Genomic Comparison of *Rickettsia honei* Strain RB<sup>T</sup> and Other *Rickettsia* Species

Dong Xin, Khalid El Karkouri, Catherine Robert, Didier Raoult, and Pierre-Edouard Fournier

Unité de Recherche sur les Maladies Infectieuses et Tropicales Emergentes, UMR CNRS 6236, IRD 198, Faculté de Médecine, Aix-Marseille Université, Marseille, France

*Rickettsia honei* strain RB<sup>T</sup> was isolated from a febrile patient on Flinders Island, Australia, in 1991 and has been demonstrated to be the agent of Flinders Island spotted fever, a disease transmitted to humans by ticks. The comparison of this 1.27-Mb genome with other *Rickettsia* genomes provides additional insight into the mechanisms of evolution in *Rickettsia* species.

The *Rickettsia* genus is composed of small, Gram-negative, obligate intracellular bacteria (1). These microorganisms underwent a reductive genomic evolution during their specialization to their intracellular lifestyle (2). Recent studies demonstrated that genome reduction was associated with virulence in rickettsiae (3). *R. honei* strain RB<sup>T</sup> was isolated from a febrile patient on Flinders Island, Australia, in 1991 and has been described as a Flinders Island spotted fever (FISF) agent which is pathogenic for humans (4). FISF is characterized by fever, headache, myalgia, transient arthralgia, maculopapular rash, and in some cases cough (6). *R. honei* strain RB<sup>T</sup> is transmitted to humans mainly by *Bothriocroton hydrosauri* (formerly *Aponomma hydrosauri*) tick bites (5). Most cases occur in summer.

The genome sequencing of *R. honei* strain RB<sup>T</sup> was performed by 454 shotgun sequencing. Briefly, the shotgun sequencing was performed using a GS-FLX Titanium sequencer (Roche, Meylan, France) and assembled into 11 contigs. All 11 contigs were part of the chromosome, with a size of 1,268,758 bp and a G+C content of 32.4%, which is similar to other rickettsial genomes. No plasmid was detected. The chromosome was predicted to encode 1,284 genes (1,595 open reading frames [ORFs]). Among these genes, 1,046 (81%) were complete, 158 (12%) were split into two or more ORFs, and 80 (6%) were present as fragments. Of the 1,284 genes, 751 (58%) encoded proteins with putative functions, and 533 (42%) encoded hypothetical proteins and proteins of unknown function. The *R. honei* genome had 3 noncontiguous rRNA genes (5S, 16S, and 23S rRNA), 33 tRNAs, and 3 other RNAs.

Phylogenically, *R. honei* is closely related to *R. rickettsii*, *R. conorii*, and *R. slovaca* (4). The 11 *R. honei* contigs exhibited an almost perfect collinearity with these 3 genomes, with the exception of an inversion of 34,133 bp, 81,501 bp and 32,367 bp, respectively. By comparison with the closest genome, that of *R. conorii*, *R. honei* missed 56 genes but had an additional 89 specific genes. Most of these differentially present genes encoded ankyrin repeat-containing proteins, leucine-rich repeats (LRRs), tetratricopeptide repeat-containing proteins, transposases, and proteins of unknown function. By comparison with *R. prowazekii*, the agent of epidemic typhus, the most pathogenic *Rickettsia* species with the smallest genome (834 genes only), *R. honei* missed only 37 genes. These genes encoded proteins of unknown function (n = 17), transposases (n = 3), ankyrin repeat-containing proteins (n = 2), transferases (n = 6), and synthetases (n = 3) and one from each of the following categories: cell surface antigens, BioY family proteins, multisubunit Na<sup>+</sup>/H<sup>+</sup> antiporters, protein kinase C inhibitors, Sec7 domain-containing proteins, site-specific recombinases, and VirD4 proteins. Thus, none of these genes is likely to play a role in rickettsial virulence. In contrast, many genes involved in the biosynthesis and regulation of biosynthesis of amino acids and nucleotides present in *R. honei* were absent from *R. prowazekii*. Further studies will be conducted to investigate whether a reduced metabolism of amino acids and/or nucleotides plays a role in the pathogenesis of rickettsiae.

Nucleotide sequence accession number. The genome sequence has been deposited in the GenBank database under accession number AJTT00000000.

ACKNOWLEDGMENT

This work did not benefit from any external funding.

REFERENCES