JB Journals.ASM.org

GENOME ANNOUNCEMENT

Genomic Comparison of *Rickettsia helvetica* and Other Rickettsia Species

Xin Dong,[a] Khalid El Karkouri,[a] Catherine Robert,[a] Frédérick Gavory,[b] Didier Raoult,[a] and Pierre-Edouard Fournier[a]

Unité de Recherche sur les Maladies Infectieuses et Tropicales Emergentes, UMR CNRS 6236, IRD 198, Faculté de Médecine, Aix-Marseille Université, Marseille, France,[a] and Genoscope, Centre National de Séquençage, Evry, France[b]

We report the complete and annotated genome sequence of *Rickettsia helvetica* strain C9P9, which was first isolated in 1979 from *Ixodes ricinus* ticks in Switzerland and is considered a human pathogen.

Rickettsia helvetica was first isolated from *Ixodes ricinus* ticks in Switzerland in 1979. It was subsequently detected in many European countries, Russia, Japan, and Thailand. *R. helvetica* has been suggested to play a role in various human diseases, but its exact pathogenesis remains uncertain.

Genomic DNA extracted from *R. helvetica* strain C9P9 grown in Vero cells was pyrosequenced using the 454 GS FLX Titanium platform (Roche, Meylan, France) (8) and assembled using the Newbler software (Roche). Potential coding sequences (CDSs) were predicted using AMIGene (5), and split genes or nonpredicted genes were detected using Artemis (http://www.sanger.ac.uk/software/Artemis/) and BLASTN (1). Assignment of protein functions was performed by searching against the RickBase, GenBank, and Pfam databases using BLASTP (1, 4, 9), while rRNAs, tRNAs, and other RNAs were identified using BLASTN or tRNAscan-SE (7). Gene orthologs were identified using OrthoMCL (6) with a BLASTP E value cutoff of 1×10^{-5} and a Markov Clustering inflation parameter default of 1.5.

Thirty-two contigs were assembled into two scaffolds. The gaps were closed using Phusion DNA polymerase and specific primers. The *R. helvetica* genome consists of a 1,369,827-bp chromosome with a G+C content of 32.2%, in the range of other rickettsial genomes, and a 47,188-bp plasmid (pRhel) with a G+C content of 32.6%. The predicted total complement of 1,135 genes (1,515 open reading frames) includes 858 complete genes, 168 split genes, and 178 genes present only as fragments. Of these, 881 were assigned putative functions and 254 encode hypothetical proteins and proteins of unknown function. *R. helvetica* also contains 33 tRNAs, a single rRNA operon with noncontiguous 16S and 23S rRNAs, and 3 other RNAs. In addition, the pRhel plasmid encodes 49 genes (59 CDSs) but no tRNA.

The *R. helvetica* chromosome exhibits a high level of synteny with its closest phylogenetic neighbor, *R. massiliae* (3), with the exception of three inversions of 36,436 bp, 48,196 bp, and 17,375 bp, respectively. *R. helvetica* has 121 genes that are absent from *R. massiliae*, and the latter has 161 genes missing from *R. helvetica*. Most of these differentially present genes encode ankyrin repeat-, leucine-rich repeat-, or tetratricopeptide repeat-containing proteins, transposases, proteins of unknown function, and Tra family proteins. The *R. helvetica* genome appears more degraded than that of *R. massiliae* (168 versus 99 split genes and 178 versus 85 fragment genes, respectively). In comparison with the *R. prowazekii* genome (2), *R. helvetica* lacks only 36 genes, including the genes for 16 proteins of unknown function, three transposases, two ankyrin repeats, four transferases, two cell surface antigens, and one each for the integration host factor β subunit, a large extracellular α-helical protein, a monovalent cation/proton antiporter, peptide chain release factor RF-2, a prolyl endopeptidase precursor, pyrroloquinoline quinine biosynthesis protein C, signal peptide peptidase SppA, and a DNA invertase Pin-like protein. Thus, none of these genes is likely to be linked to rickettsial virulence. In contrast, many genes involved in the biosynthesis and regulation of biosynthesis of amino acids and nucleotides present in *R. helvetica* were absent from *R. prowazekii*. Further studies will be conducted to investigate the pathogenesis of *R. helvetica*.

Nucleotide sequence accession number. The *R. helvetica* strain C9P9 genome sequence and the pRhel plasmid sequence have been deposited in the DDBJ/EMBL/GenBank databases under accession no. AICO00000000.

ACKNOWLEDGMENT

This work was funded by the EuroPathoGenomics Network of Excellence.

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. **Andersson SGE, et al.** 1998. The genome sequence of *Rickettsia prowazekii* and the origin of mitochondria. Nature **396**:133–140.
3. **Blanc G, et al.** 2007. Lateral gene transfer between obligate intracellular bacteria: evidence from the Rickettsia massiliae genome. Genome Res. **17**:1657–1664.
4. **Blanc G, et al.** 2007. Reductive genome evolution from the mother of *Rickettsia*. PLoS Genet. **3**:e14.
5. **Bocs S, Cruveiller S, Vallenet D, Nuel G, Medigue C.** 2003. AMIGene: annotation of microbial genes. Nucleic Acids Res. **31**:3723–3726.
6. **Li L, Stoeckert CJ, Jr, Roos DS.** 2003. OrthoMCL: identification of ortholog groups for eukaryotic genomes. Genome Res. **13**:2178–2189.
7. **Lowe TM, Eddy SR.** 1997. t-RNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. Nucleic Acids Res. **25**:955–964.
8. **Margulies M, et al.** 2005. Genome sequencing in microfabricated high-density picolitre reactors. Nature **437**:376–380.
9. **Punta M, et al.** 2012. The Pfam protein families database. Nucleic Acids Res. **40**:D290–D301.

Received 23 February 2012 **Accepted** 28 February 2012

Address correspondence to Pierre-Edouard Fournier, pierre-edouard.fournier@univmed.fr.

doi:10.1128/JB.00299-12