



# Highly Reduced Genome of the New Species *Mycobacterium uberis*, the Causative Agent of Nodular Thelitis and Tuberculoid Scrotitis in Livestock and a Close Relative of the Leprosy Bacilli

Andrej Benjak,<sup>a</sup> Charlotte Avanzi,<sup>a</sup> Yvonne Benito,<sup>b,c</sup> Franck Breyse,<sup>d</sup> Christophe Chartier,<sup>e</sup> Maria-Laura Boschioli,<sup>f</sup> Christine Fourichon,<sup>e</sup> Lorraine Michelet,<sup>f</sup> Didier Pin,<sup>g</sup> Jean-Pierre Flandrois,<sup>h</sup> Pierre Bruyere,<sup>g</sup> Oana Dumitrescu,<sup>b,c</sup> Stewart T. Cole,<sup>a,i</sup> Gerard Lina<sup>b,c</sup>

<sup>a</sup>Global Health Institute, École Polytechnique Fédérale de Lausanne, Lausanne, Switzerland

<sup>b</sup>Institut des Agent infectieux, Hôpital de la Croix Rousse, Hospices Civils de Lyon, Lyon, France

<sup>c</sup>CIRI, Centre International de Recherche en Infectiologie, Inserm U1111, Université Lyon 1, Ecole Normale Supérieure de Lyon, CNRS UMR 5308, Lyon, France

<sup>d</sup>Laboratoire d'Immunologie, Centre de Biologie Sud, Lyon, France

<sup>e</sup>BIOEPAR, INRA, Oniris, La Chantrerie, Nantes, France

<sup>f</sup>ANSES Laboratory Affairs Department, Maisons-Alfort, France

<sup>g</sup>Université de Lyon, VetAgro Sup, UPSP 2016.A104, Interactions Cellules Environnement, Marcy l'Etoile, France

<sup>h</sup>Université Lyon1, LBBE-CNRS UMR5558, Villeurbanne, France

<sup>i</sup>Institut Pasteur, Paris, France

**ABSTRACT** Nodular thelitis is a chronic enzootic infection affecting dairy cows and goats. The causative agent was recently shown to be related to the leprosy-causing bacilli *Mycobacterium leprae* and *Mycobacterium lepromatosis*. In this study, the genome of this pathogen was sequenced and analyzed. Phylogenomic analyses confirmed that the pathogen present in nodular thelitis and tuberculoid scrotitis is a distinct species related to the leprosy bacilli and *Mycobacterium haemophilum*. Because the pathogen was originally isolated from a bovine udder, it was named "*Mycobacterium uberis*." The genome of "*M. uberis*" is only 3.12 Mb in length, which represents the smallest mycobacterial genome identified so far but which is close to that of leprosy bacilli in size. The genome contains 1,759 protein-coding genes and 1,081 pseudogenes, indicative of extensive reductive evolution and likely the reason that *M. uberis* cannot be grown axenically. The pseudogenization and genome reduction in *M. uberis* seem to have been to some extent independent from the results determined for the genomes of the leprosy bacilli.

**IMPORTANCE** *M. uberis* is an emerging skin pathogen in dairy animals. Its genome underwent massive reduction and gene decay, leading to a minimal set of genes required for an obligatory intracellular lifestyle, which highly resembles the evolution of the leprosy agents *M. leprae* and *M. lepromatosis*. The genomic similarity between *M. uberis* and the leprosy bacilli can help in identifying key virulence factors of these closely related species or in identifying genes responsible for the distinct differences between thelitis or scrotitis and leprosy with respect to clinical manifestations. Specific DNA markers can now be developed for quick detection of this pathogen.

**KEYWORDS** *Mycobacterium uberis*, evolutionary biology, genome analysis, granulomatous dermatitis, veterinary pathogens

Received 25 July 2018 Accepted 12 September 2018 Published 3 October 2018

**Citation** Benjak A, Avanzi C, Benito Y, Breyse F, Chartier C, Boschioli M-L, Fourichon C, Michelet L, Pin D, Flandrois J-P, Bruyere P, Dumitrescu O, Cole ST, Lina G. 2018. Highly reduced genome of the new species *Mycobacterium uberis*, the causative agent of nodular thelitis and tuberculoid scrotitis in livestock and a close relative of the leprosy bacilli. *mSphere* 3:e00405-18. <https://doi.org/10.1128/mSphere.00405-18>.

**Editor** Christina L. Stallings, Washington University in St. Louis School of Medicine

**Copyright** © 2018 Benjak et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

Address correspondence to Stewart T. Cole, [stewart.cole@epfl.ch](mailto:stewart.cole@epfl.ch), or Gerard Lina, [gerard.lina@univ-lyon1.fr](mailto:gerard.lina@univ-lyon1.fr).

Nodular thelitis is a chronic, enzootic granulomatous dermatitis associated with acid-fast bacilli. It was originally observed in cows and was first described in France in 1963 (1) and then in Japan (2) and Switzerland (3). A similar disease, nodular tuberculoid scrotoitis, was observed in bulls and is suspected to be caused by the same pathogen. The causative agent of the bovine nodular thelitis was recently shown to be related to the leprosy-causing species *Mycobacterium leprae* and *Mycobacterium lepromatosis* (4). More recently, the same pathogen was also identified in dairy goats (5). In this study, the draft genome of this pathogen was reconstructed and analyzed to reveal a distinct mycobacterial species and for use for confirmation of its detection in nodular thelitis and tuberculoid scrotoitis.

## RESULTS AND DISCUSSION

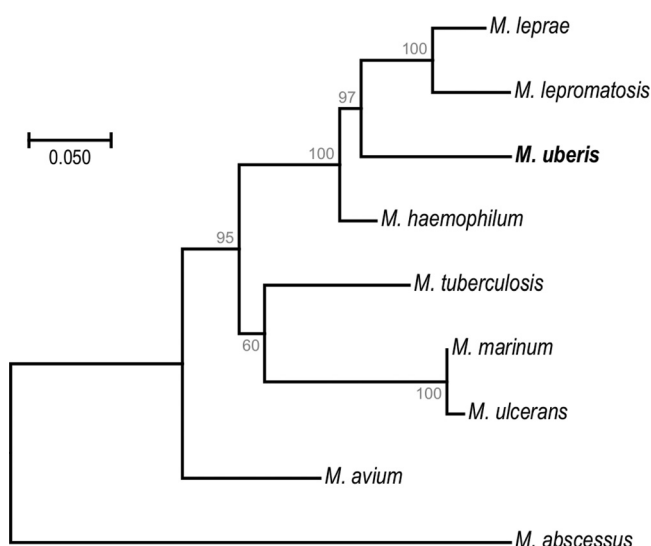
**Species name and phylogeny.** As shown in Fig. 1, the new species forms a distinct branch lying between *M. haemophilum* and the most recent common ancestor of *M. leprae* and *Mycobacterium lepromatosis*. In view of its initial identification from udder, this species is named *Mycobacterium uberis*.

***Mycobacterium uberis* genome sequence.** DNA was isolated from a skin biopsy of bovine udder with nodular thelitis and subjected to Illumina sequencing. The *de novo* sequence assembly resulted in 3,571 contigs that were larger than 1 kb and showed average coverage of over 10×. Most contigs matched sequences from a variety of bacterial species. On the basis of sequence similarity to *M. haemophilum* and *M. leprae*, we retrieved contigs that unmistakably belonged to *M. uberis*. Care was taken not to exclude any other potential *M. uberis* sequences by manually checking all the remaining contigs that displayed GC content similar to that of *M. uberis* and by repeating the analysis with another assembly program (see Text S1 in the supplemental material for details). The draft *M. uberis* genome assembly consists of 54 contigs with an average length of 58 kb and totaling 3.12 Mb. All of the contigs harbored genes that closely match those of *M. haemophilum* (85.7% average nucleotide identity) or *M. leprae* (82.4% average nucleotide identity), with no obvious outlier that would indicate an erroneous assembly with sequences from another bacterial species.

**Genome downsizing and pseudogene formation.** At a sequence length of 3.12 Mb and containing 1,081 pseudogenes, the genome of *M. uberis* is as reduced as that of *M. leprae* (Table 1). The two species share 1,318 functional protein-coding genes, which corresponds to 75% of the total number of protein-coding genes in *M. uberis* and 82% in *M. leprae*. Similarly, among the 1,309 pseudogenes in *M. leprae*, only 212 orthologs were predicted to be functional in *M. uberis*; conversely, among the 1,081 pseudogenes in *M. uberis*, only 126 orthologs are predicted to be functional in *M. leprae* (see Data Set S1 in the supplemental material).

The pseudogene content of *M. uberis* differs from that of *M. leprae*. For example, 465 pseudogenes in *M. uberis* do not have an ortholog in *M. leprae*, and 607 pseudogenes in *M. leprae* do not have an ortholog in *M. uberis*, suggesting that genome reduction was, at least in part, an independent process in each species. However, this observation does not exclude the possibility that the initial pseudogenization occurred in the ancestor of *M. uberis* and *M. leprae*. If this were the case, the pseudogenes that are present in both species would be more likely to share the same deleterious mutations, such as frameshifts and premature stop codons. We manually checked 50 random orthologous pseudogenes and identified only three pairs that shared one or more stop codons and/or frameshifts. Although some signals were lost or blurred by sequence drift, this observation suggests that some of the pseudogenization had already started in the ancestor of *M. uberis* and *M. leprae* and that the pseudogenization processes probably continued independently as the two species diverged. Note that roughly 500 genes which are missing in both species (compared to *M. haemophilum*) were likely deleted in the ancestor of *M. uberis* and *M. leprae*, which was probably adapting to a strict intracellular niche (6).

**Envelope biogenesis and other specific features.** *M. leprae* contains no methoxymycolates, probably because it has lost the MmaA2 and MmaA3 methoxy mycolic acid



**FIG 1** Phylogenetic tree of *Mycobacterium uberis* and selected mycobacterial species. The tree was created in MEGA7 from concatenated amino acid sequences (3,696 positions) of 10 proteins (DnaN, RplI, GrpE, MetG, RplY, PheT, FtsQ, HoIA, MiaA, and FtsY) (18) and inferred by using the maximum likelihood method based on the JTT matrix-based model. The tree is drawn to scale, with branch lengths measured at the number of substitutions per site. Bootstrap support values, estimated from 500 replicates, are given below each branch. *Mycobacteroides abscessus* (previously *Mycobacterium abscessus*) was used as the outgroup.

synthases (7). *M. uberis* has retained a functional MmaA3 (*M. uberis* BE\_04680 [MUBE\_04680]), which might influence the envelope composition. On the other hand, *M. uberis* has the same reduced set of five *mmpL* genes as *M. leprae*. It remains to be determined whether *M. uberis* produces a glycolipid similar or equivalent to the characteristic and highly antigenic phenolic glycolipid 1 of *M. leprae*.

A characteristic feature of *M. leprae* and *M. lepromatosis* is the presence of the horizontally acquired gene *proS*, encoding a eukaryote-like prolyl tRNA synthetase, which is both displaced and inverted with respect to the *M. tuberculosis* genome (7, 8) and is similar to those present in various members of the *Nocardiaceae* family. The same *proS* homolog is also present in *M. uberis* (MUBE\_09850) and *M. haemophilum* (B586\_RS07325) at the same genomic location, indicating that the gene was acquired by their ancestor. In addition, *M. uberis* has a cytochrome P450 (MUBE\_02130) of unknown function that is similar to those present in more distantly related mycobacterial species but that is not present in *M. leprae*, *M. lepromatosis*, or *M. haemophilum*.

**Growth.** All attempts to grow *M. uberis* have failed, a result which was expected given its highly reduced genome. As in the cases of *M. haemophilum* and *M. leprae*, *M. uberis* lacks the mycobactin synthesis gene cluster present in *Mycobacterium tubercu-*

**TABLE 1** Genomic features of *M. uberis* and close relatives *M. leprae*, *M. lepromatosis*, and *M. haemophilum*

Feature	Value			
	<i>M. uberis</i>	<i>M. leprae</i>	<i>M. lepromatosis</i>	<i>M. haemophilum</i>
Genome size (bp)	3,122,721	3,268,212	3,206,741	4,235,765
No. of protein-coding genes	1,759	1,609	1,477	3,749
No. of pseudogenes	1,081	1,309	1,334	225
% GC content (genome)	57.49	57.80	57.89	63.95
% GC content (CDS) <sup>a</sup>	59.58	60.11	60.16	64.35
% GC content (pseudogenes)	55.60	56.45	56.59	64.84
% GC content (intergenic)	53.08	54.17	54.61	60.37

<sup>a</sup>CDS, coding DNA sequence.

*losis*. Moreover, the gene coding for 50S ribosomal protein L25 (MUBE\_04325) is truncated in *M. uberis* and is probably not functional. Disruption of this gene results in growth defects in *M. tuberculosis* (9) and *Escherichia coli* (10), due to reduced efficiency of the ribosome.

**Virulence.** The ESX-1 system is the main determinant of virulence in *M. tuberculosis* and in a number of other mycobacterial pathogens (11). While *M. leprae* and *M. lepromatosis* lost some components of ESX-1 (12), *M. uberis* retained the protein-coding capacity of the entire system (see Fig. S1 in the supplemental material). However, we identified a break in the genomic synteny downstream of *espJ* (MUBE\_00800) and an insertion of a putative proline-proline-glutamate (PPE) gene (MUBE\_01185), flanked by remnants of transposases, between *espB* (MUBE\_01195) and *eccE*, (MUBE\_01180). It is not clear how these changes impact the ESX-1 system in *M. uberis*. Curiously, the structural variations occur around the same genes that lost coding capacity in *M. leprae* and *M. lepromatosis* (Fig. S1).

The ESX-5 system is the most recently evolved mycobacterial ESX system, which modulates virulence and host response, and is found only in the slow-growing mycobacterial species (13). The two *esx* genes and the flanking PE/PPE gene pair of the ESX-5 system underwent a series of duplication events that resulted in multiple copies scattered across the genome (14). It was shown that some of the paralog clusters in *M. tuberculosis* serve as accessory systems that aid in the secretion of a subset of proteins via the prototype ESX-5 system (14). Interestingly, *M. uberis* lost the core components of the prototype ESX-5 system, similarly to *M. leprae*, but has retained at least three paralog *esx* pairs.

**Drug susceptibility.** No mutations were found in the drug-resistance-conferring regions of RpoB (MUBE\_04585), FolP1 (MUBE\_01990), GyrA (MUBE\_01070), and GyrB (MUBE\_01075) (15), indicating that *M. uberis* is very likely susceptible to the antileprosy drugs rifampin, dapson, and ofloxacin. Since there are no known molecular markers for resistance, we can only presume that *M. uberis* is also susceptible to the drugs clofazimine and clarithromycin, as is *M. haemophilum* (16).

**Mycobacterium leprae cluster organisms.** A recently reported causative agent of feline leprosy, "*Candidatus* Mycobacterium lepraefelis," was found to be a close relative of *M. leprae* (17). The partial sequence of the *groEL2* (*hsp65*) gene from this pathogen is 89% to 90% identical to those of *M. uberis*, *M. haemophilum*, *M. leprae*, and *M. lepromatosis*. Phylogenetic reconstruction of the *groEL* sequence placed "*Candidatus* M. lepraefelis" between *M. uberis* and *M. leprae* (Fig. S2), so it is likely that the genomic structure of "*Candidatus* M. lepraefelis" resembles those of *M. uberis* and *M. leprae*. Efforts to close the genome sequences of the *M. leprae*-like pathogens are needed to facilitate more-detailed genomic comparisons, which, coupled with biological data, will provide further insights into the evolution and pathogenicity of this particular group of mycobacteria.

**Significance and molecular detection of *Mycobacterium uberis*.** The availability of the genome sequence of *M. uberis* allowed us to design specific PCR primers for *M. uberis*. We confirmed the presence of the bacteria in the three cases of bovine nodular thelitis and the two cases of caprine nodular thelitis reported before (4, 5), as well as in a new case of bovine nodular thelitis and two cases of caprine nodular thelitis from different farms in France (Table 2). Moreover, we detected *M. uberis* in three cases of nodular tuberculoid scrotoitis, confirming the implication of the bacterium in the two diseases. While definitive evidence is still lacking, these results strongly suggest that *M. uberis* is the causative agent of nodular thelitis and tuberculoid scrotoitis.

Early detection and diagnosis of infectious diseases are crucial in animal husbandry to prevent disease outbreaks and contamination of animal products. Molecular tools from this study can be used for routine screening of the pathogen and will facilitate epidemiological investigations.

**TABLE 2** Samples used for the PCR detection of *M. uberis*<sup>a</sup>

Sample ID	Animal	Diagnosis	Tissue	Herd	Animal ID	Reference or source
13Z000257	Goat	NT	Mammary gland	A	goat 1	5
13Z002358	Goat	NT	Inguinal lymph node	A	goat 1	5
14Z002623	Goat	NT	Unspecified tissue	A	goat 2	5
14Z002624	Goat	NT	Mammary gland	A	goat 2	5
15Z001519	Goat	NT	Mammary gland	B	goat 3	This study
15Z001520	Goat	NT	Mammary gland	B	goat 4	This study
NA	Cow	NT	Mammary gland	C	cow 1	4
NA	Cow	NT	Mammary gland	C	cow 2	4
NA	Cow	NT	Mammary gland	D	cow 4	This study
NA	Bull	TS	Scrotum	E	taurus 1	This study
NA	Bull	TS	Scrotum	E	taurus 2	This study
NA	Bull	TS	Scrotum	E	taurus 3	This study

<sup>a</sup>ID, identifier; NT, nodular thelitis; TS, tuberculoid scrotoitis; NA, not available.

## MATERIALS AND METHODS

DNA was isolated from a skin biopsy of bovine udder with nodular thelitis and subjected to Illumina sequencing, followed by sequence assembly and annotation. More details are given in Text S1 in the supplemental material.

For the PCR assay, we used BLAST to identify genomic regions in *M. uberis* with no sequence homology to any publicly available sequence. We chose a 231-bp-long intergenic region that lies within the specific genomic island in the ESX-1 locus of *M. uberis*, between the *espB* gene and *eccE1* (Fig. S1), using primers Muber6F (5'-CACCGAACCCCTTCATGTCA-3') and Muber6R (5'-CCCGGTAGTGTTGGCTTGAT-3').

**Accession number(s).** The annotated genome has been deposited at DDBJ/ENA/GenBank under accession number [QAYL00000000](https://doi.org/10.1128/mSphere.00405-18).

## SUPPLEMENTAL MATERIAL

Supplemental material for this article may be found at <https://doi.org/10.1128/mSphere.00405-18>.

**TEXT S1**, PDF file, 0.1 MB.

**FIG S1**, EPS file, 0.1 MB.

**FIG S2**, EPS file, 0.1 MB.

**DATA SET S1**, XLSX file, 1.8 MB.

## ACKNOWLEDGMENTS

We thank staff members at the Genomic Technologies Facility at the University of Lausanne, Switzerland, for Illumina sequencing and Véronique Guerin-Faublee, Pierre Guerin, and Monique Chomarat for scientific advice.

This work was supported by the Fondation Raoul Follereau and Swiss National Science Foundation grant IZRJZ3\_164174.

## REFERENCES

- Joubert L, Ferney J, Oudar J, Van H. 1963. Thélite nodulaire tuberculoïde de la vache laitière à mycobactéries atypiques scotochromogènes. *Rev Med Vet* 114:87–105.
- Shimizu K, Shirahata T, Ono T, Ueda A. 1968. Studies on skin lesion tuberculosis of the bovine udder in Japan. *Res Bull Obihiro Univ* 5:729–734.
- Rüsch P, Corboz L, Ossent P, Berchtold M, Ehrensperger F. 1984. Enzootische nodulär-ulzerierende Thelitis mit Mykobakterieninfektion bei Kühen. *Schweiz Arch Tierheilk* 126:467–478.
- Pin D, Guérin-Faubleé V, Garreau V, Breyse F, Dumitrescu O, Flandrois J-P, Lina G. 2014. Mycobacterium species related to *M. leprae* and *M. lepromatosis* from cows with bovine nodular thelitis. *Emerg Infect Dis* 20:2111–2114. <https://doi.org/10.3201/eid2012.140184>.
- Chartier C, Albaric O, Cesbron N, Despres J, Hoogveld C, Michelet L, Boschiroli M-L. 2016. Tuberculoid nodular thelitis in a dairy goat flock. *Vet J* 209:199–200. <https://doi.org/10.1016/j.tvjl.2015.12.004>.
- Weinert LA, Welch JJ. 2017. Why might bacterial pathogens have small genomes? *Trends Ecol Evol* 32:936–947. <https://doi.org/10.1016/j.tree.2017.09.006>.
- Cole ST, Eiglmeier 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, Feltwell T, Fraser A, Hamlin N, Holroyd S, Hornsby T, Jagels K, Lacroix C, Maclean J, Moule S, Murphy L, Oliver K, Quail MA, Rajandream MA, Rutherford KM, Rutter S, Seeger K, Simon S, Simmonds M, Skelton J, Squares R, Squares S, Stevens K, Taylor K, Whitehead S, Woodward JR, Barrell BG. 2001. Massive gene decay in the leprosy bacillus. *Nature* 409:1007–1011. <https://doi.org/10.1038/35059006>.
- Singh P, Benjak A, Schuenemann VJ, Herbig A, Avanzi C, Busso P, Nieselt K, Krause J, Vera-Cabrera L, Cole ST. 2015. Insight into the evolution and origin of leprosy bacilli from the genome sequence of *Mycobacterium lepromatosis*. *Proc Natl Acad Sci U S A* 112:4459–4464. <https://doi.org/10.1073/pnas.1421504112>.
- DeJesus MA, Gerrick ER, Xu W, Park SW, Long JE, Boutte CC, Rubin EJ, Schnappinger D, Ehrt S, Fortune SM, Sasseti CM, Ioerger TR. 2017. Comprehensive essentiality analysis of the *Mycobacterium tuberculosis* genome via saturating transposon mutagenesis. *mBio* 8:e02133-16. <https://doi.org/10.1128/mBio.02133-16>.
- Korepanov AP, Gongadze GM, Garber MB, Court DL, Bubunenka MG.

2007. Importance of the 5 S rRNA-binding ribosomal proteins for cell viability and translation in *Escherichia coli*. *J Mol Biol* 366:1199–1208. <https://doi.org/10.1016/j.jmb.2006.11.097>.
11. Gröschel MI, Sayes F, Simeone R, Majlessi L, Brosch R. 2016. ESX secretion systems: mycobacterial evolution to counter host immunity. *Nat Rev Microbiol* 14:677–691. <https://doi.org/10.1038/nrmicro.2016.131>.
  12. Gey Van Pittius NC, Gamielidien J, Hide W, Brown GD, Siezen RJ, Beyers AD. 2001. The ESAT-6 gene cluster of *Mycobacterium tuberculosis* and other high G+C Gram-positive bacteria. *Genome Biol* 2:RESEARCH0044.
  13. Newton-Foot M, Warren RM, Sampson SL, van Helden PD, Gey van Pittius NC. 2016. The plasmid-mediated evolution of the mycobacterial ESX (Type VII) secretion systems. *BMC Evol Biol* 16:62. <https://doi.org/10.1186/s12862-016-0631-2>.
  14. Shah S, Briken V. 2016. Modular organization of the ESX-5 secretion system in *Mycobacterium tuberculosis*. *Front Cell Infect Microbiol* 6:49. <https://doi.org/10.3389/fcimb.2016.00049>.
  15. Benjak A, Avanzi C, Singh P, Loiseau C, Girma S, Busso P, Fontes ANB, Miyamoto Y, Namisato M, Bobosha K, Salgado CG, da Silva MB, Bouth RC, Frade MAC, Filho FB, Barreto JG, Nery JAC, Bühner-Sékula S, Lupien A, Al-Samie AR, Al-Qubati Y, Alkubati AS, Bretzel G, Vera-Cabrera L, Sakho F, Johnson CR, Kodio M, Fomba A, Sow SO, Gado M, Konaté O, Stefani MMA, Penna GO, Suffys PN, Sarno EN, Moraes MO, Rosa PS, Baptista IMFD, Spencer JS, Aseffa A, Matsuoka M, Kai M, Cole ST. 2018. Phylogenomics and antimicrobial resistance of the leprosy bacillus *Mycobacterium leprae*. *Nat Commun* 9:352. <https://doi.org/10.1038/s41467-017-02576-z>.
  16. Lindeboom JA, Bruijnesteijn van Coppenraet LES, van Soolingen D, Prins JM, Kuijper EJ. 2011. Clinical manifestations, diagnosis, and treatment of *Mycobacterium haemophilum* infections. *Clin Microbiol Rev* 24:701–717. <https://doi.org/10.1128/CMR.00020-11>.
  17. O'Brien CR, Malik R, Globan M, Reppas G, McCowan C, Fyfe JA. 2017. Feline leprosy due to *Candidatus "Mycobacterium lepraefelis"*: further clinical and molecular characterisation of eight previously reported cases and an additional 30 cases. *J Feline Med Surg* 19:919–932. <https://doi.org/10.1177/1098612X17706470>.
  18. Mizuno T, Natori T, Kanazawa I, Eldesouky I, Fukunaga H, Ezaki T. 2016. Core housekeeping proteins useful for identification and classification of mycobacteria. *Microb Resour Syst* 32:25–37.
  19. Simeone R, Bottai D, Frigui W, Majlessi L, Brosch R. 2015. ESX/type VII secretion systems of mycobacteria: Insights into evolution, pathogenicity and protection. *Tuberculosis* 95:S150–S154. <https://doi.org/10.1016/j.tube.2015.02.019>.