genomeA
Journals.ASM.org

Complete Genome Sequence of the Freshwater Colorless Sulfur Bacterium *Beggiatoa leptomitiformis* Neotype Strain D-402^T

Alexey Fomenkov,[a] Tamas Vincze,[a] Margarita Y. Grabovich,[b] Galina Dubinina,[c] Maria Orlova,[b] Elena Belousova,[b] Richard J. Roberts[a]

New England BioLabs, Ipswich, Massachusetts, USA[a]; Department of Biology, Voronezh State University, Voronezh, Russia[b]; Winogradsky Institute of Microbiology, Russian Academy of Sciences, Moscow, Russia[c]

In this report, we announce the availability of a complete closed genome sequence and methylome analysis of *Beggiatoa leptomitiformis* neotype strain D-402^T (DSM 14946, UNIQEM U 779).

Received 15 October 2015 **Accepted** 23 October 2015 **Published** 10 December 2015

Citation Fomenkov A, Vincze T, Grabovich MY, Dubinina G, Orlova M, Belousova E, Roberts RJ. 2015. Complete genome sequence of the freshwater colorless sulfur bacterium *Beggiatoa leptomitiformis* neotype strain D-402^T. Genome Announc 3(6):e01436-15. doi:10.1128/genomeA.01436-15.

Address correspondence to Alexey Fomenkov, fomenkov@neb.com.

At the end of the 19th century, Sergei Winogradsky introduced the concept of chemolithotrophy when he first reported on organisms gaining energy exclusively from the oxidation of inorganic compounds (1). Members of the bacterial family *Beggiatoaceae* have gained much attention due to their ability to oxidize sulfide to elemental sulfur, which they deposit intracellularly in the form of small globules or droplets. However, due to the difficulties of isolating, purifying, and growing this bacterium *in vivo*, only a few draft genome sequences have been assembled (accession no. NZ_AHMA00000000.1, GCA_000170695.1, and GCA_000170715.1).

In this report, we announce the availability of a complete closed genome sequence of *Beggiatoa leptomitiformis* neotype strain D-402^T (DSM 14946, UNIQEM U 779). This strain was previously described based on its morphological and biochemical characteristics (2–6).

The genome was sequenced using the Pacific Biosciences (PacBio) RSII sequencing platform (7). Briefly, SMRTbell libraries were constructed from a genomic DNA sample sheared to an average size of ~10 to 20 kb using the G-tubes protocol (Covaris, Woburn, MA, USA), additionally purified using the PowerClean DNA clean-up kit (MoBio Laboratories, Inc., Carlsbad, CA), end repaired, and ligated to hairpin adapters. Incompletely formed SMRTbell templates were digested with a combination of exonuclease III and exonuclease VII (New England BioLabs, Ipswich, MA, USA). Genomic DNA fragments and SMRTbell library qualification and quantification were performed using a Qubit fluorimeter (Invitrogen, Eugene, OR) and 2100 Bioanalyzer (Agilent Technologies, Santa Clara, CA). Two SMRTbell 10- and 20-kb libraries were prepared according to the modified 10- and 20-kb PacBio sample preparation protocols and sequenced using C2 and C4 chemistry on six single-molecule real-time (SMRT) cells, with a 180-min collection protocol. Sequencing reads were processed, mapped, and assembled by the Pacific Biosciences SMRT Analysis pipeline using the HGAP3 protocol and polished using Quiver (8) to give a fully closed genome with 422× coverage. The genome size was 4,265,296 bp, and the plasmid was 6,185 bp, which together contain a total of 3,636 genes. The assembled sequences were annotated with Rapid Annotations using Subsystems Technology (RAST) (9) and the NCBI Prokaryotic Genomes Annotation Pipeline (PGAP).

Epigenetic modification at each nucleotide position was measured as kinetic variations (KVs) in the nucleotide incorporation rates, and methylated motifs were deduced from the KV data (10–12). Thirteen DNA methyltransferase recognition motifs corresponding to one m4C and nine m6A modifications were detected by direct single-molecule real-time (SMRT) sequencing, and an additional three m5C motifs were detected in Tet2-treated DNA. Matching of motifs with methyltransferase genes was carried out, and the results are shown in Table 1. They have also been deposited in REBASE (13).

TABLE 1 Summary of methyltransferases identified in *B. leptomitiformis* neotype strain D-402^T

Motif[a]	Assigned or predicted	Methylation type	Restriction modification type
Direct detection			
GATC	M.Ble402I	m6A	II
GRAGC**A**G	M.Ble402II	m6A	II
SAG**C**TS	M.Ble402III	m4C	II
AC**A**YNNNNNR**T**GT		m6A	II
CA**A**YNNNNNR**T**TG	S.Ble402ORFBP	m6A	I
C**A**GNNNNNR**T**AAT	S.Ble402ORFQP	m6A	I
CATCH**A**G		m6A	II
CGG**A**G		m6A	III
CGGTC**A**		m6A	II
DCTGG**A**TD		m6A	II
GGCTG**A**		m6A	II
GTTGN**A**G		m6A	II
TCG**A**		m6A	II
5-mC oxidation by Tet2			
GGHCC = GGNCC	M.Ble402ORFDP	5 mC	II
CCDGG = CCNGG	M.Ble402ORFLP	5 mC	II
GGCCNB = GGCC	M.Ble402ORFTP	5 mC	II

[a] Modified bases are highlighted in bold.

Nucleotide sequence accession numbers. The complete genome and plasmid sequences of the *B. leptomitiformis* neotype strain D-402^T are available in GenBank under the accession numbers CP012373 and CP012374, respectively.

ACKNOWLEDGMENTS

This project was supported by the Small Business Innovation Research Program (NIGMS) of the National Institutes of Health under award R44GM105125 to R.J.R.

R.J.R., A.F., and T.V. work for New England BioLabs, a company that sells research reagents, including restriction enzymes and DNA methylases, to the scientific community.

REFERENCES

1. **Winogradsky S**. 1887. Ueber Schwefelbacterien. Botanische Zeitung **31:** 489–507.
2. **Grabovich MY, Churikova VV, Dubinina GA, Lebedeva VY**. 1998. Mixotrophic and lithoheterotrophic growth of the freshwater filamentous sulfur bacterium *Beggiatoa leptomitiformis* D-402. Mikrobiologiya **67:** 383–388.
3. **Grabovich MY, Patritskaya VY, Muntyan MS, Dubinina GA**. 2001. Lithoautotrophic growth of the freshwater strain *Beggiatoa* D-402 and energy conservation culture under microoxic conditions. FEMS Microbiol Lett **204:**341–345.
4. **Patritskaya VY, Dubinina GA, Grabovich MY, Muntyan MS**. 2001. Lithoautotrophic growth of the freshwater colorless sulfur bacterium *Beggiatoa* "*leptomitiformis*" D-402.]. Mikrobiologiya **70:**145–150.
5. **Eprintsev AT, Falaleeva MI, Grabovich MY, Parfenova NV, Kashirskaya NN, Dubinina GA**. 2004. The role of malate dehydrogenase isoforms in the regulation of anabolic and catabolic processes in the colorless sulfur bacterium *Beggiatoa leptomitiformis* D-402. Mikrobiologiya **73:**367–371. http://dx.doi.org/10.1023/B:MICI.0000036977.79822.7a.
6. **Muntyan MS, Grabovich MY, Patritskaya VY, Dubinina GA**. 2005. Regulation of metabolic and electron transport pathways in the freshwater bacterium *Beggiatoa leptomitiformis* D-402. Mikrobiologiya **74:**388–394. http://dx.doi.org/10.1007/s11021-005-0078-8.
7. **Fomenkov A, Lunnen KD, Zhu Z, Anton BP, Wilson GG, Vincze T, Roberts RJ**. 2015. Complete genome sequence and methylome analysis of *Bacillus* strain X1. Genome Announc **3**(1)**:**e01593-14. http://dx.doi.org/10.1128/genomeA.01593-14.
8. **Chin C, Alexander DH, Marks P, Klammer AA, Drake J, Heiner C, Clum A, Copeland A, Huddleston J, Eichler EE, Turner SW, Korlach J**. 2013. Nonhybrid, finished microbial genome assemblies from long-read SMRT sequencing data. Nat Methods **10:**563–569. http://dx.doi.org/10.1038/nmeth.2474.
9. **Aziz RK, Bartels D, Best AA, DeJongh M, Disz T, Edwards RA, Formsma K, Gerdes S, Glass EM, Kubal M, Meyer F, Olsen GJ, Olson R, Osterman AL, Overbeek RA, McNeil LK, Paarmann D, Paczian T, Parrello B, Pusch GD, Reich C, Stevens R, Vassieva O, Vonstein V, Wilke A, Zagnitko O**. 2008. The RAST server: Rapid Annotations using Subsystems Technology. BMC Genomics **9:**75. http://dx.doi.org/10.1186/1471-2164-9-75.
10. **Flusberg BA, Webster DR, Lee JH, Travers KJ, Olivares EC, Clark TA, Korlach J, Turner SW**. 2010. Direct detection of DNA methylation during single-molecule, real-time sequencing. Nat Methods **7:**461–465. http://dx.doi.org/10.1038/nmeth.1459.
11. **Clark TA, Murray IA, Morgan RD, Kislyuk AO, Spittle KE, Boitano M, Fomenkov A, Roberts RJ, Korlach J**. 2012. Characterization of DNA methyltransferase specificities using single-molecule, real-time DNA sequencing. Nucleic Acids Res **40:**e29. http://dx.doi.org/10.1093/nar/gkr1146.
12. **Korlach J, Turner SW**. 2012. Going beyond five bases in DNA sequencing. Curr Opin Struct Biol **22:**251–261. http://dx.doi.org/10.1016/j.sbi.2012.04.002.
13. **Roberts RJ, Vincze T, Posfai J, Macelis D**. 2010. REBASE–a database for DNA restriction and modification: enzymes, genes and genomes. Nucleic Acids Res **38:**D234–D236. http://dx.doi.org/10.1093/nar/gkp874.