Complete Genome Sequence of *Borrelia crocidurae*

Haitham Elbir,ᵃ Grégory Gimenez,ᵃ Catherine Robert,ᵃ Sven Bergström,ᵇ Sally Cutler,ᶜ Didier Raoult,ᵃ and Michel Drancourtᵃ

Unité de Recherche sur les Maladies Infectieuses et Tropicales Emergentes (URMITE), UMR CNRS 6236 IRD 198, IFR48, Faculté de Médecine, Méditerranée Infection, Aix-Marseille-Université, Marseille, Franceᵃ; Department of Molecular Biology, Umeå University, Umeå, Swedenᵇ; and University of East London, London, United Kingdomᶜ

We announce the draft genome sequence of *Borrelia crocidurae* (strain Achema). The 1,557,560-bp genome (27% GC content) comprises one 919,477-bp linear chromosome and 638,083-bp plasmids that together carry 1,472 open reading frames, 32 tRNAs, and three complete rRNAs, with almost complete colinearity between *B. crocidurae* and *Borrelia duttonii* chromosomes.

Borrelia crocidurae is a spirochete responsible for tick-borne zoonotic relapsing fever in West Africa, where the soft tick *Ornithodoros sonrai* is the vector (2, 12). Relapsing fever mimics malaria, with which relapsing fever is often confused in the absence of laboratory confirmation (7). *B. crocidurae* remains a neglected organism that is rarely investigated using nonspecific laboratory tools.

The *B. crocidurae* Achema strain (11), grown in Barbour-Stoenner-Kelly H medium (Sigma, Saint-Quentin-Fallavier, France) supplemented with 6% rabbit serum (Eurobio, Courtaboeuf, France) at 33°C, was sequenced using pyrosequencing technology on a Roche 454 GS FLX sequencer (Roche, Boulogne-Billancourt, France). An average 24-fold genomic coverage was obtained over 22,610 paired-end reads, which were assembled using the Newbler 2.3 assembler (Roche), generating 41 scaffolds. Ordering, using the *Borrelia duttonii* genome (5) as a reference, left 39 plasmid scaffolds in draft status, whereas 2 chromosome scaffolds were closed by PCR sequencing. An in-house pipeline based on the Prodigal program was used to annotate the DNA sequences (3). tRNAs were predicted using the Aragorn program (4), and rRNAs were predicted using RNAmmer. The functional annotation of predicted genes was performed using RPS-BLAST (6) against Pfam (9) and the Cluster of Orthologous Groups (COG) database (10). We performed BLASTP analysis of *B. duttonii* open reading frames (ORFs) versus the *B. crocidurae* chromosome (1), and codon usage was determined using the E-CAI server (8).

The 1,557,560-bp linear *B. crocidurae* genome (27% GC content) carries 1,472 ORFs, 32 tRNAs, and three complete rRNAs. The 919,477-bp chromosome carries 865 ORFs, of which 79% are proteins listed in the COG database. The S200-like transposase and transposase IS605 OrfB in *B. crocidurae* disrupt the almost complete colinearity between *B. crocidurae*, *B. duttonii*, and *B. recurrentis* chromosomes. Seven copies of these elements are found in the *B. duttonii* plasmid complement, and six copies are found in the *Borrelia recurrentis* plasmid complement. Additionally, a 5-kbp *B. duttonii* duplication encoding an RNA pseudouridylate synthase family protein, MurC, and the hypothetical proteins BDU_828 and BDU_830 is absent from the *B. crocidurae* chromosome. Comparing *B. crocidurae* with *B. duttonii* revealed 771 orthologs with an average pairwise amino acid sequence identity of 95 to 100% and identical codon usage. Unlike *B. recurrentis*, *B. crocidurae* has intact *recA*, *mutS*, and *smf* genes (5). The exact number of plasmids remained undetermined due to repeat sequences and size overlap, but 607 ORFs, including a telomerase resolvase, are conserved among the three species. The *B. duttonii* pI165 and pI11 plasmids are merged as one pI173 plasmid in *B. crocidurae*. In addition, the *B. crocidurae* genome has a counterpart for *B. recurrentis* plasmid PI6, which is absent in *B. duttonii*.

Availability of the *B. crocidurae* genome sequence provides major insights into the evolution of borreliae and explains the taxonomic relationship between recurrent fever groups in Africa. This sequence will also facilitate the development of molecular tools to improve the diagnosis and characterization of relapsing fever pathogens.

Nucleotide sequence accession numbers. The complete genome sequence of the *B. crocidurae* Achema strain was deposited in GenBank under the accession numbers CP003426 (chromosome) and CP003427 to CP003465 (plasmids).

REFERENCES

1. **Altschul SF, et al.** 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. **25**:3389–3402.
2. **Cutler SJ, Abdissa A, Trape JF.** 2009. New concepts for the old challenge of African relapsing fever borreliosis. Clin. Microbiol. Infect. **15**:400–406.
3. **Hyatt D, et al.** 2010. Prodigal: prokaryotic gene recognition and translation initiation site identification. BMC Bioinformatics **11**:119.
4. **Laslett D, Canback B.** 2004. ARAGORN, a program to detect tRNA genes and tmRNA genes in nucleotide sequences. Nucleic Acids Res. **32**:11–16.
5. **Lescot M, et al.** 2008. The genome of *Borrelia recurrentis*, the agent of deadly louse-borne relapsing fever, is a degraded subset of tick-borne *Borrelia duttonii*. PLoS Genet. **4**:e1000185. doi:10.1371/journal.pgen.1000185.
6. **Marchler-Bauer A, et al.** 2002. CDD: a database of conserved domain alignments with links to domain three-dimensional structure. Nucleic Acids Res. **30**:281–283.
7. **Million M, et al.** 2009. Molecular identification of Borrelia crocidurae in a patient returning from Senegal. BMJ Case Rep. **pii**:bcr06.2008.0298. doi:10.1136/bcr.06.2008.0298.

8. **Puigbò P, Bravo IG, Garcia-Vallvé S.** 2008. E-CAI: a novel server to estimate an expected value of Codon Adaptation Index (eCAI). BMC Bioinformatics **9**:65.

9. **Sonnhammer EL, Eddy SR, Durbin R.** 1997. Pfam: a comprehensive database of protein domain families based on seed alignments. Proteins **28**:405–420.

10. **Tatusov RL, Galperin MY, Natale DA, Koonin EV.** 2000. The COG database: a tool for genome-scale analysis of protein functions and evolution. Nucleic Acids Res. **28**:33–36.

11. **Toledo A, et al.** 2010. Phylogenetic analysis of a virulent *Borrelia* species isolated from patients with relapsing fever. J. Clin. Microbiol. **48**:2484–2489.

12. **Vial L, et al.** 2006. Incidence of tick-borne relapsing fever in west Africa: longitudinal study. Lancet **368**:37–43.