



Draft Genome Sequence of *Leishmania tarentolae* Parrot Tar II, Obtained by Single-Molecule Real-Time Sequencing

Yasuyuki Goto,^a Akihiro Kuroki,^a Kengo Suzuki,^{b,c} Junya Yamagishi^{d,e}

^aLaboratory of Molecular Immunology, Department of Animal Resource Sciences, Graduate School of Agricultural and Life Sciences, The University of Tokyo, Tokyo, Japan

^beuglena Co., Ltd., Tokyo, Japan

^cMicroalgae Production Control Technology Laboratory, RIKEN, Yokohama, Japan

^dResearch Center for Zoonosis Control, Hokkaido University, Sapporo, Japan

^eGlobal Station for Zoonosis Control, GI-CoRE, Hokkaido University, Sapporo, Japan

ABSTRACT *Leishmania tarentolae* is a protozoan parasite of lizards and is non-pathogenic to mammals. Genome information for the nonpathogenic species will facilitate an understanding of the pathology caused by species pathogenic to mammals. Here, we report resequencing of the *L. tarentolae* genome with a single-molecule real-time (SMRT) sequencer to provide a more complete genome assembly.

Protozoan parasites of the genus *Leishmania* are causative agents of leishmaniasis. There are around 20 species causing the disease in humans, which can be mainly classified into three forms, namely, cutaneous leishmaniasis, mucocutaneous leishmaniasis, and visceral leishmaniasis, depending on the infecting species. Genetic comparison of *Leishmania* species causing different disease forms has facilitated understanding of the pathology and immunology of the disease (1). In addition, the genetic information is useful for identification of leishmanial targets good for the development of drugs, vaccines, or diagnostic tests. For example, we previously reported that antigens of serological significance can be identified *in silico* by searching the parasite genome for genes encoding proteins with tandem repeats (TRs) (2). To identify molecules responsible for the virulence of *Leishmania* parasites, genetic comparison between pathogenic and nonpathogenic species is a useful approach. This is why *Leishmania tarentolae*, which is found in lizards and is thought to be nonpathogenic to mammals, was sequenced, and its genome was compared with that of pathogenic *Leishmania* species (3). The *L. tarentolae* genome lacks many of the TR genes found in other *Leishmania* parasites, which represents either evolutionary loss of genes responsible for parasitism to mammals or just annotation errors due to short-read sequencing using Roche 454 and Illumina systems.

In order to elucidate the loss of TR genes in *L. tarentolae*, we resequenced the genome of *L. tarentolae* Parrot Tar II (ATCC 30267). Total DNA was isolated, using the blood and cell culture DNA midikit (Qiagen), from promastigotes grown in brain heart infusion medium supplemented with 10% heat-inactivated fetal bovine serum and 10 μ g/ml hemin. Five micrograms of DNA was fragmented to 20 kbp using a g-TUBE (Covaris, Inc.) and ligated with adaptors (SMRTBell template preparation kit v1.0) to prepare the SMRTBell structure. The resulting library was sequenced with a PacBio RS II system (Pacific Biosciences, USA) using 4 single-molecule real-time (SMRT) cells.

Using the HGAP3 assembler (with default parameters except as follows: expected genome size, 30 Gbp; target coverage, 30), a polished sequence with 510,681 reads covering 4,313,522,161 nucleotides and an assembled sequence with 35,416,496 bp in 179 contigs (genome coverage, \sim 120 \times ; maximum length, 2,700,252 bp; N_{50} ,

Citation Goto Y, Kuroki A, Suzuki K, Yamagishi J. 2020. Draft genome sequence of *Leishmania tarentolae* Parrot Tar II, obtained by single-molecule real-time sequencing. Microbiol Resour Announc 9:e00050-20. <https://doi.org/10.1128/MRA.00050-20>.

Editor Christina A. Cuomo, Broad Institute

Copyright © 2020 Goto et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

Address correspondence to Yasuyuki Goto, aygoto@mail.ecc.u-tokyo.ac.jp, or Junya Yamagishi, junya@czc.hokudai.ac.jp.

Received 8 February 2020

Accepted 28 April 2020

Published 21 May 2020

663,019 bp; GC content, 57.41%) were obtained. The genome was larger than that of *L. tarentolae* in the previous paper (30,440,719 bp) (3), *Leishmania major* (32,816,678 bp), *Leishmania infantum* (32,134,935 bp), and *Leishmania braziliensis* (32,005,207 bp) (1, 4).

Comparison between the previous and current *L. tarentolae* genomes revealed that the size difference (~5 million bp) can be partly explained by missing repeat sequences in the previous version. For example, a gene containing multiple copies of a 219-bp repeat unit can be found in all three pathogenic *Leishmania* species but is absent in the previous *L. tarentolae* genome; our data show that the repeat motif is also present in *L. tarentolae*, although it has a nonsense mutation within the repeat domain. Thus, resequencing of parasite genomes with a long-read sequencer can contribute to accurate identification of repeat motifs in the genomes, which may lead to an understanding of the evolution of TR genes and their significance for adaptation to mammals.

Data availability. These data are available under GenBank assembly accession number [GCA_009731335](https://www.ncbi.nlm.nih.gov/assembly/GCA_009731335) and SRA accession number [DRX197845](https://www.ncbi.nlm.nih.gov/sra/DRX197845).

ACKNOWLEDGMENTS

This work was supported by KAKENHI (grants 23689024 and 18H02649) from the Japan Society for the Promotion of Science and a research collaboration grant from the Research Center for Zoonosis Control, Hokkaido University. This work was also sponsored by Euglena Co., Ltd.

REFERENCES

1. Peacock CS, Seeger K, Harris D, Murphy L, Ruiz JC, Quail MA, Peters N, Adlem E, Tivey A, Aslett M, Kerhornou A, Ivens A, Fraser A, Rajandream M-A, Carver T, Norbertczak H, Chillingworth T, Hance Z, Jagels K, Moule S, Ormond D, Rutter S, Squares R, Whitehead S, Rabbinowitsch E, Arrow-smith C, White B, Thurston S, Bringaud F, Baldauf SL, Faulconbridge A, Jeffares D, Depledge DP, Oyola SO, Hilley JD, Brito LO, Tosi LRO, Barrell B, Cruz AK, Mottram JC, Smith DF, Berriman M. 2007. Comparative genomic analysis of three *Leishmania* species that cause diverse human disease. *Nat Genet* 39:839–847. <https://doi.org/10.1038/ng2053>.
2. Goto Y, Coler RN, Reed SG. 2007. Bioinformatic identification of tandem repeat antigens of the *Leishmania donovani* complex. *Infect Immun* 75: 846–851. <https://doi.org/10.1128/IAI.01205-06>.
3. Raymond F, Boisvert S, Roy G, Ritt JF, Légaré D, Isnard A, Stanke M, Olivier M, Tremblay MJ, Papadopoulou B, Ouellette M, Corbeil J. 2012. Genome sequencing of the lizard parasite *Leishmania tarentolae* reveals loss of genes associated to the intracellular stage of human pathogenic species. *Nucleic Acids Res* 40:1131–1147. <https://doi.org/10.1093/nar/gkr834>.
4. Ivens AC, Peacock CS, Worthey EA, Murphy L, Aggarwal G, Berriman M, Sisk E, Rajandream MA, Adlem E, Aert R, Anupama A, Apostolou Z, Attipoe P, Bason N, Bauser C, Beck A, Beverley SM, Bianchetti G, Borzym K, Bothe G, Bruschi CV, Collins M, Cadag E, Ciarloni L, Clayton C, Coulson RM, Cronin A, Cruz AK, Davies RM, De Gaudenzi J, Dobson DE, Duesterhoeft A, Fazelina G, Fosker N, Frasch AC, Fraser A, Fuchs M, Gabel C, Goble A, Goffeau A, Harris D, Hertz-Fowler C, Hilbert H, Horn D, Huang Y, Klages S, Knights A, Kube M, Larke N, Litvin L, Lord A, Louie T, Marra M, Masuy D, Matthews K, Michaeli S, Mottram JC, Müller-Auer S, Munden H, Nelson S, Norbertczak H, Oliver K, O'Neil S, Pentony M, Pohl TM, Price C, Purnelle B, Quail MA, Rabbinowitsch E, Reinhardt R, Rieger M, Rinta J, Robben J, Robertson L, Ruiz JC, Rutter S, Saunders D, Schäfer M, Schein J, Schwartz DC, Seeger K, Seyler A, Sharp S, Shin H, Sivam D, Squares R, Squares S, Tosato V, Vogt C, Volckaert G, Wambutt R, Warren T, Wedler H, Woodward J, Zhou S, Zimmermann W, Smith DF, Blackwell JM, Stuart KD, Barrell B, Myler PJ. 2005. The genome of the kinetoplastid parasite, *Leishmania major*. *Science* 309:436–442. <https://doi.org/10.1126/science.1112680>.