

Genome Analysis of *Moraxella catarrhalis* Strain RH4, a Human Respiratory Tract Pathogen^{∇†}

Stefan P. W. de Vries,¹ Sacha A. F. T. van Hijum,^{2,3} Wolfgang Schueler,⁴ Kristian Riesbeck,⁵ John P. Hays,⁶ Peter W. M. Hermans,^{1*} and Hester J. Bootsma¹

Laboratory of Pediatric Infectious Diseases, Radboud University Nijmegen Medical Centre, Nijmegen, Netherlands¹; Centre for Molecular and Biomolecular Informatics, Radboud University Nijmegen Medical Centre, Nijmegen, Netherlands²; NIZO Food Research, Ede, Netherlands³; Intercell AG, Vienna Biocenter 3, Vienna, Austria⁴; Medical Microbiology, Department of Laboratory Medicine, University Hospital Malmö, Lund University, Malmö, Sweden⁵; and Department of Medical Microbiology and Infectious Diseases, Erasmus MC, Rotterdam, Netherlands⁶

Received 3 February 2010/Accepted 2 May 2010

Moraxella catarrhalis is an emerging human-restricted respiratory tract pathogen that is a common cause of childhood otitis media and exacerbations of chronic obstructive pulmonary disease in adults. Here, we report the first completely assembled and annotated genome sequence of an isolate of *M. catarrhalis*, strain RH4, which originally was isolated from blood of an infected patient. The RH4 genome consists of 1,863,286 nucleotides that form 1,886 protein-encoding genes. Comparison of the RH4 genome to the ATCC 43617 contigs demonstrated that the gene content of both strains is highly conserved. *In silico* phylogenetic analyses based on both 16S rRNA and multilocus sequence typing revealed that RH4 belongs to the seroresistant lineage. We were able to identify almost the entire repertoire of known *M. catarrhalis* virulence factors and mapped the members of the biosynthetic pathways for lipooligosaccharide, peptidoglycan, and type IV pili. Reconstruction of the central metabolic pathways suggested that RH4 relies on fatty acid and acetate metabolism, as the genes encoding the enzymes required for the glyoxylate pathway, the tricarboxylic acid cycle, the gluconeogenic pathway, the nonoxidative branch of the pentose phosphate pathway, the beta-oxidation pathway of fatty acids, and acetate metabolism were present. Moreover, pathways important for survival under challenging *in vivo* conditions, such as the iron-acquisition pathways, nitrogen metabolism, and oxidative stress responses, were identified. Finally, we showed by microarray expression profiling that ~88% of the predicted coding sequences are transcribed under *in vitro* conditions. Overall, these results provide a foundation for future research into the mechanisms of *M. catarrhalis* pathogenesis and vaccine development.

Moraxella catarrhalis is an emerging human-restricted unencapsulated Gram-negative mucosal pathogen. Long considered to be a commensal of the upper respiratory tract, this bacterium has now firmly been established to be an etiological cause of otitis media (OM) and exacerbations of chronic obstructive pulmonary disease (COPD). It is the third most common cause of childhood OM after *Haemophilus influenzae* and *Streptococcus pneumoniae* (37, 64), and it is responsible for up to 20% of the cases (64, 65). Further, *M. catarrhalis* is the second most common cause of exacerbations of COPD after *H. influenzae*; it is responsible for 10 to 15% of the exacerbations and annually causes 2 to 4 million episodes in the United States (47, 60). Antibiotics are widely used for treatment of OM, but the high prevalence of this disease and the increasing incidence of antibiotic-resistant strains mean that multivalent vaccines, preferably vaccines with protective antigens for all three causative bacterial agents, must be developed (46).

M. catarrhalis is able to colonize the mucosal surfaces of the middle ear in OM patients and the lower respiratory tract in

COPD patients (31, 60). Successful colonization of the human host is a complex process which requires the expression of adhesins and the activation of metabolic pathways to overcome specific challenging environmental conditions, such as nutrient limitation (15, 53). *M. catarrhalis* also possesses several mechanisms for evasion of the host immune system (15, 53), such as the ability to withstand the action of the human complement system. Importantly, most clinical isolates obtained from OM or COPD patients are able to survive complement-mediated killing by normal human serum (66).

Various molecular typing methods, such as 16S rRNA sequencing (8) and multilocus sequence typing (MLST) (71), have shown that the species *M. catarrhalis* can be divided into two distinct phylogenetic lineages, referred to as the serosensitive and seroresistant lineages. The seroresistant, more virulent lineage contains predominantly strains that are resistant to complement-mediated killing and that are able to adhere to epithelial cells (8, 71).

Even though our understanding of the molecular pathogenesis of *M. catarrhalis* has increased over the past few years, a complete *M. catarrhalis* genome sequence would undoubtedly be a valuable resource for improving our understanding of the biology of this organism. At present, however, only a partial genome of *M. catarrhalis* strain ATCC 43617 is available (59, 68). Here, we report the first completely assembled and annotated genome sequence of *M. catarrhalis* strain RH4, which was originally isolated from the blood of an infected patient (14),

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

† Supplemental material for this article may be found at <http://jb.asm.org/>.

[∇] Published ahead of print on 7 May 2010.

and we compared the gene content of this sequence to that of the strain ATCC 43617 sequence. *In silico* phylogenetic typing revealed that the RH4 strain belongs to the seroresistant lineage. We were able to identify most of the known virulence factors and reconstructed several important biochemical pathways, including central metabolic pathways and pathways for nitrate metabolism, biosynthesis of lipooligosaccharides (LOS), peptidoglycan, type IV pilus (TFP) biosynthesis, and iron acquisition. In addition, several components that combat oxidative stress were identified. Finally, we confirmed by transcriptional profiling that most of the predicted coding sequences (CDSs) are expressed *in vitro*.

MATERIALS AND METHODS

Bacterial strain and growth conditions. *M. catarrhalis* strain RH4 was isolated from the blood of an infected patient (14). RH4 was grown on Bacto brain heart infusion (BHI) agar plates (Difco) in an atmosphere containing 5% CO₂ at 37°C or in broth at 37°C with agitation (200 rpm).

Sequencing of the *M. catarrhalis* RH4 genome. Whole-genome sequencing was performed using a hybrid strategy consisting of Roche 454 and Illumina Solexa sequencing by Agowa Genomics (Berlin, Germany). For Roche 454 sequencing, genomic DNA was extracted using a Wizard genome purification kit (Promega), after which a fragment library was prepared and sequenced using Roche standard procedures. This resulted in a total of 591,043 sequences with an average read length of 224 bp covering approximately 134 Mb, which is more than 70-fold coverage of the total genome. *De novo* assembly using the Roche 454 software Newbler resulted in 44 contigs consisting of over 500 bp, which were aligned with the 41 contigs of the ATCC 43617 strain (deposited in the NCBI database under patent WO0078968; GenBank accession numbers AX067426 to AX067466) using the gap4 assembler program (Staden Package; Roger Staden, Cambridge, United Kingdom). Gap closure by PCR and Sanger sequencing resulted in a contiguous sequence consisting of 1,863,286 bp, which was verified using Solexa sequencing as follows. Genomic DNA was isolated with a genomic DNA kit (Qiagen), a fragment library (150 to 200 bp) was prepared using Illumina's standard genomic DNA library preparation procedure, and the library was sequenced with an Illumina Genome Analyzer II. The data were analyzed using the standard Genome Analyzer pipeline, which yielded a total of 7,057,256 raw 36-bp reads. The 454-assembled genome was corrected with Solexa short read sequence data using ROAST, a tool developed in house (S. A. F. T. van Hijum, V. C. L. de Jager, B. Renckens, and R. J. Siezen, unpublished data). Briefly, Solexa reads were aligned with the assembled genome sequence by using BLAT (38). Alignment of reads with the reference sequence was allowed provided that nucleotide substitutions (single nucleotide polymorphisms [SNPs]) or gaps (small insertions or deletions [indels]) were at least 4 bp from the end of the read. Only SNPs and indels were allowed, with a sequence depth of at least six reads unanimously calling a genotype and with a maximum of one read indicating a different genotype. Altogether, three SNPs, one insertion, and one deletion were corrected in the reference genome sequence.

Annotation. Open reading frames (ORFs) and initial automated annotation were obtained using both Pedant-Pro and the Institute for Genomics Sciences (IGS) annotation engine (http://ae.igs.umaryland.edu/cgi/ae_pipeline_outline.cgi), and the minimal ORF length was 90 nucleotides (nt). The IGS-derived annotation was subjected to manual curation using the Pedant-Pro data, the DBGET database (<http://www.genome.jp/dbget/>), and data described previously. The putative origin of replication was identified using Ori-Finder (27), and bp 1 was assigned to the extreme of the CG disparity curve. Functional classification of the predicted ORFs was based on the IGS functional classification, which was manually improved with data from Clusters of Orthologous Groups (COGs) of proteins and Nonsupervised Orthologous Groups (NOGs), both obtained by searches using Signature (18).

A comparative genomic analysis of the RH4 and ATCC 43617 strains was performed by alignment of the RH4 coding sequences (CDSs) with the ATCC 43617 contigs using BLAT (38). CDSs potentially unique to ATCC 43617 were identified by alignment of the ATCC 43617 contigs with the RH4 sequence.

Pathway analysis was performed using the KEGG Pathway database (36) with KEGG orthology (KO) identifiers and was complemented by pathway reconstructions based on previously described data. KO identifiers were assigned with the web-based KEGG automatic annotation server (KAAS) (<http://www.genome.jp/kegg/kaas/>) using the bidirectional best-hit methods for a set of prokaryotic

reference genomes (organism codes prw, par, pcr, abm, aby, aci, ngo, nmc, hso, nma, acb, apl, cvi, hin, hit, hip, and pmu), manually selected based on high abundance in Pedant-Pro BlastP hits. The list acquired was complemented with KO identifiers assigned by searching against the prokaryotic reference set. Finally, the RH4 genome was analyzed for the presence of clustered regularly interspaced short palindromic repeat (CRISPR) regions using CRISPRFinder (29).

The subcellular localization of the RH4 proteins was predicted by using an extended version of the LocateP software (75) and was validated with a highly curated list of *Escherichia coli* proteins. Protein localization prediction was tailored for the Gram-negative organism *M. catarrhalis* by replacement of specific Gram-positive components with tools suitable for Gram-negative bacteria (M. Zhou, K. Mezger, S. A. F. T. van Hijum, and R. J. Siezen, unpublished data), such as BOMP (7), CELLO (73), LipoP (35), and SecretomeP (6).

A detailed annotation, including localization prediction and comparative genomics data, is shown in File S1 in the supplemental material.

In silico phylogenetic typing. The 16S rRNA type was determined as described by Bootsma et al. (8). Multilocus sequence typing (MLST) of sequence fragments of eight housekeeping genes was performed as described by Wirth et al. (71). Allelic sequences were analyzed (<http://mlst.ucc.ie/mlst/dbs/Mcatarrhalis>) and compared to the reference database containing 282 *M. catarrhalis* strains. The lipooligosaccharide (LOS) serotype was determined by using the method described by Edwards et al. (19), and *copB* typing was performed as described by Verhaegh et al. (66).

Microarray expression profiling. Bacteria were grown in BHI medium at 37°C with agitation (200 rpm), harvested by centrifugation at lag phase (optical density at 620 nm [OD₆₂₀], 0.2 to 0.3), exponential phase (OD₆₂₀, 1.2 to 1.4), and stationary phase (OD₆₂₀, 2.0 to 2.2), and treated with 2 volumes of RNaprotect bacterial reagent (Qiagen). Total RNA was extracted using an RNeasy minikit (Qiagen), after which contaminating genomic DNA was removed by treatment with DNase (DNasefree; Ambion). Total RNA was labeled essentially as described by Ouellet et al. (51). Briefly, 10 µg of RNA was incubated for 3 h with 7 µg of 5' Cy3-labeled random nonamers (TriLink Biotechnologies) and Superscript III reverse transcriptase (800 U; Invitrogen) under appropriate reaction conditions (1× first-strand buffer, 5 mM dithiothreitol [DTT], 0.33 mM deoxynucleoside triphosphates [dNTPs], 21 mM actinomycin D [Sigma Aldrich], 40 U RNaseOut). After first-strand synthesis, RNA was degraded by incubation with sodium hydroxide, which was followed by neutralization of the reaction mixture with hydrochloric acid. Labeled cDNA was purified and concentrated with CyScribe columns (GE Healthcare Life Sciences), followed by Micron-30 columns (Millipore). Two micrograms of labeled cDNA was applied in duplicate to 4x72K custom-designed NimbleGen arrays. Overnight hybridization at 42°C and subsequent washing of arrays were performed according to the manufacturer's instructions. The Nimblegen array contained 1 to 118 probes per predicted CDS with an average coverage of 15 probes per CDS, probes for both strands in the CRISPR1 and CRISPR2 regions (no specific probes could be designed for the putative CRISPR3 region), and 1,103 negative-control probes with a length distribution and G+C content similar to those of the experimental probes. Array images were acquired with a NimbleGen MS200 scanner, and images were processed with NimbleScan software using the RMA algorithm. Normalized and background-corrected probe signal intensities obtained in this way were used to calculate the median level of expression of the CDSs. An expression threshold was defined by the median of the log₂ signal intensity of the random control probes plus four times the standard deviation. The levels of expression for CDSs were classified as follows: low (log₂ signal intensity, <10), moderate (log₂ signal intensity, 10 to 12.5), high (log₂ signal intensity, >12.5), or none (4/6 replicates [biological triplicate and technical duplicate] less than the threshold value). Expression data for the RH4 predicted CDSs are shown in File S1 in the supplemental material.

Nucleotide sequence and microarray data accession numbers. The genomic sequence of *M. catarrhalis* RH4 has been deposited in the GenBank database under accession number CP002005. Microarray data have been deposited in the NCBI Gene Expression Omnibus (GEO) database (www.ncbi.nlm.nih.gov/geo/) under GEO Series accession number GSE21632.

RESULTS AND DISCUSSION

General genome features and comparative genomics. The RH4 genome (Fig. 1) consists of 1,863,286 nucleotides (nt), and the overall G+C content is 41.7%. Both the length and the G+C content are comparable to those of the unassembled,

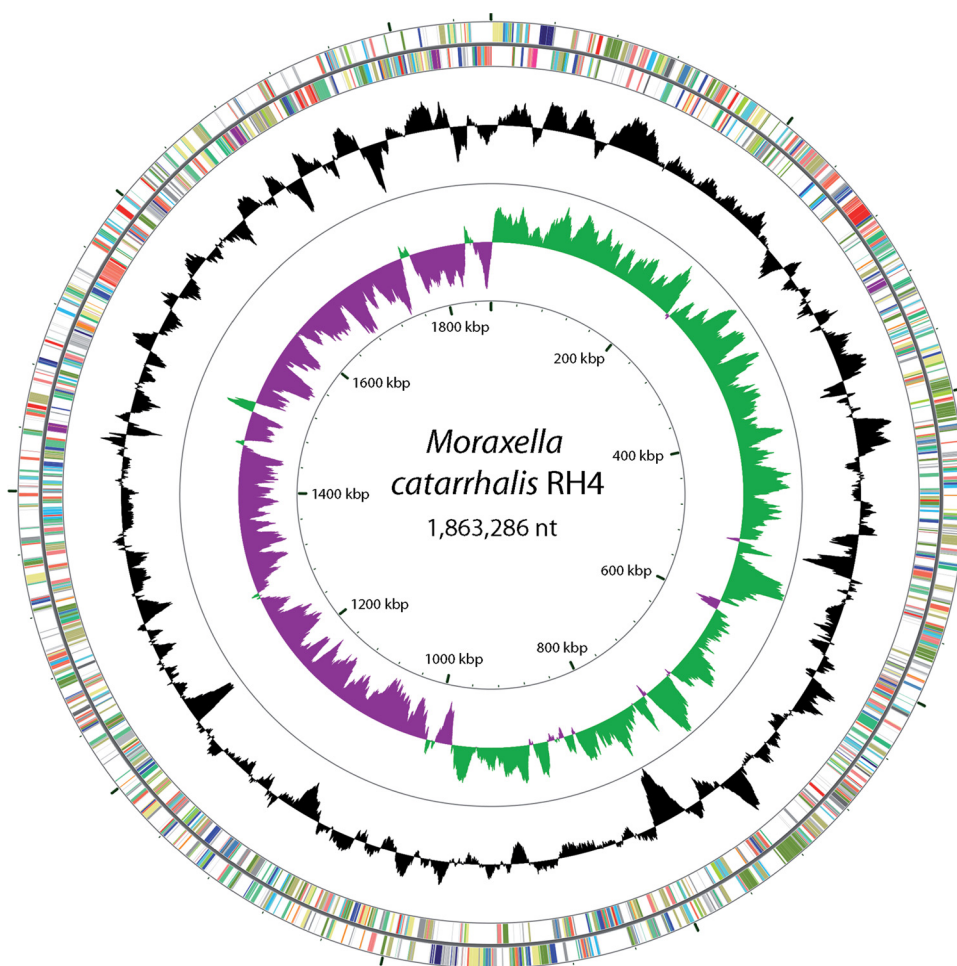


FIG. 1. Diagram of the *M. catarrhalis* RH4 genome. Gene coordinates are indicated. From the inside to the outside, the circles show the GC skew, the G+C content, the genes on the forward strand, and the genes on the reverse strand. The colors in the two outermost rings indicate the functional classes of the genes and correspond to the colors in Table S1 in the supplemental material. The diagram was generated using CGview (28).

partial ATCC 43617 genome (~1.9 Mb), which is represented by 41 contigs (59, 68). The RH4 genome is predicted to include 1,964 genes, 1,886 of which are protein-encoding genes (Table 1), similar to the 1,761 to 1,849 ORFs predicted for ATCC 43617 (59, 68). Of the 1,886 protein-encoding genes predicted for RH4, 63.6% could be assigned to a functional category based on similarity of their products to proteins in public databases, while the functions of 172 predicted proteins are unknown. The products of the remaining 515 ORFs are clas-

TABLE 1. General genome features of RH4

Feature	Value
Genome size (nt).....	1,863,286
G+C content (%).....	41.7
No. of genes.....	1,964
No. of protein-encoding ORFs.....	1,886
Avg ORF size (bp).....	850
Coding density (%).....	86.1
No. of rRNA genes.....	12
No. of tRNA genes.....	50

TABLE 2. Functional classes of the predicted RH4 proteins

Functional class	No. of proteins	% of total
Amino acid biosynthesis	77	4.0
Purines, pyrimidines, nucleosides, and nucleotides	36	1.9
Fatty acid and phospholipid metabolism	40	2.1
Biosynthesis of cofactors, prosthetic groups, and carriers	90	4.7
Central intermediary metabolism	33	1.7
Energy metabolism	140	7.4
Transport and binding proteins	148	7.8
DNA metabolism	78	4.1
Transcription	35	1.8
Protein synthesis	141	7.4
Protein fate	89	4.7
Regulatory functions	31	1.6
Signal transduction	11	0.6
Cell envelope	168	8.8
Cellular processes	61	3.2
Mobile and extrachromosomal element functions	21	1.1
Unknown function	172	9.0
Hypothetical proteins	515	27.1
Disrupted reading frame	16	0.8

TABLE 3. Comparison of *M. catarrhalis* RH4 and ATCC 43617 genomes

% RH4 CDSs in alignment	% of mismatches allowed in alignment	No. of CDSs	% of total
100	0	416	22.1
100	<20	1,252	66.4
>80	<20	125	6.6
60–80	<20	10	0.5
40–60	<20	7	0.4
0–40	<20	8	0.4
No match		68	3.6

sified as either conserved hypothetical proteins (102 ORFs) or hypothetical proteins (413 ORFs) (Table 2). The majority (~62%) of the best BlastP hit proteins were found to be proteins of *Psychrobacter* species (data not shown).

Comparative genome analysis of the RH4 genome and the ATCC 43617 contigs revealed that for 416 RH4 coding sequences (CDSs) there was an exact match in the ATCC 43617 genome (Table 3). When a maximum of 20% mismatches was allowed, 1,252 full-length RH4 CDSs were identified in the ATCC 43617 contigs. A total of 1,793 (~95%) of the predicted CDSs were found to be conserved in RH4 and ATCC 43617, and a minimum of 80% of the CDSs were covered. As expected, RH4 genes encoding highly variable virulence factors, such as members of the ubiquitous surface protein A (*uspA*) family (10), showed lower levels of homology with ATCC 43617 sequences. Further, 68 CDSs were found to be unique to RH4, 41 of which are annotated as genes encoding (conserved) hypothetical proteins. Conversely, using the preliminary annotation of the partial ATCC 43617 genome (68), 14 CDSs were found to be unique to ATCC 43617, 12 of which encode hypothetical proteins.

The RH4 genome contains genes encoding 50 tRNAs for all 20 amino acids. There are four identical rRNA operons (16S, 23S, and 5S rRNA genes) in which the 16S and 23S rRNA genes are interspersed with genes encoding tRNAs for isoleucine and alanine (Table 1). Genes encoding the RNA polymerase core subunits (α , β , β' , and ω), the σ^{70} sigma factor, and the alternative σ^{32} sigma factor were identified, as were genes encoding the transcription elongation factor GreA and the transcription termination-antitermination factors NusA, NusB, Rho, and NusG. Interestingly, the RH4 genome is predicted to encode only 32 transcription factors and 4 two-component regulatory systems and to contain 2 orphan two-component system genes. This raises the possibility that *M. catarrhalis* has other, alternative mechanisms for adapting its gene expression to changing environmental conditions, such as phase-variable expression or noncoding RNA-based regulation. Phase-variable expression has been described previously for two *M. catarrhalis* virulence factors, ubiquitous surface protein A1 (*UspA1*) (40) and *M. catarrhalis* immunoglobulin D binding protein (MID) (45), but a preliminary search for additional homopolymeric tracts known to be involved in phase variation did not result in discovery of novel candidates in the RH4 genome. Clearly, further studies to elucidate the main mechanism of transcriptional regulation in *M. catarrhalis* are required.

In silico phylogenetic analysis classified RH4 as a 16S rRNA

type 1 strain (see Fig. S1A in the supplemental material); such strains are predominantly members of the seroresistant lineage (8). MLST analysis assigned RH4 to sequence type 128 (see Fig. S1B in the supplemental material), like *M. catarrhalis* GRJ 11, which was isolated from a diseased individual in Salamanca (Spain) and also belongs to the seroresistant lineage (71). In line with this, experimental evidence also demonstrated that RH4 has a seroresistant phenotype (data not shown).

Expression profiling. Microarray expression profiling for three phases of *in vitro* growth in BHI medium (lag, exponential, and stationary) showed that 88.1% of the predicted CDSs were expressed during at least one growth phase and that 84.6% of the predicted CDSs were expressed in all three growth phases. Of the predicted CDSs for which no transcripts were detected, ~81% were annotated as (conserved) hypothetical. Further, we observed expression for 30 of the 41 (conserved) hypothetical CDSs that were not present in the ATCC 43617 strain. Expression of specific genes is discussed in detail below.

Virulence factors. Various proteins have been described as proteins that play pivotal roles in *M. catarrhalis* pathogenesis (for recent reviews, see references 15 and 53), and all but one of these proteins were found to be encoded by genes in the RH4 genome (Table 4). The ubiquitous surface proteins (*UspAs*) are among the major virulence factors; *UspA1* mediates binding to epithelial cells and extracellular matrix (ECM) components, and the mutually exclusive *UspA2* and *UspA2H* proteins have a role predominantly in immune evasion (3, 39). Determination of the modular structure of the predicted *UspA1* and *UspA2H* proteins (the strain did not possess a *uspA2* gene) revealed the presence of the VEEG-NINNY-VEEG amino acid sequence motif involved in binding to Chang conjunctival cells or fibronectin (10) in *UspA1*, whereas this motif was not present in *UspA2H* (Fig. 2). RH4 contains several other adhesins, namely, *M. catarrhalis* immunoglobulin D (IgD) binding protein (MID) (also referred to as hemagglutinin [Hag]) (22), the *M. catarrhalis* adherence protein (*McaP*) (63), and outer membrane protein (OMP) CD (33). Resistance to the action of the human complement system is an important aspect of *M. catarrhalis* virulence. Previous studies have shown that, in addition to the *UspA* proteins mentioned above, the *M. catarrhalis* proteins *CopB* (32), OMP CD (33), and OMP E (48) play a role in serum resistance, and the corresponding genes are all present in RH4. All of the virulence factors described above were found to be expressed at intermediate or high levels during all three phases of *in vitro* growth examined. Interestingly, the only known virulence locus not present in RH4 is the *mha* locus encoding filamentous hemagglutinin-like proteins involved in adhesion (4, 54). We did identify three ORFs (MCR_0770, MCR_0777, and MCR_0778) with a small region of homology (37 to 71 amino acids) to *mhaB1*, but transcriptional profiling indicated that there was either no expression (MCR_0778) or low levels of expression (MCR_0770 and MCR_0777) of these loci.

Biosynthesis of cell wall structures. A prominent surface component of *M. catarrhalis* that is generally considered to be an important virulence factor is lipooligosaccharide (LOS). Genes encoding LOS glycosyl transferases (*Lgt*), which are enzymes that catalyze the formation of core or branched oligosaccharide chains (43), as well as genes required for biogen-

TABLE 4. Major virulence factors of *M. catarrhalis* RH4^a

RH4 locus	Gene	Common name and general function of protein	Characteristics	Level of expression in growth phases in BHI medium ^b		
				Lag	Exponential	Stationary
MCR_1198	<i>uspA1</i>	Ubiquitous surface protein A1; functions in adhesion to epithelial cells and ECM components, inhibition of NF- κ B proinflammatory response via CEACAM binding, serum resistance, binding of complement regulatory proteins, and biofilm formation	9 poly(G) repeats upstream of the ORF, reported to correlate with a low level of expression ^c	+++	+++	+++
MCR_0329	<i>uspA2H</i>	Ubiquitous surface protein A2H; functions in adhesion to epithelial cells, serum resistance, and biofilm formation		+++	+++	+++
MCR_0617	<i>mid/hag</i>	Immunoglobulin D binding protein/hemagglutinin; functions in adhesion to epithelial cells and IgD binding	3 triplets of G residues at 5' end within the ORF, reported to correlate with a high level of expression ^d	+++	++	+++
MCR_0419	<i>mcaP</i>	<i>M. catarrhalis</i> adherence protein; functions in adhesion to epithelial cells		+++	+++	+++
MCR_1698	<i>ompCD</i>	Outer membrane protein CD; functions in adhesion to epithelial cells and middle ear mucin	RFLP pattern 1 (~200- and 100-nt BseJI restriction fragments), which is almost exclusively found in 16S type 1 isolates ^e	+++	+++	+++
MCR_0492	<i>copB</i>	Outer membrane protein CopB; involved in serum resistance and important for <i>in vivo</i> survival	RFLP type II (332- and 187-nt NsaI restriction fragments), which is found almost exclusively in 16S rRNA type I isolates ^e	+++	+++	++
MCR_0858	<i>ompE</i>	Outer membrane protein E; involved in serum resistance	Putative homolog of long-chain fatty acid transporter FadL	+++	+++	+++
MCR_1591	<i>bro-2</i>	Beta-lactamase family protein BRO-2		++	++	++

^a Abbreviations: ECM, extracellular matrix; ORF, open reading frame; RFLP, restriction fragment length polymorphism; nt, nucleotides.

^b ++, moderate; +++, high. See text.

^c See reference 40.

^d See reference 45.

^e See reference 66.

esis of the deoxy-D-manno-2-octulosonic acid (KDO)-lipid A moiety (for a review, see reference 55), were all identified in the RH4 genome (see Table S2 in the supplemental material). Expression was detected for all components of the pathways involved except *lgt5*, whose product catalyzes addition of the terminal α -(1 \rightarrow 4)-linked terminal galactose of the core oligosaccharide chain (70) (see Table S2 in the supplemental material). Analysis of the RH4 *lgt* locus (*lgt5*, *lgt1*, *lgt2b/c*, and *lgt3*) revealed that RH4 is a LOS type B strain (see Fig. S1C in the

supplemental material) (19); LOS type B strains are exclusively 16S rRNA type 1 isolates (66).

The peptidoglycan layer is the main target for β -lactam antibiotics that can be degraded by the BRO β -lactamases produced by *M. catarrhalis* (RH4 expresses the *bro-2* gene [Table 4]) (9). The complete set of genes required for biosynthesis of peptidoglycan (61) was identified in RH4 and was found to be expressed in all growth phases (see Table S3 in the supplemental material).

Type IV pili (TFP) have a wide variety of functions, including adhesion to epithelial cells, biofilm formation, and motility (44). Biogenesis of TFP is a complex process requiring a large set of proteins (52), and the genes encoding these proteins are present in the RH4 genome (see Table S4 in the supplemental material). Overall, low levels of expression were detected for most genes encoding components of the TFP pathway, except for *pilA*, which encodes the major pilin subunit, which on average was found to be highly expressed (see Table S4 in the supplemental material).

Protein secretion. Gram-negative bacteria transport proteins from the cytosol across the inner membrane (IM) to the periplasm via one of two protein secretion systems, either the Sec system or the Tat system (16, 41). RH4 was found to contain complete Sec machinery, as well as the main components of the Tat system (see Table S5 in the supplemental material). After deposition in the periplasmic space, proteins

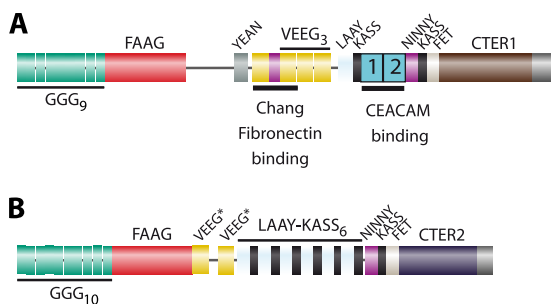


FIG. 2. Modular arrangement and functional domains of UspA1 (A) and UspA2H (B) of RH4. Both proteins are drawn to scale. The various sequence cassettes and the known functional domains are indicated. The VEEG repeats of UspA2H are indicated by asterisks because they only partially match the repeat motif. Abbreviations: CTER1, C-terminal region 1; CTER2, C-terminal region 2.

TABLE 5. Subcellular compartment distribution of the predicted RH4 proteins

Subcellular localization ^a	No. of proteins	% of total
Cytoplasm, intracellular	1,284	68.1
Cytoplasm (possibly secreted via nonclassical pathways), intracellular or secreted	35	1.9
Inner membrane, multitransmembrane	257	13.6
Inner membrane, multitransmembrane (lipid-modified N terminus)	4	0.2
Inner membrane, C terminally anchored (with CS)	1	0.1
Inner membrane (periplasm), N terminally anchored (no CS)	124	6.6
Inner membrane (periplasm), N terminally anchored (with CS)	14	0.7
Inner membrane (periplasm), lipid anchored	12	0.6
Periplasm, secreted (with CS)	22	1.2
Outer membrane, lipid anchored	53	2.9
Outer membrane, β -barrel protein (no CS)	19	1.0
Outer membrane, β -barrel protein (with CS)	23	1.2
Outer membrane (no clear SP found but have β -barrel), intracellular or β -barrel	20	1.1
Extracellular, secreted (with CS)	18	1.0

^a Abbreviations: CS, cleavage site; SP, signal peptide.

reach their final destination by other means. For instance, outer membrane lipoproteins are inserted into the outer membrane by the components of the Lol system (49) (see Table S5 in the supplemental material). Furthermore, several ORFs are predicted to encode components of resistance-nodulation-division (RND) efflux systems, such as components of the Acr and Mtr systems (50); components of both of these systems share homology with type I secretion system components (see Table S5 in the supplemental material). Strikingly, no components of the general secretory pathway, a type II secretion system, were found, although all of the evolutionarily related TFP assembly machinery (62) is present, as mentioned above. Even though a homolog of the type III secretion effector (HopJ) was identified (MCR_0582), other components of type III secretion systems were not found, nor did we identify components of type IV, V, and VI secretion systems. An overview of protein secretion components and the expression levels detected is shown in Table S5 in the supplemental material.

Protein localization prediction. A complete genome sequence provides an opportunity to discover novel vaccine targets, and surface-exposed and secreted proteins are of special interest for vaccine development. Ruckdeschel et al. used a genome mining approach on the partial genome sequence of *M. catarrhalis* ATCC 43617, which led to discovery and animal model testing of novel vaccine targets (58, 59). Extensive subcellular localization prediction revealed that 134 (7.1%) of the predicted RH4 proteins localize to the outer membrane or are secreted into the extracellular environment. The surface-exposed proteins include vaccine candidates, such as the lipid-anchored outer membrane protein Msp22 (59) and OMP E (48). In addition, we identified 35 (1.9%) proteins that are predicted to localize to the cytoplasm but may be secreted via nonclassical secretion pathways (i.e., not via the Sec or Tat pathway), including, for example, the autotransporter McaP (42). The subcellular compartment distribution of the predicted RH4 proteome is summarized in Table 5.

Central metabolic pathways. As a respiratory tract pathogen, *M. catarrhalis* is at least partially dependent on nutrients available inside its human host to fill its needs for energy and intermediates for biosynthesis of essential macromolecules. *M. catarrhalis* is reported to be incapable of utilizing exogenous carbohydrates and consequently does not produce acid from carbohydrates (12). In line with this, no genes encoding carbohydrate transport or catabolism components were found in the RH4 genome. Reconstruction of the central metabolism (Fig. 3; see Table S6 in the supplemental material) showed that while RH4 possesses an incomplete glycolytic pathway, all of the enzymes of the gluconeogenic pathway are present, indicating that carbohydrate intermediates can be synthesized. Gluconeogenesis uses phosphoenolpyruvate (PEP) as a starting substrate, which can be generated from tricarboxylic acid (TCA) cycle intermediates. The TCA cycle is supplied with acetyl coenzyme A (acetyl-CoA) via degradation of fatty acids and acetate assimilation, and the genes required for these processes were identified in RH4. The incomplete TCA cycle, which lacks both subunits of succinyl-CoA synthetase, can be bypassed by the glyoxylate pathway, which is complete in RH4. The glyoxylate pathway has an anaplerotic function and supplies the TCA cycle with oxaloacetate (acetyl-CoA acceptor molecule). Furthermore, the enzymes required for oxidative stages of the pentose phosphate pathway are not present, but transaldolase and the enzymes of the nonoxidative branch are present. As mentioned above, RH4 possesses all of the genes required for beta-oxidation of long-chain fatty acids (see Table S6 in the supplemental material) (23). Long-chain fatty acids are transported across the cell membrane by the combined action of the outer membrane protein FadL and the inner membrane-associated FadD protein, an acyl-CoA synthase. We predict that the highly conserved OMP E (48) is the FadL homolog in RH4. Two adjacent ORFs (67% identical) appear to encode homologs of FadD (aerobic) or its anaerobic counterpart, FadK (13). Interestingly, we did not detect obvious homologs of known fatty acid metabolism regulators, such as FadR, although we did identify members of the family containing such proteins, such as GntR and DeoR family regulators (23). Unfortunately, no obvious function could be assigned to these ORFs. Further, we identified two ORFs encoding proteins with homology to the 3-oxoacid CoA transferase involved in degradation of short-chain fatty acids (34). Finally, we identified all of the genes encoding the enzymes required for completion of fatty acid biosynthesis (23) (see Table S6 in the supplemental material).

Expression profiling showed that all of the gluconeogenesis, TCA cycle, fatty acid degradation, acetate assimilation, and pentose phosphate pathway genes are expressed during *in vitro* growth (see Table S6 in the supplemental material). However, expression of malate synthase, the second enzyme of the glyoxylate pathway, was not detectable in the lag and exponential phases of growth, and the levels during the stationary growth phase were low. This suggests that glyoxylate is further processed via other enzymes, e.g., glycerate dehydrogenase (MCR_1529) and phosphoglycolate phosphatase (MCR_0365), both of which were expressed at intermediate levels in all growth phases.

Nitrogen metabolism. It has been suggested that in *M. catarrhalis* a truncated denitrification pathway (reduction of ni-

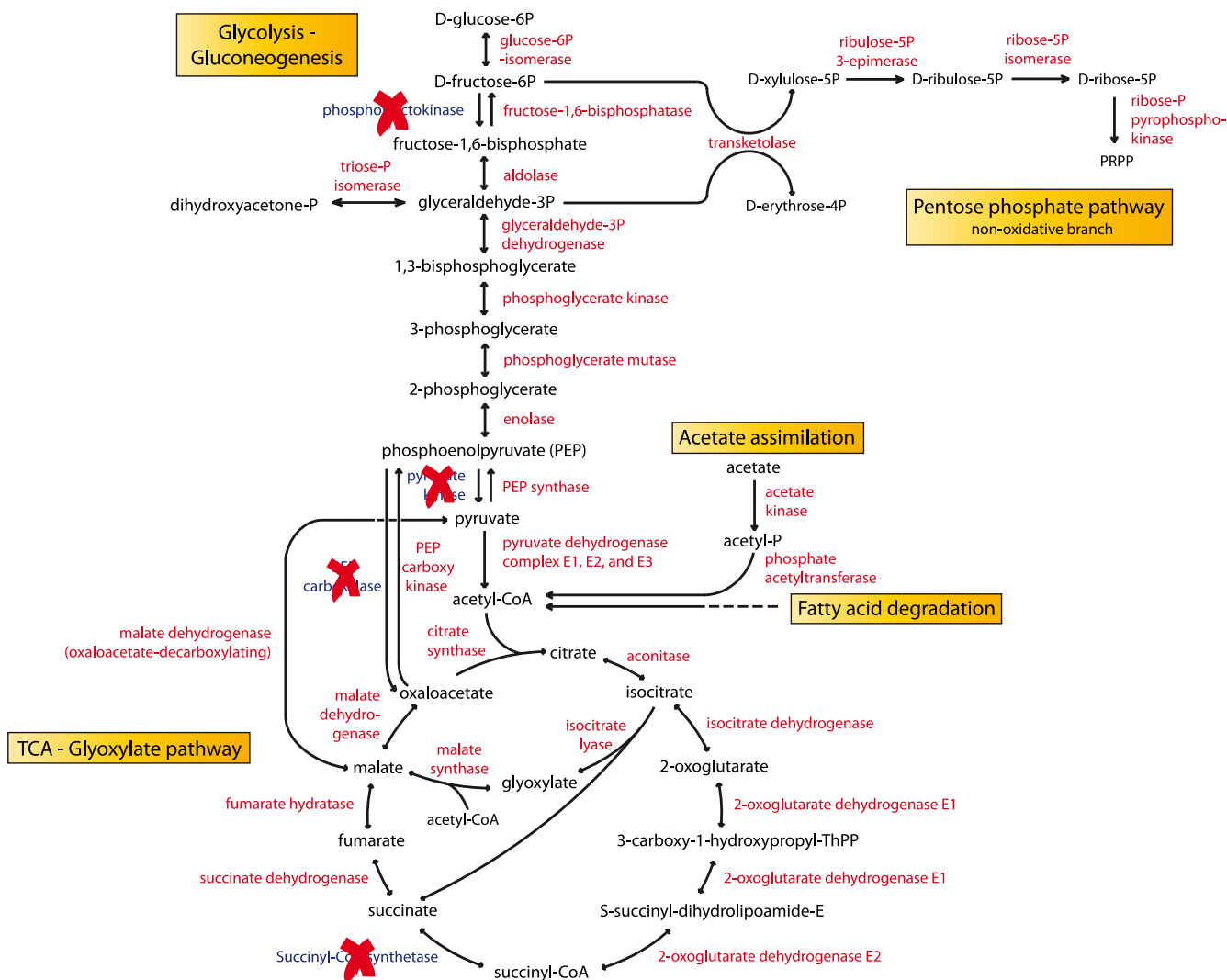


FIG. 3. Predicted central metabolism of *M. catarrhalis* RH4. The glycolysis pathway is incomplete and lacks the key enzymes phosphofruktokinase and pyruvate kinase. The enzymes of glycolysis are expected to be involved in the gluconeogenic pathway, all enzymes of which were found to be present. Fatty acid degradation and acetate assimilation are probably very important, as they supply the tricarboxylic acid (TCA) cycle with acetyl-CoA. The complete glyoxylate pathway is present and allows bypass of the TCA cycle, which lacks both subunits of succinyl-CoA synthetase.

trite to nitrous oxide) described by Wang et al. provides an alternative mechanism to generate energy under low-oxygen-tension conditions and contributes to biofilm formation and *in vivo* virulence. In diagnostic microbiological laboratories, reduction of nitrate is one of the differential tests used to confirm the identity of *M. catarrhalis* (12). The genes encoding the nitrate reductase complex (*narGHJI*), nitrite reductase (*aniA*), nitric oxide reductase (*norB*), and the *narX/narL* two-component system were identified in RH4 (see Table S7 in the supplemental material). In addition, a putative regulator of fumarate and nitrate reduction was identified; however, the true function of this fumarate and nitrate regulator (FNR)/cyclic AMP receptor protein (CRP) family member remains to be elucidated (69). In addition to the nitrate ABC transport system, encoded by *nrtABCD*, we identified two other candidates that could play a role in nitrate and nitrite transport, namely, NarK1 and a putative nitrate-nitrite transporter designated NarK2 (see Table S7 in the supplemental material). The levels

of expression of the denitrification pathway components were not uniform; e.g., the level of nitrate reductase expression ranged from undetectable to low, whereas the level of nitrite reductase expression was high during the lag and exponential phases (see Table S7 in the supplemental material).

Iron acquisition. Iron is a key nutrient that is functionally involved as a cofactor in various metabolic processes and is essential for both *M. catarrhalis* and its human host (56). The RH4 genome contains genes encoding components of many iron acquisition and transport systems (see Table S8 in the supplemental material), including all of the following *M. catarrhalis* proteins previously described as proteins involved in this process: lactoferrin binding proteins A and B (17), transferrin binding proteins A and B (74), heme utilization protein (25), *M. catarrhalis* hemoglobin utilization protein (26), CopB (1), and the main regulator of iron-responsive genes, Fur (24). In addition, iron may be acquired through degradation of heme, which is catalyzed by a heme oxygenase. An iron trans-

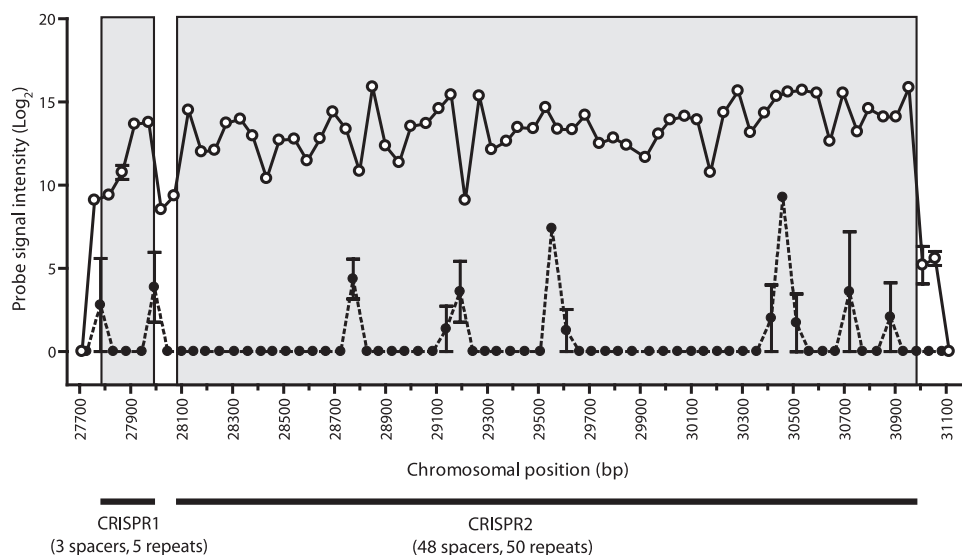


FIG. 4. Expression profiles of chromosomal CRISPR regions I (CRISPR1) and II (CRISPR2). The exponential-growth-phase \log_2 probe signal intensities of both strands (open circles, reverse-strand probes; filled circles, forward-strand probes) demonstrate that there is reverse-strand expression of both CRISPR1 and CRISPR2. The localization of CRISPR1 and CRISPR2 is indicated below the graph. Comparable results were obtained during the lag and stationary growth phases (data not shown).

port system for transport of Fe^{3+} from the periplasm to the cytosol, encoded by the *fbpABC* gene cluster, was identified based on homology to the corresponding locus in *H. influenzae* (2). Further, the *afeABCD* gene cluster was identified in RH4; this cluster has been proposed to be involved in the acquisition of chelated iron, as described previously for *Actinobacillus actinomycetemcomitans* and *Yersinia pestis*, and is regulated by Fur in these species (5, 57). Interestingly, the levels of expression of the iron binding components of the *fbpABC* and *afeABCD* ABC transporter systems were higher than the levels of expression of the other components of these systems. Two putative NRAMP homologs, which are involved in transport of Fe^{2+} and Mn^{2+} , may compete with the host divalent ion transporters of the NRAMP family (21). In addition, RH4 possesses two putative bacterioferritins, which are intracellular iron storage proteins important for preventing the presence of free iron (11), both of which were found to be expressed at high levels during all phases of growth. Overall, all of the iron acquisition and transport systems of RH4 were found to be expressed *in vitro*, but the importance of the individual pathways remains to be investigated.

Oxidative stress. Inherent in aerobic metabolism is oxidative stress caused by reactive oxygen species (ROS) (20). As mentioned above, acquisition of iron is essential for growth of *M. catarrhalis*, but it can also be harmful for the bacterium, as iron can react with hydrogen peroxide, resulting in the formation of hydroxyl radicals via the Fenton reaction (20). The well-studied superoxide dismutase-catalase system is able to counteract the effect of oxidative stress by catalyzing the conversion of superoxide to water and oxygen. RH4 contains and expresses the *sodA* and catalase genes, but it lacks the *sodB* gene (see Table S9 in the supplemental material). Catalase production is used during routine identification, but it has limited differential value because not all strains produce catalase (12). Further, RH4 was found to encode and express several putative bacte-

rial peroxiredoxins (Prx) that catalyze the reduction of peroxide, peroxyntirite, and diverse organic hydroperoxides (ROOH) (72), as well as alkyl hydroperoxide reductase-thioredoxin family members (67). Despite the presence of a large number of antioxidant genes in the RH4 genome, we could not identify homologs of known oxidative stress regulators, such as OxyR and SoxR (67).

CRISPR elements. Clustered regularly interspaced short palindromic repeats (CRISPRs) are widespread in genomes of prokaryotic organisms and are thought to be transcribed and processed into small RNAs that confer resistance to phages (30). CRISPR direct repeats are separated by nonrepetitive spacer elements and are often located near gene clusters encoding CRISPR-associated (Cas) protein family members. Two CRISPR regions, CRISPR regions I and II, and one putative repeat region (CRISPR region III) were identified in the RH4 genome. CRISPR region II, localized at nt 30997 to 28086, is in the vicinity of six genes encoding proteins with significant homology to known Cas proteins, showing high levels of similarity to the *Yersinia pestis* subtype (30). RH4 CRISPR regions I and II are characterized by an average repeat length of 28.1 nt (*Y. pestis*, 28.0 nt) and a spacer length of 32.0 nt (*Y. pestis*, 32.1 nt). Microarray analysis showed that there was constitutive expression of both CRISPR region I and CRISPR region II only from the minus strand (Fig. 4). In addition, intermediate levels of expression of the six *cas* genes were detected during all three growth phases (data not shown). The exact role of the CRISPR system in *M. catarrhalis* remains to be determined.

Conclusion. In this paper, we describe the completely assembled and annotated genome of the clinically important bacterial pathogen *M. catarrhalis* strain RH4. Comparative genomics revealed high degrees of similarity and sequence conservation for the RH4 and ATCC 43617 genomes. Like RH4, the ATCC 43617 strain is classified as a 16S rRNA type

1 strain (8), and MLST analysis showed that it also belongs to the seroresistant lineage (71). The completely assembled and annotated genome of a clinical isolate of *M. catarrhalis* described here should facilitate identification of novel (surface-exposed) vaccine targets, should provide the basis for omics-based research, such as transcriptomics and proteomics studies, and is indispensable for a complete understanding of the biology of *M. catarrhalis*.

ACKNOWLEDGMENTS

This work was supported by the OMVac project under the European Union 6th Framework Program. S.P.W.D.V. and H.J.B. were supported by a Vienna Spot of Excellence (VSOE) grant (ID337956). K.R. was supported by grants from the Anna and Edwin Berger Foundation, the Marianne and Marcus Wallenberg Foundation, and the Swedish Medical Research Council.

We thank Wolfgang Zimmermann and Steffen Krüger of Agowa Genomics for assistance with genomic sequencing, Miaomiao Zhou and Roland Siezen for performing the LocateP analyses, Tilman Todd for assistance with preparation of microarray data, and researchers of the Institute for Genome Science (IGS), particularly Michelle Gwinn Giglio, for use of their annotation engine.

REFERENCES

- Aebi, C., B. Stone, M. Beucher, L. D. Cope, I. Maciver, S. E. Thomas, G. H. McCracken, P. F. Sparling, and E. J. Hansen. 1996. Expression of the CopB outer membrane protein by *Moraxella catarrhalis* is regulated by iron and affects iron acquisition from transferrin and lactoferrin. *Infect. Immun.* **64**: 2024–2030.
- Anderson, D. S., P. Adhikari, K. D. Weaver, A. L. Crumbliss, and T. A. Mietzner. 2007. The *Haemophilus influenzae* hFbpABC Fe³⁺ transporter: analysis of the membrane permease and development of a gallium-based screen for mutants. *J. Bacteriol.* **189**:5130–5141.
- Attia, A. S., E. R. Lafontaine, J. L. Latimer, C. Aebi, G. A. Syrogiannopoulos, and E. J. Hansen. 2005. The UspA2 protein of *Moraxella catarrhalis* is directly involved in the expression of serum resistance. *Infect. Immun.* **73**: 2400–2410.
- Balder, R., J. Hassel, S. Lipski, and E. R. Lafontaine. 2007. *Moraxella catarrhalis* strain O35E expresses two filamentous hemagglutinin-like proteins that mediate adherence to human epithelial cells. *Infect. Immun.* **75**: 2765–2775.
- Bearden, S. W., and R. D. Perry. 1999. The Yfe system of *Yersinia pestis* transports iron and manganese and is required for full virulence of plague. *Mol. Microbiol.* **32**:403–414.
- Bendtsen, J. D., L. Kiemer, A. Fausboll, and S. Brunak. 2005. Non-classical protein secretion in bacteria. *BMC. Microbiol.* **5**:58.
- Berven, F. S., K. Flikka, H. B. Jensen, and I. Eidhammer. 2004. BOMP: a program to predict integral beta-barrel outer membrane proteins encoded within genomes of Gram-negative bacteria. *Nucleic Acids Res.* **32**:W394–W399.
- Bootsma, H. J., H. G. van der Heide, S. van de Pas, L. M. Schouls, and F. R. Mooi. 2000. Analysis of *Moraxella catarrhalis* by DNA typing: evidence for a distinct subpopulation associated with virulence traits. *J. Infect. Dis.* **181**: 1376–1387.
- Bootsma, H. J., H. van Dijk, J. Verhoef, A. Fleer, and F. R. Mooi. 1996. Molecular characterization of the BRO beta-lactamase of *Moraxella (Branhamella) catarrhalis*. *Antimicrob. Agents Chemother.* **40**:966–972.
- Brooks, M. J., J. L. Sedillo, N. Wagner, C. A. Laurence, W. Wang, A. S. Attia, E. J. Hansen, and S. D. Gray-Owen. 2008. Modular arrangement of allelic variants explains the divergence in *Moraxella catarrhalis* UspA protein function. *Infect. Immun.* **76**:5330–5340.
- Carrondo, M. A. 2003. Ferritins, iron uptake and storage from the bacterioferritin viewpoint. *EMBO J.* **22**:1959–1968.
- Catlin, B. W. 1990. *Branhamella catarrhalis*: an organism gaining respect as a pathogen. *Clin. Microbiol. Rev.* **3**:293–320.
- Cho, B. K., E. M. Knight, and B. O. Palsson. 2006. Transcriptional regulation of the *fad* regulon genes of *Escherichia coli* by ArcA. *Microbiology* **152**:2207–2219.
- Christensen, J. J., J. Ursing, and B. Bruun. 1994. Genotypic and phenotypic relatedness of 80 strains of *Branhamella catarrhalis* of worldwide origin. *FEMS Microbiol. Lett.* **119**:155–159.
- de Vries, S. P., H. J. Bootsma, J. P. Hays, and P. W. Hermans. 2009. Molecular aspects of *Moraxella catarrhalis* pathogenesis. *Microbiol. Mol. Biol. Rev.* **73**:389–406.
- Diessen, A. J., and N. Nouwen. 2008. Protein translocation across the bacterial cytoplasmic membrane. *Annu. Rev. Biochem.* **77**:643–667.
- Du, R. P., Q. Wang, Y. P. Yang, A. B. Schryvers, P. Chong, M. H. Klein, and S. M. Loosmore. 1998. Cloning and expression of the *Moraxella catarrhalis* lactoferrin receptor genes. *Infect. Immun.* **66**:3656–3665.
- Dutilh, B. E., Y. He, M. L. Hekkelman, and M. A. Huynen. 2008. Signature, a web server for taxonomic characterization of sequence samples using signature genes. *Nucleic. Acids. Res.* **36**:W470–W474.
- Edwards, K. J., J. M. Schwingel, A. K. Datta, and A. A. Campagnari. 2005. Multiplex PCR assay that identifies the major lipooligosaccharide serotype expressed by *Moraxella catarrhalis* clinical isolates. *J. Clin. Microbiol.* **43**: 6139–6143.
- Farr, S. B., and T. Kogoma. 1991. Oxidative stress responses in *Escherichia coli* and *Salmonella typhimurium*. *Microbiol. Rev.* **55**:561–585.
- Forbes, J. R., and P. Gros. 2001. Divalent-metal transport by NRAMP proteins at the interface of host-pathogen interactions. *Trends Microbiol.* **9**:397–403.
- Forsgren, A., M. Brant, M. Karamehmedovic, and K. Riesbeck. 2003. The immunoglobulin D-binding protein MID from *Moraxella catarrhalis* is also an adhesin. *Infect. Immun.* **71**:3302–3309.
- Fujita, Y., H. Matsuka, and K. Hirooka. 2007. Regulation of fatty acid metabolism in bacteria. *Mol. Microbiol.* **66**:829–839.
- Furano, K., and A. A. Campagnari. 2003. Inactivation of the *Moraxella catarrhalis* 7169 ferric uptake regulator increases susceptibility to the bactericidal activity of normal human sera. *Infect. Immun.* **71**:1843–1848.
- Furano, K., and A. A. Campagnari. 2004. Identification of a hemin utilization protein of *Moraxella catarrhalis* (HumA). *Infect. Immun.* **72**:6426–6432.
- Furano, K., N. R. Luke, A. J. Howlett, and A. A. Campagnari. 2005. Identification of a conserved *Moraxella catarrhalis* haemoglobin-utilization protein, MhuA. *Microbiology*. **151**:1151–1158.
- Gao, F., and C. T. Zhang. 2008. Ori-Finder: a web-based system for finding *oriC*s in unannotated bacterial genomes. *BMC Bioinform.* **9**:79.
- Grant, J. R., and P. Stothard. 2008. The CGView Server: a comparative genomics tool for circular genomes. *Nucleic. Acids. Res.* **36**:W181–W184.
- Grisso, I., G. Vergnaud, and C. Pourcel. 2007. CRISPRFinder: a web tool to identify clustered regularly interspaced short palindromic repeats. *Nucleic. Acids. Res.* **35**:W52–W57.
- Haft, D. H., J. Selengut, E. F. Mongodin, and K. E. Nelson. 2005. A guild of 45 CRISPR-associated (Cas) protein families and multiple CRISPR/Cas subtypes exist in prokaryotic genomes. *PLoS Comput. Biol.* **1**:e60.
- Hall-Stoodley, L., F. Z. Hu, A. Gieseke, L. Nistico, D. Nguyen, J. Hayes, M. Forbes, D. P. Greenberg, B. Dice, A. Burrows, P. A. Wackym, P. Stoodley, J. C. Post, G. D. Ehrlich, and J. E. Kerschner. 2006. Direct detection of bacterial biofilms on the middle-ear mucosa of children with chronic otitis media. *JAMA* **296**:202–211.
- Helminis, M. E., I. Maciver, J. L. Latimer, J. Klesney-Tait, L. D. Cope, M. Paris, G. H. McCracken, Jr., and E. J. Hansen. 1994. A large, antigenically conserved protein on the surface of *Moraxella catarrhalis* is a target for protective antibodies. *J. Infect. Dis.* **170**:867–872.
- Holm, M. M., S. L. Vanlerberg, I. M. Foley, D. D. Sledjeski, and E. R. Lafontaine. 2004. The *Moraxella catarrhalis* porin-like outer membrane protein CD is an adhesin for human lung cells. *Infect. Immun.* **72**:1906–1913.
- Jenkins, L. S., and W. D. Nunn. 1987. Genetic and molecular characterization of the genes involved in short-chain fatty acid degradation in *Escherichia coli*: the *ato* system. *J. Bacteriol.* **169**:42–52.
- Juncker, A. S., H. Willenbrock, H. G. Von, S. Brunak, H. Nielsen, and A. Krogh. 2003. Prediction of lipoprotein signal peptides in Gram-negative bacteria. *Protein Sci.* **12**:1652–1662.
- Kanehisa, M., M. Araki, S. Goto, M. Hattori, M. Hirakawa, M. Itoh, T. Katayama, S. Kawashima, S. Okuda, T. Tokimatsu, and Y. Yamaniishi. 2008. KEGG for linking genomes to life and the environment. *Nucleic. Acids. Res.* **36**:D480–D484.
- Karalus, R., and A. Campagnari. 2000. *Moraxella catarrhalis*: a review of an important human mucosal pathogen. *Microbes Infect.* **2**:547–559.
- Kent, W. J. 2002. BLAT—the BLAST-like alignment tool. *Genome Res.* **12**:656–664.
- Lafontaine, E. R., L. D. Cope, C. Aebi, J. L. Latimer, G. H. McCracken, Jr., and E. J. Hansen. 2000. The UspA1 protein and a second type of UspA2 protein mediate adherence of *Moraxella catarrhalis* to human epithelial cells in vitro. *J. Bacteriol.* **182**:1364–1373.
- Lafontaine, E. R., N. J. Wagner, and E. J. Hansen. 2001. Expression of the *Moraxella catarrhalis* UspA1 protein undergoes phase variation and is regulated at the transcriptional level. *J. Bacteriol.* **183**:1540–1551.
- Lee, P. A., D. Tullman-Ercek, and G. Georgiou. 2006. The bacterial twin-arginine translocation pathway. *Annu. Rev. Microbiol.* **60**:373–395.
- Lipski, S. L., C. Akimana, J. M. Timpe, R. M. Wooten, and E. R. Lafontaine. 2007. The *Moraxella catarrhalis* autotransporter McaP is a conserved surface protein that mediates adherence to human epithelial cells through its N-terminal passenger domain. *Infect. Immun.* **75**:314–324.
- Luke, N. R., S. Allen, B. W. Gibson, and A. A. Campagnari. 2003. Identification of a 3-deoxy-D-manno-octulosonic acid biosynthetic operon in *Moraxella catarrhalis* and analysis of a KdsA-deficient isogenic mutant. *Infect. Immun.* **71**:6426–6434.
- Luke, N. R., J. A. Jurcisek, L. O. Bakaletz, and A. A. Campagnari. 2007.

- Contribution of *Moraxella catarrhalis* type IV pili to nasopharyngeal colonization and biofilm formation. *Infect. Immun.* **75**:5559–5564.
45. Mollenkvist, A., T. Nordstrom, C. Hallden, J. J. Christensen, A. Forsgren, and K. Riesbeck. 2003. The *Moraxella catarrhalis* immunoglobulin D-binding protein MID has conserved sequences and is regulated by a mechanism corresponding to phase variation. *J. Bacteriol.* **185**:2285–2295.
 46. Murphy, T. F. 2009. Vaccine development for *Moraxella catarrhalis*: rationale, approaches and challenges. *Expert Rev. Vaccines* **8**:655–658.
 47. Murphy, T. F., A. L. Brauer, B. J. Grant, and S. Sethi. 2005. *Moraxella catarrhalis* in chronic obstructive pulmonary disease: burden of disease and immune response. *Am. J. Respir. Crit. Care Med.* **172**:195–199.
 48. Murphy, T. F., A. L. Brauer, N. Yuskiw, E. R. McNamara, and C. Kirkham. 2001. Conservation of outer membrane protein E among strains of *Moraxella catarrhalis*. *Infect. Immun.* **69**:3576–3580.
 49. Narita, S., S. Matsuyama, and H. Tokuda. 2004. Lipoprotein trafficking in *Escherichia coli*. *Arch. Microbiol.* **182**:1–6.
 50. Nikaido, H. 1996. Multidrug efflux pumps of Gram-negative bacteria. *J. Bacteriol.* **178**:5853–5859.
 51. Ouellet, M., P. D. Adams, J. D. Keasling, and A. Mukhopadhyay. 2009. A rapid and inexpensive labeling method for microarray gene expression analysis. *BMC Biotechnol.* **9**:97.
 52. Pelicic, V. 2008. Type IV pili: e pluribus unum? *Mol. Microbiol.* **68**:827–837.
 53. Perez Vidakovic, M. L., and K. Riesbeck. 2009. Virulence mechanisms of *Moraxella* in the pathogenesis of infection. *Curr. Opin. Infect. Dis.* **22**:279–285.
 54. Plamondon, P., N. R. Luke, and A. A. Campagnari. 2007. Identification of a novel two-partner secretion locus in *Moraxella catarrhalis*. *Infect. Immun.* **75**:2929–2936.
 55. Raetz, C. R., C. M. Reynolds, M. S. Trent, and R. E. Bishop. 2007. Lipid A modification systems in Gram-negative bacteria. *Annu. Rev. Biochem.* **76**:295–329.
 56. Ratledge, C., and L. G. Dover. 2000. Iron metabolism in pathogenic bacteria. *Annu. Rev. Microbiol.* **54**:881–941.
 57. Rhodes, E. R., A. P. Tomaras, G. McGillivray, P. L. Connerly, and L. A. Actis. 2005. Genetic and functional analyses of the *Actinobacillus actinomycescomitans* AfeABCD siderophore-independent iron acquisition system. *Infect. Immun.* **73**:3758–3763.
 58. Ruckdeschel, E. A., A. L. Brauer, A. Johnson, and T. F. Murphy. 2009. Characterization of proteins Msp22 and Msp75 as vaccine antigens of *Moraxella catarrhalis*. *Vaccine* **27**:7065–7072.
 59. Ruckdeschel, E. A., C. Kirkham, A. J. Lesse, Z. Hu, and T. F. Murphy. 2008. Mining the *Moraxella catarrhalis* genome: identification of potential vaccine antigens expressed during human infection. *Infect. Immun.* **76**:1599–1607.
 60. Sethi, S., and T. F. Murphy. 2008. Infection in the pathogenesis and course of chronic obstructive pulmonary disease. *N. Engl. J. Med.* **359**:2355–2365.
 61. Smith, C. A. 2006. Structure, function and dynamics in the mur family of bacterial cell wall ligases. *J. Mol. Biol.* **362**:640–655.
 62. Thanassi, D. G., and S. J. Hultgren. 2000. Multiple pathways allow protein secretion across the bacterial outer membrane. *Curr. Opin. Cell Biol.* **12**:420–430.
 63. Timpe, J. M., M. M. Holm, S. L. Vanlerberg, V. Basrur, and E. R. Lafontaine. 2003. Identification of a *Moraxella catarrhalis* outer membrane protein exhibiting both adhesin and lipolytic activities. *Infect. Immun.* **71**:4341–4350.
 64. Verduin, C. M., C. Hol, A. Fleer, H. van Dijk, and A. van Belkum. 2002. *Moraxella catarrhalis*: from emerging to established pathogen. *Clin. Microbiol. Rev.* **15**:125–144.
 65. Vergison, A. 2008. Microbiology of otitis media: a moving target. *Vaccine* **26**(Suppl. 7):G5–G10.
 66. Verhaegh, S. J., A. Streefland, J. K. Dewnarain, D. J. Farrell, A. van Belkum, and J. P. Hays. 2008. Age-related genotypic and phenotypic differences in *Moraxella catarrhalis* isolates from children and adults presenting with respiratory disease in 2001–2002. *Microbiology* **154**:1178–1184.
 67. Wang, G., P. Alamuri, and R. J. Maier. 2006. The diverse antioxidant systems of *Helicobacter pylori*. *Mol. Microbiol.* **61**:847–860.
 68. Wang, W., L. Reitzer, D. A. Rasko, M. M. Pearson, R. J. Blick, C. Laurence, and E. J. Hansen. 2007. Metabolic analysis of *Moraxella catarrhalis* and the effect of selected *in vitro* growth conditions on global gene expression. *Infect. Immun.* **75**:4959–4971.
 69. Wang, W., A. R. Richardson, W. Martens-Habbena, D. A. Stahl, F. C. Fang, and E. J. Hansen. 2008. Identification of a repressor of a truncated denitrification pathway in *Moraxella catarrhalis*. *J. Bacteriol.* **190**:7762–7772.
 70. Wilson, J. C., P. M. Collins, Z. Klipic, I. D. Grice, and I. R. Peak. 2006. Identification of a novel glycosyltransferase involved in LOS biosynthesis of *Moraxella catarrhalis*. *Carbohydr. Res.* **341**:2600–2606.
 71. Wirth, T., G. Morelli, B. Kusecek, A. van Belkum, C. van der Schee, A. Meyer, and M. Achtman. 2007. The rise and spread of a new pathogen: seroresistant *Moraxella catarrhalis*. *Genome Res.* **17**:1647–1656.
 72. Wood, Z. A., E. Schroder, H. J. Robin, and L. B. Poole. 2003. Structure, mechanism and regulation of peroxiredoxins. *Trends Biochem. Sci.* **28**:32–40.
 73. Yu, C. S., Y. C. Chen, C. H. Lu, and J. K. Hwang. 2006. Prediction of protein subcellular localization. *Proteins* **64**:643–651.
 74. Yu, R. H., and A. B. Schryvers. 1993. The interaction between human transferrin and transferrin binding protein 2 from *Moraxella* (*Branhamella*) *catarrhalis* differs from that of other human pathogens. *Microb. Pathog.* **15**:433–445.
 75. Zhou, M., J. Boekhorst, C. Francke, and R. J. Siezen. 2008. LocatEP: genome-scale subcellular-location predictor for bacterial proteins. *BMC Bioinform.* **9**:173.

AUTHOR'S CORRECTION

Genome Analysis of *Moraxella catarrhalis* Strain RH4, a Human Respiratory Tract Pathogen

Stefan P. W. de Vries, Sacha A. F. T. van Hijum, Wolfgang Schueler, Kristian Riesbeck, John P. Hays, Peter W. M. Hermans, and Hester J. Bootsma

Laboratory of Pediatric Infectious Diseases, Radboud University Nijmegen Medical Centre, Nijmegen, Netherlands; Centre for Molecular and Biomolecular Informatics, Radboud University Nijmegen Medical Centre, Nijmegen, Netherlands; NIZO Food Research, Ede, Netherlands; Intercell AG, Vienna Biocenter 3, Vienna, Austria; Medical Microbiology, Department of Laboratory Medicine, University Hospital Malmö, Lund University, Malmö, Sweden; and Department of Medical Microbiology and Infectious Diseases, Erasmus MC, Rotterdam, Netherlands

Volume 192, no. 14, pages 3574–3583, 2010. The genome sequence described is that of strain BBH18, isolated from sputum, and not of strain RH4, which was isolated from blood. Consequently, the title should read “Genome Analysis of *Moraxella catarrhalis* Strain BBH18, a Human Respiratory Tract Pathogen,” and the strain name RH4 should be replaced by BBH18 throughout the manuscript. Furthermore, in the abstract (page 3574, line 4), introduction (page 3574, column 2, line 27), and Materials and Methods (page 3575, column 1, lines 13 and 14), the phrase “isolated from blood of an infected patient (14)” should be changed to “isolated from sputum of a patient with COPD during an exacerbation (45).” The correct reference for strain BBH18 is reference 45 (A. Mollenkvist et al., *J. Bacteriol.* **185**:2285–2295, 2003).