AMERICAN SOCIETY FOR MICROBIOLOGY | Microbiology Resource Announcements

GENOME SEQUENCES

Trypanosoma cruzi Ikiakarora (TcIII) Draft Genome Sequence

Inmaculada Gómez,[a] Alberto Rastrojo,[b] Fabián Lorenzo-Díaz,[c] Francisco José Sánchez-Luque,[a] Francisco Macías,[a] Begoña Aguado,[b] Basilio Valladares,[c] José María Requena,[b] Manuel Carlos López,[a] M. Carmen Thomas[a]

[a]Instituto de Parasitología y Biomedicina López-Neyra, Consejo Superior de Investigaciones Científicas, PTS-Granada, Granada, Spain
[b]Centro de Biología Molecular Severo-Ochoa (CSIC-UAM), Consejo Superior de Investigaciones Científicas, Madrid, Spain
[c]Instituto Universitario de Enfermedades Tropicales y Salud Pública de Canarias, Universidad de La Laguna, La Laguna, Spain

ABSTRACT *Trypanosoma cruzi* shows a genetic diversity that has been associated with the variability of clinical manifestations, geographical distribution, and preferential parasite-vector interactions. In an effort to better understand this genetic variability, here, the draft genome of *T. cruzi* strain Ikiakarora (discrete typing unit TcIII), which has been associated with the sylvatic cycle, is reported.

Trypanosoma cruzi is a protozoan parasite that causes Chagas disease, one of the biggest public health problems in Latin America, with more than 10,000 deaths annually (1). Currently, the dynamics of migration of people from countries of endemicity has favored the spreading of Chagas disease to the United States, Canada, Europe, and the western Pacific. (2–4). This parasite presents a high level of genetic variability, being classified into six discrete typing units (DTUs), TcI to TcVI (5). TcIII is concentrated in South America and is associated mainly with the sylvatic cycle and the terrestrial niche (6). Here, the draft genome of *T. cruzi* strain Ikiakarora (IRHO/CO/95), which belongs to DTU III as characterized by 24Sα rRNA, miniexon, and 18S rRNA markers, is reported (7). This strain was isolated in Catatumbo (North Santander, Colombia) from the sylvatic vector *Rhodnius prolixus*, the second most frequent transmitting vector of Chagas disease (8). Analyses of the H2A gene units of different strains isolated in Colombia showed that the genome of strain Ikiakarora has a high degree of plasticity (9).

Epimastigote forms were cultivated at 28°C in liver infusion tryptose (LIT) medium supplemented with 10% fetal bovine serum. When the parasites reached log phase (10×10^6 to 20×10^6 parasites/ml), they were collected and lysed in 1% NP-40. Nuclei were lysed by addition of 1% SDS, and genomic DNA was purified by phenol-chloroform extraction and ethanol precipitation. Sequencing was carried out using Ion Torrent technology (Thermo Fischer Scientific, Inc.). Library construction, size selection, quality filtering, DNA concentration, and processing were performed as previously described (10), obtaining 3,928,712 raw reads (average read length, 254 bp).

Sequence reads were analyzed using FastQC v0.10.1 (default settings) (http://www.bioinformatics.babraham.ac.uk/projects/fastqc), and Prinseq v0.20.4 (11) was used iteratively for quality filtering using the following parameters: -derep 14, -ns_max_p 1 -ns_max_n 3 -trim_ns_left 1 -trim_ns_right 1, -trim_qual_right 20 -trim_qual_type mean -trim_qual_window 5 -trim_qual_step 1, -trim_qual_right 20 -trim_qual_type mean -trim_qual_window 1 -trim_qual_step 1, -trim_qual_left 20 -trim_qual_type mean -trim_qual_window 5 -trim_qual_step 1, -trim_qual_left 20 -trim_qual_type mean -trim_qual_window 1 -trim_qual_step 1, -lc_method entropy -lc_threshold 50, -min_qual_mean 25, and -min_len 50. The obtained 3,338,764 quality-filtered reads (average read length, 266 bp) were assembled into 11,096 contigs totaling 18,492,845 bp, with an N_{50} value of 2,193 bp and an average contig size of 1,666 bp using CLC Genomics Workbench software v8.0 (Qiagen) (length fraction, 0.90; similarity fraction, 0.97;

Citation Gómez I, Rastrojo A, Lorenzo-Díaz F, Sánchez-Luque FJ, Macías F, Aguado B, Valladares B, Requena JM, López MC, Thomas MC. 2020. *Trypanosoma cruzi* Ikiakarora (TcIII) draft genome sequence. Microbiol Resour Announc 9:e00453-20. https://doi.org/10.1128/MRA.00453-20.

Editor Jason E. Stajich, University of California, Riverside

Address correspondence to M. Carmen Thomas, mcthomas@ipb.csic.es.

Received 22 April 2020
Accepted 9 June 2020
Published 2 July 2020

minimum contig length, 500 bp). The longest contig was 33,607 bp, and the G+C content was 48.72%. To assess genome assembly completeness, BUSCO (Benchmarking Universal Single-Copy Orthologs) v4.0.5 analysis (12) (parameters: -m genome, -l euglenozoa_odb10) was performed on the assembled genome using the Euglenozoa odb10 orthologue set ($n = 130$). A total of 106 complete BUSCOs (81.6%) and 102 single-copy BUSCOs (78.5%) were identified from 130 BUSCO-searched groups.

Data availability. The *T. cruzi* Ikiakarora assembled genome has been deposited in GenBank (accession number WWPZ00000000), and the raw reads have been deposited in the SRA (accession numbers SRR11235102, SRR11235103, and SRR11235104 and BioProject accession number PRJNA595095).

ACKNOWLEDGMENTS

We are grateful to Almudena López-Barajas (IPBLN) for her technical collaboration in the identification of the lineage of strain Ikiakarora.

This work was supported by grant SAF2016-80998-R from the Programa Estatal I+D+i (MINECO) and the Network of Tropical Diseases Research RICET (RD16/0027/0001, RD16/0027/0005, and RD16/0027/0008) and FEDER.

This work is part of the Ph.D. thesis of student Inmaculada Gómez at the University of Granada in the Biochemistry and Molecular Biology Program.

REFERENCES

1. Anonymous. 2011. Working to overcome the global impact of neglected tropical diseases—summary. Wkly Epidemiol Rec 86:113–120.
2. Bern C, Montgomery SP. 2009. An estimate of the burden of Chagas disease in the United States. Clin Infect Dis 49:e52–e54. https://doi.org/10.1086/605091.
3. Gascon J, Bern C, Pinazo MJ. 2010. Chagas disease in Spain, the United States and other non-endemic countries. Acta Trop 115:22–27. https://doi.org/10.1016/j.actatropica.2009.07.019.
4. Guerri-Guttenberg RA, Grana DR, Ambrosio G, Milei J. 2008. Chagas cardiomyopathy: Europe is not spared! Eur Heart J 29:2587–2591. https://doi.org/10.1093/eurheartj/ehn424.
5. Zingales B, Andrade SG, Briones MR, Campbell DA, Chiari E, Fernandes O, Guhl F, Lages-Silva E, Macedo AM, Machado CR, Miles MA, Romanha AJ, Sturm NR, Tibayrenc M, Schijman AG, Second Satellite M, Second Satellite Meeting. 2009. A new consensus for Trypanosoma cruzi intraspecific nomenclature: second revision meeting recommends TcI to TcVI. Mem Inst Oswaldo Cruz 104:1051–1054. https://doi.org/10.1590/s0074-02762009000700021.
6. Zingales B, Miles MA, Campbell DA, Tibayrenc M, Macedo AM, Teixeira MM, Schijman AG, Llewellyn MS, Lages-Silva E, Machado CR, Andrade SG, Sturm NR. 2012. The revised Trypanosoma cruzi subspecific nomenclature: rationale, epidemiological relevance and research applications. Infect Genet Evol 12:240–253. https://doi.org/10.1016/j.meegid.2011.12.009.
7. Brisse S, Verhoef J, Tibayrenc M. 2001. Characterisation of large and small subunit rRNA and mini-exon genes further supports the distinction of six Trypanosoma cruzi lineages. Int J Parasitol 31:1218–1226. https://doi.org/10.1016/s0020-7519(01)00238-7.
8. Rodriguez P, Montilla M, Nicholls S, Zarante I, Puerta C. 1998. Isoenzymatic characterization of Colombian strains of Trypanosoma cruzi. Mem Inst Oswaldo Cruz 93:739–740. https://doi.org/10.1590/s0074-02761998000600008.
9. Thomas MC, Olivares M, Escalante M, Maranon C, Montilla M, Nicholls S, Lopez MC, Puerta C. 2000. Plasticity of the histone H2A genes in a Brazilian and six Colombian strains of Trypanosoma cruzi. Acta Trop 75:203–210. https://doi.org/10.1016/S0001-706X(00)00061-9.
10. Gómez I, Rastrojo A, Sanchez-Luque FJ, Lorenzo-Díaz F, Macías F, Valladares B, Aguado B, Requena JM, López MC, Thomas MC. 2020. Draft genome sequence of Trypanosoma cruzi B. M. López strain (TcIa), isolated from a Colombian patient. Microbiol Resour Announc 9:e00031-20. https://doi.org/10.1128/MRA.00031-20.
11. Schmieder R, Edwards R. 2011. Quality control and preprocessing of metagenomic datasets. Bioinformatics 27:863–864. https://doi.org/10.1093/bioinformatics/btr026.
12. Seppey M, Manni M, Zdobnov EM. 2019. BUSCO: assessing genome assembly and annotation completeness. Methods Mol Biol 1962:227–245. https://doi.org/10.1007/978-1-4939-9173-0_14.