Genome Sequence of *Corynebacterium ulcerans* Strain FRC11

Leandro de Jesus Benevides,[a] Marcus Vinicius Canário Viana,[a] Diego César Batista Mariano,[a] Flávia de Souza Rocha,[a] Priscilla Carolinne Bagano,[a] Edson Luiz Folador,[a] Felipe Luiz Pereira,[b] Fernanda Alves Dorella,[b] Carlos Augusto Gomes Leal,[b] Alex Fiorini Carvalho,[b] Siomar de Castro Soares,[b] Adriana Carneiro,[c] Rommel Ramos,[c] Edgar Badell-Ocando,[d] Nicole Guiso,[d] Artur Silva,[c] Henrique Figueiredo,[b] Vasco Azevedo,[a] Luis Carlos Guimarães[a]*

Laboratory of Cellular and Molecular Genetics (LGCM), Federal University of Minas Gerais, Belo Horizonte, Minas Gerais, Brazil[a]; National Reference Laboratory for Aquatic Animal Diseases (AQUACEN), Ministry of Fisheries and Aquaculture, Federal University of Minas Gerais, Belo Horizonte, Minas Gerais, Brazil[b]; Laboratory of Polymorphic DNA (LPDNA), Federal University of Para, Belém, Brazil[c]; Institut Pasteur, Unité de Prévention et Thérapies Moléculaires des Maladies Humaines, National Centre of Reference of Toxigenic Corynebacteria, Paris, France[d]

* Present address: Luis Carlos Guimarães, Departamento de Biologia Geral, ICB/UFMG, Pampulha, Belo Horizonte, Minas Gerais, Brazil.

Here, we present the genome sequence of *Corynebacterium ulcerans* strain FRC11. The genome includes one circular chromosome of 2,442,826 bp (53.35% G+C content), and 2,210 genes were predicted, 2,146 of which are putative protein-coding genes, with 12 rRNAs and 51 tRNAs; 1 pseudogene was also identified.

Address correspondence to Vasco Azevedo, vasco@icb.ufmg.br.

*C*orynebacterium ulcerans is a bacterium that presents catalase-positive, nitrate-negative, and urease-positive biochemical proprieties (1). This bacterium belongs to the *Actinobacteria* class, which includes the genera *Corynebacterium*, *Mycobacterium*, *Nocardia*, and *Rodococcus*, collectively termed the CMNR group. This is a very heterogeneous group; however, most of the species share particular characteristics, such as (i) a specific organization of the cell wall, which is mainly composed of peptidoglycans, arabinogalactans, and mycolic acids, and (ii) high G+C content (2–4).

Although *C. ulcerans* has increasing medical and veterinary importance, little is known about its lifestyle and associated virulence factors (5). The sequencing of more *C. ulcerans* genomes of both toxigenic and nontoxigenic strains will help in the identification of distinctive features of strains from human and animal sources (6). In addition, the data generated by newly sequenced genomes are helpful for identifying antibiotic and vaccine targets by way of a comparative analysis (7).

Nowadays, only seven complete genomes and two drafts are available in the National Center for Biotechnology Information (NCBI) database (http://www.ncbi.nlm.nih.gov/genome/). This scenario shows that more genomic knowledge is required in order to better characterize the virulence mechanisms of this emergent pathogen.

In the current study, we present the genome sequence of *C. ulcerans* strain FRC11, isolated from a 74-year-old human a with leg ulcerans infection in Toulouse, France. This strain was first identified as *Corynebacterium pseudotuberculosis* (8), but recent analysis shows that it belongs to *C. ulcerans*.

The sequencing, assembly, and annotation of this strain were performed by the teams from the Laboratory of Cellular and Molecular Genetics (LGCM) and the National Reference Laboratory for Aquatic Animal Diseases (AQUACEN), both located at the Federal University of Minas Gerais, Belo Horizonte, Minas Gerais, Brazil, and the Laboratory of Polymorphic DNA (LPDNA) at the Federal University of Pará, Belém, Pará, Brazil.

The platform used for sequencing was the Ion Torrent Personal Genome Machine (PGM) system (Life Technologies), using a fragment library. The quality of the raw data was analyzed using the Web tool FastQC (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/). The assembly was done using the Simple Manager for Bacterial Assemblies (SIMBA) interface (http://ufmg-simba.sourceforge.net). The reads with good quality were assembled using a *de novo* strategy with the software MIRA 4.0 (9).

The assembly produced a total of 30 contigs, with a coverage of 179.14× and an N_{50} contig length of 236.335. Additionally, a scaffold was created using the CONTIGuator 2 software (10), using the genome sequence of *C. ulcerans* strain 0102 (accession no. NC_018101.1) (11) as a reference. The gap closure was performed automatically using SIMBA and manually using the CLC Genomics Workbench 7 software.

The genome was automatically annotated using Rapid Annotations using Subsystems Technology (RAST) (12). The manual curation of the annotation was performed using the Artemis software (13) and the UniProt database (http://www.uniprot.org). The CLC Genomics Workbench 7 software was used to correct indel errors in the regions of homopolymers.

The genome includes one circular chromosome of 2,442,826 bp (53.35% G+C content), and 2,210 genes were predicted, 2,146 of which are putative protein-coding genes, with 12 rRNAs and 51 tRNAs; 1 pseudogene was also identified.

Nucleotide sequence accession number. This genome has been deposited in GenBank under the accession no. CP009622.

REFERENCES

1. **Riegel P, Funke G.** 2000. Corynebacterium et bactéries apparentées, p 993–1019. *In* Freney J, Renaud F, Hansen W (ed), Précis de Bacteriologie Clinique. Edtions Alexandre Lacassagne, Paris, France.

2. **Dorella FA, Pacheco LG, Oliveira SC, Miyoshi A, Azevedo V.** 2006. *Corynebacterium pseudotuberculosis*: microbiology, biochemical properties, pathogenesis and molecular studies of virulence. Vet Res **37**:201–218. http://dx.doi.org/10.1051/vetres:2005056.

3. **Hassan SS, Guimarães LC, Pereira UDP, Islam A, Ali A, Bakhtiar SM, Ribeiro D, Rodrigues Dos Santos A, Soares SDC, Dorella F, Pinto AC, Schneider MP, Barbosa MS, Almeida S, Abreu V, Aburjaile F, Carneiro AR, Cerdeira LT, Fiaux K, Barbosa E, Diniz C, Rocha FS, Ramos RTJ, Jain N, Tiwari S, Barh D, Miyoshi A, Müller B, Silva A, Azevedo V.** 2012. Complete genome sequence of *Corynebacterium pseudotuberculosis* biovar ovis strain p54b96 isolated from antelope in South Africa obtained by rapid next generation sequencing technology. Stand Genomic Sci **7**:189–199. http://dx.doi.org/10.4056/sigs.3066455.

4. **Soares SC, Silva A, Trost E, Blom J, Ramos R, Carneiro A, Ali A, Santos AR, Pinto AC, Diniz C, Barbosa EGV, Dorella Fa, Aburjaile F, Rocha FS, Nascimento KKF, Guimarães LC, Almeida S, Hassan SS, Bakhtiar SM, Pereira UP, Abreu VA, Schneider MP, Miyoshi A, Tauch A, Azevedo V.** 2013. The pan-genome of the animal pathogen *Corynebacterium pseudotuberculosis* reveals differences in genome plasticity between the biovar ovis and equi strains. PLoS One **8**:e53818. http://dx.doi.org/10.1371/journal.pone.0053818.

5. **LaPointe P, Wei X, Gariépy J.** 2005. A role for the protease-sensitive loop region of Shiga-like toxin 1 in the retrotranslocation of its A1 domain from the endoplasmic reticulum lumen. J Biol Chem **280**:23310–23318. http://dx.doi.org/10.1074/jbc.M414193200.

6. **Trost E, Al-Dilaimi A, Papavasiliou P, Schneider J, Viehoever P, Burkovski A, Soares SC, Almeida SS, Dorella Fa, Miyoshi A, Azevedo V, Schneider MP, Silva A, Santos CS, Santos LS, Sabbadini P, Dias AA, Hirata R, Jr, Mattos-Guaraldi AL, Tauch A.** 2011. Comparative analysis of two complete *Corynebacterium ulcerans* genomes and detection of candidate virulence factors. BMC Genomics **12**:383. http://dx.doi.org/10.1186/1471-2164-12-383.

7. **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. World J Biol Chem **5**:161–168.

8. **Guiso N.** 2008. Rapport Annuel d'activité 2008. Institut Pasteur, Paris, France. http://www.pasteur.fr/ip/resource/filecenter/document/01s-00004f-0q7/ra-cnr-coryne-2008.pdf.

9. **Chevreux B, Wetter T, Suhai S.** 1999. Genome sequence assembly using trace signals and additional sequence information, p 45–56. *In* Computer science and biology. Proceedings of the German Conference on Bioinformatics, GCB '99. GCB, Hannover, Germany.

10. **Galardini M, Biondi EG, Bazzicalupo M, Mengoni A.** 2011. CONTIGuator: a bacterial genomes finishing tool for structural insights on draft genomes. Source Code Biol Med **6**:11. http://dx.doi.org/10.1186/1751-0473-6-11.

11. **Sekizuka T, Yamamoto A, Komiya T, Kenri T, Takeuchi F, Shibayama K, Takahashi M, Kuroda M, Iwaki M.** 2012. *Corynebacterium ulcerans* 0102 carries the gene encoding diphtheria toxin on a prophage different from the *C. diphtheriae* NCTC 13129 prophage. BMC Microbiol **12**:72. http://dx.doi.org/10.1186/1471-2180-12-72.

12. **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. http://dx.doi.org/10.1186/1471-2164-9-75.

13. **Rutherford K, Parkhill J, Crook J, Horsnell T, Rice P, Rajandream MA, Barrell B.** 2000. Artemis: sequence visualization and annotation. Bioinformatics **16**:944–945. http://dx.doi.org/10.1093/bioinformatics/16.10.944.