Genome sequence of *Moraxella catarrhalis* RH4, an isolate of seroresistant lineage.

Aldert Zomer\textsuperscript{1,2,†}, Stefan P.W. de Vries\textsuperscript{1,†}, Kristian Riesbeck\textsuperscript{3}, Andreas L. Meinke\textsuperscript{4}, Peter W.M Hermans\textsuperscript{1}, Hester J. Bootsma\textsuperscript{1*}

\textsuperscript{1}Laboratory of Pediatric Infectious Diseases, Radboud University Medical Centre, Nijmegen, The Netherlands, \textsuperscript{2}Centre for Molecular and Biomolecular Informatics, Nijmegen Centre for Molecular Life Sciences, Radboud University Medical Centre, Nijmegen, The Netherlands, \textsuperscript{3}Medical Microbiology, Department of Laboratory Medicine Malmö, Skåne University Hospital, Lund University, Malmö, Sweden, \textsuperscript{4}Intercell AG, Vienna, Austria

\textsuperscript{†}These authors contributed equally to this work.

* Corresponding author. Mailing address: Laboratory of Pediatric Infectious Diseases, Radboud University Medical Centre, P.O. Box 9101 (Route 224), 6500 HB Nijmegen, The Netherlands. Phone: 31-24-3666406. Fax: 31-24-3666352. E-mail: H.Bootsma@cukz.umcn.nl
ABSTRACT

Here we report the annotated genome sequence of *Moraxella catarrhalis* strain RH4, a seroresistant lineage strain isolated from the blood of an infected patient. This genome sequence will allow us to gain further insight into the genetic diversity of clinical *M. catarrhalis* isolates and will facilitate study of *M. catarrhalis* pathogenesis.

GENOME ANNOUNCEMENT

The Gram-negative diplococcus *Moraxella catarrhalis* is an emerging human-restricted respiratory tract pathogen. It is the third most common cause of childhood otitis media, and is frequently associated with exacerbations of chronic obstructive pulmonary disease (COPD) in adults. We reported the first completely assembled and annotated *M. catarrhalis* genome in 2010 (6) of strain BBH18 (erroneously referred to as RH4 at the time), a sputum isolate from a COPD patient during an exacerbation (8). In 2011, an additional 10 genome sequences of clinical *Moraxella* isolates were published (5) and compared to the BBH18 genome and the partial genome sequence of strain ATCC 43617 (10). This indicated a modest diversity in gene content and chromosomal organization between these isolates. Here we present the annotated genome sequence of the clinically relevant RH4 strain, which was originally isolated from the blood of an infected patient (4).

The draft genome sequence of *M. catarrhalis* RH4 was obtained using Illumina 50 bp paired-end technology (total of 13,826,736 reads, with 700x coverage). Reads were assembled with the Ray assembler software program (3), resulting in a total of 31 contigs (>100 bp in size). Contigs were ordered using the program Projector 2 (9) with the BBH18 sequence as a scaffold and the correct order was verified by gap-spanning PCR's. When possible, gaps were filled or corrected by Sanger sequencing of PCR products, followed by use of GapFiller (2).
After manual contig assembly, 9 contigs covering a total of 1,836,691 bp were obtained, which is within the size range of the reported genomes (1.78 to 1.96 Mbp).

The RH4 genome sequence was annotated using the RAST (rapid annotations using subsystems technology) server (1) and manually corrected for errors in open reading frame (ORF) calling. The total genome has a G+C content of 41.6% and is composed of 1,904 genes, including 1,845 protein-encoding genes, 4 rRNA operons, and at least 43 tRNA’s. RH4 has a novel sequence variant for the abcZ allele and thus a novel multilocus sequence type (MLST), but clearly belongs to the seroresistant lineage (11). The RH4 genome contains the bro-1 β-lactamase gene and all of the major known M. catarrhalis virulence factors, among which UspA1, UspA2H, MID/Hag, and, in contrast to, for instance, BBH18, a complete mha locus. Compared to all Moraxella genomes published to date, the RH4 genome contains 10 unique genes, 8 of which are located consecutively on a 10.1 kb fragment. In addition to 4 putative restriction-modification protein-encoding genes, this cluster contains a putative CiaB-encoding gene, where the ortholog in Campylobacter jejuni is involved in internalization into mammalian cells (7).

**Nucleotide sequence accession number.**

This whole-genome shotgun project has been deposited in DDBJ/EMBL/GenBank under the accession no. AMSO00000000. The version described in this article is the first version, AMSO01000000.1.

**Acknowledgements**

This work was supported by the Vienna Spot of Excellence (VSOE) grant (ID337956), the Anna and Edwin Berger foundation (to K.R.) and the Swedish Medical Research Council (grant number 521-2010-4221; to K.R.).
REFERENCES


