10-17-2013

Draft genome sequences of Burkholderia cenocepacia ET12 lineage strains K56-2 and BC7

John J. Varga
University of Virginia

Liliana Losada
The J. Craig Venter Institute, Rockville, MD

Adrian M. Zelazny
National Institutes of Health

Maria Kim
The J. Craig Venter Institute, Rockville, MD

Jamison McCorrison
The J. Craig Venter Institute, Rockville, MD

See next page for additional authors

Recommended Citation

Varga, J.J., Losada, L., Zelazny, A.M., Kim, M., McCorrison, J., Brinkac, L., Sampaio, E.P., Greenberg, D.E., Singh, I., Heiner, C., Ashby, M., Nierman, W.C., Holland, S.M., Goldberg, J.B. (2013). Draft genome sequences of Burkholderia cenocepacia ET12 lineage strains K56-2 and BC7. Genome Announcements, 1(5), e00841-13.

Authors
John J. Varga, Liliana Losada, Adrian M. Zelazny, Maria Kim, Jamison McCorrison, Lauren Brinkac, Elizabeth P. Sampaio, David E. Greenberg, Indresh Singh, Cheryl Heiner, Meredith Ashby, William C. Nierman, Steven M. Holland, and Joanna B. Goldberg

genomeA
Journals.ASM.org

Draft Genome Sequences of *Burkholderia cenocepacia* ET12 Lineage Strains K56-2 and BC7

John J. Varga,[a,b,c] Liliana Losada,[c] Adrian M. Zelazny,[d,e] Maria Kim,[c] Jamison McCorrison,[c] Lauren Brinkac,[c] Elizabeth P. Sampaio,[d] David E. Greenberg,[e,f] Indresh Singh,[c] Cheryl Heiner,[g] Meredith Ashby,[g] William C. Nierman,[c,h] Steven M. Holland,[e] Joanna B. Goldberg[a,b]

Department of Microbiology, Immunology, and Cancer Biology, University of Virginia Health System, Charlottesville, Virginia, USA[a]; Department of Pediatrics and the Center for Cystic Fibrosis Research, Emory University School of Medicine, Children's Healthcare of Atlanta, Inc., Atlanta, Georgia, USA[b]; The J. Craig Venter Institute, Rockville, Maryland, USA[c]; Microbiology Service, Department of Laboratory Medicine, Clinical Center, National Institutes of Health, Bethesda, Maryland, USA[d]; Laboratory of Clinical Infectious Diseases, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, Maryland, USA[e]; University of Texas Southwestern Medical School, Dallas, Texas, USA[f]; Pacific Biosciences, Menlo Park, California, USA[g]; The George Washington University, Washington, DC, USA[h]

J.J.V. and L.L. contributed equally.

The *Burkholderia cepacia* complex (BCC) is a group of closely related bacteria that are responsible for respiratory infections in immunocompromised humans, most notably those with cystic fibrosis (CF). We report the genome sequences for *Burkholderia cenocepacia* ET12 lineage CF isolates K56-2 and BC7.

Received 17 September 2013 **Accepted** 25 September 2013 **Published** 17 October 2013

Citation Varga JJ, Losada L, Zelazny AM, Kim M, McCorrison J, Brinkac L, Sampaio EP, Greenberg DE, Singh I, Heiner C, Ashby M, Nierman WC, Holland SM, Goldberg JB. 2013. Draft genome sequences of *Burkholderia cenocepacia* ET12 lineage strains K56-2 and BC7. Genome Announc. 1(5):e00841-13. doi:10.1128/genomeA.00841-13.

Address correspondence to Joanna B. Goldberg, joanna.goldberg@emory.edu.

The *Burkholderia cepacia* complex (BCC) consists of 17 genetically related but phenotypically distinct betaproteobacterial species. BCC species can be found in a variety of niches and are capable of infecting numerous hosts. The most well-known member of the BCC is *Burkholderia cenocepacia*, a pathogen of onions and immunocompromised humans (1). We report the genome sequences of *B. cenocepacia* strains BC7 (2) and K56-2 (3), members of the highly transmissible genomovar III ET12 lineage (4, 5). While a genome sequence exists for the ET12 lineage *B. cenocepacia* strain J2315 (6), BC7 was isolated from a patient with "cepacia syndrome" and K56-2 is less antibiotic resistant, making it more amenable to genetic manipulation.

Genomic DNA was prepared with the DNeasy blood and tissue kit (Qiagen), according to the manufacturer's instructions. A combination of the Roche 454 GS-FLX Titanium 8-kb mate-pair libraries (~12× coverage) and 100-bp Illumina fragment reads (50× coverage) were used for sequence determinations. All reads were used to generate hybrid assemblies with Celera Assembler version 7.0 (CA7.0) (7). Almost all of the intrascaffold breaks in these assemblies were estimated at <20 bases, suggesting that they are ambiguities likely resulting from the high content of short tandem repeats in the G+C-rich regions of the genome. Thus, we sought additional PacBio reads to obtain a better consensus of the K56-2 genome; short- and long-length PacBio reads (<2,000 and >2,000, respectively) were used to generate highly accurate, long, preassembled reads using the HGAP algorithm (Pacific Biosciences), which were combined with 454 reads and input into the CA7.0. K56-2 assembled into 14 scaffolds containing 19 contigs. The final contig sequences were then run through the consensus polisher program, Quiver (PacBio), resulting in 3,516 base changes and 624 base insertions or deletions.

The BC7 assembly was initially in 7 scaffolds and 785 contigs. Since additional sequencing was not pursued for this genome, a reference-guided gap closure pipeline was employed to resolve hundreds of gaps found throughout the assembly. The gap sequences were predicted from the closed-genome reference (J2315) and used to recruit and locally assemble reads into the gaps to merge the adjacent contigs. The resulting assembly is 296 contigs in 7 scaffolds.

As in J2315, each genome has 3 chromosomes and 1 plasmid (6). The chromosomes in BC7 and K56-2 have very similar sizes to those reported in J2315 (3.83 Mb, 3.19 Mb, and 0.88 Mb), except for chromosome 1 in K56-2, which has an estimated size of 3.67 Mb, due to the absence of the large duplication in J2315. The presence of the plasmid was previously detected in K56-2, BC7, and J2315 (8). Sequence data confirm the presence of the plasmid in BC7 and K56-2, with practically no differences between the three strains except for the presence of an additional copy of an insertion element in the J2315 plasmid, pBCJ2315 (6). The genomes were annotated using the annotation pipeline of the J. Craig Venter Institute (JCVI) (http://www.jcvi.org) and submitted to GenBank. Sequence data indicate that K56-2 and BC7 have similar gene contents to that of J2315, with 7,714 and 7,930 open reading frames (ORFs), respectively.

Nucleotide sequence accession numbers. The *B. cenocepacia* BC7 whole-genome shotgun project has been deposited at DDBJ/EMBL/GenBank under the accession no. ALIZ00000000. The version described in this paper is the second version, ALIZ01000000.

The *B. cenocepacia* K56-2 whole-genome shotgun project has been deposited at DDBJ/EMBL/GenBank under the accession no. ALJA00000000. The version described in this paper is the second version, ALJA01000000.

ACKNOWLEDGMENTS

This project has been funded in part with federal funds from the National Institute of Allergy and Infectious Diseases, National Institutes of Health, Department of Health and Human Services, under contract no. HHSN272200900007C and N01-AI30071. PacBio kindly contributed sequence reads for *B. cenocepacia* K56-2. J.J.V. was supported in part by the National Institutes of Health grant no. 5T32AI055432 awarded to the University of Virginia.

REFERENCES

1. **Mahenthiralingam E, Urban TA, Goldberg JB.** 2005. The multifarious, multireplicon *Burkholderia cepacia* complex. Nat. Rev. Microbiol. **3:**144–156.
2. **Sajjan US, Corey M, Karmali MA, Forstner JF.** 1992. Binding of *Pseudomonas cepacia* to normal human intestinal mucin and respiratory mucin from patients with cystic fibrosis. J. Clin. Invest. **89:**648–656.
3. **Darling P, Chan M, Cox AD, Sokol PA.** 1998. Siderophore production by cystic fibrosis isolates of *Burkholderia cepacia*. Infect. Immun. **66:**874–877.
4. **LiPuma JJ, Spilker T, Gill LH, Campbell PW, III, Liu L, Mahenthiralingam E.** 2001. Disproportionate distribution of *Burkholderia cepacia* complex species and transmissibility markers in cystic fibrosis. Am. J. Respir. Crit. Care Med. **164:**92–96.
5. **Livesley MA, Baxter IA, Lambert PA, Govan JR, Weller PH, Lacey DE, Allison DG, Giwercman B, Hoiby N.** 1998. Subspecific differentiation of *Burkholderia cepacia* isolates in cystic fibrosis. J. Med. Microbiol. **47:**999–1006.
6. **Holden MT, Seth-Smith HM, Crossman LC, Sebaihia M, Bentley SD, Cerdeño-Tárraga AM, Thomson NR, Bason N, Quail MA, Sharp S, Cherevach I, Churcher C, Goodhead I, Hauser H, Holroyd N, Mungall K, Scott P, Walker D, White B, Rose H, Iversen P, Mil-Homens D, Rocha EP, Fialho AM, Baldwin A, Dowson C, Barrell BG, Govan JR, Vandamme P, Hart CA, Mahenthiralingam E, Parkhill J.** 2009. The genome of *Burkholderia cenocepacia* J2315, an epidemic pathogen of cystic fibrosis patients. J. Bacteriol. **191:**261–277.
7. **Miller JR, Delcher AL, Koren S, Venter E, Walenz BP, Brownley A, Johnson J, Li K, Mobarry C, Sutton G.** 2008. Aggressive assembly of pyrosequencing reads with mates. Bioinformatics **24:**2818–2824.
8. **Engledow AS, Medrano EG, Mahenthiralingam E, LiPuma JJ, Gonzalez CF.** 2004. Involvement of a plasmid-encoded type IV secretion system in the plant tissue watersoaking phenotype of *Burkholderia cenocepacia*. J. Bacteriol. **186:**6015–6024.