Draft Genome Sequence of Mycobacterium ulcerans S4018 Isolated from a Patient with an Active Buruli Ulcer in Benin, Africa

Stanimir Kambarev, Stéphane Corvec, Annick Chauty, Estelle Marion, Laurent Marsollier, Frédéric Pecorari

AMERICAN SOCIETY FOR MICROBIOLOGY | genomeAnnouncements™

PROKARYOTES

Draft Genome Sequence of *Mycobacterium ulcerans* S4018 Isolated from a Patient with an Active Buruli Ulcer in Benin, Africa

Stanimir Kambarev,[a] Stéphane Corvec,[b,c] Annick Chauty,[d] Estelle Marion,[e] Laurent Marsollier,[e] Frédéric Pecorari[a]

CRCINA, Inserm, CNRS, Université d'Angers, Université de Nantes, Nantes, France[a]; CRCINA, Inserm, Université d'Angers, Université de Nantes, Nantes, France[b]; Service de Bactériologie et Hygiène Hospitalière, CHU de Nantes, Nantes, France[c]; CDTUB de Pobè, Pobè, Benin[d]; CRCINA, Inserm, Université de Nantes, Université d'Angers, Angers, France[e]

ABSTRACT Currently, there are only two publicly available genomes of *Mycobacterium ulcerans*—the causative agent of the neglected, but devastating, tropical disease Buruli ulcer. Here, we report the draft genome sequence of isolate S4018, recovered from an active cutaneous lesion of a patient with Buruli ulcer in Benin, Africa.

Received 1 March 2017 **Accepted** 2 March 2017 **Published** 27 April 2017

Citation Kambarev S, Corvec S, Chauty A, Marion E, Marsollier L, Pecorari F. 2017. Draft genome sequence of *Mycobacterium ulcerans* S4018 isolated from a patient with an active Buruli ulcer in Benin, Africa. Genome Announc 5:e00248-17. https://doi.org/10.1128/genomeA.00248-17.

Address correspondence to Frédéric Pecorari, frederic.pecorari@univ-nantes.fr.

The neglected, but devastating, tropical disease Buruli ulcer is the third most common mycobacteriosis worldwide after tuberculosis and leprosy. It is associated with extensive subcutaneous necrosis caused by the human pathogen *Mycobacterium ulcerans* (1). Surprisingly, since the sequencing of the first genome of this species in 2007 (2), there has been only one other genome submitted to the public nucleotide databases (draft sequence, strain Harvey, accession number JAOL01000000). Here, we report the draft genome of isolate S4018, recovered from a cutaneous lesion of patient with Buruli ulcer in Benin, Africa.

The isolate was cultivated on Lowenstein-Jensen medium for 5 months. Prior to DNA extraction, it was grown on Middlebrook 7H10 agar and enriched with an OADC (oleic acid, albumin, dextrose, and catalase) supplement. Cells were then scraped from the agar and washed with wash solution (0.3 M sucrose, 50 mM Tris pH 8.0, and 10 mM EDTA). Every 50 mg of cells were treated with 3.5 mM lysozyme and 0.4 mM RNase A in 20-μL volume overnight at 37°C, followed by centrifugation. The pellets were resuspended in 1 × Tris/EDTA (TE) buffer, containing 36 nM proteinase K and 1% SDS. After incubation for 1 h at 50°C and another centrifugation, the pellets were resuspended in lysis solution (6% guanidine hydrochloride, 1% Tween 20, and 1% Nonidet P-40) and incubated for 1 h at 37°C. Finally, the cells were completely lysed with a bead beater, and the aqueous phase of the lysates was extracted with chloroform. DNA was precipitated with isopropanol/3 M sodium acetate on ice. The precipitates were washed with 70% ethanol, dried, and dissolved in 1 × TE buffer. DNA was fragmented (200 to 300 bp) using a Bioruptor Standard (Diagenode) device. Sequencing libraries were prepared from 1 μg of fragments using a NEBNext Ultra DNA library prep kit for Illumina (NEB) and sequenced on a MiSeq sequencer (Illumina). *De novo* assembly was performed with Velvet version 1/2/10 (3) and VelvetOptimiser version 2.2.5 (4) from 4,059,460 high-quality paired-ends reads (150 bp). The obtained contigs were reordered with Mauve version 1/2/10 (5) against the complete genome of strain Agy99 (2). Finally, the assembly was annotated through the NCBI Prokaryotic Genome Annotation Pipeline (6).

The draft sequence consists of 503 contigs with an average coverage of 39× and an N_{50} of 37 kb. It has a G+C content of 66% and a total length of 5,399,220 bp. Annotation revealed 4,914 coding sequences, 1,035 pseudogenes, and 51 RNA genes. We attributed the high number of contigs to the characteristic accumulation of the IS*2404* and IS*2606* mobile elements in the genome of *M. ulcerans* (2, 7), since it has been shown that such repeated elements may interfere with the sequence contiguity of assemblies from short reads (<1,000 bp) (8, 9). Therefore, even though Illumina is one of the most established platforms, more recent "third-generation" approaches performing at longer read lengths, such as the approach with the PacBio platform (10), might be more suitable for sequencing *M. ulcerans* genomes.

Accession number(s). The draft genome of *M. ulcerans* S4018 sequenced under this project has been deposited at DDBJ/EMBL/GenBank under the accession number MDUB00000000. The version described in this paper is the first version, MDUB01000000.

ACKNOWLEDGMENTS

We acknowledge Tim Stinear and Jessica Porter from the University of Melbourne for their technical advice, and we are most grateful to the GenoBIRD Core Facility for its technical support.

This work was supported by the ARMINA consortium (Alliance de Recherche sur les Maladies Infectieuses Nantes-Angers) of La Région des Pays de la Loire, France (grant no. 201209680), and Fondation Raoul Follereau.

REFERENCES

1. Johnson PDR, Stinear T, Small PLC, Pluschke G, Merritt RW, Portaels F, Huygen K, Hayman JA, Asiedu K. 2005. Buruli ulcer (*M. ulcerans* infection): new insights, new hope for disease control. PLoS Med 2:e108. https://doi.org/10.1371/journal.pmed.0020108.
2. Stinear TP, Seemann T, Pidot S, Frigui W, Reysset G, Garnier T, Meurice G, Simon D, Bouchier C, Ma L, Tichit M, Porter JL, Ryan J, Johnson PDR, Davies JK, Jenkin GA, Small PLC, Jones LM, Tekaia F, Laval F, Daffé M, Parkhill J, Cole ST. 2007. Reductive evolution and niche adaptation inferred from the genome of *Mycobacterium ulcerans*, the causative agent of Buruli ulcer. Genome Res 17:192–200. https://doi.org/10.1101/gr.5942807.
3. Zerbino DR, Birney E. 2008. Velvet: algorithms for de novo short read assembly using de Bruijn graphs. Genome Res 18:821–829. https://doi.org/10.1101/gr.074492.107.
4. Zerbino DR. 2010. Using the Velvet *de novo* assembler for short-read sequencing technologies. Curr Protoc Bioinformatics Chapter 11:Unit 11.5.
5. Darling ACE, Mau B, Blattner FR, Perna NT. 2004. Mauve: multiple alignment of conserved genomic sequence with rearrangements. Genome Res 14:1394–1403. https://doi.org/10.1101/gr.2289704.
6. Angiuoli SV, Gussman A, Klimke W, Cochrane G, Field D, Garrity G, Kodira CD, Kyrpides N, Madupu R, Markowitz V, Tatusova T, Thomson N, White O. 2008. Toward an online repository of standard operating procedures (SOPs) for (meta)genomic annotation. OMICS 12:137–141. https://doi.org/10.1089/omi.2008.0017.
7. Stinear T, Ross BC, Davies JK, Marino L, Robins-Browne RM, Oppedisano F, Sievers A, Johnson PDR, Marino LUI, Robins-browne ROYM. 1999. Identification and characterization of IS 2404 and IS 2606: two distinct repeated sequences for detection of *Mycobacterium ulcerans* by PCR. J Clin Microbiol 37:1018–1023.
8. Kingsford C, Schatz MC, Pop M. 2010. Assembly complexity of prokaryotic genomes using short reads. BMC Bioinformatics 11:21. https://doi.org/10.1186/1471-2105-11-21.
9. Barbosa EG, Aburjaile FF, Ramos RT, Carneiro AR, Le Loir Y, Baumbach J, Miyoshi A, Silva A, Azevedo V. 2014. Value of a newly sequenced bacterial genome. Biol Chem 5:161–168. https://doi.org/10.4331/wjbc.v5.i2.161.
10. Rhoads A, Au KF. 2015. PacBio sequencing and its applications. Genomics Proteomics Bioinformatics 13:278–289. https://doi.org/10.1016/j.gpb.2015.08.002.