Draft Genome Sequence of *Ezakiella peruensis* Strain M6.X2, a Human Gut Gram-Positive Anaerobic Coccus

Awa Diop,[a] Khoudia Diop,[a] Enora Tomei,[a] Didier Raoult,[a,b] Florence Fenollar,[a] Pierre-Edouard Fournier[a]

[a]Unité de Recherche sur les Maladies Infectieuses et Tropicales Emergentes, Aix-Marseille Université, UM 63, CNRS UMR7278, IRD 198, INSERM U1095, Assistance Publique-Hôpitaux de Marseille, Institut Hospitalo-Universitaire Méditerranée-Infection, Faculté de Médecine, Marseille, France

[b]Special Infectious Agents Unit, King Fahd Medical Research Center, King Abdulaziz University, Jeddah, Saudi Arabia

ABSTRACT We report here the draft genome sequence of *Ezakiella peruensis* strain M6.X2[T]. The draft genome is 1,672,788 bp long and harbors 1,589 predicted protein-encoding genes, including 26 antibiotic resistance genes with 1 gene encoding vancomycin resistance. The genome also exhibits 1 clustered regularly interspaced short palindromic repeat region and 333 genes acquired by horizontal gene transfer.

Ezakiella peruensis is the type and only species of the genus *Ezakiella*, created in 2015 (1). *E. peruensis* occupies a unique position in an undefined family within the phylum *Firmicutes* (1). This microorganism is a Gram-positive anaerobic coccus. Gram-positive anaerobic cocci include many commensal species of humans and animals and also some human pathogens (2). The type strain M6.X2[T] was isolated from a fecal sample of a healthy individual residing in a coastal traditional community in Peru (1). It is nonmotile and non-spore forming. Here, we present the annotated draft genome sequence of *E. peruensis* strain M6.X2[T] (DSM 27367 = NBRC 109957 = CCUG 64571), which we obtained from the DSMZ collection.

Genomic DNA of *E. peruensis* strain M6.X2[T] was sequenced using a MiSeq sequencer with the mate-pair strategy (Illumina, Inc., San Diego, CA, USA). DNA was quantified by a Qubit assay with a high-sensitivity kit (Life Technologies, Carlsbad, CA, USA) at 38.4 ng/μl. The 576,285 high-quality paired-end reads were trimmed and then assembled using the SPAdes assembler program (3). The draft genome sequence was annotated using Prokka software (4). Functional annotation was achieved using the BLASTp algorithm (5) against the Clusters of Orthologous Groups (COGs) database and the Rapid Annotations using Subsystems Technology (RAST) web server (6). Ribosomal RNAs (5S, 16S, and 23S rRNAs) were predicted using RNAmmer software (7).

The genome was 1,672,788-bp long, assembled in five scaffolds (seven contigs) with a G+C content of 36.9%. Overall, 1,589 protein-coding sequences were identified, including 1,165 (73.31%) protein-coding genes that had orthologs in the COGs database, 1,052 of which were assigned a putative function. A total of 46 tRNA loci and 1 rRNA operon (16S, 5S, and 23S rRNA) were identified in the genome. Strain M6.X2[T] exhibited 26 genes associated with antibiotic resistance and toxic compounds, including one *vanW* gene encoding vancomycin resistance. No toxin/antitoxin module or bacteriocin-associated gene was identified. The genome of *E. peruensis* harbored 1 clustered regularly interspaced short palindromic repeat locus of 763 bp with 12 repeats (mean repeat length = 36 bp). We also detected 333 putative genes acquired by horizontal gene transfer, including 209 from bacteria within the order *Clostridiales*.

Accession number(s). The 16S rRNA and genome sequences from *Ezakiella peruensis* strain M6.X2[T] are available in GenBank under accession numbers KJ469554 and OCSL00000000, respectively.

REFERENCES

1. Patel NB, Tito RY, Obregón-Tito AJ, O'Neal L, Trujillo-Villaroel O, Marin-Reyes L, Troncoso-Corzo L, Guija-Poma E, Hamada M, Uchino Y, Lewis CM, Lawson PA. 2015. *Ezakiella peruensis* gen. nov., sp. nov. isolated from human fecal sample from a coastal traditional community in Peru. Anaerobe 32:43–48. https://doi.org/10.1016/j.anaerobe.2014.12.002.

2. Ulger-Toprak N, Liu C, Summanen PH, Finegold SM. 2010. *Murdochiella asaccharolytica* gen. nov., sp. nov., a Gram-stain-positive, anaerobic coccus isolated from human wound specimens. Int J Syst Evol Microbiol 60:1013–1016. https://doi.org/10.1099/ijs.0.015909-0.

3. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Prjibelski AD, Pyshkin AV, Sirotkin AV, Vyahhi N, Tesler G, Alekseyev MA, Pevzner PA. 2012. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. J Comput Biol 19:455–477. https://doi.org/10.1089/cmb.2012.0021.

4. Seemann T. 2014. Prokka: rapid prokaryotic genome annotation. Bioinformatics 30:2068–2069. https://doi.org/10.1093/bioinformatics/btu153.

5. Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, Madden TL. 2009. BLAST+: architecture and applications. BMC Bioinformatics 10:421. https://doi.org/10.1186/1471-2105-10-421.

6. 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. https://doi.org/10.1186/1471-2164-9-75.

7. Lagesen K, Hallin P, Rødland EA, Staerfeldt H-H, Rognes T, Ussery DW. 2007. RNAmmer: consistent and rapid annotation of ribosomal RNA genes. Nucleic Acids Res 35:3100–3108. https://doi.org/10.1093/nar/gkm160.