AMERICAN SOCIETY FOR MICROBIOLOGY | Microbiology Resource Announcements

GENOME SEQUENCES

Metagenome-Assembled Genome Sequences of a Biofilm Derived from Marsberg Copper Mine

Sania Arif,[a] Heiko Nacke,[b] Michael Hoppert[a]

[a]Department of General Microbiology, Institute of Microbiology and Genetics, Georg-August-Universität, Göttingen, Germany
[b]Department of Genomic and Applied Microbiology, Institute of Microbiology and Genetics, Georg-August University, Göttingen, Germany

ABSTRACT We sequenced the metagenome of a biofilm collected near a leachate stream of the Marsberg copper mine (Germany) and reconstructed eight metagenome-assembled genomes. These genomes yield copper resistance through Cu(I) oxidation via multiple copper oxidases and extrusion through copper-exporting P-type ATPases.

The historic Marsberg copper mine (51°27′12.6″N, 8°51′42.1″E) offers ambient natural conditions for the enrichment of heavy metal-resistant consortia under the influence of copper-rich (acidic) sulfidic mine waters at low temperature (10°C) (1). In February 2018, a biofilm near copper-rich leachate was aseptically collected from rock samples by using a sterile scalpel. Microbial DNA was extracted using the DNeasy PowerSoil kit (Qiagen, Venlo, Netherlands) according to the manufacturer's protocol. The purified DNA from the biofilm sample (designated MB1) was used to generate Illumina paired-end sequencing libraries with the Nextera DNA sample preparation kit (Illumina, San Diego, CA, USA); the libraries were sequenced by employing the MiSeq reagent kit v.3 and a MiSeq instrument as stated by the manufacturer (Illumina). Default parameters were used for all software unless otherwise specified. Quality trimming of reads was performed by employing fastp v.0.20.1 (qualified quality phred score, 20; minimal read length, 50 bp) (2) and yielded 18,235,972 paired-end reads. Base correction in overlapping regions (the correction option was selected concerning fastp-based quality trimming using default parameters; this option allows identification of overlapping regions of each pair of reads, and mismatched base pairs in these regions can be corrected if one base shows high quality and the other very low quality) and removal of the automatically detected adapter sequences were performed. Low-quality bases at the 5′ and 3′ ends of reads were trimmed once the mean quality score within a sliding window of 4 dropped below 20. Sequences were *de novo* assembled into 53,638 contigs of ≥1,000 bp via metaSPAdes v.3.14.0 (3). Binning was performed using MaxBin v.2.2.7 (minimum contig length, 1,000 bp; minimum probability for binning, 0.50) (4). Application of CheckM v.1.1.2 revealed eight relatively complete metagenome-assembled genomes (MAGs) (completeness, ≥89%; contamination rate, ≤10%) (5). Each MAG was annotated using Prokka v.1.14.5 (6). Subsequently, Prokka output was analyzed by using the Pathway Tools software v.23.5 (7) with the MetaCyc database v.23.5 (8). MAGs were classified taxonomically using GTDB-Tk v.1.0.2 and the Genome Taxonomy Database (GTDB) (release 89) (9, 10).

The MAGs were classified as members of *Actinobacteria* (Mberg 009), "*Candidatus* Binatota" (Mberg 010 and 011), *Chlorobacteria* (Mberg 002, 006, 008, and 019), and *Deinococcus-Thermus* (Mberg 015). Functional analysis revealed the presence of genes for copper-sensing transcriptional repressors CsoR and RicR, copper-exporting P-type ATPases such as ActP, CptA, and CopA, and oxidation enzymes, multicopper oxidases (MCOs), involved in copper homeostasis in all MAGs. The detoxification pathways for reactive oxygen species, toxins, and antibiotic compounds involve superoxide dismutase

Citation Arif S, Nacke H, Hoppert M. 2021. Metagenome-assembled genome sequences of a biofilm derived from Marsberg copper mine. Microbiol Resour Announc 10:e01253-20. https://doi.org/10.1128/MRA.01253-20.

Editor John J. Dennehy, Queens College

Address correspondence to Sania Arif, sania.arif@stud.uni-goettingen.de.

Received 20 November 2020
Accepted 17 December 2020
Published 14 January 2021

(SOD) and peroxidases, which degrade superoxide anion radicals, and mycothiol-mediated detoxification through the enzyme Mca with thiols (11–13). All MAGs also include genes encoding aromatic compound degradation enzymes to generate ATP, which could potentially be used by copper-ATPase transporters.

Data availability. Raw sequencing data are available at the NCBI Sequence Read Archive (SRA) under accession number SRR12886061. The metagenome assembly is available at GenBank under accession number JADEYI000000000. The MAGs are available at GenBank under accession numbers JADMIG000000000, JADMIH000000000, JADMII000000000, JADMIJ000000000, JADMIK000000000, JADMIL000000000, JADMIM000000000, and JADMIN000000000. Prokka-based annotations of the eight MAG contigs are publicly available at the Göttingen Research Online Database (https://doi.org/10.25625/ODCARY).

ACKNOWLEDGMENTS

The provision of a Deutscher Akademischer Austauschdienst (DAAD) doctoral research grant is acknowledged. We also acknowledge the Open Access Publication Funds of the University of Göttingen.

We thank Petra Ackermann, Gerhard Rosenkranz, and the Marsberger Heimatbund eV (Marsberg) for support with respect to the sample collection at Marsberg Kilianstollen.

REFERENCES

1. Emmerich HP, Heydemann A. 1987. Sekundärmineralbildung aus Grubenwässern im Kupferbergwerk Niedermarsberg [Secondary mineral formation from mine water in the Niedermarsberg copper mine]. Der Aufschluss 38:149–156. (In German.)
2. Chen S, Zhou Y, Chen Y, Gu J. 2018. fastp: an ultra-fast all-in-one FASTQ preprocessor. Bioinformatics 34:i884–i890. https://doi.org/10.1093/bioinformatics/bty560.
3. Nurk S, Meleshko D, Korobeynikov A, Pevzner PA. 2017. metaSPAdes: a new versatile metagenomic assembler. Genome Res 27:824–834. https://doi.org/10.1101/gr.213959.116.
4. Wu Y-W, Simmons BA, Singer SW. 2016. MaxBin 2.0: an automated binning algorithm to recover genomes from multiple metagenomic datasets. Bioinformatics 32:605–607. https://doi.org/10.1093/bioinformatics/btv638.
5. Parks DH, Imelfort M, Skennerton CT, Hugenholtz P, Tyson GW. 2015. CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. Genome Res 25:1043–1055. https://doi.org/10.1101/gr.186072.114.
6. Seemann T. 2014. Prokka: rapid prokaryotic genome annotation. Bioinformatics 30:2068–2069. https://doi.org/10.1093/bioinformatics/btu153.
7. Karp PD, Latendresse M, Paley SM, Krummenacker M, Ong QD, Billington R, Kothari A, Weaver D, Lee T, Subhraveti P, Spaulding A, Fulcher C, Keseler IM, Caspi R. 2016. Pathway Tools version 19.0 update: software for pathway/genome informatics and systems biology. Brief Bioinform 17:877–890. https://doi.org/10.1093/bib/bbv079.
8. Caspi R, Billington R, Keseler IM, Kothari A, Krummenacker M, Midford PE, Ong WK, Paley S, Subhraveti P, Karp PD. 2020. The MetaCyc database of metabolic pathways and enzymes: a 2019 update. Nucleic Acids Res 48:D445–D453. https://doi.org/10.1093/nar/gkz862.
9. Chaumeil P-A, Mussig AJ, Hugenholtz P, Parks DH. 2020. GTDB-Tk: a toolkit to classify genomes with the Genome Taxonomy Database. Bioinformatics 36:1925–1927. https://doi.org/10.1093/bioinformatics/btz848.
10. Parks DH, Chuvochina M, Chaumeil P-A, Rinke C, Mussig AJ, Hugenholtz P. 2019. Selection of representative genomes for 24,706 bacterial and archaeal species clusters provide a complete genome-based taxonomy. bioRxiv 771964. https://doi.org/10.1101/771964.
11. Newton GL, Buchmeier N, Fahey RC. 2008. Biosynthesis and functions of mycothiol, the unique protective thiol of *Actinobacteria*. Microbiol Mol Biol Rev 72:471–494. https://doi.org/10.1128/MMBR.00008-08.
12. Newton GL, Av-Gay Y, Fahey RC. 2000. A novel mycothiol-dependent detoxification pathway in mycobacteria involving mycothiol *S*-conjugate amidase. Biochemistry 39:10739–10746. https://doi.org/10.1021/bi000356n.
13. Broxton CN, Culotta VC. 2016. SOD enzymes and microbial pathogens: surviving the oxidative storm of infection. PLoS Pathog 12:e1005295. https://doi.org/10.1371/journal.ppat.1005295.