Received: 12 February 2018 | Revised: 15 April 2018 | Accepted: 19 April 2018

DOI: 10.1002/oa.2673

RESEARCH ARTICLE

WILEY

The genome sequence of a SNP type 3K strain of *Mycobacterium leprae* isolated from a seventh-century Hungarian case of lepromatous leprosy

T. A. Mendum[1] | G. M. Taylor[1] | H. D. Donoghue[2] | H. Wu[1] | C. Szalontai[3] | A. Marcsik[4] | E. Molnár[4] | G. Pálfi[4] | G. R. Stewart[1]

[1] Department of Microbial Sciences, Faculty of Health and Medical Sciences, University of Surrey, Guildford, UK

[2] Centre for Clinical Microbiology, Division of Infection and Immunity, Royal Free Campus, University College London, London, UK

[3] Salisbury Ltd., Budaörs, Hungary

[4] Department of Biological Anthropology, University of Szeged, Szeged, Hungary

Correspondence
G Michael Taylor, Department of Microbial Sciences, Faculty of Health and Medical Sciences, University of Surrey, Guildford GU2 7XH, UK.
Email: gm.taylor@surrey.ac.uk

Funding information
British Academy, Grant/Award Number: SG150114

Abstract

We report on a *Mycobacterium leprae* genome isolated from the remains of an individual with lepromatous leprosy that were excavated from a seventh-century Hungarian cemetery. We determined that the genome was from a single nucleotide polymorphism (SNP) type 3K0 *M. leprae* strain, a lineage that diverged early from other *M. leprae* lineages. This is one of the earliest 3K0 *M. leprae* genomes to be sequenced to date. A number of novel SNPs as well as SNPs characteristic of the 3K0 lineage were confirmed by conventional polymerase chain reaction and Sanger sequencing. Recovery of accompanying human DNA from the burial was poor, particularly when compared with that of the pathogen. Modern 3K0 *M. leprae* strains have only been isolated from East Asia and the Pacific, and so these findings require new scenarios to describe the origins and routes of dissemination of leprosy during antiquity that have resulted in the modern phylogeographical distribution of *M. leprae*.

KEYWORDS

ancient DNA, Branch 0, *Mycobacterium leprae*, phylogeny, whole genome sequencing

1 | INTRODUCTION

The spread of leprosy around the world in antiquity has previously been inferred by characterizing the infrequent single nucleotide polymorphisms (SNPs) of modern strains of *Mycobacterium leprae*, collected from across the world (Monot et al., 2005). More recently, the genotyping of isolates recovered from archaeological human remains displaying skeletal evidence of lepromatous leprosy (LL) has provided unique opportunities to refine our understanding of the distribution of *M. leprae* strains seen in extant populations, as well as allowing analysis of strains from regions where the disease is no longer present. In this respect, we have previously studied archaeological cases from various parts of Britain and Europe (Donoghue et al., 2015; Inskip et al., 2015; Taylor et al., 2013) as well as modern cases from regions where the disease is still common (Monot et al., 2009).

One of the cases identified previously was burial KD271, a seventh-century male individual whose remains were excavated from Kiskundorozsma (Szeged) in Hungary (Pálfi & Molnár, 2009). Study of a limited number of informative SNPs indicated that the strain of *M. leprae* present in KD271 was a type 3K on the original scheme of 16 genotypes (ranging from 1A–4P) proposed by Monot et al. (2009).

With the application of whole genome sequencing (WGS) to both modern and ancient cases of leprosy (Mendum et al., 2014; Schuenemann et al., 2013), it has become clear that SNP type 3K isolates belong on a separate lineage of the *M. leprae* phylogenetic tree, now designated Branch 0. Although a number of modern examples of this lineage have been studied, all from East Asian locations (Benjak et al., 2018; Schuenemann et al., 2013), no ancient isolates have previously been available for such genomic analysis. The current study presents WGS analysis of the strain from burial KD271 and provides a rare opportunity to study what is one of the oldest cases of *M. leprae* to be sequenced. This is likely to represent a strain close to the most recent common ancestor of extant strains that is predicted to have existed only 3–4,000 years ago (Schuenemann et al., 2013). As

previous biomolecular tests applied to KD271 have appeared in diverse sources, a summary of the conventional genotyping is brought together here, along with some novel observations. These serve both to provide context and as a useful comparator for the WGS study.

2 | MATERIALS AND METHODS

2.1 | Sampling

The subject of the current study, burial KD271, was a mature male individual judged to be between 50 and 60 years of age at the time of death. The burial was excavated from the cemetery of Kiskundorozsma Szeged in Hungary, and the diagnosis of leprosy was made on osteological grounds (Pálfi & Molnár, 2009). For the present genomic study, a 50-mg sample of bone from the palate of KD271 was ground into a fine powder in a sterilised pestle and mortar.

A second burial, Sk12, an adult male, without osteological signs of leprosy, acted as a control for the WGS protocol. Bone was sampled from around the vomer region (Schuenemann et al., 2013).

2.2 | DNA extraction

DNA was extracted using an in-house version of the Boom method (Boom et al., 1990). In this procedure, bone powder (50 mg) was incubated in 1 ml of 1 × Tris-EDTA buffer, containing 40 mAU/ml of proteinase K at 37 °C for 48 hr with occasional mixing. The sample was then centrifuged at 3,000 rpm for 3 min, and the supernatant was transferred into five volumes of 6 M guanidinium thiocyanate (GUSCN, product G9020 from US Biologicals, Salem, MA) containing 1% Triton X-100 and buffered in 1 × Tris-EDTA buffer adjusted to pH 6.5 with 3 M sodium acetate, pH 5.5. Bone powder was mixed with the GUSCN buffer on a mixing wheel for 1 hr at 4 °C. The samples were then subjected to 3 freeze–thaw cycles. Bone powder was removed by centrifugation at 12,000 rpm for 5 min, and the supernatant was transferred to a sterile 1.5-ml Eppendorf tube. Pre-washed silica suspension (40 µl of 0.5–10 µm, Sigma-Aldrich, S5631) was added and kept in contact for 3 hr to maximise recovery of fragmented DNA. After centrifugation, silica was further washed twice with 1 ml aliquots of GUSCN extraction buffer, followed by three washes with 75% ethanol and finally with 1 ml of acetone. After thorough drying of the silica pellet, DNA residues were eluted in 60 µl HPLC grade water at 55 °C. This was then subdivided into 2 × 30 µl aliquots and stored in low retention plastic tubes to minimise loss of DNA through repeated freeze–thawing events.

2.3 | *Mycobacterium leprae* screening

Before undertaking WGS, we screened for evidence of *M. leprae* DNA in the new extract using a polymerase chain reaction (PCR) for the RLEP multi-copy element. This method amplifies a 78 bp amplicon with product monitored with a specific 6-fluorescein amidite (6-FAM)-FAM-labelled hybridisation probe. Details of this and the primer sequences and conditions have been previously reported (Inskip et al., 2015).

2.4 | Variable number tandem repeat and SNP genotyping

The PCR methods for the variable number tandem repeat (VNTR) and individual SNP loci amplification and sequencing have been previously reported (Taylor et al., 2009; Taylor & Donoghue, 2011). The *rpoT* locus (*ML1022*) was PCR amplified and sequenced for the present study using the primers and conditions described in Taylor et al., 2009.

2.5 | *Mycobacterium lepromatosis* screening

The opportunity was taken to screen the new extract for any evidence of the second leprosy agent, *Mycobacterium lepromatosis*. The primers were modifications of those described by Singh et al., 2015, 5-CTGT TCGTGAGGTACCGGTGAAA and 5′-GTTCGGCCGGAGTGTAGGTGT TA. These amplify a 135 bp fragment from the *hem*N gene, present in *M. lepromatosis* but absent in *M. leprae*. The PCR reagents and conditions were as described previously for *M. leprae* specific primers (Inskip et al., 2015), except that an annealing temperature of 56 °C was used.

2.6 | Screening of extract for *Mycobacterium tuberculosis* complex DNA

The KD271 extract was also tested for the presence of *Mycobacterium tuberculosis* (MTB) complex organisms using a real-time PCR method for the IS*1081* repetitive element (Taylor, Murphy, Hopkins, Rutland, & Chistov, 2007).

2.7 | Conventional PCR amplification and sequencing of newly identified SNPs in the KD271 isolate of *Mycobacterium leprae*

Scrutiny of the WGS data from KD271 (Table 3) revealed 16 novel polymorphic SNPs not found in other sequenced strains. Conventional PCR amplification and sequencing was undertaken to confirm the WGS findings for three randomly selected SNPs. Sequences of primers and amplicon sizes are given in Table S1.

2.8 | Human DNA

An attempt was made to confirm the sex of individual KD271 using a PCR method based on a polymorphism in the amelogenin gene (Mannucci, Sullivan, Ivanov, & Gill, 1994). This generates two bands from males of 106 and 112 bp (AMELX and AMELY products, respectively), and a single AMELX product of 106 bp from females.

2.9 | WGS of *Mycobacterium leprae* genome from skeleton KD271

DNA from skeleton KD271 was enriched for *M. leprae* sequences using microarrays and sequenced as described previously for cases Sk2, Sk8, and the control Sk12 from medieval Winchester (Schuenemann et al., 2013) but with only a single round of microarray-based enrichment. The raw sequence files were deposited in the Sequence Read Archive database with Submission ID SAMN08093649. After initial quality controls and alignment to the

M. leprae TN reference genome (Schuenemann et al., 2013), SNPs were identified and accepted if they had a read depth of 3 or greater; more than 75% of reads agreed; an alignment quality score (MQ) ≥ 30; and an absence of reads in the control sample Sk12.

Phylogenies were generated by aligning all SNPs from selected leprosy genomes (Data S1) using both the maximum likelihood and neighbour-joining functions of MEGA7 (Kumar et al., 2016) and the Tamura 3-parameter model (as determined to be the best fit by the model selection function of MEGA7). Equivalent loci in *M. lepromatosis* were determined by aligning *M. leprae* TN, *M. leprae* Br423, and *M. lepromatosis* FJ924 (Han et al., 2015) with Mauve v2.4.0 (Darling, Mau, Blattner, & Perna, 2004).

3 | RESULTS

3.1 | Paleopathological lesions in KD271

The diagnosis of LL in individual KD271 was originally made on palaeopathological grounds. The lesions have been described in detail previously (Donoghue et al., 2005; Molnár, Marcsik, Bereczki, & Donoghue, 2006), so only a summary is given here. The skull shows signs of the rhinomaxillary syndrome, with erosion and widening of the nasal margins and resorption of the anterior nasal spine (Figure 1 and inset). There is loss of bone from around the alveolar region on the maxillary process, and the upper incisor teeth have been lost ante mortem with remodelling of the tooth sockets (Figure 1). There is also pitting on the nasal surface of the palate due to the disease. Pitting due to periostitis is also present on some tarsal bones and on the surfaces of the tibiae and fibulae. Subperiosteal exostoses seen on the fibulae are more evident on the distal third of the bone shafts (Figure 2).

3.1.1 | Burial Sk12

This individual was found to be free of any macroscopic signs of the rhinomaxillary syndrome or other indications of leprosy on the hands, feet, or distal lower limbs (Taylor et al., 2013).

3.2 | Screening of KD271 extract for *Mycobacterium leprae* DNA

The extract prepared from the palatal region of the skull tested positive using the RLEP PCR probe method (Figure 3). The control case was negative for *M. leprae* DNA (not shown).

FIGURE 2 The fibulae from KD271. KD271 fibulae showing subperiosteal exostoses on the distal third of the bone shafts [Colour figure can be viewed at wileyonlinelibrary.com]

3.3 | SNP genotyping and multiple loci VNTR analysis

The results of SNP and VNTR typing are summarised in Table 1. Figure S1 shows conventional Sanger sequencing of six informative SNP loci used in the Monot typing scheme (2009).

3.4 | Screening for *Mycobacterium lepromatosis* DNA

No PCR products were amplified from the extracts using the *hemN* primers specific for *M. lepromatosis*; therefore, there is no evidence that this individual was co-infected with this pathogen. To show that this PCR method can amplify *M. lepromatosis*, an appropriately sized

FIGURE 1 The skull of KD271. Frontal view of skull from the burial (less mandible) showing evidence of rhinomaxillary syndrome, notably loss of several incisors, resorption of the anterior nasal spine, and widening of the piriform aperture. Inset highlights loss of anterior nasal spine [Colour figure can be viewed at wileyonlinelibrary.com]

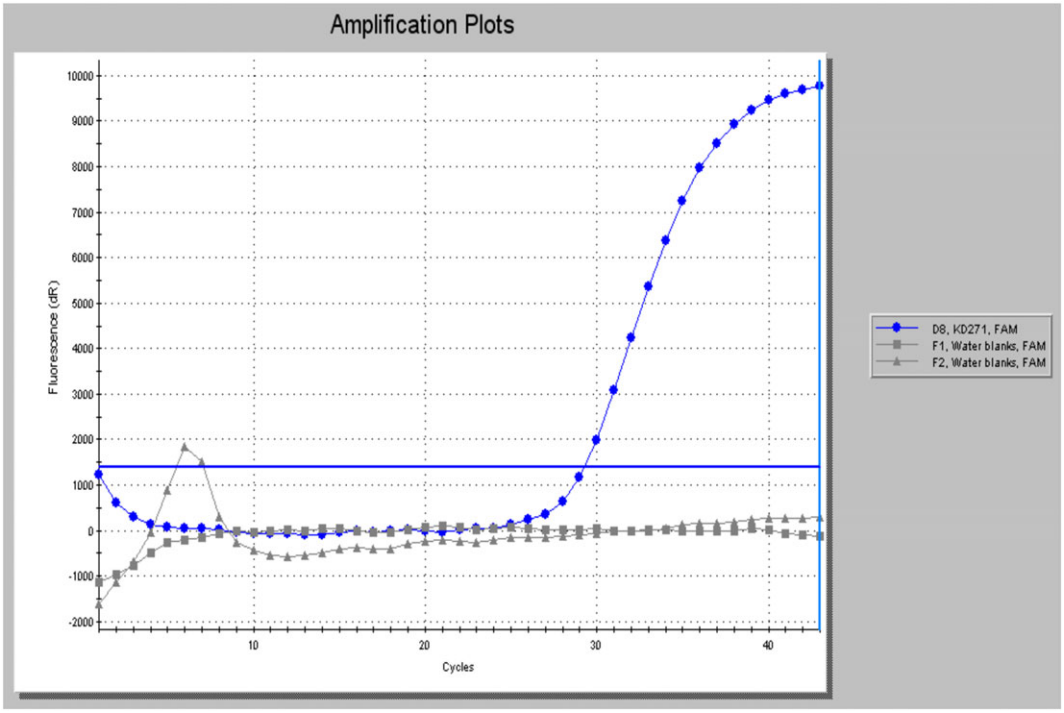

FIGURE 3 Confirmation of *M. leprae* DNA in extracts from KD271. Real-time amplification of RLEP polymerase chain reaction product (78 bp) monitored with a dual-labelled fluorescent hybridisation probe. KD271 is shown in blue and negative controls in grey [Colour figure can be viewed at wileyonlinelibrary.com]

TABLE 1 Comparison of selected informative SNPs and VNTR loci determined by conventional and whole genome sequencing. Regions with insufficient or poor coverage in the WGS are indicated as not determined (nd)

SNP locus[1]	Amplicon size (bp)	Nucleotide base	SNP typing inference	Nucleotide base by WGS
14,676 1,642,879 2,935,693	136 122 107	C T C	Type 3	nd T nd
413,903	120	G	3I–3K	G
591,858	107	C	3I–3L	C
1,133,495	121	G	3J–3M	G
2,312,066	120	G	3K–3M	G
7,614 1,113,926	109 117	C A	Not 3I	nd A
Overall			3K	
VNTR loci		Copies		Copies by WGS
AGA(20) 2,785,364–2,785,494	Variable	16		nd
GTA(9) 2,583,816–2,583,839	Variable	24		nd
21–3, *ML0058* 72,683–73,686	96	2		2
rpoT, (*sigA*) *ML1022*	91	3		3

Note. SNP = single nucleotide polymorphism; VNTR = variable number tandem repeat; WGS = whole genome sequencing.

amplicon (135 bp) was PCR amplified from partially purified *M. lepromatosis* DNA harvested from infected tissue of a Scottish red squirrel (*Sciurus vulgaris*; Figure S2).

3.5 | Screening for MTB complex DNA

The extract prepared from the palate of KD271 was negative for evidence of *M. tuberculosis* complex DNA.

3.6 | Human DNA

No PCR products were obtained using the amelogenin PCR, which probably reflects the extremely fragmented nature of DNA in this skeleton.

3.7 | Genome sequencing of KD271

DNA extracts from skeletons KD217 and Sk12 were enriched for *M. leprae* sequences, PCR amplified and sequenced. Details of the

read depths, their percentage alignment to the *M. leprae* TN reference genome, percentage genome coverage, and average read length are given in Table 2. Samples from the KD271 burial gave an average read depth of 11.94, whereas the Sk12 control sample's average read depth was 0.39% and less than 1% of reads aligned to the *M. leprae* genome. As observed previously (Mendum et al., 2014; Schuenemann et al., 2013), the alignments were punctuated with regions with pan-genus similarity that had large numbers of highly heterogeneous reads that presumably derive from environmental mycobacteria. The KD271 sequences had short read lengths (Sawyer, Krause, Guschanski, Savolainen, & Pääbo, 2012) typical of ancient DNA. These data confirm that there was little or no cross-contamination between samples, or from modern *M. leprae* DNA during the sample preparation.

The WGS sequencing was in broad agreement with Sanger sequencing of targeted SNPs (Table 1), so validating both conventional typing and genome sequencing methods. However, the WGS had insufficient coverage for three of the nine Sanger-determined SNP loci. Analysis of VNTR genotypes in KD271 was limited to the 21–3 locus (*ML0058*) and to the *rpoT* locus (*ML1022*), as insufficient numbers of reads spanned the necessary regions for the other two VNTR loci, GTA(9) and AGA(20). Sixteen novel SNPs were identified that have not been found in other published strains, including the 3K0 strains, S9, S10, CM-1, or Kyoto-1. None of the SNPs in coding regions are likely to have phenotypic effects. To validate the WGS data, three randomly chosen, newly identified SNPs were additionally confirmed by conventional PCR and Sanger sequencing (Table 3).

3.8 | Phylogenetic analysis of KD271

Phylogenetic analyses were consistent between both maximum likelihood and neighbour-joining methods (Figures 4 and Figure S3). Both placed KD271 within the 3K0 lineage, branching from the common lineage before the S10/Kyoto-1 branch but after the S9/CM-1 branch (Avanzi et al., 2015; Honap et al., 2018; Mendum et al., 2014; Schuenemann et al., 2013). The branch lengths of KD271 were shorter than for the modern strains as would be expected for an ancient genome. KD271 was found to exhibit 26 SNPs, 24 of which were unique to the 3K0 group (S9, S10, CM-1, and Kyoto-1), six of these SNPs were uniquely shared with S10 and Kyoto-1. In contrast, no SNPs were uniquely shared with S9 or CM-1.

3.8.1 | *rpoT* locus

The presence of only three copies of GACATC in the *rpoT* tandem repeat distinguishes KD271, as well as S9 and S10, all of which have earlier branch points, from the modern isolates from Japan, Korea, and parts of China that result from an apparently recent radiation of 3K0 strains, and all contain four copies of the hexanucleotide (Kai et al., 2013; Weng et al., 2013).

3.8.2 | *ML0411* locus

Polymorphism was noted in *ML0411*. Locus *ML0411* is a serine rich, 45 kDa antigen (408 aa) recognised by B cells of the immune system. It is a member of the PPE protein gene families of pathogenic mycobacteria, having the characteristic Pro-Pro-Glu motif at the N

TABLE 2 WGS sequencing and alignment statistics for KD271. Sequence reads were quality controlled and aligned to M. leprae TN and human genomes

Total number of reads	Percentage of reads aligning to the *Mycobacterium leprae* TN genome	Percentage of reads aligning to the human genome	Average *M. leprae* read depth	Average length of aligned reads	Percentage genome coverage≥3 reads, MQ score ≥ 30
52,430,302	17.1%	0.9%	11.9	88.6 bp	83.8%

TABLE 3 SNPs Unique to KD271. SNPs confirmed using conventional PCR and Sanger sequencing

Position (TN)	Loci	Base (TN/KD271)	Amino acid (TN/KD271)	Function
18,136	ML0014	C/T	A/A	Probable conserved membrane protein (pseudogene)[1]
19,223	ML0016	C/T	L/L	*pknB*
250,837	ML0185	C/T	T/I	Probable conserved transmembrane protein
2,127,288	Intergenic	C/T		
2,223,039	Intergenic	C/T		
2,235,810	ML1852	C/T	V/M	Probable conserved transmembrane protein (pseudogene)[1]
2,297,562	ML1914	T/A	I/N	*lpqN*
2,564,872	Intergenic	T/C		
2,731,405	Intergenic	G/A		
2,881,087	ML2410	G/A	A/A	Probable conserved transmembrane protein
3,021,965	ML2536	C/T	V/V	ESX conserved component *eccB3*, possible membrane protein
3,108,100	ML2603	G/T	P/P	Possible lysophospholipase[1]
3,158,822	ML2639	G/A	L/L	Probable aldehyde dehydrogenase, (NAD+) dependent
3,256,354	ML2700	G/A	A/T	Probable conserved transmembrane protein
3,261,877	Intergenic	G/A		
3,268,026	Intergenic	G/A		

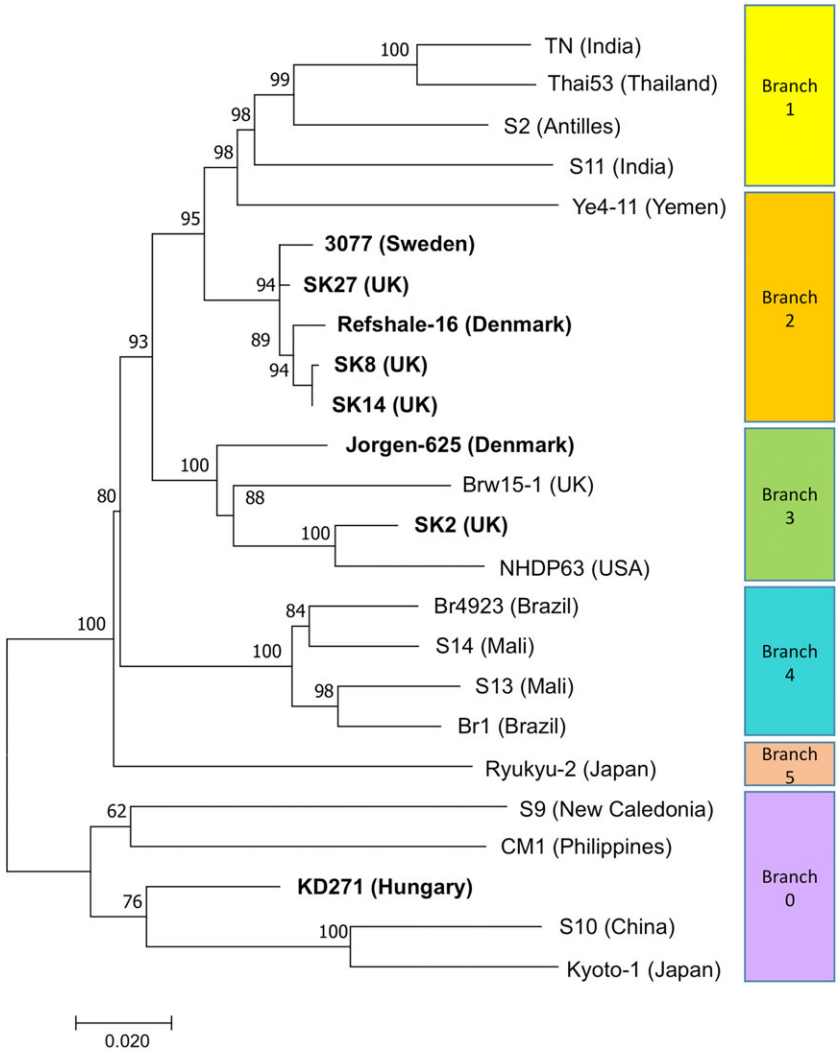

FIGURE 4 Phylogenetic relationships between selected modern (regular text) and ancient (bold text) *Mycobacterium leprae* strains. The phylogeny was inferred by the maximum likelihood method of MEGA7 (Kumar et al., 2016) and the Tamura 3-parameter model. The tree with the highest log likelihood value is shown. Bootstrap percentages from 1,000 replicates are shown next to the branches. The scale indicates the number of substitutions per site. All positions with less than 90% site coverage were eliminated. *M. lepromatosis* was used as an outgroup (not shown). CM1 and Br15-1 are derived from a cynomolgus macaque and a red squirrel, respectively [Colour figure can be viewed at wileyonlinelibrary.com]

TABLE 4 *ML0411*: Nucleotide bases in KD271 compared to the TN reference strain (SNP Type 1) and other 3K0 branch members

ML0411 position (Kai et al., 2013)	TN loci	TN	S9	S10	3K0 strains, for example, Kyoto-1 (Kai et al., 2013)	KD271 base
−41st	508,714	G	G	G	T	No coverage
276th	509,030	C	C	A	A	A
424th	509,178	G	A	A	A	A
571st	509,325	C	G	G	G	G
657th	509,411	C	C	C	C	C
671st	509,425	C	C	T	T	C
1097th	509,851	G	T	T	T	T
1153rd	509,907	G	G	G	G	No coverage

termini of the proteins. These proteins, along with the related PE family, are likely to give rise to antigenic variation and may modulate the immune response. *ML0411* is the single most variable gene in the genome of *M. leprae*. Comparison of the KD271 polymorphic SNPs in *ML0411* (Table 4) with 3K0 strains from North-east Asia (Kai et al., 2013) is compatible with its phylogenetic position as an early member of the 3K0 lineage.

3.8.3 | Subdivision of the Branch 0 lineage

Recent work shows that the type 3K strains of Branch 0 may be further subdivided into 3K0 and 3K1 lineages (shown as branch 5 in Figure 4), depending on specific SNP subsets (Avanzi et al., 2015). These workers report that the 3K0 lineage displays a total of 20 specific SNPs, and the newly described branch 3K1 demonstrates 23 specific SNPs. The 3K common branch is characterised by C at position 711,197 and G at SNPs 563,796 and 57,633. Inspection of the data from KD271 shows the isolate is consistent with 3K0 rather than 3K1 lineage (Table 5).

4 | DISCUSSION

The subject of the current report, KD271, was a mature male individual showing skeletal evidence of LL. The remains were excavated from

TABLE 5 Distinguishing SNPs of the 3KI and 3KO strains. Three SNP loci associated with the 3K1 lineage and the equivalent positions in the 3KO burials KD271, S9, S10 and reference strains TN and Br4923. After Avanzi et al., 2015

Genome	ML0585c Pseudogene *qor* SNP 711,197	ML0466 Hypothetical protein SNP 563,796	ML0046c Pseudogene *espJ* SNP 57,633
3K1 branch			
Ryukyu-2	C	G	G
3K0 branch			
S9	C	C	T
S10	C	C	T
CM-1	C	C	T
KD271	C	C	T
Other strains			
TN	T	G	T
Br 4923	T	G	T

Note. SNP = single nucleotide polymorphism.

the cemetery of Szeged-Kiskundorozsma-Daruhalom dűlő II, Hungary in the summer of 2003. On the basis of the associated grave goods (Mészáros, Paluch, & Szalontai, 2005; Paluch & Szalontai, 2004; Szalontai, 2012), the burial ground was in use in the early/middle Avar transition period, in the late seventh century. The cemetery contained 94 individuals, and skeletal signs of leprosy were noted in eight of these. This presumptive diagnosis was previously confirmed by ancient DNA analysis in three cases, including the subject of the current work, KD271 (Molnár et al., 2006). Subsequently, this case was partially genotyped for a later study of leprosy using conventional PCR and Sanger sequencing targetting phylogenetically informative SNP loci (Monot et al., 2009). This showed the isolate to be a type 3K strain of *M. leprae*.

Since those initial investigations, the understanding of the phylogeography of leprosy has improved considerably, with the recognition of an additional lineage (Branch 0) into which the 3K strains may be placed (Schuenemann et al., 2013) and which branches earlier from the common lineage than the other groups. To extend our understanding of how KD271 and the other leprous remains at Szeged-Kiskundorozsma-Daruhalom dűlő II relate to other ancient and modern *M. leprae* strains, the genome of KD271 was sequenced. This genome is, to date, one of the oldest Branch 0 *M. leprae* genomes to have been studied and so is likely to represent a strain more similar to the notional most recent common ancestor (predicted to have existed ~3,000–4,000 years ago). From recent and ongoing studies on modern Branch 0 isolates, it is evident that there is a greater heterogeneity within this lineage than first realised so that the type 3K strains may be further subdivided into 3K0 and 3K1, depending on specific SNP subsets (Avanzi et al., 2015). The genome sequencing, phylogenetic analysis, and SNP-specific sequencing all confirm that KD271 falls within the currently known monophyletic *M. leprae* radiation and does indeed represent a 3K strain, being placed on an early branch of the 3K0 lineage.

This position is reflected in the SNPs of the hypervariable gene, *ML0411*, which, although not identical to modern strains, was found to be consistent with its place on the 3K0 lineage.

The isolate affecting KD271, along with S9 and S10, the two other strains with early branch points in the 3K0 lineage, possessed only three copies of the *rpoT* hexanucleotide GACATC, rather than the four copies typical of modern strains as described by Avanzi et al. (2015) and Kai et al., 2013. When first reported, this VNTR polymorphism was suggested as one, albeit limited, way of examining differences in the *M. leprae* genome (Matsuoka et al., 2000). The locus may therefore be helpful in defining those strains mentioned above with a more recent radiation. However, the relevance of the *rpoT* locus is difficult to assess without further investigation of strains isolated from various regions of India (Lavania et al., 2007; Lavania et al., 2009) where strains with four copies have been described. This is because the recently described second agent of leprosy, *M. lepromatosis* (Han et al., 2008), also contains four copies of this sequence, as reported for cases in Western and Central Mexico (Han, Sizer, Velarde-Félix, Frias-Castro, & Vargas-Ocampo, 2012). So in studies where the *rpoT* primers recognise both pathogens, and where it was the only typing method used (Lavania et al., 2007; Lavania et al., 2009), further testing will be needed to distinguish between *M. lepromatosis* and the 3K lineage of *M. leprae* with four copies of the *rpoT* hexanucleotide.

In the earlier study by Monot et al. (2009), two further LL cases from another Hungarian burial ground were also found to be positive for *M. leprae*. These were burials 222 and 503, a 45- to 50-year-old male and a 30- to 35-year-old female, respectively. These were both excavated from the 10th- to 11th-century Eastern Hungarian cemetery at Püspökladány-Eperjesvölgy. They were found to be SNP types 3K and 3M, respectively. Interestingly, they were both subsequently shown to be co-infected with *M. tuberculosis* (Donoghue, Marcsik, Molnár, Paluch, & Szalontai, 2005); hence, we checked burial KD271 for any signs of MTB complex DNA but found none. The lack of either MTB complex or *M. lepromatosis* mycobacterial DNA is important for the accurate interpretation of WGS data.

The presence of 3K0 strains in seventh-century Hungarian remains and of a 3K strain from the 10th centrury is consistent with two contrasting scenarios for the origins of geographical distibution of 3K *M. leprae* strains. The global distribution of 3K0 and 3K1 strains is today restricted to regions of the Western Pacific such as Japan (except Okinawa), Korea, China, The Philippines, New Caledonia and Indonesia, among others (Avanzi et al., 2015; Honap et al., 2018; Kai et al., 2013; Monot et al., 2009; Weng et al., 2013). This could indicate that the 3K lineage originated in Northern or Eastern Asia. The

presence of two type 3K cases (KD271 and 222) in early medieval Hungary would then suggest a route of dissemination from Asia to Central Europe, perhaps via trade links or migrations. This would be consistent with what is known of the origins of the Panonian Avars, who are believed to have reached the Hungarian plain from the Eurasian steppe in the late sixth to early ninth centuries (Curta, 2006). The other possibility is that Europe was a centre of dissemination of the ancestral 3K0 and related strains, some of which later became less common or even absent from Europe but persisted in East Asia and the Pacific. Determining the likelihood of each of these scenarios will require more sampling and characterization of both ancient and modern strains.

5 | CONCLUSIONS

Ancient DNA analysis is a powerful approach for understanding past human diseases such as leprosy. In particular, it allows us to obtain strain typing data from geographical locations where the disease may no longer be found, to compare ancient with modern strain distributions and to assist with understanding earlier human migrations. Recognition of greater diversity within 3K lineage strains (Branch 0), and its recognition as the deepest lineage, has come from WGS studies applied to both ancient and modern cases. Burial KD271 represents one of the earliest examples of this archaic lineage to be studied to date.

ACKNOWLEDGEMENTS

We thank Dr. Charlotte Avanzi of The Global Health Institute, École Polytechnique Fédérale de Lausanne, 1015 Lausanne, Switzerland, for the generous gift of *Mycobacterium lepromatosis* genomic DNA, partially purified from red squirrel tissue. We also thank Dr. Tanvi Honap of the Department of Anthropology, University of Oklahoma for making available at short notice her sequence variant data. This study was supported in part by a Small Research Grant, reference SG150114, from the British Academy.

ORCID

G. M. Taylor http://orcid.org/0000-0002-4215-3916

REFERENCES

Avanzi, C., Benjak, A., Kai, M., Busso, P., Singh, P., Matsuoka, M., & Cole, S. T. (2015). *Analysis of* Mycobacterium leprae *strains from Japan: New trends in phylogeography*. Copenhagen, Denmark: Eposter EP009 at 25th ECCMID. 25-28th April

Benjak, A., Avanzi, C., Singh, P., Loiseau, C., Girma, S., Busso, P., … Cole, S. T. (2018). Phylogenomics and antimicrobial resistance of the leprosy bacillus *Mycobacterium leprae*. *Nature Communications*, *9*(1), 352. https://doi.org/10.1038/s41467-017-02576-z

Boom, R., Sol, C. J. A., Salimans, M. M. M., Jansen, C. L., Wertheim-van Dillen, P. M. E., & van der Noordaa, J. (1990). Rapid and simple method for purification of nucleic acids. *Journal of Clinical Microbiology*, *28*, 495–503.

Curta, F. (2006). Southeastern Europe in the Middle Ages, 500–1250. In *Cambridge medieval textbooks*. Cambridge: Cambridge University Press. ISBN 978-0-521-81539-0

Darling, A. C., Mau, B., Blattner, F. R., & Perna, N. T. (2004). Mauve: Multiple alignment of conserved genomic sequence with rearrangements. *Genome Research*, *14*, 1394–1403. https://doi.org/10.1101/gr.2289704

Donoghue, H. D., Marcsik, A., Matheson, C., Vernon, K., Nuorala, E., Molto, J. E., … Spigelman, M. (2005). Co-infection of *Mycobacterium tuberculosis* and *Mycobacterium leprae* in human archaeological samples: A possible explanation for the historical decline of leprosy. *Proceedings of the Royal Society B*, *272*, 389–394. https://doi.org/10.1098/rspb.2004.2

Donoghue H. D., Marcsik A., Molnár E., Paluch T., Szalontai C.. 2005. Lepra nyomai a kiskundorozsmai avar temetőből Előzetes beszámoló. In Hadak útján" – Népességek és iparok a népvándorlás korában, 171-186. [Osteological signs of leprosy from the Avar period series of Kiskundorozsma. Preliminary report. In: In the way of the armies– Population and industries in the Migration Period] pp 171-186.

Donoghue, H. D., Taylor, G. M., Marcsik, A., Molnár, E., Palfi, G., Pap, I., … Spigelman, M. (2015). A migration-driven model for the historical spread of leprosy in medieval Eastern and Central Europe. *Infection, Genetics and Evolution*, *31*, 250–256. https://doi.org/10.1016/j.meegid.2015.02.001

Han, X. Y., Mistry, N. A., Thompson, E. J., Tang, H.-L., Khanna, K., & Zhang, L. (2015). Draft genome sequence of new leprosy agent "*Mycobacterium lepromatosis*". *Genome Announcements*, *3*, e00513–e00515. https://doi.org/10.1128/genomeA.00513-15

Han, X. Y., Seo, Y.-H., Sizer, K. C., Schoberle, T., May, G. S., Spencer, J. S., … Nair, R. G. (2008). A new Mycobacterium species causing diffuse lepromatous leprosy. *American Journal of Clinical Pathology*, *130*, 856–864. https://doi.org/10.1309/AJCPP72FJZZRRVMM

Han, X. Y., Sizer, K. C., Velarde-Félix, J. S., Frias-Castro, L. O., & Vargas-Ocampo, F. (2012). The leprosy agents *Mycobacterium lepromatosis* and *Mycobacterium leprae* in Mexico. *International Journal of Dermatology*, *51*, 952–959. https://doi.org/10.1111/j.1365-4632.2011.05414.x

Honap, T. P., Pfister, L.-A., Housman, G., Mills, S., Tarara, R. P., Suzuki, K., … Stone, A. C. (2018). *Mycobacterium leprae* genomes from naturally infected nonhuman primates. *PLoS Neglected Tropical Diseases*, *12*(1), e0006190. https://doi.org/10.1371/journal.pntd.0006190

Inskip, S. A., Taylor, G. M., Zakrewski, S., Mays, S. A., Pike, A. W., Llewellyn, G., … Stewart, G. R. (2015). Osteological, biomolecular and geochemical analysis of an early Anglo-Saxon case of lepromatous leprosy. *PLoS One*, *13*(10), e0124282. https://doi.org/10.1371/journal.pone.0124282

Kai, M., Nakata, N., Matsuoka, M., Sekizuka, T., Kuroda, M., & Makino, M. (2013). Characteristic mutations found in the *ML0411* gene of *Mycobacterium leprae* isolated in Northeast Asian countries. *Infection, Genetics and Evolution*, *19*, 200–204. https://doi.org/10.1016/j.meegid.2013.07.014

Kumar, S., Stecher, G., & Tamura, K. (2016). MEGA7: Molecular evolutionary genetics analysis version 7.0. *Molecular Biology and Evolution*, *33*, 1870–1874. https://doi.org/10.1093/molbev/msw054. Epub 2016 Mar 22

Lavania, M., Katoch, K., Singh, H., Das, R., Gupta, A. K., Sharma, R., … Katoch, V. M. (2007). Predominance of three copies of tandem repeats in *rpoT* gene of *Mycobacterium leprae* from Northern India. *Infection, Genetics and Evolution*, *7*, 627–631. https://doi.org/10.1016/j.meegid.2007.05.011

Lavania, M., Lal, R., Joseph, G., Darlong, J., Abraham, S., Nanda, N. K., & Jadhav, R. S. (2009). Genotypic analysis of *Mycobacterium leprae* strains from different regions of India on the basis of *rpoT*. *Indian Journal of Leprosy*, *81*, 119–124.

Mannucci, A., Sullivan, K. M., Ivanov, P. L., & Gill, P. (1994). Forensic application of a rapid and quantitative DNA sex test by amplification of the X-Y homologous gene amelogenin. *International Journal of Legal Medicine*, *106*, 190–193. https://doi.org/10.1007/BF01371335

Matsuoka, M., Maeda, S., Kai, M., Nakata, N., Chae, G. T., Gillis, T. P., … Kashiwabara, Y. (2000). *Mycobacterium leprae* typing by genomic diversity and global distribution of genotypes. *International Journal of Leprosy*, *68*, 121–128.

Mendum, T. A., Schuenemann, V. J., Roffey, S., Taylor, G. M., Wu, H., Singh, P., … Stewart, G. R. (2014). Mycobacterium leprae genomes from a British medieval leprosy hospital: Towards understanding an ancient epidemic. *BMC Genomics*, *15*, 270. https://doi.org/10.1186/1471-2164-15-270

Mészáros P., Paluch T., Szalontai C. S. 2005. Avar kori temetők Kiskundorozsma határában. Előzetes beszámoló az M5 autópályán feltárt lelőhelyről. – MKCsM (2004) Szeged, pp. 145–162.

Molnár, E., Marcsik, A., Bereczki, Z., & Donoghue, H. D. (2006). Pathological cases from the 7th century in Hungary. In *16th European meeting of the Paleopathology Association*. Fira, Santorini: Greece.

Monot, M., Honore, N., Garnier, T., Araoz, R., Coppee, J.-Y., Lacroix, C., … Cole, S. T. (2005). On the origin of leprosy. *Science*, *308*, 1040–1042. https://doi.org/10.1038/ng.477

Monot, M., Honoré, N., Garnier, T., Zidane, N., Sherafi, D., Paniz-Mondolfi, A., et al. (2009). Phylogeography of Leprosy. *Nature Genetics*, *41*, 1282–1289. https://doi.org/10.1126/science/1109759

Pálfi, G., & Molnár, E. (2009). The paleopathology of specific infectious diseases from Southeastern Hungary: A brief overview. *Acta Biologica Szegediensis*, *53*, 111–116. (http://www.sci.u-szeged.hu/ABS)

Paluch, T., & Szalontai, C. S. (2004). Kiskundorozsma-Daruhalom dűlő II. In I. Kisfaludy (Ed.), *Régészeti kutatások Magyaroszagon 2003* (pp. 293–329). Hungary, 2003: Archaeological Investigations.

Sawyer, S., Krause, J., Guschanski, K., Savolainen, V., & Pääbo, S. (2012). Temporal patterns of nucleotide misincorporations and DNA fragmentation in ancient DNA. *PLoS One*, *7*(3), e34131. https://doi.org/10.1371/journal.pone.0034131

Schuenemann, V. J., Singh, P., Mendum, T. A., Krause-Kyora, B., Jäger, G., Bos, K. I., … Krause, J. (2013). Genome-wide comparison of medieval and modern *Mycobacterium leprae*. *Science*, *341*, 179–183. https://doi.org/10.1126/science.1238286

Singh, P., Benjak, A., Schuenemann, V. J., Herbig, A., Avanzi, C., Busso, P., … Cole, S. T. (2015). Insight into the evolution and origin of leprosy bacilli from the genome sequence of *Mycobacterium lepromatosis*. *Proceedings of the National Academy of Science U S a.*, *112*, 4459–4464. https://doi.org/10.1073/pnas.1421504112

Szalontai C. 2012. Ismét az avar kori lepráról. — Again, about the leprosy in the Avar period. In Pető Zs. (ed) „Hadak Útján". A Népvándorláskor Fiatal kutatóinak XX. konferenciája. Budapest pp. 149–161.

Taylor, G. M., Blau, S., Mays, S. A., Monot, M., Lee, O. Y.-C., Minnikin, D. E., … Rutland, P. C. (2009). *Mycobacterium leprae* genotype amplified from an archaeological case of lepromatous leprosy in Central Asia. *Journal of Archaeological Science*, *36*, 2408–2414. https://doi.org/10.1016/j.jas.2009.06.026

Taylor, G. M., & Donoghue, H. D. (2011). Variable nucleotide tandem repeat (VNTR) typing of *Mycobacterium leprae* isolates amplified from European archaeological human remains with lepromatous leprosy. *Microbes and Infection*, *13*, 923–929. https://doi.org/10.1016/j.micinf.2011.05.003

Taylor, G. M., Murphy, E., Hopkins, R., Rutland, P. C., & Chistov, Y. (2007). First report of *Mycobacterium bovis* DNA in archaeological human remains. *Microbiology*, *153*, 1243–1249. https://doi.org/10.1099/mic.0.2006/002154-0

Taylor, G. M., Tucker, K., Butler, R., Pike, A. W. G., Lewis, J., Roffey, S., … Stewart, G. R. (2013). Detection and strain typing of ancient *Mycobacterium leprae* from a medieval leprosy hospital. *PLoS One*, *8*, e62406. https://doi.org/10.1371/journal.pone.0062406

Weng, X., Xing, Y., Liu, J., Wang, Y., Ning, Y., Li, M., … Heiden, J. V. (2013). Molecular, ethno-spatial epidemiology of leprosy in China: Novel insights for tracing leprosy in endemic and non-endemic provinces. *Infection, Genetics and Evolution*, *14*, 361–368. https://doi.org/10.1016/j.meegid.2012.12.009

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.

How to cite this article: Mendum TA, Taylor GM, Donoghue HD, et al. The genome sequence of an SNP type 3K strain of *Mycobacterium leprae* isolated from a seventh-century Hungarian case of lepromatous leprosy. *Int J Osteoarchaeol*. 2018;28:439–447. https://doi.org/10.1002/oa.2673