

Draft Genome Sequence of Tubulinosema ratisbonensis , a Microsporidian Species Infecting the Model Organism Drosophila melanogaster

Valérie Polonais, Sebastian Niehus, Ivan Wawrzyniak, Adrien Franchet, Christine Gaspin, Abdel Belkorchia, Matthieu Matthieu.Reichstadt@inrae.Fr Reichstadt, Caroline Belser, Karine Labadie, Arnaud Couloux, et al.

▶ **To cite this version:**

Valérie Polonais, Sebastian Niehus, Ivan Wawrzyniak, Adrien Franchet, Christine Gaspin, et al.. Draft Genome Sequence of Tubulinosema ratisbonensis , a Microsporidian Species Infecting the Model Organism Drosophila melanogaster. Microbiology Resource Announcements, 2019, 8 (31), 10.1128/MRA.00077-19 . hal-02276377

HAL Id: hal-02276377
https://hal.science/hal-02276377

Submitted on 21 Mar 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

AMERICAN SOCIETY FOR MICROBIOLOGY | Microbiology Resource Announcements

GENOME SEQUENCES

Draft Genome Sequence of *Tubulinosema ratisbonensis*, a Microsporidian Species Infecting the Model Organism *Drosophila melanogaster*

Valérie Polonais,[a] Sebastian Niehus,[b] Ivan Wawrzyniak,[a] Adrien Franchet,[b] Christine Gaspin,[c] Abdel Belkorchia,[a] Matthieu Reichstadt,[d,e] Caroline Belser,[f] Karine Labadie,[f] Arnaud Couloux,[f] Frédéric Delbac,[a] Eric Peyretaillade,[a] Dominique Ferrandon[b]

[a]Université Clermont Auvergne, CNRS, Laboratoire Microorganismes: Génome et Environnement, Clermont-Ferrand, France
[b]Université de Strasbourg CNRS, M3I UPR 9022, Strasbourg, France
[c]Université de Toulouse, INRA, UR 875 Unité Mathématiques et Informatique Appliquées de Toulouse, Auzeville, Castanet-Tolosan, France
[d]INRA, Saint-Genès-Champanelle, France
[e]UMR1213 Herbivores, Clermont Université, VetAgro Sup, Clermont-Ferrand, France
[f]Genoscope, Institut de Biologie François-Jacob, Commissariat à l'Energie Atomique (CEA), Université Paris-Saclay, Evry, France

ABSTRACT We present the draft genome sequence of *Tubulinosema ratisbonensis*, a microsporidium species infecting *Drosophila melanogaster*. A total of 3,013 protein-encoding genes and an array of transposable elements were identified. This work represents a necessary step to develop a novel model of host-parasite relationships using the highly tractable genetic model *D. melanogaster*.

Tubulinosema ratisbonensis belongs to the *Microsporidia*, a group of obligate intracellular parasites that group near the base of the fungal radiation (1). *T. ratisbonensis* is a pathogen of *Drosophila melanogaster*, responsible for fly culture collapse (2). This species presents a unique opportunity to study and better understand specific host-parasite interactions and also to decipher the adaptation capacities of this parasite. For example, a recent study demonstrated that phosphatidic acid is a limiting host metabolite for the proliferation of *T. ratisbonensis* in *Drosophila* flies (3).

The *T. ratisbonensis* strain isolated from *D. melanogaster* was propagated in the MRC-5 human cell line (4, 5). Spores were isolated and purified from host cells as previously described (6). *T. ratisbonensis* genomic DNA was extracted according to Katinka et al. (7), and two mate-paired libraries of 5-kb nebulized fragments were constructed to perform pyrosequencing as described by Aury et al. (8). 454 sequencing was used to remove low-quality sequences and generated 1,031,738 useful reads with an average size of 250 bp. In addition, genomic DNA was also randomly fragmented by sonication to produce an Illumina library according to Choulet et al. (9). An in-house quality control process based on the FastX package (http://www.genoscope.cns.fr/externe/fastxtend/) was applied to the reads that passed the Illumina quality filters defined by Alberti et al. (10). The 44,119,481 obtained single reads with an average length of 100 bp were assembled with pyrosequencing data using Newbler version 2.7 (11) with default parameters. To improve genome assembly, total RNAs were extracted from *Drosophila* S2R+ cells after 12 hours and 6 days postinfection with *T. ratisbonensis* using an RNeasy kit (Qiagen) according to the manufacturer's protocol. Transcriptome sequencing (RNA-seq) Illumina libraries were then obtained as described by Lee et al. (12), and the 54,845,250 single-read sequences with an average length of 54 bp were assembled using the IDBA-UD assembler version 1.1.1 (13). Finally, *T. ratisbonensis* contigs obtained from whole-genome sequencing and RNA-seq approaches were assembled using the CAP3 program from the Geneious platform (14).

Citation Polonais V, Niehus S, Wawrzyniak I, Franchet A, Gaspin C, Belkorchia A, Reichstadt M, Belser C, Labadie K, Couloux A, Delbac F, Peyretaillade E, Ferrandon D. 2019. Draft genome sequence of *Tubulinosema ratisbonensis*, a microsporidian species infecting the model organism *Drosophila melanogaster*. Microbiol Resour Announc 8:e00077-19. https://doi.org/10.1128/MRA.00077-19.

Editor Vincent Bruno, University of Maryland School of Medicine

Copyright © 2019 Polonais et al. This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International license.

Address correspondence to Eric Peyretaillade, eric.peyretaillade@uca.fr, or Dominique Ferrandon, D.Ferrandon@ibmc-cnrs.unistra.fr.

V.P. and S.N. contributed equally to this study.

Received 18 February 2019
Accepted 5 July 2019
Published 1 August 2019

Downloaded from https://journals.asm.org/journal/mra on 07 February 2022 by 130.79.37.131.

The resulting draft genome's total size is 7,555,974 nucleotides (nt) spread over 952 contigs with an overall GC content of 23.2%. The N_{50} contig length is 12,152 bp, and the largest contig is composed of 59,712 bp. From these assembly data, a total of 3,013 coding DNA sequences (CDSs) were predicted using a dedicated microsporidian annotation pipeline (E. Peyretaillade, M. Reichstadt, S. Rimour, I. Wawrzyniak, and F. Delbac, unpublished data) combining homology search, transcriptional regulatory motif detection, and GLIMMER gene prediction (15, 16). Functional annotation conducted using Blast2GO software (17) with default parameters revealed that 2,099 CDSs (70%) present at least one ortholog in microsporidia, and 44.5% can be assigned to putative functions consistent with other microsporidium genomic data (18). Noncoding rRNA and tRNA genes were detected using RNammer (19) and tRNAScan-SE version 1.23 (20), respectively. Thus, six 5S rRNAs, a contig including the 16S-23S rRNA unit, and 55 tRNA genes were annotated. Additional noncoding RNAs were identified using a comparative analysis following a previously described approach (21) and *Anncaliia algerae* genomic data. Transposable elements (TEs) were obtained from tblastx analysis (E value cutoff of 1e-5) using TE sequences previously identified in numerous microsporidian species (22). An exhaustive search of TEs has also been carried out with TransposonPSI software (http://transposonpsi.sourceforge.net/) to identify highly variable subclasses of TEs. TE family affiliation was also checked using RepeatMasker (http://www.repeatmasker.org/). A total of 160 sequences were identified as TEs and were classified among the following 3 major families: long terminal repeat (LTR) retrotransposon (50.6%), DNA Mariner (20.6%), and DNA Merlin (20.6%). Among the 95 complete transposase sequences, 67.4% present a CCC/GGG-like signal upstream of the transcriptional initiation signals (16), suggesting that these sequences were under the control of *T. ratisbonensis* transcriptional machinery.

Data availability. Annotated contig sequences have been deposited at DDBJ/ENA/GenBank under the accession number RCSS00000000. The version described in this paper is version RCSS01000000. All the data produced for this project are available in NCBI databases under the accession numbers PRJEB31246 and PRJNA494796.

ACKNOWLEDGMENTS

This work was supported by the BEELOSS ANR project (ANR-12-BSV3-0020).

We thank the AuBi platform group (https://www.france-bioinformatique.fr/fr/plateformes/aubi) and the Mésocentre of Clermont Auvergne University (https://mesocentre.uca.fr/), C. Franzen (University of Regensburg, Germany) for the gift of the *T. ratisbonensis* strain used in this study, and L. Troxler for preprocessing the RNA-seq data.

REFERENCES

1. Bass D, Czech L, Williams BAP, Berney C, Dunthorn M, Mahe F, Torruella G, Stentiford GD, Williams TA. 2018. Clarifying the relationships between Microsporidia and Cryptomycota. J Eukaryot Microbiol 65:773–782. https://doi.org/10.1111/jeu.12519.
2. Niehus S, Giammarinaro P, Liegeois S, Quintin J, Ferrandon D. 2012. Fly culture collapse disorder: detection, prophylaxis and eradication of the microsporidian parasite Tubulinosema ratisbonensis infecting Drosophila melanogaster. Fly (Austin) 6:193–204. https://doi.org/10.4161/fly.20896.
3. Franchet A, Niehus S, Caravello G, Ferrandon D. 2019. Phosphatidic acid as a limiting host metabolite for the proliferation of the microsporidium Tubulinosema ratisbonensis in Drosophila flies. Nat Microbiol 4:645–655. https://doi.org/10.1038/s41564-018-0344-y.
4. Franzen C, Fischer S, Schroeder J, Bleiss W, Schneuwly S, Scholmerich J, Salzberger B. 2005. In vitro cultivation of an insect Microsporidian Tubulinosema ratisbonensis in mammalian cells. J Eukaryot Microbiol 52:349–355. https://doi.org/10.1111/j.1550-7408.2005.00043x.
5. Franzen C, Fischer S, Schroeder J, Scholmerich J, Schneuwly S. 2005. Morphological and molecular investigations of Tubulinosema ratisbonensis gen. nov., sp. nov. (Microsporidia: Tubulinosematidae fam. nov.), a parasite infecting a laboratory colony of Drosophila melanogaster (Diptera: Drosophilidae). J Eukaryot Microbiol 52:141–152. https://doi.org/10.1111/j.1550-7408.2005.04-3324.x.
6. Belkorchia A, Biderre C, Militon C, Polonais V, Wincker P, Jubin C, Delbac F, Peyretaillade E, Peyret P. 2008. In vitro propagation of the microsporidian pathogen Brachiola algerae and studies of its chromosome and ribosomal DNA organization in the context of the complete genome sequencing project. Parasitol Int 57:62–71. https://doi.org/10.1016/j.parint.2007.09.002.
7. Katinka MD, Duprat S, Cornillot E, Méténier G, Thomarat F, Prensier G, Barbe V, Peyretaillade E, Brottier P, Wincker P, Delbac F, El Alaoui H, Peyret P, Saurin W, Gouy M, Weissenbach J, Vivarès CP. 2001. Genome sequence and gene compaction of the eukaryote parasite Encephalitozoon cuniculi. Nature 414:450–453. https://doi.org/10.1038/35106579.
8. Aury JM, Cruaud C, Barbe V, Rogier O, Mangenot S, Samson G, Poulain J, Anthouard V, Scarpelli C, Artiguenave F, Wincker P. 2008. High quality draft sequences for prokaryotic genomes using a mix of new sequencing technologies. BMC Genomics 9:603. https://doi.org/10.1186/1471-2164-9-603.
9. Choulet F, Alberti A, Theil S, Glover N, Barbe V, Daron J, Pingault L,

Downloaded from https://journals.asm.org/journal/mra on 07 February 2022 by 130.79.37.131.

Sourdille P, Couloux A, Paux E, Leroy P, Mangenot S, Guilhot N, Le Gouis J, Balfourier F, Alaux M, Jamilloux V, Poulain J, Durand C, Bellec A, Gaspin C, Safar J, Dolezel J, Rogers J, Vandepoele K, Aury JM, Mayer K, Berges H, Quesneville H, Wincker P, Feuillet C. 2014. Structural and functional partitioning of bread wheat chromosome 3B. Science 345:1249721. https://doi.org/10.1126/science.1249721.

10. Alberti A, Poulain J, Engelen S, Labadie K, Romac S, Ferrera I, Albini G, Aury J-M, Belser C, Bertrand A, Cruaud C, Da Silva C, Dossat C, Gavory F, Gas S, Guy J, Haquelle M, Jacoby E, Jaillon O, Lemainque A, Pelletier E, Samson G, Wessner M, Bazire P, Beluche O, Bertrand L, Besnard-Gonnet M, Bordelais I, Boutard M, Dubois M, Dumont C, Ettedgui E, Fernandez P, Garcia E, Aiach NG, Guerin T, Hamon C, Brun E, Lebled S, Lenoble P, Louesse C, Mahieu E, Mairey B, Martins N, Megret C, Milani C, Muanga J, Orvain C, Payen E, Perroud P, et al. 2017. Viral to metazoan marine plankton nucleotide sequences from the Tara Oceans expedition. Sci Data 4:170093. https://doi.org/10.1038/sdata.2017.93.
11. Miller JR, Koren S, Sutton G. 2010. Assembly algorithms for next-generation sequencing data. Genomics 95:315–327. https://doi.org/10.1016/j.ygeno.2010.03.001.
12. Lee KZ, Lestradet M, Socha C, Schirmeier S, Schmitz A, Spenle C, Lefebvre O, Keime C, Yamba WM, Bou Aoun R, Liegeois S, Schwab Y, Simon-Assmann P, Dalle F, Ferrandon D. 2016. Enterocyte purge and rapid recovery is a resilience reaction of the gut epithelium to pore-forming toxin attack. Cell Host Microbe 20:716–730. https://doi.org/10.1016/j.chom.2016.10.010.
13. Peng Y, Leung HC, Yiu SM, Chin FY. 2012. IDBA-UD: a de novo assembler for single-cell and metagenomic sequencing data with highly uneven depth. Bioinformatics 28:1420–1428. https://doi.org/10.1093/bioinformatics/bts174.
14. Kearse M, Moir R, Wilson A, Stones-Havas S, Cheung M, Sturrock S, Buxton S, Cooper A, Markowitz S, Duran C, Thierer T, Ashton B, Meintjes P, Drummond A. 2012. Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. Bioinformatics 28:1647–1649. https://doi.org/10.1093/bioinformatics/bts199.
15. Peyretaillade E, Goncalves O, Terrat S, Dugat-Bony E, Wincker P, Cornman RS, Evans JD, Delbac F, Peyret P. 2009. Identification of transcriptional signals in Encephalitozoon cuniculi widespread among Microsporidia phylum: support for accurate structural genome annotation. BMC Genomics 10:607. https://doi.org/10.1186/1471-2164-10-607.
16. Peyretaillade E, Parisot N, Polonais V, Terrat S, Denonfoux J, Dugat-Bony E, Wawrzyniak I, Biderre-Petit C, Mahul A, Rimour S, Goncalves O, Bornes S, Delbac F, Chebance B, Duprat S, Samson G, Katinka M, Weissenbach J, Wincker P, Peyret P. 2012. Annotation of microsporidian genomes using transcriptional signals. Nat Commun 3:1137. https://doi.org/10.1038/ncomms2156.
17. Gotz S, Garcia-Gomez JM, Terol J, Williams TD, Nagaraj SH, Nueda MJ, Robles M, Talon M, Dopazo J, Conesa A. 2008. High-throughput functional annotation and data mining with the Blast2GO suite. Nucleic Acids Res 36:3420–3435. https://doi.org/10.1093/nar/gkn176.
18. Pombert JF, Haag KL, Beidas S, Ebert D, Keeling PJ. 2015. The Ordospora colligata genome: evolution of extreme reduction in microsporidia and host-to-parasite horizontal gene transfer. mBio 6:e02400-14. https://doi.org/10.1128/mBio.02400-14.
19. Lagesen K, Hallin P, Rodland EA, Staerfeldt HH, 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.
20. Lowe TM, Eddy SR. 1997. tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. Nucleic Acids Res 25:955–964. https://doi.org/10.1093/nar/25.5.955.
21. Belkorchia A, Pombert JF, Polonais V, Parisot N, Delbac F, Brugere JF, Peyret P, Gaspin C, Peyretaillade E. 2017. Comparative genomics of microsporidian genomes reveals a minimal non-coding RNA set and new insights for transcription in minimal eukaryotic genomes. DNA Res 24:251–260. https://doi.org/10.1093/dnares/dsx002.
22. Parisot N, Pelin A, Gasc C, Polonais V, Belkorchia A, Panek J, El Alaoui H, Biron DG, Brasset E, Vaury C, Peyret P, Corradi N, Peyretaillade E, Lerat E. 2014. Microsporidian genomes harbor a diverse array of transposable elements that demonstrate an ancestry of horizontal exchange with metazoans. Genome Biol Evol 6:2289–2300. https://doi.org/10.1093/gbe/evu178.

Downloaded from https://journals.asm.org/journal/mra on 07 February 2022 by 130.79.37.131.