in a fully resolved genome sequence consisting of a single circular chromosome of 4,016,386 bp and two plasmids, pAB17961-1 (9,395 bp) and pAB17961-2 (6,667 bp). The G+C content was 39.1%. For *A. baumannii* strain AB09-003, the resulting assembly consisted of 17 contigs with a total length of 3,952,207 bp (N_{50}, 3,902,515) and a G+C content of 39.1%. The majority of the sequence (98.7%) is present in a single contig of 3,902,513 bp, and the remaining 16 contigs contain approximately 50 kb that could not be fully resolved due to multiple repetitive genetic elements.

Data availability. The completed genome sequences and raw sequence data for *A. baumannii* strain ATCC 17961 are available through NCBI under BioProject accession no. PRJNA437968, GenBank accession no. CP065432 (chromosome sequence), CP065433 (plasmid pAB17961-1), and CP065434 (plasmid pAB17961-2), and Sequence Read Archive (SRA) accession no. SRX3783012 (Illumina raw sequence data) and SRX9401252 (MinION raw sequence data). The data for *A. baumannii* strain AB09-003 are available under BioProject accession no. PRJNA615908, GenBank accession no. JADPLY000000000 (genome sequence), and SRA accession no. SRX8013862 (Illumina raw sequence data) and SRX9408185 (MinION raw sequence data).

ACKNOWLEDGMENTS

REFERENCES

1. Wong D, Nielsen TB, Bonomo RA, Pantapalangkoor P, Luna B, Spellberg B. 2017. Clinical and pathophysiological overview of *Acinetobacter* infections: a century of challenges. Clin Microbiol Rev 30:409–447. https://doi.org/10.1128/CMR.00058-16.

2. Tacconelli E, Carrara E, Savoldi A, Harbarth S, Mendelson M, Monnet DL, Pulcini C, Kahlmeter G, Kluytmans J, Carmeli Y, Ouellette M, Outterson K, Patel J, Cavaleri M, Cox EM, Houchens CR, Grayson ML, Hansen P, Singh N, Theuretzbacher U, Magrini N, WHO Pathogens Priority List Working Group. 2018. Discovery, research, and development of new antibiotics: the WHO priority list of antibiotic-resistant bacteria and tuberculosis. Lancet Infect Dis 18:318–327. https://doi.org/10.1016/S1473-3099(17)30753-3.

3. Mitchell PD, Burrell RG. 1964. Serology of the *Mima-Herellea* group and the genus *Moraxella*. J Bacteriol 87:900–909. https://doi.org/10.1128/JB.87.4.900-909.1964.

4. Baumann P, Doudoroff M, Stanier RY. 1968. A study of the *Moraxella* group. II. Oxidative-negative species (genus *Acinetobacter*). J Bacteriol 95:1520–1541. https://doi.org/10.1128/JB.95.5.1520-1541.1968.

5. Wilson K. 1997. Preparation of genomic DNA from bacteria, p 2.4.1–2.4.5. *In* Ausubel FM, Brent R, Kingston RE, Moore DD, Seidman JG, Smith JA, Struhl K (ed), Current protocols in molecular biology, vol 1. John Wiley & Sons, Inc., Hoboken, NJ.

6. 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.

7. Wick RR, Judd LM, Gorrie CL, Holt KE. 2017. Unicycler: resolving bacterial genome assemblies from short and long sequencing reads. PLoS Comput Biol 13:e1005595. https://doi.org/10.1371/journal.pcbi.1005595.

8. Todd SM, Settlage RE, Lahmers KK, Slade DJ. 2018. *Fusobacterium* genomics using MinION and Illumina sequencing enables genome completion and correction. mSphere 3:e00269-18. https://doi.org/10.1128/mSphere.00269-18.

9. Mukherjee S, Huntemann M, Ivanova N, Kyrpides NC, Pati A. 2015. Large-scale contamination of microbial isolate genomes by Illumina PhiX control. Stand Genomic Sci 10:18. https://doi.org/10.1186/1944-3277-10-18.

10. Haft DH, DiCuccio M, Badretdin A, Brover V, Chetvernin V, O'Neill K, Li W, Chitsaz F, Derbyshire MK, Gonzales NR, Gwadz M, Lu F, Marchler GH, Song JS, Thanki N, Yamashita RA, Zheng C, Thibaud-Nissen F, Geer LY, Marchler-Bauer A, Pruitt KD. 2018. RefSeq: an update on prokaryotic genome annotation and curation. Nucleic Acids Res 46:D851–D860. https://doi.org/10.1093/nar/gkx1068.