The Complete Genome Sequence of the Emerging Pathogen *Mycobacterium haemophilum* Explains Its Unique Culture Requirements

JoAnn M. Tufariello,1,2 Christopher A. Kerantzas,3 Catherine Vilchèze,1,4,5 R. Brent Calder,1,6 Eric K. Nordberg,7 Jack A. Fischer,8 Travis E. Hartman,9 Eva Yang,3 Timothy Driscoll,1 Laura E. Cole,10 Robert Sebra,9 Shahina B. Maqbool,9 Alice R. Wattam,9 William R. Jacobs, Jr.,9,10

Department of Microbiology and Immunology,1 and Howard Hughes Medical Institute;2 Division of Computational Genetics and Department of Genetics,3 Albert Einstein College of Medicine, Bronx, New York, USA;4 Virginia Bioinformatics Institute, Virginia Tech University, Blacksburg, Virginia, USA;5 Department of Biology (Microbial Metagenomics), West Virginia University, Morgantown, Virginia, USA;6 Department of Genetics and Genomic Sciences, Icahn School of Medicine, Mount Sinai, New York, New York, USA;7 Epigenomics Shared Facility, Albert Einstein College of Medicine, Bronx, New York, USA;8

ABSTRACT  *Mycobacterium haemophilum* is an emerging pathogen associated with a variety of clinical syndromes, most commonly skin infections in immunocompromised individuals. *M. haemophilum* exhibits a unique requirement for iron supplementation to support its growth in culture, but the basis for this property and how it may shape pathogenesis is unclear. Using a combination of Illumina, PacBio, and Sanger sequencing, the complete genome sequence of *M. haemophilum* was determined. Guided by this sequence, experiments were performed to define the basis for the unique growth requirements of *M. haemophilum*. We found that *M. haemophilum*, unlike many other mycobacteria, is unable to synthesize iron-binding siderophores known as mycobactins or to utilize ferri-mycobactins to support growth. These differences correlate with the absence of genes associated with mycobactin synthesis, secretion, and uptake. In agreement with the ability of heme to promote growth, we identified genes encoding heme uptake machinery. Consistent with its propensity to infect the skin, we show at the whole-genome level the genetic closeness of *M. haemophilum* with *Mycobacterium leprae*, an organism which cannot be cultivated in vitro, and we identify genes uniquely shared by these organisms. Finally, we identify means to express foreign genes in *M. haemophilum*. These data explain the unique culture requirements for this important pathogen, provide a foundation upon which the genome sequence can be exploited to improve diagnostics and therapeutics, and suggest use of *M. haemophilum* as a tool to elucidate functions of genes shared with *M. leprae*.

IMPORTANT  *Mycobacterium haemophilum* is an emerging pathogen with an unknown natural reservoir that exhibits unique requirements for iron supplementation to grow in vitro. Understanding the basis for this iron requirement is important because it is fundamental to isolation of the organism from clinical samples and environmental sources. Defining the molecular basis for *M. haemophilum*’s growth requirements will also shed new light on mycobacterial strategies to acquire iron and can be exploited to define how differences in such strategies influence pathogenesis. Here, through a combination of sequencing and experimental approaches, we explain the basis for the iron requirement. We further demonstrate the genetic closeness of *M. haemophilum* and *Mycobacterium leprae*, the causative agent of leprosy which cannot be cultured in vitro, and we demonstrate methods to genetically manipulate *M. haemophilum*. These findings pave the way for the use of *M. haemophilum* as a model to elucidate functions of genes shared with *M. leprae*.

---

Mycobacterium haemophilum is a slow-growing acid-fast bacterium of the nontuberculous mycobacteria group that prefers a lower temperature for growth (30°C to 32°C) and requires iron supplementation, typically either hemin or high concentrations of ferric ammonium citrate, for its growth (1–3). Like other mycobacterial species that grow optimally at lower temperatures, such as *Mycobacterium marinum* and *Mycobacterium ulcerans*, *M. haemophilum* is most associated with cutaneous infections of the extremities (1, 4). Less common clinical manifestations of *M. haemophilum* infection include skeletal infections, such as osteomyelitis and septic arthritis, pulmonary infections, ophthalmologic involvement, and disseminated disease (1, 5–11). These cutaneous and extracutaneous manifestations occur primarily in the setting of severe immune compromise (12). Some case reports...
have also described a leprosy-like presentation of M. haemophilum in immunocompromised individuals (13, 14). M. haemophilum has also been reported to cause cervicofacial lymphadenitis in immunocompetent children (15) and adults (16).

Infection with M. haemophilum occurs worldwide and, although it has been until now an infrequently isolated pathogen, the incidence appears to be rising since its initial discovery in 1978 (2). This likely reflects the emergence of severe immune compromise due to HIV/AIDS or in the settings of bone marrow/solid organ transplantation, hematologic malignancy, or treatment with disease-modifying agents for rheumatologic disorders (8, 12, 17–22), all of which predispose to infection with this opportunistic pathogen. At present there is no standardized protocol for in vitro susceptibility testing (1), although recommendations for a disk agar elution method for M. haemophilum are included in a recent Clinical and Laboratory Standards Institute (CLSI) document (23). Further, it is unknown how predictive the results are of clinical outcomes. Optimal treatment regimens also remain undefined (1); however, there is a general consensus that patients should receive a combination of agents that appear to be active in vitro, such as clarithromycin, ciprofloxacin, and a rifamycin, for a prolonged duration (~12 to 24 months) which is tailored to the individual’s presentation and degree of immune compromise (1, 12, 24, 25). For immunocompetent patients with lymph node involvement, surgical excision may be sufficient (1, 12). The natural habitat/reservoir of M. haemophilum as well as its mode of transmission also remain poorly defined. While many nontuberculous mycobacteria are found in soil and aquatic habitats, attempts to recover M. haemophilum from environmental samples have often proven unsuccessful, although this may be attributed in part to the slow growth of the organism and its fastidious growth requirements (1).

Very little genetic information has been available for M. haemophilum. This limits understanding of its unique growth requirements and how these properties shape its pathogenesis. We therefore determined the complete genome sequence of the M. haemophilum DSM 46634 type strain and performed accompanying experiments that implicated the absence of the mycobactin synthesis machinery, which in other mycobacteria plays a critical role in iron acquisition, as a major factor in the iron dependence of this organism. Further analysis demonstrated a close evolutionary relationship with Mycobacterium leprae and identified both genetic features conserved with other mycobacteria and also unique features that reflect the specific biology of this emerging pathogen. The data also provide groundwork for the development of improved diagnostic tests that will facilitate understanding of the pathogenesis, epidemiology, and mode of transmission of this emerging pathogen. Finally, as a first step toward experimental use of this genetic information, we identify means to introduce and express foreign genetic material in M. leprae.

RESULTS AND DISCUSSION

In vitro growth phenotypes. M. haemophilum requires iron supplementation for its growth. We explored the effect of different iron sources by propagating M. haemophilum in heme-containing 7H9 liquid medium and then washing and plating the bacteria onto various solid media. Unable to grow on oleic acid-albumin-dextrose-catalase (OADC)-enriched Middlebrook 7H10 agar without other supplements, M. haemophilum grew well on 7H10 medium containing 100 μM hemin (Fig. 1A). Growth was also well-supported on 7H10 agar supplemented with human hemoglobin and on 5% sheep blood agar. Growth was somewhat less abundant on chocolate agar (in which blood cells have been lysed by heating) (Fig. 1A). These data are consistent with prior observations (2, 5, 26, 27).

Mycobacterium tuberculosis and other mycobacteria produce siderophores, small iron-chelating compounds, termed mycobactins and carboxymycobactins, which bind and facilitate uptake of environmental ferric iron (Fe3+). In one prior report, several M. haemophilum clinical isolates failed to grow on Lowenstein-Jensen (LJ) medium supplemented with mycobactin at 2 μg/ml (28), but earlier studies on one of the same isolates (prior to its identification as M. haemophilum) reported that mycobactin did support growth (29). We observed that M. haemophilum failed to form colonies on 7H10 agar supplemented with 2 μg/ml mycobactin (from Mycobacterium avium subsp. paratuberculosis), but it did form colonies with 100 μM hemin (Fig. 1A). We also compared the growth properties of M. haemophilum with mc6230, an M. tuberculosis H37Rv ΔRD1 ΔpanCD auxotrophic strain that
requires pantothenate supplementation (7H10-pan) and can be used under biosafety level 2 (BSL2) containment (30, 31), with those of a ΔmbtB mutant strain generated in the mc²6230 background which lacks an essential component of the mycobactin synthesis pathway (32). Whereas mc²6230 grew well under all conditions tested, the ΔmbtB mutant failed to grow on 7H10-pan but did grow on 7H10-pan supplemented with 100 μM hemin, 2 μg/ml mycobactin J, and 200 ng/ml M. smegmatis carboxymycobactin (Fig. 1B). In contrast, M. haemophilum was unable to grow with either mycobactin J or carboxymycobactin supplementation, although again, growth was abundant on 100 μM hemin (Fig. 1B). These results suggest a defect(s) in the uptake and/or utilization of siderophore-bound iron by M. haemophilum.

In several experiments, inoculation of M. haemophilum onto 7H10 medium supplemented with 15 mg/ml ferric ammonium citrate (FAC) resulted in absent or very suboptimal growth (data not shown). Others have observed growth of M. haemophilum on LJ medium supplemented with 15 mg/ml FAC but not on 7H10 medium with the same concentration of FAC (28, 33). Conversely, although 7H10 supplemented with hemin supports growth, it has been observed that LJ with hemin does not (27, 28, 33). These findings suggest that interactions occur between particular iron complexes and other components of the media.

Overview of the M. haemophilum genome. To gain further insights into its growth requirements, the complete M. haemophilum genome sequence was determined. The genome consists of a single circular chromosome of 4,235,765 bp with an average GC content of 63.9% and encodes an estimated 3,964 genes (NCBI Prokaryotic Genome Annotation Pipeline). Pertinent characteristics of the genome in comparison with other mycobacterial species are listed in Table S1 in the supplemental material. Functional assignments using database comparisons were obtained for approximately two-thirds of coding sequences, with the remainder encoding hypothetical proteins (RAPID Annotation Using Subsystem Technology, or RAST [National Microbial Pathogen Data Resource]). Gene numbers referred to in the text are those assigned by RAST under Genome ID 1202450.9. A representation of the M. haemophilum chromosome is shown in Fig. 2, including the orientation of the coding sequences, percent GC content, and GC skew.

Mycobactin synthesis: mbt-1 (mbtA-mbtf) gene cluster. In M. tuberculosis, the salicylate-based mycobactin siderophores are synthesized by enzymes encoded in two distinct gene clusters, mbt-1 (mbtA-mbtf) and mbt-2 (mbtK-mbtN) (34, 35). One striking feature of the M. haemophilum genome is the absence of close homologues of almost all of the 14 genes comprising these two mycobactin synthesis clusters (Fig. 3A and B). For the mbtA-mbtf cluster, close homologues of 7 of the 10 genes—mbtG, mbtF, mbtE, mbtD, mbtC, mbtB, and mbtA—appear to be missing in M. haemophilum (Fig. 3A). An open reading frame (ORF) encoding a putative protein with significant homology to MbtH, found at one end of the mbtA-mbtf locus in H37Rv, does appear to be present [with 78% identity to H37Rv MbtH (Rv2377c) over 71 amino acids]. However, the product of this gene may not function in mycobactin synthesis, as homologues in other bacteria, such as actinomycetes and a rapid-growing mycobacterium, Mycobacterium smegmatis, have been implicated in other pathways, such as synthesis of cell wall glycopeptidolipids (36–38).

For the remainder of the mbt-1 gene cluster, close homologues of mbtF and mbtB, found at the opposite end of the H37Rv gene cluster, are also present in M. haemophilum (Fig. 3A). However, the M. haemophilum mbtJ and mbtI homologues are in a locus physically distant from the mbtH homologue, separated by greater than 600 kb. MbtI is a putative acetyl hydrolase whose precise role in mycobactin synthesis remains unclear (35). MbtI, the first enzyme in the mycobactin synthesis pathway, is a salicylate synthase (39). M. haemophilum does not possess a full-length mbtI gene but does contain an open reading frame encoding a product with homology to the C-terminal portion of MbtI (88% identity to the C-terminal 100 amino acids of the H37Rv MbtI). This region also shares homology with TrpE (anthranilate synthase, which catalyzes the first committed step in tryptophan biosynthesis) and other enzymes which act on chorismate in a variety of metabolic pathways (39). Therefore, this gene might participate in a different metabolic pathway.

Mycobactin synthesis: mbt-2 (mbtK-mbtN) gene cluster. Although M. tuberculosis mbt-1 encodes enzymes necessary for synthesis of the mycobactin backbone, it lacks genes encoding enzymes that attach the fatty acyl chain. These components of the siderophore biosynthetic machinery lie within the mbt-2 (mbtK-mbtN) gene cluster (34). The entire mbt-2 gene cluster appears to be absent in M. haemophilum (Fig. 3B). Together with the absence of mbtG, mbtF, mbtE, mbtD, mbtC, mbtB, and mbtA homologues, the absence of all four genes comprising the mbt-2 cluster makes it highly unlikely that M. haemophilum is capable of synthesizing mycobactin siderophores. Interestingly, the only mycobacterial species with similarly extensive deletions of the mbt-1 and mbt-2 regions are M. leprae (40) and the recently identified and closely related M. lepromatosis (41, 42), both of which have genomes characterized by extreme reductive evolution.

Functional analysis of mycobactin synthesis. To experimentally address production of mycobactin or carboxymycobactin by M. haemophilum, we cultured the bacteria under iron-replete conditions followed by transfer to iron-limited medium containing the radiolabeled mycobactin precursor [7,14C]salicylic acid. The iron-limited medium should, based on data from M. tuberculosis (43), increase expression of the mycobactin synthesis machinery. M. tuberculosis mc²6230, which has intact mycobactin synthesis, secretion, and uptake systems, was grown under similar conditions and served as a control. We tested two growth conditions: chelated 7H9 (see the supplemental text describing further materials and methods) in the presence and absence of 10 μM hemin. We expected that the latter might afford better growth while still imposing iron restriction. Cell-associated and secreted siderophores extracted from cell pellets and culture filtrates were analyzed by thin-layer chromatography (TLC). Because M. haemophilum failed to grow as well as mc²6230 in the iron-limited medium, even with the addition of 10 μM hemin, the two M. haemophilum cultures were pooled for siderophore extraction. TLC analysis revealed an absence of mycobactin or carboxymycobactin production by M. haemophilum under the growth conditions tested, in either the cell pellet or the culture supernatant (Fig. 4). For mc²6230, both mycobactin and carboxymycobactin were present in the cell pellet, with smaller amounts in the culture supernatant. These findings are the first direct evidence for lack of mycobactin/carboxymycobactin production by M. haemophilum. Due to relatively poor growth of M. haemophilum compared with mc²6230 in the iron-limited medium, a significantly smaller cell mass was analyzed; however, despite equal radioactive label on material being run on the TLC, there was still a complete absence
of mycobactin and carboxymycobactin detected in the *M. haemophilum* samples, and the only radiolabeled species associated with the cell pellet was [14C]salicylate. The suboptimal growth of *M. haemophilum* in iron-limited medium is, in itself, consistent with a lack of siderophore production by the strain. The *M. haemophilum* genome, like that of *M. tuberculosis*, appears to lack a gene cluster homologous to that present in the rapid-growing *M. smegmatis* which encodes the synthesis of the non-salicylate-containing peptidic exochelin siderophores (44–46); however, the possibility that *M. haemophilum* synthesizes an alternative siderophore is not excluded.

**Mycobactin uptake and secretion: IrtA/B and MmpL/MmpS proteins.** Adjacent to the mbt-2 gene cluster in *M. tuberculosis* H37Rv are the iron-regulated transporter A (*irtA*; Rv1348) and

*FIG 2* Circular representation of the *M. haemophilum* chromosome. Concentric rings, numbered from outer to inner rings, show the following: 1, scale (in megabases) starting at the top center and running clockwise; 2, forward coding sequences (CDS), in dark green; 3, reverse CDS, in light green; 4, GC content (percentage) as a line plot, with axes from 20% (inner) to 80% (outer), with lines every 15%, where the dark orange background indicates a GC content above 50%; 5, GC skew of genes as a line plot symmetrical about a skew of 1. The plot was generated using the CIRCOS circular genome data visualization program and the genome sequence as submitted for annotation by RAST.
irtB (Rv1349) genes, which are cotranscribed (47) and, along with the mbt-1 and mbt-2 clusters, are part of the IdeR regulon, with their transcription repressed under iron-replete conditions (43, 48). irtA and irtB appear to encode an ABC transporter which participates in the uptake and utilization of siderophore-bound iron and enables growth under iron-deficient conditions (49).

Homologues of irtA and irtB are lacking in M. haemophilum (Fig. 3B), again similar to M. leprae. Recently, members of the mycobacterial membrane protein large (MmpL) and small (MmpS) families have been found to play a role in the secretion of mycobacterial siderophores (50). M. haemophilum encodes close counterparts of MmpL4 (Mh0573) and MmpS4 (Mh0574), located immediately adjacent to one another. However, although there are additional mmpL and mmpS genes, we were unable to identify equally close homologues of MmpL5 and MmpS5, which contribute to siderophore synthesis and secretion in M. tuberculosis (50). Notably, MmpS4 is also required for export of cell surface glycopeptidolipids in M. smegmatis (51).

Therefore, M. haemophilum MmpS4/MmpL4 homologues may participate in transport of substances other than mycobactins.

Heme uptake, utilization, and biosynthesis. Recently, a heme acquisition system involving a secreted heme-binding protein (Rv0203) and the transmembrane proteins MmpL11 (Rv0202c) and MmpL3 (Rv0206c) was identified for M. tuberculosis (52). M. haemophilum encodes homologues of Rv0203 (Mh0273), MmpL11 (Mh0272), and MmpL3 (Mh0276), located in close proximity to one another, similar to their counterparts in H37Rv. M. haemophilum also encodes a homologue (Mh3902) of the
**FIG 4** *M. haemophilum* does not produce mycobactin. Synthesis and release of mycobactin was assessed by culturing *M. tuberculosis* strain mc²6230 or *M. haemophilum* in iron-deficient media, conditions that induced mycobactin synthesis for *M. tuberculosis*. [¹⁴C]salicylate was added to cultures to label newly synthesized mycobactin, and extracts of bacterial pellets and culture supernatants were analyzed by TLC by loading equal activity (counts per minute). Lane 1, mc²6230 cell pellet; lane 2, mc²6230 supernatant; lane 3, mc²6230 supernatant; lane 4, *M. haemophilum* cell pellet; lane 5, *M. haemophilum* supernatant; lane 6, [¹⁴C]salicylate. Indicated are the expected locations of mycobactin (Mb) and carboxymycobactin (cMb) from *M. smegmatis*, based on migration of purified nonradioactive standards (data not shown).

*M. tuberculosis* MhuD (mycobacterial heme utilization, degrader; Rv3592), which binds to and degrades heme (53). Therefore, our genome sequence data reveal intact heme uptake and utilization systems in *M. haemophilum*, in agreement with the results of the growth studies reported here. This in contrast to *M. leprae*, which, although it encodes homologues of MmpL11 and MmpL3, appears to lack a counterpart to the secreted heme-binding protein Rv0203 (40), suggesting a possible deficiency in the scavenging of heme-iron. *M. haemophilum* also appears to possess close homologues of *M. tuberculosis* H37Rv genes that encode putative heme biosynthetic pathway components (see Table S2 in the supplemental material), and again this differs from *M. leprae*, which is missing the hemN gene (42), and therefore may be impaired in heme biosynthesis.

**ESX loci including ESX-3.** Type VII secretion systems (T7SSs) encoded within ESX loci permit transport of proteins across the lipid-rich cell envelope to the cell surface or external environment. The ESX loci are large gene clusters which code for the ATP-dependent secretory apparatuses needed to export members of the 6-kDa early secretory antigenic target (ESAT-6)/culture filtrate protein 10 (CFP-10) protein families (EssA/EssB of the ESX-1 region and their paralogues), as well as a number of additional/associated substrates which remain to be fully defined. *M. tuberculosis* possesses five ESX systems, encoded within gene clusters designated ESX-1 to ESX-5, several of which play important roles in *M. tuberculosis* pathogenesis and intracellular survival (54). *M. haemophilum* possesses homologues of all five ESX systems (see Table S3 in the supplemental material).

The ESX-3 locus of *M. haemophilum* is of particular interest, because it has been implicated in iron and zinc homeostasis in *M. tuberculosis* (55–57). *M. tuberculosis* ess-3 expression is regulated by the availability of iron and zinc, under control of both the IdeR repressor and the zinc uptake regulator Zur/FurB (Rv2359) (58). ESX-3 is likely necessary for the uptake and/or utilization of iron bound to mycobactin (57), although involvement in alternative mycobacterial metal ion uptake processes has also been suggested (59, 60). The absence of genes required for mycobactin synthesis and mycobactin uptake/export in *M. haemophilum* makes a largely intact ESX-3 locus in *M. haemophilum* noteworthy (Fig. 5). Of the 11 ORFs in the *M. tuberculosis* ESX-3 region, 10 have close homologues in the corresponding region of *M. haemophilum*, including genes encoding both EssG and EssH, which are believed to be secreted as a heterodimeric complex (57, 61), similar to the EssA-EssB secreted effectors of the ESX-1 region.

All ESX loci with the exception of the smallest and most archaic, ESX-4, possess a pe-pppe gene pair adjacent to the ess gene pair (54); these are named for the conserved proline-glutamic acid (PE) and proline-proline-glutamic acid (PPE) residues found near the N termini of the encoded proteins. The only gene which is absent from its expected location within the ESX-3 region of *M. haemophilum* is ppe4 (Rv0286), the ppe component of the pe-pppe pair present within the *M. tuberculosis* H37Rv ESX-3 locus (Fig. 5). The significance of this absence with regard to the iron-dependent growth phenotype of the organism and the inability of mycobactin to support growth (Fig. 1A and B) is unclear, because the functions of the pe and ppe gene families remain largely obscure. It is possible that this deficiency is compensated by ppe homologues encoded elsewhere in the genome. Notably, *M. haemophilum* does possess a ppe gene immediately upstream of the ESX-3 locus.

Although the *M. haemophilum* genome shares with *M. tuberculosis* the presence of five ESX loci, there are some differences from the corresponding *M. tuberculosis* ESX regions. The primary difference is that several of the pe and ppe genes are absent from *M. haemophilum*, such as an absent pe35 homologue within the ESX-1 locus and an absent ppe4 in the ESX-3 locus, although some of the regions differ in other respects as well (see Table S3 in the supplemental material). For ESX-2, homologues of all 12 of the corresponding *M. tuberculosis* genes are present, but the region is discontinuous, similar to findings for *M. avium* subsp. *paratuberculosis* (62).

**Genes present in *M. haemophilum* but absent from *M. tuberculosis* H37Rv that may be related to iron acquisition.** Comparison of the *M. haemophilum* genome with other mycobacterial species revealed a number of genes present in *M. haemophilum* but absent from *M. tuberculosis* H37Rv. These included several genes with potential roles in iron acquisition. One, Mh1776, is annotated as ferrous iron transport protein B (*feoB*) and encodes a protein with significant homology to an integral component of the Feo inner membrane transport system, which functions in the uptake of ferrous (Fe²⁺) iron (63, 64). Although this system has been best studied in enteric Gram-negative bacteria such as *Escherichia coli* and *Salmonella* species (65, 66), genes encoding the Feo system have been identified in more than 40% of sequenced bacterial genomes (64).

In *E. coli*, the Feo system contains three genes encoded in the locus *feoABC*, although in other bacterial genomes, *feoC* is often absent (64). This is the situation for *M. haemophilum*, which con-
FeoB in a number of nontuberculous mycobacteria, including M. haemophilum, bears two copies of a gene (Mh0536) with homology to ATP-binding cassette (ABC) transmembrane transport systems of the iron/siderophore/heme/vitamin B₁₂ type (67). The small hydrophilic FeoA protein appears to interact with FeoB and to be required for transport (68). The presence of FeoA and FeoB suggests that M. haemophilum may encounter conditions (anaerobic and/or acidic) in its in vivo niche that increase the availability of ferrous iron, or perhaps that the bacterium is capable of reducing extracellular ferric ions to promote uptake by this system. Although absent in M. tuberculosis, BLAST analysis identifies homologues of FeoA and FeoB in a number of nontuberculous mycobacteria, including M. kansasi, M. marinum, M. genavensis, and M. xenopi.

M. haemophilum also possesses a second natural resistance-associated macrophage protein (Nramp) family member. Nramps were initially identified in eukaryotes as integral membrane proteins that function as pH-dependent high-affinity divalent metal ion transporters (69–71). Many bacteria also possess Nramp family members (72, 73), including M. tuberculosis, which has a single homologue (74–76). This is in contrast to M. leprae which, even with its small genome size, bears two copies of a mntH/nramp homologue (40); our sequencing results demonstrated that this is also true for M. haemophilum (Mh0166 and Mh3510), which occupies a somewhat similar in vivo niche. These conserved Nramp homologues are candidate iron uptake factors.

Iron uptake in bacteria is both critical and complex, characterized by the presence of multiple transport systems which may be utilized under different environmental conditions (77). In addition to the systems described above, it is likely that M. haemophilum encodes additional gene products involved in iron uptake. For instance, M. haemophilum encodes a locus (Mh0534-Mh0536) with homology to ATP-binding cassette (ABC) transport systems of the iron/siderophore/heme/vitamin B₁₂ type (78), including putative permease (Mh0536) and ATP-binding (Mh0535) components as well as a putative substrate binding protein (Mh0534) homologous to proteins that play roles in iron uptake in other bacteria. Interestingly, this locus is absent in both M. tuberculosis and M. leprae. Future studies examining growth in the presence of ferrous iron and transcriptional responses of M. haemophilum under iron-replete versus iron-limited conditions and aerobic versus hypoxic conditions may help clarify the contributions of the gene products described above and identify additional genes involved in novel iron uptake pathways. Ultimately, it may also be informative to compare in vitro gene expression profiles to those obtained during intracellular and in vivo growth, to help elucidate mechanisms of iron acquisition in the context of infection.

Evolutionary relationships of M. haemophilum. Previous phylogenetic analyses based on limited available sequence data, such as rpoB gene comparisons, have suggested a close evolutionary relationship between M. haemophilum and M. leprae (79, 80). Our complete genome sequence data confirm the closeness of this relationship. A phylogenetic tree constructed by whole-chromosome comparison to 14 selected mycobacterial genomes using the BLAST algorithm revealed the closest neighbors to M. haemophilum DSM 44634/ATCC 29548 to be the M. leprae strains TN and B4923 and the recently identified M. lepromatosis (Fig. 6). This is also consistent with the genetic similarities detailed above and with analyses of fatty acid compositions and phenolic glycolipid contents of the organisms (81–83). However, at ~4.23 Mb, the genome of M. haemophilum is substantially larger than the ~3.27-Mb genomes of M. leprae and M. lepromatosis. The latter have both undergone extensive pseudogenization, of about approximately 50% of the genome (40, 42), while for M. haemophilum pseudogenes account for only ~7% of the total (NCBI Prokaryotic Genome Annotation Pipeline) (see Table S1 in the supplemental material). This suggests that the described reductive evolution of M. leprae and M. lepromatosis was not shared with the
common ancestor, but instead began after divergence of *M. haemophilum* from the other two species.

**Expression of foreign genes in *M. haemophilum***. Exploiting the genome sequence of and developing *M. haemophilum* into a model system could expand our understanding of its pathophysiology and allow new insights into *M. leprae*, which cannot be cultivated in vitro. Therefore, we explored whether we could transform *M. haemophilum* by electroporation, using protocols established for other slow-growing mycobacteria. Transformations with “empty vector” plasmid DNAs demonstrated abundant recovery of colonies when we used 100 μg/ml kanamycin (for pMV261-kan and pMV361-kan) or 100 μg/ml hygromycin (for pMV361-hyg) as selection agents (Fig. 7A). The no-DNA control transformations yielded no colonies at ~3 weeks of incubation, although with extended incubation small colonies appeared on the kanamycin-supplemented plates, indicating that higher antibiotic concentrations may be preferable. We noted the *attB* site (integration site for the pMV361 vectors) of *M. haemophilum* was identical to that of *M. smegmatis*. Integration of pMV361-kan and pMV361-hyg was confirmed by PCR across the *attB* locus and sequencing of products (data not shown). We also transformed *M. haemophilum* with an episomal plasmid expressing the tdTomato fluorescent protein under control of the *M. marinum* G13 promoter, a strong constitutive mycobacterial promoter that yields high expression levels (84). We found tdTomato to be very highly expressed, as indicated by the intense pink color of the colonies, which deepened further with extended incubation (Fig. 7A). Fluorescence of transformants was confirmed by microscopy (Fig. 7B), revealing a characteristic mycobacterial cording. These data provide the first demonstration that *M. haemophilum* can be genetically manipulated with molecular tools established for other mycobacterial species and establish a foundation upon which the sequence data obtained here can be exploited to study this unique human pathogen.

**Conclusions.** Fundamental to understanding *M. haemophilum* as a pathogen are defining the basis for its unique growth requirements and its relationships to other mycobacteria. Our sequence and experimental data provide insight into the requirement of *M. haemophilum* for specific iron supplements, which may contribute to its modest virulence in immunocompetent individuals. The sequence data provided here may also facilitate diagnostics. Current molecular diagnostic approaches for detection of *M. haemophilum* in clinical samples utilize PCR-based methods (85, 86) and rely on very limited sequence information. Because there may be genetic diversity among *M. haemophilum* isolates from different geographic areas (87–89), diagnostics might be improved by designing additional PCR-based, hybridization-based, and antigen-based tests for *M. haemophilum* by using information derived from the complete genome sequence. This information may also enable improved epidemiological and environmental testing, expanding the molecular tools

---

**FIG 6** Phylogenetic tree of selected mycobacterial species. A phylogenetic tree was constructed for 15 mycobacterial genomes, including *M. haemophilum*, and one outgroup genome as described in Materials and Methods. The scale (0.025) indicates substitutions per aligned position.

**FIG 7** Successful transformation of *M. haemophilum*. (A) *M. haemophilum* was transformed with an integrating vector conferring hygromycin resistance (pMV361-hyg), an episomal vector conferring kanamycin resistance (pMV261-kan), an integrating vector conferring kanamycin resistance (pMV361-kan), or an episomal vector conferring kanamycin resistance and expressing the red fluorescent protein tdTomato from the *M. marinum* G13 promoter (pYUB1169). Colonies were photographed after 24 days of incubation at 30°C, while colonies harboring the tdTomato-expressing construct were photographed again at 33 days postincubation, to demonstrate intensification of the pink color. No-DNA controls demonstrating sensitivity of *M. haemophilum* to kanamycin and hygromycin at the indicated concentrations are also shown. In addition to antibiotic supplements, all plates also contained 100 μM hemin. (B) *M. haemophilum* transformed with the tdTomato fluoresces red. Images were obtained with a Nikon Eclipse TiE microscope using a CFI Plan Apo VC 100×/1.40 numerical aperture objective and a BrightLine Ex56240 filter set. The fluorescent image is shown on the left and a differential interference contrast image is on the right.

---

Tufariello et al.
available to evaluate diversity. Our findings also demonstrate a close genetic relationship between M. haemophilum and M. leprae, the latter an important human pathogen that cannot be propagated in vitro. The successful transformation of M. haemophilum coupled with the genomic sequence will facilitate studies to elucidate the features that contribute to the pathogenic properties of M. haemophilum and allow its use as a surrogate to functionally characterize gene products uniquely shared by M. haemophilum and M. leprae.

MATERIALS AND METHODS

Bacterial strains. The Mycobacterium haemophilum type strain (DSM 44634; ATCC 29548) was cultured at 30°C in Middlebrook 7H9 broth supplemented with 0.2% glycerol, 10% OADC enrichment (Becton, Dickinson), 0.05% tyloxapol or Tween 80, and 100 µM hemin (bovine [Sigma]; 7H9-hemin) or as described above. Comparative growth studies utilized mc6230, a ΔmbtB ΔpanCD derivative of M. tuberculosis H37Rv which is approved for use in BSL2 facilities at the Albert Einstein College of Medicine (31), and an mbtB deletion mutant in the mc6230 background (see Text S1 in the supplemental material). Media for growth of mc6230 and mc6230 ΔmbtB were supplemented with 24 µg/ml n-panthothenate (Sigma).

Genome sequencing, assembling, and annotation. Genomic DNA isolated from cultures of M. haemophilum by a cetyltrimethylammonium bromide-chloroform-based method (90) was sequenced by the Epigenomics Shared Facility (ESF) at the Albert Einstein College of Medicine, using MiSeq (Illumina), and at the Department of Genetics and Genomic Sciences, Icahn School of Medicine at Mount Sinai, using PacBio methodologies. Details of genome sequencing, assembling, and annotation are provided in Text S1 in the supplemental material.

Phylogenetic analysis. A phylogenetic tree was constructed for 15 mycobacterial genomes, including that of M. haemophilum and one outgroup genome. Further details concerning our materials and methods are provided in Text S1 in the supplemental material.

Protocol for plating on various media types. M. haemophilum bacteria cultured in 7H9-hemin were pelleted by centrifugation and washed in phosphate-buffered saline (PBS) containing 0.05% tyloxapol or 0.05% Tween 80 (PBS-T), following which equal inocula were plated onto the various media being tested. Details are provided in the Text S1 in the supplemental material.

Transformation of M. haemophilum. M. haemophilum transformation was performed according to methods developed for M. tuberculosis (90), except M. haemophilum was grown at 30°C. See Text S1 in the supplemental material.

Examination of mycobactin and carboxymycobactin synthesis and secretion based on use of a radiolabeled precursor. M. haemophilum and strain mc6230 were cultured in 7H9 medium without hemin in order to deplete iron stores, following which bacteria were transferred to iron-limited 7H9 medium (treated overnight with the iron chelator Chelex-100; details are provided in Text S1 in the supplemental material). Bacteria were inoculated at a starting optical density at 600 nm of 0.3 in 5 ml of medium in the presence of 1 µCi/ml [7-14C]salicylic acid (PerkinElmer) and grown with shaking for 20 days at 37°C for mc6230 and at 30°C for M. haemophilum. For M. haemophilum, cultures were prepared in duplicate, one lacking iron sources and one containing 10 µM hemin. Siderophores were extracted according to the method of Wells et al. (50) (see Text S1 for further details regarding our materials and methods). Extracts from cell pellets and culture supernatants were resuspended in chloroform and spotted on TLC silica gel 60F254 plates, with samples normalized based on equivalent radioactivity levels.

Nucleotide sequence accession numbers. Sequence data were deposited in the NCBI repository under BioProject PRJNA171821 (Mycobacterium haemophilum DSM 44634; ATCC 29548). The complete genome sequence is available under accession number CP011883. Short illumina reads, illumina mate-pair reads, Pacific Biosciences SMRT reads, and Sanger sequencing information were deposited in the NCBI Sequence Read Archive (SRA) under accession numbers SRX1059681, SRX1059688, SRX1059690, and SRX1059710, respectively. Genome annotations were deposited on the RAST website (http://ra.st/nmpdr.org/) (91, 92), first under Genome ID 1202450.3 and later updated to begin the sequence at dnaA, under Genome ID 1202450.9. The M. haemophilum gene names referenced in the text are those assigned based on the RAST annotation, Genome ID 1202450.9.

SUPPLEMENTAL MATERIAL

Supplemental material for this article may be found at http://mbio.asm.orglookup/suppl/doi:10.1128/mBio.01313-15/-/DCSupplemental.

Text S1, DOCX file, 0.1 MB.
Table S1, DOCX file, 0.03 MB.
Table S2, DOCX file, 0.02 MB.
Table S3, DOCX file, 0.04 MB.

ACKNOWLEDGMENTS

This work was supported by National Institutes of Health grants AI26170 and AI098925 to W.R.I.

We thank David Reynolds, Director of the Genomics Core Facility at the Albert Einstein College of Medicine, for helpful discussions on genome finishing and Colin Ratledge for generously providing the M. smegmatis mycobactin.

REFERENCES

1. Lindeboom JA, Brujinestein van Coppenraet LES, van Soolingen D, Prins JM, Kuiper EJ. 2011. Clinical manifestations, diagnosis, and treatment of Mycobacterium haemophilum infections. Clin Microbiol Rev 24:701–717. http://dx.doi.org/10.1111/j.1365-2391.2006.05332.x
2. Sompolinsky D, Lagziel A, Naveh D, Yanklevitz T. 1978. Mycobacterium haemophilum sp. nov., a new pathogen of humans. Int J Syst Bacteriol 28:67–75. http://dx.doi.org/10.1099/00207713-28-1-67
3. Sompolinsky D, Lagziel A, Rosenberg I. 1979. Further studies of a new pathogenic mycobacterium (M. haemophilum sp. nov.). Can J Microbiol 25:217–226. http://dx.doi.org/10.1139/m79-033
4. Davis BR, Brumbach J, Sanders WJ, Wolinsky E. 1982. Skin lesions caused by Mycobacterium haemophilum. Ann Intern Med 97:723–724. http://dx.doi.org/10.7326/0003-4819-97-5-723
5. Saabolle MA, Kiehn TE, White MH, Rudinska MF, Armstrong D. 1996. Mycobacterium haemophilum microbiology and expanding clinical and geographic spectra of disease in humans. Clin Microbiol Rev 9:435–447.
6. Fairhurst RM, Kubak BM, Pegues DA, Morishige JD, Han HF, Haley JC, Kobashigawa JA. 2002. Mycobacterium haemophilum infections in heart transplant recipients: case report and review of the literature. Am J Transplant 2:476–479. http://dx.doi.org/10.1046/j.1600-6143.2002.00174.x
7. Hirsch R, Miller S, Kazi S, Cate T, Reveille J. 1996. Human immunodeficiency virus-associated atypical mycobacterial skeletal infections. Semin Arthritis Rheum 25:347–356. http://dx.doi.org/10.1016/S0049-0172(96)80020-5
8. White MH, Papadopoulos EB, Small TN, Kiehn TE, Armstrong D. 1995. Mycobacterium haemophilum infections in bone marrow transplant recipients. J Transplant 60:957–960.
9. Elsayed S, Read R. 2006. Mycobacterium haemophilum osteomyelitis: case report and review of the literature. BMC Infect Dis 6:70. http://dx.doi.org/10.1186/1471-2334-6-70
10. Sogani J, Ivandzic J, Phillips CD. 2014. Chiasmatisis caused by Mycobacterium haemophilum in an immunocompromised adult. Clin Imaging 38:277–279. http://dx.doi.org/10.1016/j.clinimag.2014.02.012
11. Ducharlet K, Murphy C, Tan S, Dwyer KM, Goodman D, Abolins C, Daffy JR, Langham RG. 2014. Recurrent Mycobacterium haemophilum in a renal transplant recipient. Nephrology 19(Suppl 1):14–17. http://dx.doi.org/10.1111/nep.12193
12. Shah M, Sebti A, Kiehn T, Massarella S, Sepkowitz K. 2001. Mycobacterium haemophilum masquerading as leprosy in a renal transplant patient. Case Rep Nephrol 38:387–390. http://dx.doi.org/10.1111/nep.12193
13. Elsayed S, Read R. 2006. Mycobacterium haemophilum osteomyelitis: case report and review of the literature. BMC Infect Dis 6:70. http://dx.doi.org/10.1186/1471-2334-6-70
14. Sogani J, Ivandzic J, Phillips CD. 2014. Chiasmatisis caused by Mycobacterium haemophilum in an immunocompromised adult. Clin Imaging 38:277–279. http://dx.doi.org/10.1016/j.clinimag.2014.02.012
15. Copeland NK, Arora NS, Ferguson TM. 2013. Mycobacterium haemophilum masquerading as leprosy in a renal transplant patient. Case Rep Dermatol Med 2013:793127. http://dx.doi.org/10.1155/2013/793127
16. Ishii K, Ishii N, Nakanao K, Nakano K, Saito I, Ashina A. 2015. Mycobacterium haemophilum infection with prominent facial manifestations
compromised mice against experimental tuberculosis. Vaccine 24: 6309–6320. http://dx.doi.org/10.1016/j.vaccine.2006.05.097.

32. De Voss JJ, Rutter K, Schroeder BG, Su H, Zhu Y, Barry CE, III. 2000. The salicylate-derived mycobacterin siderophores of Mycobacterium tuberculosis are essential for growth in macrophages. Proc Natl Acad Sci U S A 97:1257–1262. http://dx.doi.org/10.1073/pnas.97.3.1252.

33. Mezo A, Dennis J, McCarthy SW, Dawson DJ. 1979. Unusual mycobac-
teria in 5 cases of opportunistic infections. Pathology 11:377–384. http://dx.doi.org/10.3109/0301379090391401.

34. Kritikha R, Marathe U, Saxena P, Ansari MZ, Mohanty D, Gokhale RS. 2006. A genetic locus required for iron acquisition in Mycobacterium tuberculosis. Proc Natl Acad Sci U S A 103:2069–2074. http://dx.doi.org/10.1073/pnas.0507924103.

35. Quadri LEN, Sello J, Keating TA, Weinreb PH, Walsh CT. 1998. Identification of a Mycobacterium tuberculosis gene cluster encoding the biosynthetic enzymes for assembly of the virulence-conferring sidero-
phore mycobactin. Chem Biol 5:631–645. http://dx.doi.org/10.1016/S1074-5521(98)90291-5.

36. Chavadi SS, Stirratt RL, Edupuguri UR, Vergnolle O, Sadhanandan G, Marchiano E, Martin C, Qiu W, Soll C. Quadri LEN. 2011. Muta-
tional and phylogenetic analyses of the mycobacterial mbt gene cluster. J Bacteriol 193:5905–5913. http://dx.doi.org/10.1128/JB.05811-11.

37. Felnagle EA, Barkei JJ, Park H, Podevels AM, McMahon MD, Drott DW, Thomas MG. 2010. MtbH-like proteins as integral components of bacterial nonribosomal peptide synthetases. Bio Chem 49:8615–8617. http://dx.doi.org/10.1002/bjc.21258.

38. Eason E, Chavadi S, Mohandas P, Edupuguri UR, Angala SK, Chat-
terjee D, Quadri LEN. 2012. Production of mycobacterial cell wall glyco-
peptidolipids requires a member of the MtbH-like protein family. BMC Microbiol 12:118. http://dx.doi.org/10.1186/1471-2180-12-118.

39. Harrison AJ, Yu M, Gardenborg T, Middditch MJ, Ramsay BJ, Baker EN, Lott JS. 2006. The structure of MbtH from Mycobacterium tuberculosis, the first enzyme in the biosynthesis of the siderophore mycobactin, reveals it to be a salicylate syr.

40. Cole ST, Eiglmeyer K, Parkhill J, James KD, Thomson NR, Wheeler PR, Honoré N, Garnier T, Churcher C, Harris D, Mungall K, Basham D, Brown D, Chillingworth T, Connor R, Davies RM, Devlin K, Duthoy S, Felltw T, Fraser A, Hamlin N, Holroyd S, Hornsby T, Jagels K, Lacroix C, Maclean J, Mogile S, Murphy L, Olive K, Quail MA, Rajandrae M, Rutherford KM, Rutter S, Seeger K, Simon S, Sim-
monds M, Skelton J, Squares R, Squares S, Stevens K, Taylor K, Whitehead S, Woodward JR, Barrett BG. 2001. Massive gene decay in the leprosy bacillus. Nature 409:1007–1011. http://dx.doi.org/10.1038/35059006.

41. Han XY, Sot Y, Sizer KC, Schoberele T, May GS, Spencer JS, Li W, Nair B. 2008. A new mycobacterium species causing diffuse leprosy laspor 

42. Land J, Eglemeier K, Parkhill J, James KD, Thomson NR, Wheeler PR, Honoré N, Garnier T, Churcher C, Harris D, Mungall K, Basham D, Brown D, Chillingworth T, Connor R, Davies RM, Devlin K, Duthoy S, Felltw T, Fraser A, Hamlin N, Holroyd S, Hornsby T, Jagels K, Lacroix C, Maclean J, Mogile S, Murphy L, Olive K, Quail MA, Rajandrae M, Rutherford KM, Rutter S, Seeger K, Simon S, Sim-
monds M, Skelton J, Squares R, Squares S, Stevens K, Taylor K, Whitehead S, Woodward JR, Barrett BG. 2001. Massive gene decay in the leprosy bacillus. Nature 409:1007–1011. http://dx.doi.org/10.1038/35059006.

43. Singh P, Benjak A, Schuenemann VJ, Herbig A, Avanzi C, Busso P, Nieselt K, Kraus T, Vera-Cabrera L, Cole ST. 2015. Insight into the evolution and origin of leprosy bacilli from the genome sequence of myco-
bacterium leprosum. Proc Natl Acad Sci U S A 112:4459–4464. http://dx.doi.org/10.1073/pnas.1315041112.

44. Gold B, Rodríguez GM, Marras SAE, Pentecost M, Smith I. 2001. The Mycobacterium tuberculosis IdeR is a duality functional regulator that controls transcription of genes involved in iron acquisition, iron storage and survival in macrophages. Mol Microbiol 42:851–865. http://dx.doi.org/10.1046/j.1365-2958.2001.02684.x.

45. Fiss EH, Yu S, Jacobs WR, Jr. 1994. Identification of genes involved in the sequestration of iron in mycobacteria: the ferric exochelin biosynthetic and uptake pathways. Mol Microbiol 14:557–569. http://dx.doi.org/10.1111/j.1365-2958.1994.tb01218.x.

46. Yu S, Fiss E, Jacobs WR, Jr. 1998. Analysis of the exochelin locus of Mycobacterium smegmatis: biosynthesis genes have homology with genes of the peptide synthetase family. J Bacteriol 176:4676–4685.

47. Zhu W, Arceneaux JEL, Beggs ML, Byers BR, Eisenach KD, Lundrigan RD. 1998. Exochelin genes in Mycobacterium smegmatis: identification of an ABC transporter and two non-ribosomal peptide synthetase genes. Mol Microbiol 29:629–639. http://dx.doi.org/10.1111/j.1365-2958.1998.00691.x.

48. Farhana A, Kumar S, Rathore SS, Ghosh PC, Ehtesham NZ, Tyagi AK, Haasnain SE. 2008. Mechanistic insights into a novel exportor-

importer systerm of Mycobacterium tuberculosis unravel its role in
patients by PCR-restriction endonuclease analysis. J Clin Microbiol 42: 3336–3338. http://dx.doi.org/10.1128/JCM.42.7.3336-3338.2004.

87. Bruijnesteijn van Coppenraet LES, Savelkoul PHM, Buffing N, van der Bijl MW, Woudenberg J, Lindeboom JA, Kiehn TE, Haverkort F, Samra Z, Kuijper EJ. 2009. Amplified fragment length polymorphism analysis of human clinical isolates of Mycobacterium haemophilum from different continents. Clin Microbiol Infect 15:924–930. http://dx.doi.org/10.1111/j.1469-0691.2009.02798.x.

88. Kikuchi K, Bernard EM, Kiehn TE, Armstrong D, Riley LW. 1994. Restriction fragment length polymorphism analysis of clinical isolates of Mycobacterium haemophilum. J Clin Microbiol 32:1763–1767.

89. Yakrus MA, Straus WL. 1994. DNA polymorphisms detected in Mycobacterium haemophilum by pulsed-field gel electrophoresis. J Clin Microbiol 32:1083–1084.

90. Larsen MH, Biermann K, Tandberg S, Hsu T, Jacobs WR. 2007. Genetic manipulation of Mycobacterium tuberculosis. Curr Protoc Microbiol 6:A:10A.2:10A.2.1–10A.2.21.

91. Aziz RK, Bartels D, Best AA, DeJongh M, Disz T, Edwards RA, Formsma K, Gerdes S, Glass EM, Kubal M, Meyer F, Olsen GJ, Olson R, Osterman AL, Overbeek RA, McNeil JK, Paarmann D, Paczian T, Parrello B, Pusch GD, Reich C, Stevens R, Vassieva O, Vonstein V, Willke A, Zagnitko O. 2008. The RAST server: rapid annotations using subsystems technology. BMC Genomics 9:75. http://dx.doi.org/10.1186/1471-2164-9-75.

92. Overbeek R, Olson R, Pusch GD, Olsen GJ, Davis JJ, Disz T, Edwards RA, Gerdes S, Parrello B, Shukla M, Vonstein V, Wattam AR, Xia F, Stevens R. 2014. The SEED and the rapid annotation of microbial genomes using subsystems technology (RAST). Nucleic Acids Res 42: D206–D214. http://dx.doi.org/10.1093/nar/gkt1226.