FEMS Microbiology Reviews, 2025, 49, fuaf012
DOI: 10.1093/femsre/fuaf012
Advance access publication date: 25 April 2025
Review article

Genetic variability, genotyping, and genomics of *Mycobacterium leprae*

Afzal Ansari[1,2], Roopendra Kumar[1], Suman Kumar Ray[1], Aarti Patel[1], Purna Dwivedi[1,3], Arup Ghosh[1], Edson Machado[4], Philip N. Suffys[4,‡], Pushpendra Singh[1,2,5,*,‡]

[1]ICMR-National Institute of Research in Tribal Health, Jabalpur 482003, Madhya Pradesh, India
[2]Academy of Scientific and Innovative Research (AcSIR), Ghaziabad 201002, India
[3]The Maharaja Sayajirao University of Baroda, Vadodara 390002, Gujarat, India
[4]Laboratory of Molecular Biology Applied to Mycobacteria - Oswaldo Cruz Institute, Fiocruz, Rio de Janeiro, Brazil
[5]Model Rural Health Research Unit, Badoni, Datia 475686, Madhya Pradesh, India
*Corresponding author: ICMR-National Institute of Research in Tribal Health, Jabalpur 482003, Madhya Pradesh, India. E-mail: Pushpendra.S@icmr.gov.in
‡Senior author
Editor: [Tomasz Jagielski]

Abstract

Leprosy, caused by *Mycobacterium leprae* and *Mycobacterium lepromatosis*, remains a significant global health issue despite a tremendous decline in its worldwide prevalence in the last four decades. *Mycobacterium leprae* strains possess very limited genetic variability, making it difficult to distinguish them using traditional genotyping tools. Successful genome sequencing of a considerable number of *M. leprae* strains in the recent past has allowed development of improved genotyping tools for the molecular epidemiology of leprosy. Comparative genomics has identified distinct *M. leprae* genotypes and revealed their characteristic genomic markers. This review summarizes the progress made in *M. leprae* genomics, with special emphasis on the development of genotyping schemes. Further, an updated genotyping scheme is introduced that also includes the newly reported genotypes 1B_Bangladesh, 1D_Malagasy, 3K-0/3K-1, 3Q and 4N/O. Additionally, genotype-specific markers (single nucleotide polymorphisms, Insertion/Deletion) have been incorporated into the typing scheme for the first time to enable differentiation of closely related strains. This will be particularly useful for geographic regions where *M. leprae* strains characterized by a small number of genotypes are predominant. The detailed compilation of genomic markers will also enable accurate identification of *M. leprae* genotypes, using targeted analysis of variable regions. Such markers are good candidates for developing artificial intelligence-based algorithms for classifying *M. leprae* genomic datasets.

Keywords: *Mycobacterium leprae*; genotyping; genomic variability; genetic diversity; comparative genomics; phylogeography

Introduction

Leprosy is regarded as one of the earliest infectious diseases ever known to mankind (Bennett et al. 2008, Santacroce et al. 2021). Leprosy primarily affects peripheral nerves, skin, mucosal membranes, and eyes (Bhat and Prakash 2012, Eichelmann et al. 2013). *Mycobacterium leprae* and *Mycobacterium lepromatosis* are the two known causative agents of leprosy (Han et al. 2014, Deps and Collin 2021). There is a general perception that leprosy is mostly a disease of the past. However, global statistics reveal that ~200 000 new cases are recorded annually worldwide. The latest report from the World Health Organization (WHO) recorded 182 815 new leprosy cases worldwide in 2023, which reflects a 5% increase from 2022. Looking at the data for the last 10 years (2014 to 2023), the total number of cases detected annually has declined by 14.6%. In India, leprosy remains a significant health issue, with cases rising to 135 485 in 2016, then suddenly dropping to 65 147 in 2020 due to under-detection during the COVID-19 lockdown. This number increased to 107 851 by 2023. This fluctuation mirrors patterns in countries like Brazil, South Sudan, and Sri Lanka. The COVID-19 pandemic further complicated leprosy programs globally, reducing detection rates significantly. Since then, increased case reporting has been observed, emphasizing the need for sustained efforts, particularly in priority regions where 95.7% of new cases are reported (WHO 2024).

Looking at historical and documentary evidence of leprosy shows that the earliest recorded description of leprosy is in the 1550 bce in Egyptian papyri and 600 bce in the Indian *Sushruta Samhita* (Robbins et al. 2009, Kohler et al. 2017, Pfrengle et al. 2021, Vedam et al. 2021). The earliest skeletal remains of leprosy were found in Hungary (~5700 years old) and in India (~4000 years old) (Robbins et al. 2009, Kohler et al. 2017, Roberts 2018). It should be noted, however, that no *M. leprae* DNA could be detected in these specimens (Haas et al. 2000, Suzuki et al. 2010), most likely due to poor preservation of the DNA in these samples (Suzuki et al. 2014).

Mycobacterium leprae remains uncultivated on any artificial medium, making it difficult to perform simple microbiological investigations such as drug susceptibility testing or cloning and expression (Maeda et al. 2001, Williams and Gillis 2004, Scollard et al. 2006, Avanzi et al. 2020). The desperate efforts to cultivate *M. leprae* have remained futile. The genomics-based clues into the biology of this bacterium have therefore been invaluable. The genomic sequencing of *M. leprae* TN, a strain from Tamil Nadu, India (Cole et al. 2001) was, therefore, a significant milestone that offered new promises to almost all fields of leprosy research and

Received 13 November 2024; **revised** 13 February 2025; **accepted** 24 April 2025

even beyond. Molecular and immunodiagnostics of leprosy bacillus, efforts for designing vaccines, and identification of potential drug targets, etc., are some of the examples of the fields where the genomic insights were very helpful (Cole et al. 2000, 2001, Eiglmeier et al. 2001a, 2001b, Vissa and Brennan 2001). Therefore, the genomics of leprosy bacilli successfully demonstrates the value of genomic tools when traditional microbiological procedures are not applicable or practicable, such as in the case of uncultivated organisms, archived samples, or where biohazard- and/or transportation-related concerns preclude the possibility of cultivation (Avanzi et al. 2020, Sharma et al. 2020).

There have been significant advances in our understanding of the genome composition and association with the biology of leprosy bacilli in the last two decades. A large number of *M. leprae* strains have been sequenced and their comparative genomic analysis has revealed many new genomic markers that have helped to advance further in the field the molecular epidemiology of leprosy (Vissa and Brennan 2001, Monot et al. 2005, 2009, da Silva Rocha et al. 2011, Kuruwa et al. 2012, Lavania et al. 2018, Das et al. 2020), as well as molecular drug susceptibility testing (MDST) (Maeda et al. 2001, Cambau et al. 2002, Singh and Cole 2011, Benjak et al. 2018, Cambau et al. 2018, Lavania et al. 2018). There have also been significant advances in the field of the comparative genomics, phylogeny, and phylogeography of *M. leprae* as more genome sequences became available for comparison, allowing the development of different genotyping schemes to differentiate *M. leprae* strains (Brosch et al. 2000, Matsuoka et al. 2004, Singh et al. 2015, Urban et al. 2024). Even although the single nucleotide polymorphism (SNP)-typing scheme of *M. leprae* strains was proposed based on a very limited number of strains, it has remained valid over time, and most of the conclusions and classifications are still valid (Singh and Cole 2011, Williams and Gillis 2012, Avanzi et al. 2016, 2020). However, there is a need for an update in the genotyping scheme, as more *M. leprae* genomes from different regions of the world have become available and their comparative genomic analysis has helped in identifying newer subtypes of *M. leprae*, which are not included in the existing methods of *M. leprae* genotyping.

This review aims to summarize these updates in genotyping schemes and presents an overview of genomic variability in *M. leprae* strains. In addition, we have also compiled the genomic variability data to identify genomic markers specific to various genotypes, including the ones which have been identified and described in recent years.

Discovery of the leprosy bacillus

Discovery of *Mycobacterium leprae*

The Norwegian doctor Daniel Cornelius Danielssen (1815–94) is regarded as an early researcher in the field of leprosy research. Danielssen was determined to find the source of leprosy because he thought the condition was inherited and was not a communicable disease. Danielssen's protégé and eventual son-in-law, Gerhard Armauer Hansen, who began working with Danielssen in 1868, had the opinion that microbes were the root cause of the leprosy disease. Despite Danielssen's refusal to accept the bacillus as the causative agent of leprosy, Hansen later identified *M. leprae* as the primary agent of the disease, due to which leprosy is also known as Hansen's disease. Hansen discovered the bacterium using histological examination, staining methods, and microscopy. The disease's neurological symptoms are explained by the bacterium's affinity for nerve tissues, which Hansen also observed. After the leprosy bacillus was discovered, early scientists were curious as to why *M. leprae* had a neurological affinity. Before this discovery, leprosy was primarily believed to be a genetic disease. Therefore, identifying a microbial cause for leprosy was a revolutionary finding. Indeed, leprosy is one of the earliest human diseases to be associated with any bacterial agent (Festskrift 1973, Jay 2000, Marmor 2002, Grzybowski et al. 2013, Hegde et al. 2015). This finding paved the way for the development of diagnostic tests, detection of pathogen transmission, and enabled research into therapeutic interventions. With this discovery, a significant focus on cultivating *M. leprae* in the artificial medium also began, which has so far remained futile. This inability has been a major obstacle in carrying out research into the basic biology of this organism (Rambukkana 2001, Levy and Baohong 2006, Lahiri and Adams 2016).

A collection of clinical and histological studies from leprosy patients spanning the whole clinical range led to the discovery that *M. leprae* specifically targets Schwann cells in peripheral nerves (Drevets et al. 2004, Jessen et al. 2015, Hess and Rambukkana 2019). The ability of leprosy bacilli to adapt to intracellular parasitism, long-term survival inside its host cells, and peripheral nerve predilection, are some of the peculiar features of their biology. By their affinity for Schwann cells of peripheral nerves, leprosy bacilli avail substantial advantages in survival and pathogenesis: it gives access to bacteria to circulate in an area that the host's immune system is unable to reach (Ooi et al. 2004, Tapinos and Rambukkana 2005, Scollard 2008, Scollard et al. 2015, Mietto et al. 2020). It also permits the bacillus to multiply indefinitely because the Schwann cell lacks antimicrobial cellular machinery.

The successful sequencing of the *M. leprae* genome has also provided new insights into the fundamental aspects of the biology of this bacteria (Cole et al. 2001, Eiglmeier et al. 2001). The genomes of *Mycobacterium tuberculosis* H37Rv (Cole et al. 1998) and *M. leprae* TN were compared (Cole et al. 1998, Brosch et al. 2000, Cole et al. 2001, Cole 2002), and it was found that the *M. leprae* genome is much smaller than that of *M. tuberculosis*, suggesting a loss of ∼1000 genes (Table 1). Nevertheless, the remaining functional genes of *M. leprae* are still capable of allowing the bacterium to adapt to intracellular parasitism and long-term survival (Oré et al. 2001, Suzuki et al. 2006, Singh and Cole 2011).

Genomics features of leprosy bacilli and other related mycobacterial pathogens were retrieved from the National Center for Biotechnology Information (NCBI) database, where genome annotations have been updated using the NCBI Prokaryotic Genome Annotation Pipeline. In Table 1, the genome assembly size of the MRHRU-G-235 strain is shown to be ∼80 Kb smaller compared with other complete genomes of *M. leprae* (TN, Br4923, and Kyoto-2). This is most probably due to incomplete assembly/coverage of the *M. leprae* strain MRHRU-235-G.

As can be seen from Table 1, the genome size of *M. leprae* strains and *M. lepromatosis* strains are ∼3.2 Mb while other related mycobacterial pathogens have genome sizes of >4 Mb with the only exception being the genome assembly of *Mycobacterium uberis* Jura strain where the currently assembled genome is represented in the form of 54 contigs totaling to 3.1 Mb (Benjak et al. 2018). *Mycobacterium uberis*is the closest mycobacterial relative of leprosy bacilli, *M. leprae*, and *M. lepromatosis* (Pin et al. 2014) and also shows evidence of reductive evolution. However, a closed circular genome of *M. uberis* is not yet available and thus the actual genome size cannot be stated accurately at present.

The total number of genes in the *M. leprae* and *M. lepromatosis* species is ∼3000, with ∼2300 protein coding genes and ∼700

Table 1. The genome annotation of leprosy bacilli and other related mycobacterial pathogens (source: NCBI RefSeq database, August 2024).

Mycobacterium species		Genomic features										
		Genome size (bp)	Total genes	Protein coding genes (CDS)	CDS pseudogenes	rRNA genes	tRNA genes	Other non-coding RNA genes	% GC content	No. contigs	NCBI Genome Assembly ID	References
M. leprae strains	TN	3 268 203	2898	2164	683	3	45	3	57.8	1	ASM19585v1	Cole et al. 2001
	Br4923	3 268 071	2951	2131	769	3	45	3	57.8	1	ASM2668v1	Monot et al. 2009
	Mycobacterium leprae 3 125 609	3 266 420	2941	2198	692	3	45	3	57.8	122	ASM165349v1	NCBI ^
	Mycobacterium leprae 7 935 681	3 266 556	3116	2354	711	3	45	3	57.8	118	ASM164883v1	NCBI ^
	Kyoto-2	3 268 122	3027	2296	680	3	45	3	57.8	1	ASM358472v1	Kia et al. 2013
	MRHRU-235-G	3 187 112#	2974	2246	677	3	45	3	57.8	1	ASM325377v1	NCBI ^
M. lepromatosis strains	FJ924	3 271 694	3115	2387	676	3	46	3	57.89	1	ASM97526v2	Silva et al. 2022
	Mx1-22A	3 206 741	3025	2276	697	3	46	3	57.89	127	ASM96635v1	Singh et al. 2015
M. uberis Jura		3 122 721	3033	2382	601	3*	44	3	57.5	54	ASM340870v1	Benjak et al. 2018
M. haemophilum DSM 44 634		4 235 765	4042	3721	270	3	45	3	63.9	1	ASM34043v3	Tufariello at al. 2015
M. tuberculosis H37Rv		4 411 532	4008	3906	30	3	45	2	65.5	1	ASM19595v2	Cole et al. 1998

The mycobacterium strains for which assembly is not with complete circular chromosomes are underlined and their contig numbers are mentioned. Abbreviations: ^: Genome assembly not published. *: The 16S rRNA gene was earlier shown in the list of protein-coding genes (Benjak et al. 2018) and thus only two rRNA genes were mentioned in the original annotation. This has been corrected in the above table. #: MRHRU-235-G genome assembly size (3.18 Mb) seems to be incomplete as it is much smaller than the previously reported genomes of *M. leprae* strains. %GC: The percentage of guanine (G) and cytosine (C) bases in genome.

pseudogenes. A comparable number has been observed for *M. uberis*. The other related mycobacterium pathogens *M. tuberculosis* and *Mycobacterium haemophilum*have ∼4000 genes each, with 3721 and 3900 protein-coding genes, while the pseudogene number is only 30 and 270, respectively. This indicates evidence of reductive evolution and genome downsizing in *M. leprae*, *M. lepromatosis*, and *M. uberis*. In all these mycobacterial species, the rRNA, tRNA, and other non-coding RNA genes numbers remain identical. Likewise, the GC content of these reductively evolved species is ∼57.8%, while in other mycobacterial pathogens it is ∼65% (Table 1).

Pseudogenes and intergenic regions, dormant reading frames that have functional equivalents in other mycobacteria, and regulatory sequences make up 49.5% of the total genome of *M. leprae*. Our knowledge of the physiology, pathophysiology, and genetics of *M. leprae* has improved because of the availability of genome sequence analysis and annotations. This knowledge is also helpful for developing molecular epidemiology tools and improved diagnostics for tracking drug-resistance trends and transmission dynamics (Singh and Cole 2011, Williams and Gillis 2012, Avanzi et al. 2016, 2020). As per the current curated annotation available at the NCBI (accessed on 15 August 2024), the genomic features of the reference strains of *M. leprae* (genome assembly ID ASM19585v1), *M. lepromatosis* (ASM97526v2), and other closely related mycobacterial pathogens are shown in Table 1.

Evidence of reductive evolution and genome downsizing is visible for both the leprosy bacilli (*M. leprae* and *M. lepromatosis*) as well as for *M. uberis*, which is a closely related uncultivated mycobacterial species (Table 1). The loss of several metabolic pathways and loss of redundancy are characteristic features of these pathogens. Reductive evolution is typically distinguished by the large number of pseudogenes, a build-up of insertion sequence (IS) elements, and a decreased G + C content (Oré et al. 2001, Gomez-Valero et al. 2007). These traits may indicate evolution via an evolutionary bottleneck and the minimum gene set for maintenance of the pathogenic mycobacterium might have been spontaneously determined by this process. Interestingly, the G + C content of the genes belonging to *M. leprae* is higher (60.1%) than that of the pseudogenes (56.5%) and the remaining genome (54.5%). Cytosine deamination in non-coding areas may have resulted in more neutral G + C content in leprosy bacilli and *M. uberis*. In other mycobacteria species, high G + C content appears to be driven by active gene codon preference (Cole et al. 2001).

Mycobacterium lepromatosis, a newly identified causative agent of leprosy

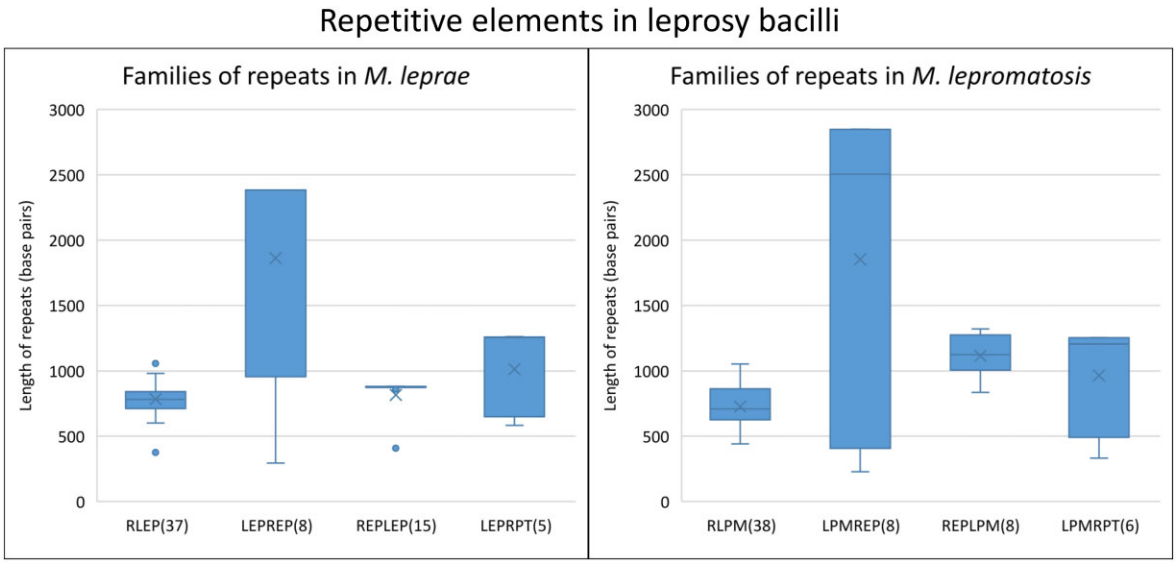
Leprosy has been exclusively associated with *M. leprae* since 1873. However, a new mycobacterial species named *M. lepromatosis* was found to be the causative agent in two Mexican patients affected with leprosy and that died with diffuse lepromatous leprosy in 2008 (Han et al. 2008). The initial report described the sequences of 20 genes of *M. lepromatosis* including 16S rRNA, *rpoB*, and *rpoT* genes (Han et al. 2009), and revealed ∼90% similarity with the corresponding *M. leprae* sequences. This is in stark contrast to the very high level of genome conservation between *M. leprae* strains (>99.995%). The first independent confirmation of the existence of *M. lepromatosis* was then described in a Mexican patient (Vera-Cabrera et al. 2011) and this sample was later used for whole genome sequencing (WGS) that yielded 126 contigs and a genome assembly size of ∼3.20 Mb (Singh et al. 2015). Upon comparing the sequences of the *M. lepromatosis* Mx1-22 strain with the *M. leprae* TN genome, 90.9% identity at the genome level was observed. It was interesting to note that the genetic identity between both species was relatively higher in the protein-coding sequences (93%) compared with the pseudogenes (82%) (Han et al. 2009, Singh et al. 2015), indicating that the SNPs are more easily tolerated in non-coding regions or pseudogenes in contrast to those in the coding regions.

Upon comparative genomics analysis of *M. leprae* and *M. lepromatosis*, species-specific genomic regions were also detected, e.g. *HemW* (earlier annotated as *HemN*) (Dwivedi et al. 2021, Sharma et al. 2021), which is present in *M. lepromatosis* (but absent in *M. leprae*) and has served as a specific diagnostic marker and allowed for screening of *M. lepromatosis* in the archived samples (Singh et al. 2015). Later on, a simple sequencing independent assay for simultaneous detection and differentiation of *M. leprae* and *M. lepromatosis* was developed that was based on a 45 base deletion in the *rpoT* gene of *M. leprae* compared with that in *M. lepromatosis* (Dwivedi et al. 2021), which can be visualized by simple PCR-agarose gel electrophoresis. However, the *rpoT* and *hemW* genes are present in a single copy in the genomes of both species. Therefore, identification of a multicopy repetitive sequence (similar to the RLEP repetitive element in *M. leprae*) could increase sensitivity during the screening of *M. lepromatosis* and will be useful in determining its true prevalence and distribution, which is possible only upon availability of a closed circular genome of *M. lepromatosis*.

Subsequently, a major development took place when *M. lepromatosis* was isolated from a Costa Rican patient and successfully passaged in nude mice (Sharma et al. 2020). This strain, named *M. lepromatosis* NHDP-385, was also analyzed by WGS, which yielded a second genome composed of 77 large and 26 small contigs and comparable genome size (3.26 Mb), as previously reported for other *M. leprae* and *M. lepromatosis* strains. Following the discovery of the first repetitive region associated with *M. lepromatosis* (initially named as RLPM, which was later on renamed as LPMREP) by WGS, a sensitive and species-specific real-time qPCR technique was developed and validated (Sharma et al. 2020). Targeting these sequences, the pathogen was found in 42% of leprosy patients (15/36) in Mexico and among 4% of patients in the USA (3/72). Co-infection with *M. leprae* and *M. lepromatosis* was observed among 11% of patients in Mexico (Sharma et al. 2020). Previously also, high co-infection rates have been reported based on the primers targeting the species-specific polymorphisms in the 16S rRNA gene. However, in most of these initial reports, the resultant amplicons were not sequenced, and detection was primarily based on the appearance of the target amplicon using *M. leprae* or *M. lepromatosis* specific primers.

This very much awaited development took place when the complete circular genome of *M. lepromatosis* FJ924 was described in 2022. Having this complete circular genome also allowed the identification of four families of repetitive sequences in *M. lepromatosis* FJ924, which were named RLPM (38 copies), LPMREP (eight copies), REPLPM (eight copies), and LPMREP (six copies) (Silva et al. 2022). This nomenclature follows the same pattern as for *M. leprae*. The copy number, average length, and percentage identity between conserved parts of these elements are shown in Fig. 1. The knowledge about copy numbers and genomic locations of these repeats has also allowed the development of more sensitive diagnostics.

As per the available sequencing data at NCBI, *M. lepromatosis* is distributed mainly in Mexico, Singapore, Canada, Myanmar, and Brazil (Virk et al. 2017). In addition to the reports of *M. lepromatosis* in human samples from different countries, a very unusual



Repetitive elements								
M. leprae				Repetitive elements similarity (%)	*M. lepromatosis*			
Four main repetitive elements	Repeat name	Copy number of repeats	Average repeats size (bases)		Four main repetitive elements	Repeat name	Copy number of repeats	Average repeats size (bases)
A	RLEP	**37**	700	75%	A	RLPM	**38**	700
B	LEPREP	8	**2400**	79–83%	B	LPMREP	8	**2800**
C	REPLEP	**15**	**900**	74%	C	REPLPM	**8**	**1100**
D	LEPRPT	**5**	1200	87%	D	LPMRPT	**6**	1200

Figure 1. Families of repeats in *M. leprae* and *M. lepromatosis*, and the percentage of sequence similarity between them. There are four families of repetitive elements in both *M. leprae* and *M. lepromatosis*. The similarity between repeat groups (A, B, C, and D) in both species is very low (75–87%) and repeat copy number and length size (bases) also vary a lot. *Mycobacterium leprae* multicopy repeats REPLEP have very few differences in repeat size compared with the multicopy repeats REPLPM of *M. lepromatosis* .

finding, which was the detection of *M. lepromatotis* in red squirrels (*Sciurius vulgaris*) in Scotland, has been described (Meredith et al. 2014, Avanzi et al. 2016). *Mycobacterium lepromatosis* has also been found in squirrels with and without lesions resembling leprosy in England, Ireland, and Scotland (Avanzi et al. 2016). Until now, *M. lepromatosis* has not been discovered in armadillos, in contrast to *M. leprae* (Truman et al. 2011, Monsalve-Lara et al. 2024). Comparative genomics and Bayesian analysis of the Mexican strain of *M. lepromatosis* isolated from patients and the British strain of *M. lepromatosis* isolated from red squirrels showed large genetic differences and a divergence period of ~30 000 years, implying that these two lineages have remained quite distant from each other during their evolutionary journeys (Silva et al. 2022). The illnesses among red squirrels appear to cause little concern to public health, as there are no reports of leprosy infections among native inhabitants in the UK (Avanzi et al. 2016).

Upon data mining of the raw fastq sequences while writing this review article, we observed that the RLPM repetitive element sequences are also present in previously described *M. lepromatosis* genomes from other patients (Mx1-22, PL-02, and NHDP-385), as well as in the red squirrel-derived *M. lepromatosis* strains (namely, IR-25 and IR-36), and their count is consistently higher compared with the fastq reads corresponding to the single copy genes such as *rpoB* and *gyrA*. This is an important and encouraging observation, as it implies that the RLPM is present in multiple copies in all currently known strains of *M. lepromatosis*. This is similar to the RLEP copy counts in *M. leprae* strains.

Genomics and evolutionary journey of the leprosy bacilli

Leprosy presents as a range of diverse signs and symptoms with different bacillary loads that frequently cause painful immunological reactions. Multidrug therapy (MDT) is recommended by the WHO for the treatment of leprosy. All forms of leprosy, including the condition known as Lucio's phenomenon, were believed to be only caused by *M. leprae* until *M. lepromatosis* was described in diffuse lepromatous leprosy cases by Han and colleagues in 2008. MDT is equally effective for infections caused by *M. leprae* and *M. lepromatosis* (Sharma et al. 2020).

Repetitive elements in the leprosy bacilli

The *M. leprae* genome has four families of scattered repeats: RLEP (37 copies); REPLEP (15 copies); LEPREP (eight copies); and LEPRPT (five copies) (Fig. 1). Additionally, there are about 26 dead IS elements. These repetitive sequences make up ~2% of the TN genome, and their recombination activities are primarily responsible for the downsizing of the *M. leprae* genome (Cole et al. 2001), but the repeats seem to have lost their ability to transpose. Genomic evidence suggests that these repeats likely underwent recombination events that contributed to the *M. leprae* genome downsizing. Comparative genomics of *M. leprae* strains later confirmed the lack of variety found in early RLEP-based restriction fragment length polymorphism investigations. Consequently, it appears that the sequences at the 3' end of genes and often within

Table 2. Identification of nine base insertions in the *M. leprae* TN strain genome.

Genomic position	Genomic variant type	*M. leprae* TN genome, GenBank id AL450380.1 version 1, genome size 3 268 203	All other *M. leprae* genomes (*)
384 393	INDEL	GC	GCC
944 191	INDEL	CA	CAA
958 228	INDEL	ACC	ACCC
1 533 315	INDEL	C	CG
1 841 279	INDEL	CG	CGG
1 849 026	INDEL	GCC	GCCC
1 912 456	INDEL	AC	ACC
2 893 091	INDEL	GT	GTT
3 100 774	INDEL	GCCC	GCCCC

M. leprae TN genome, version 2 (ID AL450380.2) with genome size 3 268 212 updated version present in the Mycobrowser database (https://mycobrowser.epfl.ch/) (Kapopoulou et al. 2011). Abbreviation: InDel: Insertion/Deletion.

Table 3. The types of genomic variations observed upon comparing *M. leprae* genomes (Monot et al. 2009, Schuenemann et al. 2013, Benjak et al. 2018, Avanzi et al. 2020, Hockings et al. 2021, Urban et al. 2024).

Genomic polymorphism and their effects	Number of SNPs
Intergenic region	1167
Missense variant	1352
Synonymous variant	884
Non-coding transcript variant	1340
Start lost	3
Stop gained	22
Stop lost	2
Total SNPs	4770

pseudogenes include RLEP, indicating that these are the remaining transposons that are no longer able to conduct transposition (Williams et al. 1990). Additionally, in the sections that these repeating elements flank, the chromosomal rearrangements (inversions, translocations, and gene deletions) caused by recombination account for the loss of synteny (discontinuities in gene order, in contrast to that in the *M. tuberculosis* genome) (Cole et al. 2001).

Genomic variability and genetic diversity

Despite genome degradation and the presence of many pseudogenes, *M. leprae* strains have exceptionally low genomic variability. The second *M. leprae* strain to be sequenced for its entire genome was *M. leprae* Br4923, isolated from a Brazilian patient. It contains 3268071 bp and differed in 185 locations only, including 31 variable-number tandem repeats (VNTR) regions, eight insertion-deletion (InDel) events, and 146 SNPs. Following that, whole genome re-sequencing of two strains, namely Thai53 (Thailand) and NHDP63 (USA), was completed using Illumina technology and were compared with the TN and Br4923 reference strains (Monot et al. 2009).

This comparison revealed that the original submission of the *M. leprae* TN genome had missed nine bases at various positions (Monot et al. 2009) (Table 2). Hence, the size of the *M. leprae* TN genome was determined to be 3268212 bp and subsequent work had described the SNPs as per the new coordinates (Truman et al. 2011), which are also available from the Mycobrowser database (Kapopoulou et al. 2011). However, it was observed that this could lead to inconsistencies in the SNP numbering system (GenBank ID AL450380.1, genome size 3268203 bp) and can create difficulties while comparing the SNP positions described earlier, e.g. the SNP positions described in a recent paper follow the updated numbering system (Jouet et al. 2023). Therefore, to maintain uniformity with the previous literature, the originally described genome coordinate numbering system can be uniformly followed. These additional nine bases missed from the original genome assembly could be useful for validating the bioinformatic data analysis pipelines, particularly those for determining the InDels and can serve as an inbuilt quality control measure. For all the newly sequenced *M. leprae* genomes, these positions should appear as InDels at the positions mentioned in Table 2.

The comparative genomic analysis of *M. leprae* genomes available at the time of this analysis revealed ~4770 SNPs. These SNPs are present in genes, pseudogenes, repetitive regions, or in the intergenic regions and are either non-synonymous or synonymous variants. The classification of these SNPs into various categories is shown in Table 3. Importantly, 22 SNPs led to a gain of premature stop gain while two SNPs also led to the loss of a start codon and one SNP led to the loss of a stop codon, which could potentially lead to pseudogenization. In addition, the InDel mutations are very likely to disrupt the entire protein sequence of a gene and thus could result in a pseudogene formation. One well-characterized example is the pseudogenization of ML0825c due to the insertion of a "T" at position 978 585 in the strains belonging to the genotypes 4O and 4P. The ML0825c has been described as a repressor of 32 genes, many of which are related to virulence, such as *esxG* and *esxH*. In an important investigation of comparing the growth kinetics of the two different genotypes (with and without this frameshift deletion), the difference in growth rate was observed (Sharma et al. 2018), implying the possibility of the existence of pathological variants in leprosy bacilli. In a recent study comparing the morphotypes of three different strains belonging to different genotypes (1A, 3I, and 4N), it has been shown that, despite remarkable genome conservation, a few genomic changes between different strains may still cause changes in the phenotype such as their growth rates, intracellular viability, and interaction with the host cells, which can alter the host immune response and thus may contribute to differences in clinical manifestations (Gomes et al. 2024). More studies concerning this will be helpful in determining the role of the pathogen genotype in diverse clinical presentations, which so far has been mostly attributed to host immunogenetics only.

Genome sequencing of *M. leprae* strains has been done successfully from different countries representing various time periods including ancient DNA studies from leprosy skeletons samples (Supplementary Table S1). Considering that such genomes are highly decayed, one might expect higher variability in the genome from different locations. Nonetheless, genomic variability was also minimal between those strains. However, according to detailed phylogeographic investigations into *M. leprae* strains collected from around the world, Monot and colleagues showed that all known *M. leprae* strains are clonal and possess only a handful of SNPs (Monot et al. 2005, 2009) or varying amounts of tandem repeats in their genome. This makes it very difficult to differentiate different strains of *M. leprae* by simple PCR-based tests.

After having undergone a massive pseudogenization and genome downsizing event some 20 million years ago, The Most

Recent Common Ancestor diverged into *M. leprae* and *M. lepromatosis* ~13.9 million years ago (Singh et al. 2015). However, it seems that there has been a remarkable level of genome conservation since then, as seen from the comparative genomic analysis of the extinct and extant strains of *M. leprae* (Schuenemann et al. 2013, Avanzi et al. 2020).

Findings of the genomic analysis of *M. leprae* strains, the genotyping scheme, and MDST assays for *M. leprae* have been developed. Sequencing has become more affordable now owing to significant advances in Next Generation Sequencing (NGS) technologies and library preparation methods. However, bioinformatics analysis of the sequencing datasets (for Sanger and NGS) still requires significant expertise. Currently, while curated databases or bioinformatics pipelines are available for other mycobacterial species, there are no such tools available for *M. leprae*.

A wide range of molecular epidemiology techniques are being developed to identify potential human and non-human reservoirs, the source of infection, and to distinguish between relapse cases and re-infection cases. The source and dynamics of transmission can also be better understood with the use of strain molecular typing, and the modes of transmission of *M. leprae* are still not fully understood. However, early detection of persons affected with leprosy is crucial for the successful implementation of MDT, the prevention of disabilities, and the control of spread. Nowadays, there are more opportunities to investigate the molecular epidemiology of leprosy via genome-derived information of polymorphic positions. However, WGS still remains a challenging and expensive approach for routine application. Therefore, the variable positions that are useful for molecular epidemiology or molecular drug susceptibility testing can be investigated in a targeted manner wherein many such loci are targeted. This approach has been used successfully in Deeplex Myc-Lep technology (Jouet et al. 2023). In this approach, a total of 43 loci are included: 18 SNPs and 11 VNTR loci for molecular strain typing and the remaining loci for detecting drug resistance determining regions, including the well-known mutations in *rpoB* (associated with rifampicin resistance), in *folP1* (associated with dapsone resistance), and in *gyrA* and *gyrB* (associated with fluoroquinolones resistance) genes. In addition, a few exploratory loci such as *ctpC* and *ctpI* (Singh and Cole 2011) have been included in this assay as these genes are annotated as efflux proteins and were suspected to play a role in resistance in the phenotypically rifampicin-resistant strain Airaku-3, which had wild-type *rpoB* sequences, although no such confirmatory link has been shown so far. Additionally, the Endonuclease III gene *nth* (ML2301) involved in base excision DNA repair mechanism is also included in the Deeplex assay as it has been associated with the emergence of mutations in other genes (*rpoB, folP1*, and *gyrA*) (Benjak et al. 2018, Cambau et al. 2018). However, *nth* mutations were not found in 17 strains possessing mutations in *gyrA* or *folP1* genes in a recent study in *M. leprae* strains from central India. By contrast, mutations (including both synonymous and non-synonymous changes) were found in seven samples possessing wild-type sequence in *folP1*, *gyrA*, and *rpoB* (Sharma et al. 2024). Thus, the role of *nth* needs further evaluation in drug resistance in strains representing various geographical regions.

Such targeted sequencing using NGS technologies offers great promise as they allow genotyping of *M. leprae* directly in clinical samples and the species differentiation from *M. lepromatosis* (Marijke Braet et al. 2022, Jouet et al. 2023). This assay also includes well-established drug resistance associated markers and newly proposed loci, allowing the detection of mutations associated with rifampicin (*rpoB/ctpC/ctpI*), dapsone (*folP1*) and fluoroquinolone (*gyrA/gyrB*), for dapsone-resistance, for ofloxacin and the newly proposed *atpE* for bedaquiline. This ability to accurately predict (hetero)resistance to major anti-leprosy drugs such as rifampicin, dapsone, and fluoroquinolones in a single assay provides a significant advancement in leprosy diagnosis and clinical decision-making on patient treatment. The assay can be updated adding or removing loci as per the need and this assay also contributes to the epidemiological tracking of leprosy by recognition of strains, lineages and identification of mixed infections, essential for molecular epidemiology and detecting transmission of resistant strains (Marijke Braet et al. 2022, Jouet et al. 2023). The currently available Deeplex assay includes the *hsp65* gene for mycobacterial species identification. Inclusion of RLEP (for *M. leprae*) and RLPM (for *M. lepromatosis*) can be considered in subsequent updates in this assay to further enhance the analytical sensitivity and ability to detect mixed infection.

Molecular epidemiological tools based on genomic variability in *M. leprae*

The determination of prevalent genotypes of a pathogen using various genotyping tools or molecular epidemiological approaches has been helpful in improving the understanding of pathogen transmission. For *M. leprae*, the SNP-based genotyping scheme was developed at the time when detailed genomic information of only two reference strains (*M. leprae* TN and *M. leprae* Br4923) was available (Monot et al. 2005, 2009). Their comparison identified polymorphic locations that were then evaluated in strains from different locations of the globe. Based on the other two later obtained genomes, the number of polymorphic loci was expanded. The current genotyping system for *M. leprae* was initially based on SNP analysis on just three SNPs (14676, 1642875, and 2935685) located in RNC1, ML1378, and ML2462 in the genome, classifying all known strains of *M. leprae* into four major SNP types (SNP types 1, 2, 3, and 4). These markers have been shown to correlate with the geographical distribution of *M. leprae* strains. Further subtyping into 16 groups was proposed (1A-4P). Subsequently, when more genomes became available for comparison, there were some refinements, e.g. the merging of a few genotypes while a few new subtypes have been identified, taking the current total to 19 (depicted in Figs 3, 4, and 5) (Avanzi et al. 2020).

However, the main structure of this genotyping scheme largely still stands. Since the description of SNP-subtyping (Monot et al. 2009), it took ~15 years before the number of publicly available *M. leprae* genomes substantially increased to 328. From published reports, the most common genotypes in the world among these strains are subtypes 1D, 3I, and 4N, although the representation of different strains and geographical areas is not uniform. Molecular epidemiology research does not yet employ WGS as a routine procedure and limited information is available about polymorphic markers to explore their genetic diversity and transmission in a global context. To completely understand the diversity of genetic variation in *M. leprae* strains, additional research in this area is necessary. This paper therefore also describes the evolution of leprosy molecular epidemiology and the role played by WGS-derived data in expediting the identification of genomic markers and mutations associated with various genotypes and drug resistance.

Development of genotyping schemes for *Mycobacterium leprae* through comparative genomic analysis

The development of genotyping schemes for *M. leprae* has significantly advanced through comparative genomic analysis. Recent

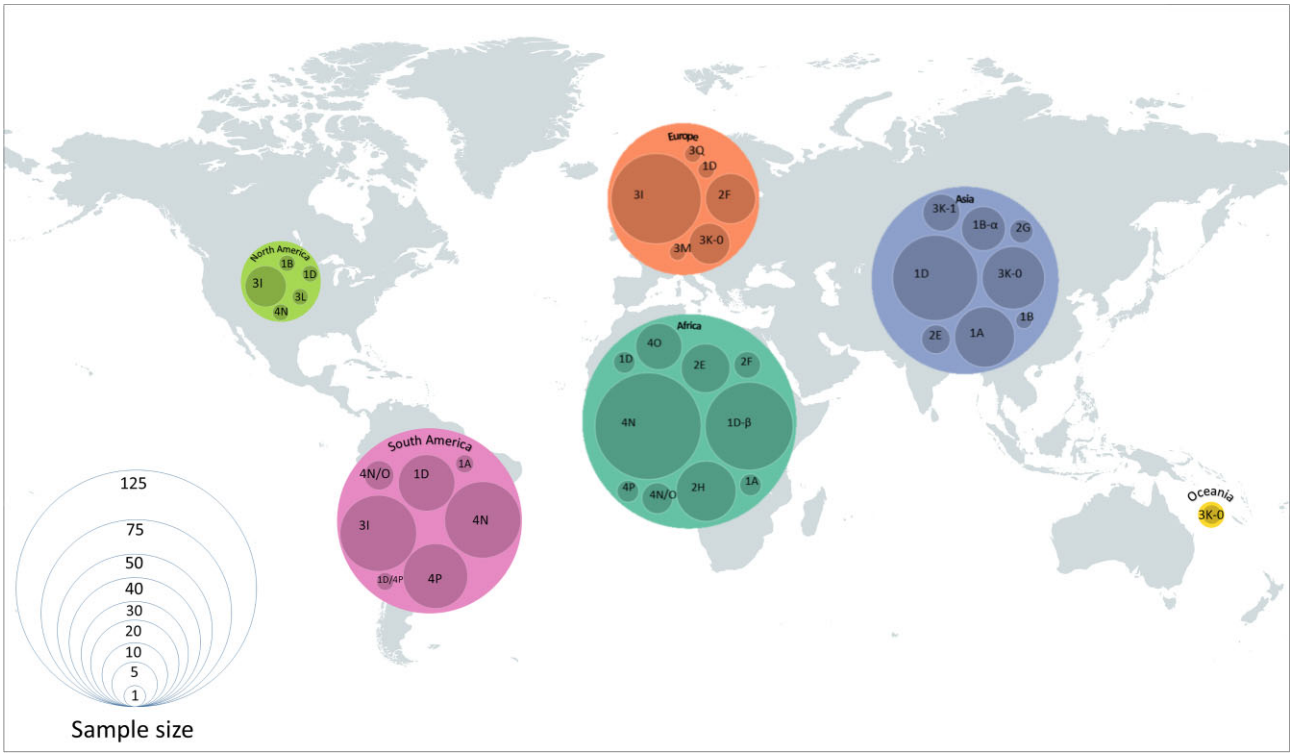


Figure 2. *Mycobacterium leprae* genomes representing various genotypes from different continents. The size of circles corresponds to the number of samples of a particular genotype and colors represent the different continents. α: Bangladesh; β: Malagasy.

advancements in WGS have enabled researchers to analyze *M. leprae* strains directly from clinical specimens, which is crucial given the organism's inability to be cultured *in vitro*, containing considerable number of SNPs in genes, intergenic regions, and pseudogenes that are instrumental for strain differentiation and understanding evolutionary relationships among strains (Dwivedi et al. 2024).

VNTRs are short, repetitive DNA sequences that vary in number among different strains. In *M. leprae*, specific VNTR loci have been identified as useful markers for genotyping. The addition and integration of VNTRs to SNP-based genotyping has refined the genotyping schemes, allowing for a more comprehensive understanding of strain diversity and transmission dynamics. For instance, studies have identified distinct VNTR patterns that can differentiate between closely related strains, aiding in epidemiological tracking (Ho Tinh Tam et al. 2023, Dwivedi et al. 2024).

It was observed that individually, either SNP- or VNTR-based methods have their inherent limitations. For example, SNPs are not frequent enough for differentiation of closely related strains, while some of the VNTRs could be so variable that *M. leprae* DNA from two different lesions of the same patient may show variations in some loci. Therefore, careful selection of VNTRs is important and adding such VNTR loci to the SNP patterns can further help in discriminating closely related strains. In some situations, this discrimination can even occur without the need for identification at the SNP level. On the contrary, despite the use of these highly discriminatory loci, if one observes identity on the genome-scale level, this is an indication of a recent common origin. The power of combining the SNP- and VNTR-based genotyping schemes was realized during demonstrating the zoonotic link between human and armadillo leprosy (Truman et al. 2011, Sharma et al. 2015, da Silva et al. 2018, da Silva Ferreira et al. 2020).

Some of the VNTR loci have received particular attention during studies on genetic diversity among *M. leprae* isolates, including AC8b, AC9, AC8a, and GTA9. A study on leprosy patients conducted in Brazil utilized these VNTR loci to analyze nasal secretions and skin biopsies from patients, revealing significant variability in VNTR patterns that could indicate different transmission pathways within the community (Fontes et al. 2017, Lima et al. 2018). However, a few VNTR loci are known to be hypervariable (Monot et al. 2009) and the inclusion of selected VNTR loci can be helpful for short range transmission studies where SNP differences are often minimal. This was further demonstrated in a study by Fontes et al. (2017) by combining SNP and VNTR typing. They observed different genotypes of *M. leprae* in nasal swabs and skin biopsies in a few cases when VNTRs with highest resolution were included. The authors also tried to associate demographic data with molecular typing but with limited success (Fontes et al. 2017).

The application of VNTR typing is crucial for understanding leprosy transmission networks. By analyzing the genetic profiles of various isolates, researchers can identify clusters of infection and track the spread of specific strains within populations. This capability is essential for public health efforts aimed at controlling leprosy outbreaks and understanding the disease's epidemiology in endemic regions (Ho Tinh Tam et al. 2023). As genomic techniques advance, the refinement of VNTR typing in light of newer genome datasets will likely enhance its utility in leprosy research. Combining VNTR data with other genomic markers, such as SNPs, may provide a more comprehensive understanding of *M. leprae* diversity and evolution. VNTR analysis represents a pow-

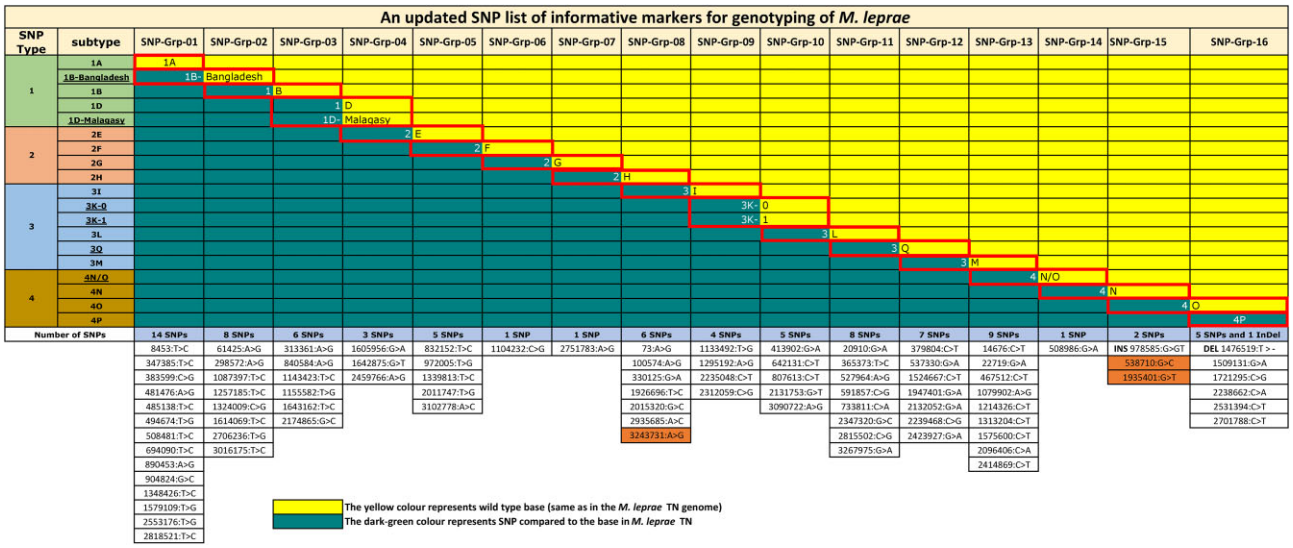


Figure 3. Updated informative markers for genotyping in *Mycobacterium leprae*. Starting from the third column, the header row indicates SNP-groups (SNP-Grp-01 to 16) whose SNPs are mentioned at the bottom of the column. The left-side column shows SNP types (1–4) and their 19 subtypes (1A-4P). The SNPs mentioned at the bottom represent the full list of SNPs for every SNP group. A red rectangle is shown in each row corresponding to all known 19 genotypes. For the genotype on the terminal sides of this scheme, i.e. 1A and 4P, the rectangle covers only one cell representing one SNP-Grp. Genotype 1A can be identified by a single SNP from all wild-type positions (shown in yellow color for SNP-Grp-1), as is the case for genotype 4P (all alternative nucleotides, shown in green color in SNP-Grp 16). The corresponding rectangle for the remaining genotypes includes two adjoining SNP-Grps: the first possessing alternative bases (dark green color) at the corresponding positions, while the second SNP-Grp exhibits wild-type bases (yellow color). A single marker per SNP group from the corresponding rectangle is considered to be sufficient for identifying the respective genotype. The SNPs shown in orange highlighted cells are newly identified genomic markers added in this figure. The total number of SNPs for every SNP-Grp is mentioned at the start of the SNP list at the bottom of each column. The subtypes that are newly incorporated in this genotyping scheme are underlined. The genomic positions are based on original reference genome *M. leprae* TN strain (GenBank ID AL450380.1 version 1, genome size 3268203).

erful approach for identifying genomic markers in *M. leprae*, contributing significantly to our understanding of leprosy transmission and epidemiology (Shinde et al. 2009, Fontes et al. 2017, Lima et al. 2018).

Figure 2 illustrates the genetic diversity of *M. leprae* based on an analysis of 328 genomes from 50 countries, as detailed in Supplementary Table S1. The genotypes encompass all four SNP types (1–4) and 19 subtypes (1A-4P), with subtypes 1D, 3I, and 4N being the most prevalent, collectively representing 50% of the total genomes analyzed.

Geographic distribution of these subtypes reveals that in Asia, the predominant subtypes are 1A, 1D, and 3K (3K-0 and 3K-1), while in Africa the subtypes 1D-Malagasy, 2E, 2H, 4N, and 4O are primarily found. In the South American countries, the subtypes 1D, 3I, 4N, and 4P are common. In North America, based on the data from patients detected in the USA, the subtype 3I is frequently identified, particularly in the autochthonous cases and in the strains derived from armadillos. The human strains from Europe are only represented by skeletal samples, which exhibit genotypes 2F, 3I, and 3K-0. In addition, one patient sample from Surinam exhibited a mixed infection of subtypes 1D and 4P (Fig. 2 and Supplementary Table S1). The detection of such mixed genotype infection is possible only through NGS-based approaches and reflects the power of genomic technologies. In the recent past, hundreds of *M. leprae* genomes have been successfully sequenced and their genomic datasets made freely available. However, the identification of correct SNP-types/subtypes and drug resistance-related mutations from millions of small reads is not only difficult but can be error prone as it also requires understanding of the genotyping scheme. There have been several updates in recent years as new subtypes have been reported in recent studies. An updated genotyping scheme including these new subtypes is therefore required.

Updated SNP-genotyping scheme and genotype-specific markers

There have been several studies where SNP-genotyping has been successfully used for strain classification including the recent WGS-based descriptions of new SNP-subtypes. These subtypes are defined by the help of genomic markers in the form of SNP groups (SNP-Grp), which include informative SNPs and InDels. In the context of *M. leprae* genotyping, the informative markers are described as the SNPs/InDels that follow a "staircase pattern", as shown in Fig. 3. The ladder-like representation in this figure illustrates that a row featuring a combination of a green cell (indicating a base similar to *M. leprae* Br4923) immediately followed by a yellow cell defines its corresponding SNP subtype (listed in the second column of that row). The first informative marker (corresponding to subtype 1A) is at position 8453, in which subtype 1A strains possess a "T" nucleotide, consistent with the reference genome (*M. leprae* TN), represented in yellow. By contrast, all other subtypes (1B-4P) exhibit a "C" nucleotide, shown in dark green. Over the years, there have been several new studies, and genomes of several *M. leprae* strains have been sequenced that have revealed new subtypes. However, the genotyping scheme has not been updated to include these newly detected subtypes. Hence, this review has aimed to detail these updates and compile a new genotyping scheme as detailed in Fig. 3.

(A)

Genotype (number of strains)	Best five genotype-specific SNPs for direct identification of *M. leprae* genotype				
1A (16)	8453:T	494674:T	508481:T	890453:A	2818521:T
1B_Bangladesh (7)	14502:G>A	62630:C>A	70738:G>A	501955:C>A	572401:C>T
1D (36)	639580:A>G	3016895:C>A	3262657:C>T	492614:C>A (26/36)	3070616:A>C (26/36)
1D_Malagasy (32)	163438:C>T	228773:T>G	317057:G>T	480629:C>T	616172:G>C
2E (13)	48732:C>T	421565:C>T	848261:G>A	1974673:C>T	2061910:G>A
2F (10)	13907:C>T	537503:G>T	1604320:A>G	3076110:G>A	3257070:G>A
2G (2)	9297:C>T	16094:C>A	20906:G>A	27227:G>A	29739:A>G
2H (15)	24864:G>A	31049:G>A	31839:G>A	44180:T>C	58052:C>T
3I (40	883:G>A	7614:C>T	193848:C>T	604325:T>C	1017407:G>A
3K-0 (17)	509178:G>A	563796:G>C	1262010:T>C	1530718:A>G	1828126:A>G
3K-1 (5)	15439:G>A	26545:G>A	57633:T>G	62545:G>T	165853:C>A
4N (66)	305597:G>C	511079:A>C	1037678:C>T	1237129:C>T	1730876:C>T
4O (9)	886998:T>G	2105674:T>C	2105745:C>T	2311044:G>T	2501116:C>T
4P (8)	1509131:G>A	1721295:C>G	2238662:C>A	2531394:C>T	2701788:C>T

(B)

Genotype (number of strains)	SNPs uniquely present in strains belonging to these genotypes.				
1B (2)	4423:G>A (1/2)	12709:A>G (1/2)	21266:G>A (1/2)	43827:C>G (1/2)	230575:G>T (1/2)
3L (1)*	23299:G>A	46096:A>G	56421:C>T	76343:C>T	86867:G>A
3Q(1)*	116056:C>T	184256:G>T	507793:T>G	1003977:C>G	1136631:G>A
3M (1)*	681512:C>G	931817:G>A	1007582:G>A	1012293:C>T	1674957:A>G
4N/O (6)	45159:A>G (2/6)	47405:C>T (2/6)	223657:C>T (2/6)	321320:G>T (2/6)	433909:A>G (2/6)

Gray color indicates that these SNPs are not present in all strains of that genotype

Figure 4. (A) The best five genomic markers specific to each SNP-genotype of *M. leprae* strains. These SNPs can be used for decision trees for automated and accurate identification of genotypes from the SNP list obtained from whole genome sequencing of *M. leprae* strains. This can help with accurate and reproducible identification of sub-genotypes of *M. leprae*. **(B)** SNPs uniquely present in strains belonging to these genotypes. *The genotypes 3L, 3Q, and 3M are represented by only a single strain for which genome information is available. That is why some or all of the SNPs uniquely present in these strains (Fig. 4b) might be strain-specific instead of genotype-specific. Likewise, the genotypes 1B and 4N/O are represented by two and six strains, respectively. However, SNPs specific to these two genotypes are not known.

Identification of new subtypes in light of recent WGS studies and their inclusion in the genotyping scheme

The WGS of >300 *M. leprae* strains has been completed. Approximately 4770 SNPs have been identified by comparative genomic analysis of these strains (Monot et al. 2009, Schuenemann et al. 2013, Benjak et al. 2018, Avanzi et al. 2020, Hockings et al. 2021, Urban et al. 2024). These genomes also include the newly described subtypes such as 1B-Bangladesh, 1D-Malagasy, 3K-0, 3K-1, 3Q, and 4N/O. This has been addressed in the present genotyping scheme by incorporating additional markers (referred to as subtype-specific markers). The updated genotyping scheme shown in Fig. 3 includes newly described subtypes while the subtypes 1C and 3 J have been omitted as these cluster within the subtypes 1D and 3K, respectively. Genotype 3Q has been reported in an ancient DNA sample BEL024 (from Belarus) (Pfrengle et al. 2021). Upon looking at the SNP genomic markers of this strain using the genomic loci mentioned in Fig. 3, it appears to be an intermediate strain between 3L and 3M. Hence this 3Q strain has been incorporated in the typing scheme accordingly as per its SNP pattern.

Identification of subtype-specific genomic markers

In Fig. 3, there are 16 groups of SNPs (SNP-Grp), totalling 86 SNPs. It is difficult to utilize all SNP markers mentioned in Fig. 3 using a PCR-based approach, and also not necessary as just one SNP per group can provide the required information. Using these SNP groups, the subtypes 1D cannot be differentiated from 1D-Malagasy. Likewise, 3K-0 and 3K-1 remain undifferentiated using these SNP-group markers. Therefore, the use of subtype-specific markers can be helpful. This has been compiled for all the subtypes in Fig. 4. Figure 4A lists the best five SNP markers specific to a particular subtype. Such markers can differentiate between 1D and 1D-Malagasy, and 3K-0 and 3K-1 also. Such markers are useful for automated bioinformatics pipelines or targeted NGS approaches such as NGS and the Deeplex technology. However, for manual and PCR-based genotyping, a simple version of this list will be useful. Therefore, a list representing just one SNP specific to each subtype is compiled in Fig. 5.

However, the genotypes 3L, 3Q, and 3M are represented currently by only one strain each. The subtypes 1B and 4N/O are represented by only two and six strains, respectively. Because these subtypes are represented by only one or a very few strains, it is difficult to determine which markers are associated with that subtype and which ones are likely to be strain-specific, therefore these subtypes are listed separately in Fig. 4b. Such subtype-specific markers need to be evaluated in more samples belonging to these genotypes. Fig. 4b displays the five subtypes for which it is difficult to determine specific SNPs, as a limited number of genomes belonging to these SNP-types are available for comparison at present.

Genotype (number of strains)	Genotype-specific SNP for direct identification of a sub-genotype of *M. leprae*																
	8453 T>C	14502 G>A	639580 A>G	163438 C>T	48732 C>T	13907 C>T	9297 C>T	24864 G>A	883 G>A	509178 G>A	15439 G>A	23299 G>A	116056: C>T	681512 C>G	305597 G>C	886998 T>G	1509131 G>A
1A (16)	T	G	A	C	C	C	C	G	G	G	G	G	C	C	G	T	G
1B-Bangladesh (7)	C	A	A	C	C	C	C	G	G	G	G	G	C	C	G	T	G
1B (2)*	C	G	A	C	C	C	C	G	G	G	G	G	C	C	G	T	G
1D (36)	C	G	G	C	C	C	C	G	G	G	G	G	C	C	G	T	G
1D-Malagasy (32)	C	G	A	T	C	C	C	G	G	G	G	G	C	C	G	T	G
2E (13)	C	G	A	C	T	C	C	G	G	G	G	G	C	C	G	T	G
2F (10)	C	G	A	C	C	T	C	G	G	G	G	G	C	C	G	T	G
2G (2)	C	G	A	C	C	C	T	G	G	G	G	G	C	C	G	T	G
2H (15)	C	G	A	C	C	C	C	A	G	G	G	G	C	C	G	T	G
3I (40	C	G	A	C	C	C	C	G	A	G	G	G	C	C	G	T	G
3K-0 (17)	C	G	A	C	C	C	C	G	G	A	G	G	C	C	G	T	G
3K-1 (5)	C	G	A	C	C	C	C	G	G	G	A	G	C	C	G	T	G
3L (1)	C	G	A	C	C	C	C	G	G	G	G	A	C	C	G	T	G
3Q(1)	C	G	A	C	C	C	C	G	G	G	G	G	T	C	G	T	G
3M (1)	C	G	A	C	C	C	C	G	G	G	G	G	C	G	G	T	G
4N/O (6)*	C	G	A	C	C	C	C	G	G	G	G	G	C	C	G	T	G
4N (66)	C	G	A	C	C	C	C	G	G	G	G	G	C	C	C	T	G
4O (9)	C	G	A	C	C	C	C	G	G	G	G	G	C	C	G	G	G
4P (8)	C	G	A	C	C	C	C	G	G	G	G	G	C	C	G	T	A

Figure 5. Genotype-specific SNPs in *M. leprae*. The number of genomes currently known for each subtype is indicated in parentheses. It should be noted that the markers indicated for the subtypes represented by only one strain (3L, 3Q, and 3M) could be strain-specific also. ∗represents the genotypes (1B and 4N/O) for which currently genotype-specific markers are not known.

A simple description of genotype-specific markers is compiled in Fig. 5, which represents one genomic marker specific to a particular genotype, enabling its direct detection. For example, the second column in Fig. 5 lists specific SNP "14 502:G > A", where "A" is present in only 1B-Bangladesh strains at position 14 502 while the remaining genotypes have a "G" at this position. These SNP positions shown in Fig. 5 can identify a predominant genotype in a particular geographical region, as e.g. >75% of the strain present in north and central India belong to genotype 1D (Kuruwa et al. 2012, Lavania et al. 2015, Das et al. 2020). This genotype can be directly identified by SNP 639580: a "G" can confirm the presence of 1D subtype while all other genotypes exhibit a "A" at this position. In most of the studies, InDels are not being reported and both their inclusion and improved and standardized InDels identification pipelines are therefore needed as these could be helpful for finer resolution by targeted analysis using Deeplex technology.

Compilation of updated genotyping scheme using informative and subtype-specific markers

The complete list of SNPs in each SNP-group is mentioned in Fig. 3. However, for PCR-based investigations, only one SNP per SNP-group is sufficient to provide the desired information. Therefore, a summary version representing only the first SNP of each SNP-group is depicted in Fig. 6. It also includes the use of subtype-specific SNPs (highlighted in blue in the last column). Thus, Fig. 6 includes the summary of Figs 3, 4, and 5 and is proposed as the updated genotyping scheme.

It includes only one SNP position from every SNP-Grp (shown in Fig. 3) and only the selected SNPs from Fig. 4a, which are needed for further differentiation of *M. leprae* subtypes. The three SNPs shown on the left side are used for identification of SNP types 1–4 and SNP positions mentioned on the right side are used for identification of SNP subtypes (1A-4P). The last column on the right side was added newly in this scheme to differentiate the genotypes 1D and 1D-Malagasy from each other (163 438:C > T). Likewise, the SNP types 3K-0 and 3K-1 can be differentiated based upon the genotype-specific SNP marker (509 178:G > A). Such subtype-specific SNPs are shown in blue color in Fig. 6.

In this review, we have summarized all recent studies in which new genotypes of *M. leprae* were described. This has been presented as a comprehensive list of the SNPs in various figures. We have kept the same color schemes as in the previous publications (Monot et al. 2009, Truman et al. 2011) to keep the pattern consistent and easily understandable. However, there are important and novel updates in the revised genotyping scheme compiled in this review article. To the best of our knowledge and understanding, the SNP-genotyping scheme compiled in this review article is the most updated one. It has taken out the genotypes that are now merged with other genotypes such as 1C and 3 J, and the genotype 3 K, which is now split into 3K-0 and 3K-1. Thus, out of 16 genotypes initially described, only 13 remain, while six new genotypes have been included in this new updated genotyping scheme (1B-Bangladesh, 1D-Malagassy, 3K-0, 3K-1, 3Q, 4 N/O).

The characteristic genomic markers of all genotypes including the newly described ones have been compiled for the first time in this article. This novelty aspect of the present review article is highlighted by underlining the newly described genotypes in the second column of Figs 3 and 6. Additionally, the genotype-specific markers for all these 19 SNP-genotypes are being presented for the first time through this article (Fig. 5). Also, this updated version of the SNP-genotyping scheme, presented in Fig. 6, has made use of informative markers (an up-to-date list is shown in Fig. 3 as SNP-Grps 1 to 16), as well as genotype-specific markers (Fig. 5) to provide higher resolution between closely related genotypes.

Concluding remarks

The genomic and evolutionary analysis of *M. leprae* and *M. lepromatosis* offers significant insights into the molecular epidemiology and genetic diversity of leprosy bacilli, underscoring the high level of genome conservation despite substantial pseudogenization and downsizing events. *Mycobacterium leprae* strains demon-

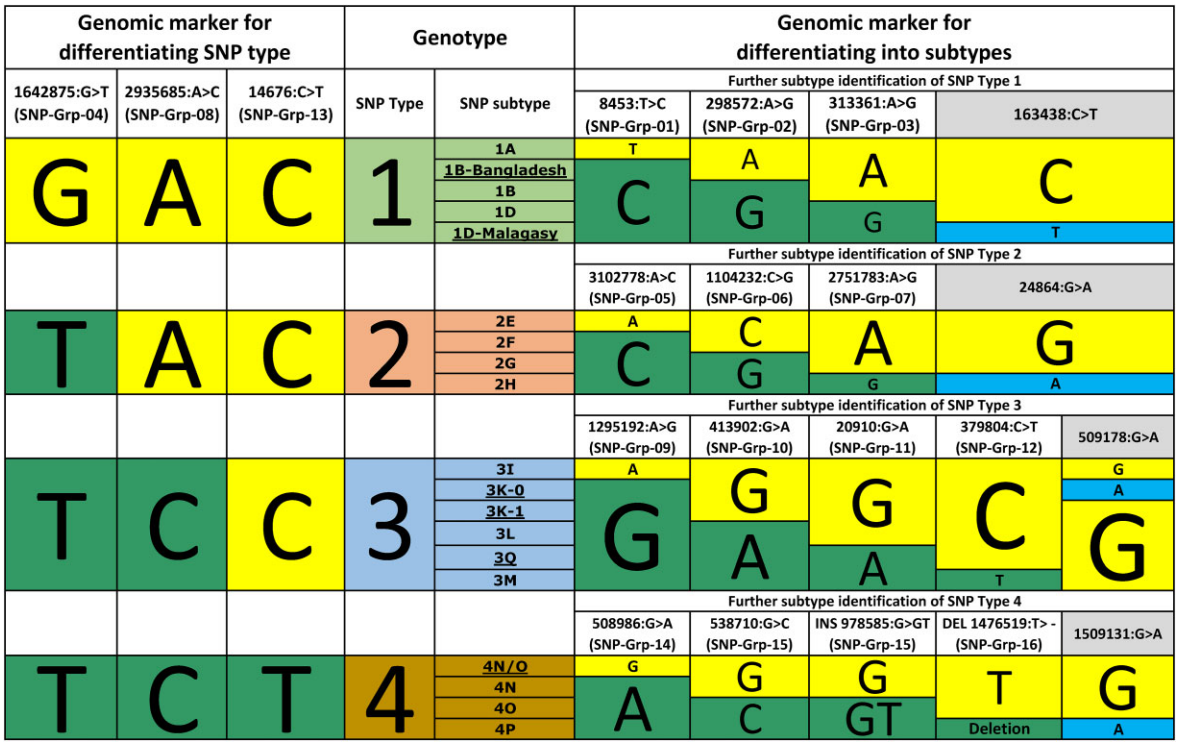


Figure 6. Updated SNP-genotyping scheme of *M. leprae*. Wild-type bases are in yellow while the green indicates the SNPs compared with the *M. leprae* TN reference strain. The blue color indicates the SNPs restricted to only one genotype. The SNP genomic positions are based on the reference genome *M. leprae* TN strain (GenBank ID AL450380.1 version 1, genome size 3268203).

strate minimal variability, limiting the use of conventional PCR-based SNP genotyping and targeted sequencing approaches. The availability of more genomes of *M. leprae* has enabled finer resolution of the *M. leprae* genotyping scheme. The updated genotyping scheme as shown in this article can be useful for further molecular epidemiological applications and thus can help in understanding the transmission dynamics of *M. leprae*. The integration of SNPs, InDels, and VNTRs has significantly advanced the genotyping of *M. leprae*, enabling better definition of bacterial population structure, detailed tracking of strains, and recognition of transmission patterns. While SNPs alone are of limited value in differentiating closely related strains due to their low substitution rate, VNTRs provide greater variability, which is particularly useful for distinguishing closely related strains circulating within a region. Careful selection of VNTR loci, in conjunction with SNP-based approaches, enhances the resolution needed for identifying and tracing specific genotypes across different regions and hosts. This has substantial value from the One Health perspective also, as shown in previous studies where the zoonotic link between human and armadillo leprosy was established (Truman et al. 2011, Sharma et al. 2015).

The updated genotyping scheme, which categorizes *M. leprae* strains into 19 subtypes across four primary SNP-Types, offers a more refined understanding of strain diversity worldwide. Additionally, a defined geographic distribution of specific subtypes reveals characteristic regional genetic patterns, which can be valuable while delineating *M. leprae* transmission. By continuing to refine VNTR selection and exploring new SNPs and InDel markers, researchers can further improve existing genotyping schemes, contributing valuable insights into *M. leprae*'s evolutionary history, disease spread, and drug resistance. Such updated lists of genomic variants can be a useful resource for updating the targeted amplification/capture-based NGS assays such as Deeplex Myc-Lep. Additionally, specific markers for less characterized strains (e.g. 1B, 4N/O) should be studied further, preferably by including additional representative strains belonging to these subtypes. The updated genotyping scheme can allow precise identification of *M. leprae* strains, while the genomics markers associated with MDST can provide clinically relevant information about the choice of effective drugs. Considering all these difficulties while classifying strains after WGS, it is evident that suitable bioinformatics tools/algorithms/browsers should be developed to accurately identify the *M. leprae* genotypes from NGS datasets to minimize the possibilities of wrong classifications manually (Economou et al. 2013a, 2013b). The use of bioinformatics algorithms to accurately and rapidly fetch this information from NGS datasets is possible only when an input database of updated genomic markers is created and made accessible to those interested. As genomic sequencing and bioinformatics tools improve, the use of WGS in refining the global genotyping framework and expanding our understanding of leprosy epidemiology and evolution will also increase. Therefore, future research focusing on additional VNTR loci, SNPs, and InDels will enhance the accuracy and resolution of *M. leprae* genotyping, ultimately contributing to public health efforts in tracking and controlling leprosy.

Acknowledgements

We thank the members of the Publication Screening Committee and other colleagues for their constructive criticism and inputs. The authors are also thankful to colleagues at the VRDL lab (DHR) at ICMR-NIRTH and the clinician and collaborators/colleagues at the MRHRU Badoni, Government Medical College Datia and NSCB Medical College Jabalpur.

Supplementary data

Supplementary data are available at *FEMSRE* online.

Conflict of Interest Statement

The authors declare that they have no conflict of interest.

Funding

Our research group is supported by Indian Council of Medical Research (ICMR), Department of Health Research (DHR), Government of India, R2STOP Canada, Turing Foundation and Leprosy Research Initiative (LRI) Netherlands. The authors are also thankful to previous researchers (cited in this article), who have made their genomic and NGS datasets publicly available through online databases, which were retrieved and mined for identification of genomic markers mentioned in this paper.

References

Avanzi C, Busso P, Benjak A *et al.* Transmission of drug-resistant leprosy in Guinea-Conakry detected using molecular epidemiological approaches. *Clin Infect Dis* 2016;**63**:1482–4.

Avanzi C, del-Pozo J, Benjak A *et al.* Red squirrels in the British Isles are infected with leprosy bacilli. *Science* 2016;**354**:744–7.

Avanzi C, Lécorché E, Rakotomalala FA *et al.* Population genomics of mycobacterium leprae reveals a new genotype in Madagascar and the Comoros. *Front Microbiol* 2020;**11**:711.

Avanzi C, Maia RC, Benjak A *et al.* Emergence of mycobacterium leprae rifampin resistance evaluated by whole-genome sequencing after 48 years of irregular treatment. *Antimicrob Agents Chemother* 2020;**64**:10–1128.

Avanzi C, Singh P, Truman RW *et al.* Molecular epidemiology of leprosy: an update. *Infect Genet Evol* 2020;**86**:104581.

Benjak A, Avanzi C, Singh P *et al.* Phylogenomics and antimicrobial resistance of the leprosy bacillus mycobacterium leprae. *Nat Commun* 2018;**9**:352.

Bennett BH, Parker DL, Robson M. Leprosy: steps along the journey of eradication. *Public Health Rep* 2008;**123**:198–205.

Bhat RM, Prakash C. Leprosy: an overview of pathophysiology. *Interdiscip Perspect Infect Dis* 2012;**2012**:1.

Brosch R, Gordon SV, Eiglmeier K *et al.* Comparative genomics of the leprosy and tubercle bacilli. *Res Microbiol* 2000;**151**:135–42.

Cambau E, Bonnafous P, Perani E *et al.* Molecular detection of rifampin and ofloxacin resistance for patients who experience relapse of multibacillary leprosy. *Clin Infect Dis* 2002;**34**:39–45.

Cambau E, Saunderson P, Matsuoka M *et al.* Antimicrobial resistance in leprosy: results of the first prospective open survey conducted by a WHO surveillance network for the period 2009-15. *Clin Microbiol Infect* 2018;**24**:1305–10.

Cole ST, Brosch R, Parkhill J *et al.* Deciphering the biology of mycobacterium tuberculosis from the complete genome sequence. *Nature* 1998;**393**:537–44.

Cole ST, Eiglmeier K, Parkhill J *et al.* Massive gene decay in the leprosy bacillus. *Nature* 2001;**409**:1007–11.

Cole ST, Honore N, Eiglmeier K. Preliminary analysis of the genome sequence of mycobacterium leprae. *Lepr Rev* 2000;**71** Suppl: S162-164; discussion S164–167.

Cole ST, Supply P, Honoré N. Repetitive sequences in Mycobacterium leprae and their impact on genome plasticity. *Lepr Rev* 2001;**72**:449–61.

Cole ST. Comparative and functional genomics of the mycobacterium tuberculosis complex. *J Microbiology* 2002;**148**:2919–28.

da Silva Ferreira J, de Carvalho FM, Vidal Pessolani MC *et al.* Serological and molecular detection of infection with mycobacterium leprae in Brazilian six banded armadillos (Euphractus sexcinctus). *Comp Immunol Microbiol Infect Dis* 2020;**68**:101397.

da Silva MB, Portela JM, Li W *et al.* Evidence of zoonotic leprosy in Pará, Brazilian Amazon, and risks associated with human contact or consumption of armadillos. *PLoS Negl Trop Dis* 2018;**12**:e0006532.

da Silva Rocha A, Cunha D, Santos AA *et al.* Genotyping of mycobacterium leprae from Brazilian leprosy patients suggests the occurrence of reinfection or of bacterial population shift during disease relapse. *J Med Microbiol* 2011;**60**:1441–6.

Das M, Diana D, Wedderburn A *et al.* Molecular epidemiology and transmission dynamics of leprosy among multicase families and case-contact pairs. *Int J Infect Dis* 2020;**96**:172–9.

Deps P, Collin SM. Mycobacterium lepromatosis as a second agent of Hansen's disease. *Front Microbiol* 2021;**12**:698588.

Drevets DA, Leenen PJ, Greenfield RA. Invasion of the central nervous system by intracellular bacteria. *Clin Microbiol Rev* 2004;**17**:323–47.

Dwivedi P, Sharma M, Ansari A *et al.* Genetic diversity of mycobacterium leprae: need to move towards genome-wide approaches. *Indian J Med Res* 2024;**159**:121–9.

Dwivedi P, Sharma M, Patel P *et al.* Simultaneous detection and differentiation between mycobacterium leprae and mycobacterium lepromatosis using novel polymerase chain reaction primers. *J Dermatol* 2021;**48**:1936–9.

Economou C, Kjellström A, Lidén K *et al.* Ancient-DNA reveals an Asian type of mycobacterium leprae in medieval Scandinavia. *J Archaeolog Sci* 2013a;**40**:465–70.

Economou C, Kjellström A, Lidén K *et al.* Corrigendum to "Ancient-DNA reveals an Asian type of mycobacterium leprae in medieval Scandinavia" [J. Archaeol. Sci. 40 (1) (2013) 465–470]. *J Archaeolog Sci* 2013b;**40**:2867.

Eichelmann K, González González SE, Salas-Alanis JC *et al.* Leprosy. An update: definition, pathogenesis, classification, diagnosis, and treatment. *Actas Dermo-Sifiliográficas. (English Edition)* 2013;**104**:554–63.

Eiglmeier K, Parkhill J, Honore N *et al.* The decaying genome of mycobacterium leprae. *Lepr Rev* 2001a;**72**:387–98.

Eiglmeier K, Simon S, Garnier T *et al.* The integrated genome map of mycobacterium leprae. *Lepr Rev* 2001b;**72**:462–9.

Festskrift C. Daniel Cornelius Danielssen 1815-1894. *Int J Lepr* 1973;**41**:172–5.

Fontes ANB, Lima L, Mota RMS *et al.* Genotyping of mycobacterium leprae for better understanding of leprosy transmission in Fortaleza, Northeastern Brazil. *PLoS Negl Trop Dis* 2017;**11**:e0006117.

Gomes TA, da Silva TP, Machado E *et al.* Genomic and phenotypic variations among Thai-53 and mycobacterium leprae clinical isolates: implications for leprosy pathogenesis and research. *Pathogens* 2024;**13**:986.

Gomez-Valero L, Rocha EP, Latorre A *et al.* Reconstructing the ancestor of Mycobacterium leprae: the dynamics of gene loss and genome reduction. *Genome Res* 2007;**17**:1178–85.

Grzybowski A, Sak J, Pawlikowski J *et al.* Gerhard Henrik Armauer Hansen (1841-1912)–the 100th anniversary of the death of the discoverer of Mycobacterium leprae. *Clin Dermatol* 2013;**31**:653–5.

Haas CJ, Zink A, Pálfi G *et al.* Detection of leprosy in ancient Human skeletal remains by molecular identification of mycobacterium leprae. *Am J Clin Pathol* 2000;**114**:428–36.

Han XY, Aung FM, Choon SE *et al.* Analysis of the leprosy agents mycobacterium leprae and mycobacterium lepromatosis in four countries. *Am J Clin Pathol* 2014;**142**:524–32.

Han XY, Seo YH, Sizer KC *et al.* A new mycobacterium species causing diffuse lepromatous leprosy. *Am J Clin Pathol* 2008;**130**:856–64.

Han XY, Sizer KC, Thompson EJ *et al.* Comparative sequence analysis of mycobacterium leprae and the new leprosy-causing mycobacterium lepromatosis. *J Bacteriol* 2009;**191**:6067–74.

Hegde SP, Shenoy MM, Pinto M *et al.* Leprosy: chronicles of a disabling disease. *Sciences* 2015;**3**:346–9.

Hess S, Rambukkana A. Cell biology of intracellular adaptation of mycobacterium leprae in the peripheral nervous system. *Microbiol Spectr* 2019;**7**. 227–45.

Ho Tinh Tam C, Phuc Nhu Ha N, Hoang Bach N. Genotyping of mycobacterium leprae strains in south central coast and central highlands of Vietnam. *Iran J Microbiol* 2023;**15**:201–7.

Hockings KJ, Mubemba B, Avanzi C *et al.* Leprosy in wild chimpanzees. *Nature* 2021;**598**:652–6.

Jay V. The legacy of Armauer Hansen. *Arch Pathol Lab Med* 2000;**124**:496–7.

Jessen KR, Mirsky R, Lloyd AC. Schwann cells: development and role in nerve repair. *Cold Spring Harb Perspect Biol* 2015;**7**:a020487.

Jouet A, Braet SM, Gaudin C *et al.* Hi-plex deep amplicon sequencing for identification, high-resolution genotyping and multidrug resistance prediction of mycobacterium leprae directly from patient biopsies by using Deeplex Myc-Lep. *EBioMedicine* 2023;**93**:104649.

Kai M., Nakata N., Matsuoka M., Sekizuka T., Kuroda M., Makino M.. Characteristic mutations found in the ML0411 gene of Mycobacterium leprae isolated in Northeast Asian countries. *Infection, Genetics and Evolution* 2013;**19**:200–204.

Kapopoulou A, Lew JM, Cole ST. The MycoBrowser portal: a comprehensive and manually annotated resource for mycobacterial genomes. *Tuberculosis* 2011;**91**:8–13.

Kohler K, Marcsik A, Zadori P *et al.* Possible cases of leprosy from the Late Copper age (3780-3650 cal BC) in Hungary. *PLoS One* 2017;**12**:e0185966.

Kuruwa S, Vissa V, Mistry N. Distribution of mycobacterium leprae strains among cases in a rural and urban population of Maharashtra, India. *J Clin Microbiol* 2012;**50**:1406–11.

Lahiri R, Adams LBJI. Cultivation and viability determination of mycobacterium leprae. *International textbook of leprosy* 2016.

Lavania M, Jadhav R, Turankar RP *et al.* Genotyping of mycobacterium leprae strains from a region of high endemic leprosy prevalence in India. *Infect Genet Evol* 2015;**36**:256–61.

Lavania M, Singh I, Turankar RP *et al.* Molecular detection of multidrug-resistant mycobacterium leprae from Indian leprosy patients. *J Glob Antimicrob Resist* 2018;**12**:214–9.

Levy L, Baohong JL. The mouse foot-pad technique for cultivation of mycobacterium leprae. *Lepr Rev* 2006;**77**:5–24.

Lima LNC, Frota CC, Suffys PN *et al.* Genotyping comparison of mycobacterium leprae isolates by VNTR analysis from nasal samples in a Brazilian endemic region. *Pathogen Glob Health* 2018;**112**:79–85.

Maeda S, Matsuoka M, Nakata N *et al.* Multidrug resistant mycobacterium leprae from patients with leprosy. *Antimicrob Agents Chemother* 2001;**45**:3635–9.

Marijke Braet S, Jouet A, Aubry A *et al.* Investigating drug resistance of Mycobacterium leprae in the Comoros: an observational deep-sequencing study. *Lancet Microbe* 2022;**3**:e693–700.

Marmor MF. The ophthalmic trials of G. H. A. Hansen. *Surv Ophthalmol* 2002;**47**:275–87.

Matsuoka M, Zhang L, Budiawan T *et al.* Genotyping of mycobacterium leprae on the basis of the polymorphism of TTC repeats for analysis of leprosy transmission. *J Clin Microbiol* 2004;**42**:741–5.

Meredith A, Del Pozo J, Smith S *et al.* Leprosy in red squirrels in Scotland. *Vet Rec* 2014;**175**:285–6.

Mietto BS, de Souza BJ, Rosa PS *et al.* Myelin breakdown favours mycobacterium leprae survival in Schwann cells. *Cellular Microbiology* 2020;**22**:e13128.

Monot M, Honore N, Garnier T *et al.* Comparative genomic and phylogeographic analysis of mycobacterium leprae. *Nat Genet* 2009;**41**:1282–9.

Monot M, Honore N, Garnier T *et al.* On the origin of leprosy. *Science* 2005;**308**:1040–2.

Monsalve-Lara J, Drummond M, Romero-Alvarez D *et al.* Prevalence of mycobacterium leprae and mycobacterium lepromatosis in roadkill armadillos in Brazil. *Acta Trop* 2024;**258**:107333.

Ooi WW, Srinivasan JJM. Leprosy and the peripheral nervous system: basic and clinical aspects. *Muscle & Nerve* 2004;**30**:393–409.

Oré N, Tekaia FR, Ti AT *et al.* The decaying genome of mycobacterium leprae. *Lepr Rev* 2001;**72**:387–98.

Pfrengle S, Neukamm J, Guellil M *et al.* Mycobacterium leprae diversity and population dynamics in medieval Europe from novel ancient genomes. *BMC Biol* 2021;**19**:220.

Pin D, Guérin-Faublée V, Garreau V *et al.* Mycobacterium species related to M. leprae and M. lepromatosis from cows with bovine nodular thelitis. *Emerg Infect Dis* 2014;**20**:2111–4.

Rambukkana A. Molecular basis for the peripheral nerve predilection of Mycobacterium leprae. *Curr Opin Microbiol* 2001;**4**:21–27.

Robbins G, Tripathy VM, Misra VN *et al.* Ancient skeletal evidence for leprosy in India (2000 B.C.). *PLoS One* 2009;**4**:e5669.

Roberts CI. The bioarchaeology of leprosy: learning from the past. *International textbook of leprosy* 2018;**6**:1–26.

Santacroce L, Del Prete R, Charitos IA *et al.* Mycobacterium leprae: a historical study on the origins of leprosy and its social stigma. *Infez Med* 2021;**29**:623–32.

Schuenemann VJ, Singh P, Mendum TA *et al.* Genome-wide comparison of medieval and modern Mycobacterium leprae. *Science* 2013;**341**:179–83.

Scollard DM, Adams LB, Gillis TP *et al.* The continuing challenges of leprosy. *Clin Microbiol Rev* 2006;**19**:338–81.

Scollard DM, Truman RW, Ebenezer GJJC. Mechanisms of nerve injury in leprosy. *Clinics in Dermatology* 2015;**33**:46–54.

Scollard DM. The biology of nerve injury in leprosy. *Lepr Rev* 2008;**79**:242–53.

Sharma M, Dwivedi P, Joshi V *et al.* Novel mutations found in Mycobacterium leprae DNA repair gene nth from central India. *J Infect Chemother* 2024;**30**:531–5.

Sharma M, Gupta Y, Dwivedi P *et al.* Mycobacterium lepromatosis MLPM_5000 is a potential heme chaperone protein HemW and mis-annotation of its orthologues in mycobacteria. *Infect Genet Evol* 2021;**94**:105015.

Sharma R, Singh P, Loughry WJ *et al.* Zoonotic leprosy in the southeastern United States. *Emerg Infect Dis* 2015;**21**:2127–34.

Sharma R, Singh P, McCoy RC *et al.* Isolation of mycobacterium lepromatosis and development of molecular diagnostic assays to distinguish mycobacterium leprae and M. lepromatosis. *Clin Infect Dis* 2020;**71**:e262–9.

Sharma R, Singh P, Pena M *et al.* Differential growth of mycobacterium leprae strains (SNP genotypes) in armadillos. *Infect Genet Evol* 2018;**62**:20–26.

Shinde V, Newton H, Sakamuri RM *et al.* VNTR typing of mycobacterium leprae in South Indian leprosy patients. *Lepr Rev* 2009;**80**:290–301.

Silva FJ, Santos-Garcia D, Zheng X *et al.* Construction and analysis of the complete genome sequence of leprosy agent mycobacterium lepromatosis. *Microbiol Spectr* 2022;**10**:e0169221.

Singh P, Benjak A, Schuenemann VJ *et al.* Insight into the evolution and origin of leprosy bacilli from the genome sequence of mycobacterium lepromatosis. *Proc Natl Acad Sci USA* 2015;**112**:4459–64.

Singh P, Cole ST. Mycobacterium leprae: genes, pseudogenes and genetic diversity. *Future Microbiol* 2011;**6**:57–71.

Suzuki K, Nakata N, Bang PD *et al.* High-level expression of pseudogenes in Mycobacterium leprae. *FEMS microbiology letters* 2006;**259**:208–14.

Suzuki K, Saso A, Hoshino K *et al.* Paleopathological evidence and detection of mycobacterium leprae DNA from archaeological skeletal remains of Nabe-kaburi (head-covered with iron pots) burials in Japan. *PLoS One* 2014;**9**:e88356.

Suzuki K, Takigawa W, Tanigawa K *et al.* Detection of mycobacterium leprae DNA from archaeological skeletal remains in Japan using whole genome amplification and polymerase chain reaction. *PLoS One* 2010;**5**:e12422.

Tapinos N, Rambukkana A. Insights into regulation of human Schwann cell proliferation by Erk1/2 via a MEK-independent and p56Lck-dependent pathway from leprosy bacilli. *Proceedings of the National Academy of Sciences* 2005;**102**:9188–93.

Truman RW, Singh P, Sharma R *et al.* Probable zoonotic leprosy in the southern United States. *N Engl J Med* 2011;**364**:1626–33.

Tufariello JoAnn M., Kerantzas Christopher A., VilchÃäze Catherine, *et al.* The Complete Genome Sequence of the Emerging Pathogen Mycobacterium haemophilum Explains Its Unique Culture Requirements. *mBio* 2015;**6**:e01313–01315.

Urban C, Blom AA, Avanzi C *et al.* Ancient mycobacterium leprae genome reveals medieval English red squirrels as animal leprosy host. *Curr Biol* 2024;**34**:2221–30. e2228.

Vedam R, Pansare TA, Narula J. Contributions of ancient Indian knowledge to modern medicine and cardiology. *Indian Heart J* 2021;**73**:531–4.

Vera-Cabrera L, Escalante-Fuentes WG, Gomez-Flores M *et al.* Case of diffuse lepromatous leprosy associated with "mycobacterium lepromatosis." *J Clin Microbiol* 2011;**49**:4366–8.

Virk A, Pritt B, Patel R *et al.* Mycobacterium lepromatosis lepromatous leprosy in US citizen who traveled to disease-endemic areas. *Emerg Infect Dis* 2017;**23**:1864–6.

Vissa VD, Brennan PJ. The genome of mycobacterium leprae: a minimal mycobacterial gene set. *Genome Biol* 2001;**2**: Reviews1023.

WHO. *Global Leprosy (Hansen disease) Update, 2023: Elimination of Leprosy Disease Is Possible—Time to Act!* Vol. WER No 37, 2024, 99, 501–521. 2024. pp.21.

Williams DL, Gillis TP, Booth RJ *et al.* The use of a specific DNA probe and polymerase chain reaction for the detection of mycobacterium leprae. *J Infect Dis* 1990;**162**:193–200.

Williams DL, Gillis TP. Molecular detection of drug resistance in Mycobacterium leprae. *Lepr Rev* 2004;**75**:118–30.

Williams DL, Gillis TPJL. Drug-resistant leprosy: monitoring and current status. *Lepr Rev* 2012;**83**:269–81.

Received 13 November 2024; **revised** 13 February 2025; **accepted** 24 April 2025