genomeA
Journals.ASM.org

Draft Genome Assemblies of *Enterobacter aerogenes* CDC 6003-71, *Enterobacter cloacae* CDC 442-68, and *Pantoea agglomerans* UA 0804-01

T. D. Minogue,[a] H. E. Daligault,[b] K. W. Davenport,[b] K. A. Bishop-Lilly,[c,d] D. C. Bruce,[a] P. S. Chain,[b] S. R. Coyne,[a] O. Chertkov,[b] T. Freitas,[b]* K. G. Frey,[c,d] J. Jaissle,[b] G. I. Koroleva,[e] J. T. Ladner,[e] G. F. Palacios,[e] C. L. Redden,[c,d] Y. Xu,[b] S. L. Johnson[b]

Diagnostic Systems Division (DSD), United States Army Medical Research Institute of Infectious Diseases (USAMRIID), Fort Detrick, Maryland, USA[a]; Los Alamos National Laboratory (LANL), Los Alamos, New Mexico, USA[b]; Naval Medical Research Center (NMRC), Frederick, Fort Detrick, Maryland, USA[c]; Henry M. Jackson Foundation, Bethesda, Maryland, USA[d]; Center for Genome Sciences (CGS), USAMRIID, Fort Detrick, Maryland, USA[e]

* Present address: T. Freitas, Med Fusion, Lewisville, Texas, USA.

The *Enterobacteriaceae* are environmental and enteric microbes. We sequenced the genomes of two *Enterobacter* reference strains, *E. aerogenes* CDC 6003-71 and *E. cloacae* CDC 442-68, as well as one near neighbor used as an exclusionary reference for diagnostics, *Pantoea agglomerans* CDC UA0804-01. The genome sizes range from 4.72 to 5.55 Mbp and have G+C contents from 54.6 to 55.1%.

Received 10 September 2014 **Accepted** 17 September 2014 **Published** 23 October 2014

Citation Minogue TD, Daligault HE, Davenport KW, Bishop-Lilly KA, Bruce DC, Chain PS, Coyne SR, Chertkov O, Freitas T, Frey KG, Jaissle J, Koroleva GI, Ladner JT, Palacios GF, Redden CL, Xu Y, Johnson SL. 2014. Draft genome assemblies of *Enterobacter aerogenes* CDC 6003-71, *Enterobacter cloacae* CDC 442-68, and *Pantoea agglomerans* UA 0804-01. Genome Announc. 2(5):e01073-14. doi:10.1128/genomeA.01073-14.

Address correspondence to S. L. Johnson, shannonj@lanl.gov.

Members of the family *Enterobacteriaceae* are notable opportunistic pathogens, with some species accounting for up to 5% of documented hospital-acquired infections. *Enterobacteriaceae* includes both saprophytic soil and commensal enteric microbes. Many have become troublesome nosocomial pathogens, with high resistance to β-lactam antimicrobials (1). In this study, we sequenced two reference strains of *Enterobacter* (*E. aerogenes* ATCC 29904 and *E. cloacae* ATCC 13047), as well as a near neighbor and plant commensal, *Pantoea agglomerans* UA 0804-01 (1, 2).

High-quality genomic DNA was extracted from purified isolates of each strain using Qiagen Genomic-tip 500 at the United States Army Medical Research Institute of Infectious Diseases-Diagnostic Systems Division (USAMRIID-DSD). Specifically, 100-ml bacterial cultures were grown to stationary phase and nucleic acid extracted as per the manufacturer's recommendations. These draft genomes included a combination of Illumina (3) and 454 technologies (4). For each genome, we constructed and sequenced a 100-bp Illumina library (300- to 733-fold genome coverage) and a long-insert paired-end 454 library (6,936- to 7,528-bp inserts at 14- to 35-fold genome coverage). The 454 paired-end data were assembled in Newbler (5), and those consensus sequences were computationally shredded into 2-kbp overlapping fake reads (shreds). The Illumina sequencing data were assembled with Velvet (6), and the consensus sequence was computationally shredded into 1.5-kbp overlapping shreds. All data were additionally assembled with AllPaths (5), and the consensus sequences were computationally shredded into 5-kbp overlapping shreds. We integrated the 454 Newbler consensus shreds, the Illumina Velvet consensus shreds, AllPaths consensus shreds, and the 454 paired-end library read pairs using the Phrap parallel version (High Performance Software, LLC). Possible misassemblies were corrected using in-house scripts and/or manual editing in Consed (7–9).

Automatic annotation for each genome utilized an Ergatis-based (10) workflow at the Los Alamos National Laboratory (LANL), with minor manual curation. Table 1 lists the basic genome statistics and annotation counts. The final annotated assemblies are released in NCBI, and the raw data can be provided upon request, with a preliminary review of the annotated genomes and located β-lactam resistance genes in each.

TABLE 1 Basic assembly and annotation statistics for the three *Enterobacteriaceae* genomes

Species	Strain	Alt ID[a]	ATCC	Accession no.	Length (bp)	No. of:					G+C content (%)
						CDSs	Scaffolds	Contigs	rRNAs	tRNAs	
Enterobacter aerogenes	CDC 6003-71	EAX	ATCC 29904	JOUC00000000	5,084,129	4,692	1	9	19	84	55.1
Enterobacter cloacae	CDC 442-68	EBB	ATCC 13047	JPPR00000000	5,551,574	5,393	4	43	14	85	54.6
Pantoea agglomerans	CDC UA0804-01	EAY	NA[b]	JOUD00000000	4,720,650	4,494	2	9	17	83	54.8

[a] Alt ID, alternative identification.
[b] NA, not applicable.

Nucleotide sequence accession numbers. The NCBI accession numbers are JOUC00000000 for *E. aerogenes* CDC 6003-71, JPPR00000000 for *E. cloacae* CDC 442-68, and JOUD00000000 for *P. agglomerans*.

ACKNOWLEDGMENTS

Funding for this effort was provided by the Defense Threat Reduction Agency's Joint Science and Technology Office (DTRA J9-CB/JSTO).

This text is approved by LANL for unlimited release (LA-UR-14-26055).

The views expressed in this article are those of the authors and do not necessarily reflect the official policy or position of the Department of the Navy, Department of Defense, or the U.S. Government.

REFERENCES

1. **Sanders WE, Sanders CC.** 1997. *Enterobacter* spp.: pathogens poised to flourish at the turn of the century. Clin. Microbiol. Rev. **10:**220–241.
2. **Cruz AT, Cazacu AC, Allen CH.** 2007. *Pantoea agglomerans*, a plant pathogen causing human disease. J. Clin. Microbiol. **45:**1989–1992. http://dx.doi.org/10.1128/JCM.00632-07.
3. **Bennett S.** 2004. Solexa Ltd. Pharmacogenomics **5:**433–438. http://dx.doi.org/10.1517/14622416.5.4.433.
4. **Margulies M, Egholm M, Altman WE, Attiya S, Bader JS, Bemben LA, Berka J, Braverman MS, Chen Y-J, Chen Z, Dewell SB, Du L, Fierro JM, Gomes XV, Godwin BC, He W, Helgesen S, Ho CH, Irzyk GP, Jando SC, Alenquer ML, Jarvie TP, Jirage KB, Kim JB, Knight JR, Lanza JR, Leamon JH, Lefkowitz SM, Lei M, Li J, Lohman KL, Lu H, Makhijani VB, McDade KE, McKenna MP, Myers EW, Nickerson E, Nobile JR, Plant R, Puc BP, Ronan MT, Roth GT, Sarkis GJ, Simons JF, Simpson JW, Srinivasan M, Tartaro KR, Tomasz A, Vogt KA, Volkmer GA, Wang SH, Wang Y, Weiner MP, Yu P, Begley RF, Rothberg JM.** 2005. Genome sequencing in microfabricated high-density picolitre reactors. Nature **437:**376–380. http://dx.doi.org/10.1038/nature03959.
5. **Butler J, MacCallum I, Kleber M, Shlyakhter IA, Belmonte MK, Lander ES, Nusbaum C, Jaffe DB.** 2008. ALLPATHS: *de novo* assembly of whole-genome shotgun microreads. Genome Res. **18:**810–820. http://dx.doi.org/10.1101/gr.7337908.
6. **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.
7. **Ewing B, Green P.** 1998. Base-calling of automated sequencer traces using Phred. II. Error probabilities. Genome Res. **8:**186–194.
8. **Ewing B, Hillier L, Wendl MC, Green P.** 1998. Base-calling of automated sequencer traces using Phred. I. Accuracy assessment. Genome Res. **8:**175–185. http://dx.doi.org/10.1101/gr.8.3.175.
9. **Gordon D, Abajian C, Green P.** 1998. Consed: a graphical tool for sequence finishing. Genome Res. **8:**195–202. http://dx.doi.org/10.1101/gr.8.3.195.
10. **Hemmerich C, Buechlein A, Podicheti R, Revanna KV, Dong Q.** 2010. An Ergatis-based prokaryotic genome annotation Web server. Bioinformatics **26:**1122–1124. http://dx.doi.org/10.1093/bioinformatics/btq090.