BIOSCIENCES BIOTECHNOLOGY RESEARCH ASIA, June 2017. Vol. **14**(2), p. 783-784

Draft Genome Sequence of *Cellulosimicrobium aquatile* 3bp, A Glucosidase Active Bacterium and the Type Strain of a Novel Species

Vishnuvardhan Reddy Sultanpuram# and Thirumala Mothe

Microbial Ecology Laboratory, Department of Applied Biosciences, Mahatma Gandhi University, Anneparthy, Yellareddygudem (PO), Nalgonda-508254, Telangana, India.

http://dx.doi.org/10.13005/bbra/2508

(Received: 29 April 2017; accepted: 10 May 2017)

A novel strain of bacterium designated as 3bp, a Gram-stain positive, non-spore forming, non-motile, yellow-pigmented, rod shaped and coccoid aerobic strain was isolated from Panagal reservoir, Nalgonda, Telangana, India, represents a novel species from the phylum *Actinobacteria*. Here we report the draft genome of the strain 3bp. It comprises ~ 4.31771 Mb with the G + C content of 74.73 %, and a total of 3796 protein-coding genes.

Keywords: *Cellulosimicrobium aquatile,* glucosidase active, Whole genome sequencing.

The strain *Cellulosimicrobium aquatile* 3bp was recently described as a novel species of the genus *Cellulosimicrobium*. The cells are Gram-stain positive, non-spore forming, non-motile, yellow-pigmented, rod shaped and coccoid aerobic in nature. The species showed positive for catalase activity and oxidase activities (Sultanpuram *et al*., 2015).

Sequencing of the type strain is part of Genomic Encyclopedia of Type Strains, Phase III: the (KMG-III) project (Whitman *et al*., 2015). Project information is deposited in the Genomes on Line Database (GOLD) (Reddy *et al*., 2015), and the high quality draft genome sequence is deposited in GenBank and in the Integrated Microbial Genomes database (IMG) (Markowitz *et al*., 2014). Draft sequencing, initial gap closure and annotation were performed by the DOE Joint Genome Institute (JGI) using state-of-the-art sequencing technology (Mavromatis *et al*., 2015).

MATERIALS AND METHODS

An Illumina standard shotgun library was constructed and sequenced using the Illumina HiSeq 2500-1TB. Illumina paired-end reads of average 250-bp length and total read size of ~ 548 Mbp (clean data) were assembled using Velvet (version 1.1.04)(Zerbino and Birney, 2008), simulated paired end reads were created from Velvet contigs using wgsim and simulated read pairs were reassembled using Allpaths-LG (version r42328) (MacCallum *et al*., 2009). Protein-coding genes were identified using Prodigal (Hyatt *et al*., 2010), as part of the DOE-JGI genome annotation pipeline (Huntemann *et al*., 2015). Additional gene prediction analysis and manual functional annotation were performed within the Integrated Microbial Genomes (IMG) platform, which provides tools for analyzing and reviewing the structural and functional annotations of genomes in a comparative context (Markowitz *et al*., 2014). Genome annotation procedures are detailed in Markowitz *et al*. (2014) and references therein. Briefly, the predicted CDSs were translated and used to search the NCBI nonredundant database,

---

* To whom all correspondence should be addressed.
E-mail: svr.micro@gmail.com

UNIProt, TIGRFam, Pfam, KEGG, COG and InterPro databases. Transfer RNA genes (51 in number) were identified using the tRNAScan-SE tool and other noncoding RNAs (8 in number) were found using INFERNAL.

RESULTS AND DISCUSSIONS

The assembled draft genome contained 4,317,712 bps with 74.73 % GC content. The analysis of the draft genome of *Cellulosimicrobium aquatile* 3bp predicted 3796 protein coding sequences (CDS) of which 77.97 % were assigned to recognized functional genes. Ribosomal RNA genes (4 in number) were predicted using hmmsearch against the custom models generated for each type of rRNA.

The genome also harboured a cluster of genes coding for glucosidase activities. These genes include, â-glucosidase (770 aa; locus tag BXA05_RS01230; accession number FTM101000001), glucan, 1,3-â-glucosidase (526 aa, locug tag BXA05_RS01565; accession number FTM1000001), â-glucosidase (997 aa; locus tag BXA05_RS03520; accession number FTM101000001), â-glucosidase (391 aa; locus tag BXA05_RS04915; accession number FTM101000002), á-glucosidase (588 aa; locus tag BXA05_RS15140; accession number FTM101000006), â-glucosidase (516 aa; locus tag BXA05_RS15215; accession number FTM101000006) and á-glucosidase (496 aa; locus tag BXA05_RS15890; accession number FTM101000006). This draft genome sequence will further help in understanding the genetic potential of *Cellulosimicrobium aquatile* for glucosidase activity.

Nucleotide sequence accession numbers

The draft sequence of *C. aquatile* strain 3bp obtained in this Whole Genome Shotgun project has been deposited at GenBank under the accession no. FTMI00000000. The Genomes on Line Database (GOLD) ID is PRJEB18885.

ACKNOWLEDGMENTS

We are thankful to Prof. William Whitman for proposing the genomic sequencing of this organism under KMG-III. This organism was isolated and described in Department of Science & Technology, Government of India, FTSYS project grants (SB/FT/LS-115/2012-13 and SB/FT/LS-320/2012-13, respectively).

REFERENCES

1. Huntemann M, Ivanova NN, Mavromatis K, Tripp HJ, Paez-Espino D, Palaniappan K, *et al*. The standard operating procedure of the DOE-JGI Microbial Genome Annotation Pipeline (MGAP v.4). *Stand Genomic Sci.* 2015; **10**:86.
2. Hyatt D, Chen GL, Locascio PF, Land ML, Larimer FW, Hauser LJ. Prodigal: prokaryotic gene recognition and translation initiation site identification. *BMC Bioinformatics* 2010; **11**:119.
3. MacCallum I, Przybylski D, Gnerre S, Burton J, Shlyakhter I, Gnirke A, *et al*. ALLPATHS 2: small genomes assembled accurately and with high continuity from short paired reads. *Genome Biol.* 2009; **10**:R103.
4. Markowitz VM, Chen IM, Palaniappan K, Chu K, Szeto E, Pillay M, *et al*. IMG 4 version of the integrated microbial genomes comparative analysis system. *Nucleic Acids Res.,* 2014; **42**(Database issue):D560–7.
5. Mavromatis K, Land M, Brettin TS, Quest DJ, Copeland A, Clum A, *et al*. The fast changing landscape of sequencing technologies and their impact on microbial assemblies and annotation. *PLoS One,* 2012; **7**:e48837.
6. Reddy TBK, Thomas A, Stamatis D, Bertsch J, Isbandi M, Jansson J, *et al*. The Genomes On Line Database (GOLD) v.5: a metadata management system based on a four level (meta) genome project classification. *Nucleic Acids Res.*, 2015; **43**:D1099.
7. Vishnuvardhan Reddy Sultanpuram, Thirumala Mothe, Sasikala, Chintalapti, Ventata Ramana, Chintalapati, *Cellulosimicrobium aquatile* sp. nov., isolated from Panagal reservoir, Nalgonda, India. *Antonie van Leewenhoek,* 2015; **108**: 1357-1364.
8. Whitman WB, Woyke T, Klenk HP, Zhou Y, Lilburn TG, Beck BJ, *et al*. Genomic encyclopedia of bacterial and archaeal type strains, phase III: the genomes of soil and plant associated and newly described type strains. *Stand Genomic Sci.,* 2015; 10: 26.
9. Zerbino DR, Birney E. Velvet: algorithms for de novo short read assembly using de bruijn graphs. *Genome Res.,* 2008; **18**:821–829.