Mycobacterium leprae diversity and population dynamics in medieval Europe from novel ancient genomes

Saskia Pfrengle, Judy Neukamm, Meriam Guellil, Marcel Keller, Martyna Molak, Charlotte Avanzi, Nataliya Y. Berezina, Alexander P. Buzhilova, Dmitry S. Korobov, Stian Suppersberger Hamre, Vitor M. J. Matos, Maria T. Ferreira, Laura González-Garro, Sofia N. Wasterlain, Célia Lopes, Ana Luisa Santos, Nathalie Antunes-Ferreira, Vitória Duarte, Ana Maria Silva, Linda Melo, Natasa Sarkic, Lehti Saag, Kristiina Tambets, Philippe Busso, Stewart T. Cole, Alexei Aivasovich, Charlotte A. Roberts, Alison Sheridan, Craig Cessford, John Robb, Johannes Krause, Christiana L. Scheib, Sarah A. Inskip, and Verena J. Schuenemann

Abstract

Background: Hansen’s disease (leprosy), widespread in medieval Europe, is today mainly prevalent in tropical and subtropical regions with around 200,000 new cases reported annually. Despite its long history and appearance in historical records, its origins and past dissemination patterns are still widely unknown. Applying ancient DNA approaches to its major causative agent, Mycobacterium leprae, can significantly improve our understanding of the disease’s complex history. Previous studies have identified a high genetic continuity of the pathogen over the last 1500 years and the existence of at least four M. leprae lineages in some parts of Europe since the Early Medieval period.

© The Author(s). 2021 Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/. The Creative Commons Public Domain Dedication waiver (http://creativecommons.org/publicdomain/zero/1.0/) applies to the data made available in this article, unless otherwise stated in a credit line to the data.
Background

Hansen’s disease (leprosy), caused by infection with *Mycobacterium leprae* or *Mycobacterium lepromatosis*, is one of the oldest recorded diseases known to human-kind. Its notoriety relates both to its potential to cause extreme physical manifestations of infection, which include damage to the peripheral nervous system, mucosal membranes, skin, and ultimately the extremities [1], and its misattribution to a disease in biblical texts [2, 3]. The earliest probable descriptions of the disease are from Egyptian papyri from 1550 BCE and the Sushruta Samhita (600 BCE) from India [4]. More reliable accounts of Hansen’s disease are found in ancient Greek and Roman literature [5] from the first century CE onwards [6]. This geographically focused information led some historians to suggest that the disease may have originated in Africa [7], although most agree on a likely origin in Asia, possibly in the region of today’s India [8]. It was thought to have travelled west during the conquests of Alexander the Great (fourth century BCE) or through trading and likely then diffused around the Mediterranean basin and into Western Europe with the expansion of the Roman Empire (200 BCE–600 CE) [9].

Geospatial analyses of archeological skeletons of individuals with Hansen’s disease have done little to change this narrative. Until recently, the earliest individuals with evidence for infection have been dated to the 2nd millennium BCE in India [8], fourth to third centuries BCE in Italy [10], third century BCE in Egypt [11], and first century CE in Israel [12, 13]. However, the recent identification of two possible, albeit genetically unconfirmed cases from Bronze Age Hungary (4th millennium BCE) [14] and Early Bronze Age Scotland (late 3rd millennium to early 2nd millennium BCE) [1, 15] throws a simple eastern Asian origin hypothesis into question. In addition, there is evidence from 3rd to 2nd millennium BCE Pakistan [16] as well as possible evidences from Nubia, 2300 BCE [17], Iran, 6200-5700 cal BCE [18], and Turkey dated to 2300 BCE [19], that all need further investigation.

Archeological and historical sources all demonstrate that Hansen’s disease was widespread in Europe by the Middle Ages, being increasingly identified from the Roman and early Medieval periods (200 BCE–600 CE). The period from 1000 to 1300 CE saw the increased foundation of leprosaria across the continent [20]. These charitable institutions, often running under monastic rules, were set up to receive and support individuals who had “leprosy,” which not only included people with Hansen’s disease, but likely encompassed those with other conditions that medieval people also diagnosed as “leprosy” [21, 22]. They also received individuals with other diagnosed diseases, increasingly so in the late medieval period [21, 22]. Despite popular belief, people were not forced to live in these institutions when diagnosed, and could even be expelled, but they often paid for their residence and would receive spiritual and practical support [23]. Modern excavations of leprosarium cemeteries show tens to several hundreds of people buried in them often without skeletal evidence of Hansen’s disease [24–28]. Of the archeological cemeteries not related to a leprosarium where Hansen’s disease has been identified in skeletons, the majority show that infected people were buried in the common manner for their
location and time period [1]. For medieval Europe, this includes, for example, Norwich, England [29]; Kirk Hill, Scotland [30, 31]; multiple sites in Schleswig, Denmark/Germany [32]; Seville [33] and Gijon in Spain [34]; Beja, Portugal [35], and Kaldus, Poland [36]—see also [1] for a global view.

For as yet unclear reasons, the disease prevalence began to decline in Europe from the fourteenth century, although pockets of infection remained until the nineteenth century, e.g., in Scandinavia [1] or even until the twentieth century in Spain [37]. Currently accepted hypotheses for its decline include cross-immunity offered by tuberculosis infection [38], or a loss of susceptible hosts due to the rise of other competing infectious diseases (plague, tuberculosis, etc.), or changes in hygienic practices including the construction of leprosaria [26]. For an overview of potential causes for the decline, see [1], and references within. While autochthonous cases are rarely reported in Europe today, the disease remains a significant social challenge in Brazil, India, and Indonesia with up to 200,000 new cases globally per year [39]. In 2020, the World Health Organization cites 177,175 “registered cases” and 202,185 “new cases” for the end of 2019, of which 71% were reported from South and East Asia [40].

The understanding of *M. leprae*’s evolutionary history benefits from the genetic investigation of archeological human remains (skeletons or preserved bodies such as mummies). Initially, PCR-based analyses identified four major SNP (single-nucleotide polymorphism) types from 1 to 4, which allowed an assessment of the phylogeographic distribution of archeological evidence within a framework of modern strain distribution [41]. Later, the four major SNP types were resolved into 16 subgroups from A to P [42]. At this point of studying modern and ancient *M. leprae*, the distribution of SNP types correlated with geographic location and could largely be explained by major population movements [42, 43]. Prior to the introduction of the branch system and the SNP subtyping, it was argued that while the ancestral SNP type 2 strains originated in Africa, Hansen’s disease spread westwards giving rise to SNP type 3 and eastwards resulting in SNP type 1 [42]. Further, it was suggested that SNP type 4 emerged in West Africa [42]. Due to technical improvements in the field of ancient DNA research, including next-generation sequencing and targeted DNA enrichment methods, the first complete genomes of *M. leprae* from archeological human samples were reconstructed in 2013 [44]. Investigation of the phylogegetic resolution of the ancient genomes initially clustered *M. leprae* genomes into five major branches (0-4) [44]. Later studies including modern and ancient *M. leprae* genomes [45, 46] improved the resolution of branch 2, which was split into branches 2E, 2H, 2F, identified a completely new fifth branch, further named branch 5 as well as correlating SNP subtyping and branching system [45–47]. Both classifications are now used based on the resolution needed for phylogenetic and transmission in ancient and modern *M. leprae* studies [46].

Genome-wide analyses of ancient and modern *M. leprae* DNA strains made it clear that there is no longer a simple correlation between the diverse bacterial genomes and their geographic origins. Hence, questions concerning the origin of leprosy are still unresolved. Ancient DNA studies have uncovered a high genetic diversity of *M. leprae* strains in medieval Europe [44, 45], revealing the possibility that some *M. leprae* strains had spread worldwide from the European continent. For example, it is assumed that Europeans spread branch 3 genomes to the Americas in the sixteenth century and later, through contact via the slave trade [48], where they still exclusively persist [44]. More intriguingly, the ten previously reconstructed ancient genomes demonstrated that nearly all major branches of *M. leprae* seen today, including the most basal (branch 0) currently associated with modern East Asian samples [45, 47], were present in medieval Europe. Furthermore, lineages from three different branches were identified in people buried in the same medieval cemetery at St Jørgen (Odense, Denmark) [45]. Unfortunately, the low number of sequenced genomes from medieval Europe, including from leprosaria, make it difficult to assess how widespread such high local diversity was, or whether it was unique to this particular site. Uncovering such high diversity raises important questions about whether the origin of the disease may in fact be in Eurasia, but the lack of ancient data from outside Europe and the resulting potential sampling bias do not yet allow precise conclusions [45]. Even the recently published oldest genome from an ancient Egyptian mummy from Abusir el-Meleq, thereby representing the only ancient *M. leprae* genome outside Europe, does not further refine the localization of leprosy’s origin [49]. Furthermore, the identification of *M. leprae* in modern red squirrels in Britain [50], genetically closely related to a branch 3 strain isolated in a fifth-sixth century male from the Essex/Cambridgeshire border, eastern England, highlights the possibility that there may be important animal reservoirs that could add a further layer of complexity to the identified diversity [51]. In addition, there are still large parts of Europe for which there is ample historical and archeological evidence for the disease, but for which we have little to no information on the genetic variation of the strains present there.

Here, we address these gaps by examining 41 individuals including 39 with osteological or historical contextual evidence of Hansen’s disease, and two were
exclusively identified as positive for leprosy by genetic examination. These individuals originated from 20 archeological sites across Europe dating from the sixth to the twentieth centuries CE including areas for which previously no genome-wide data existed: Belarus, Iberia, Russia, and Scotland. To further assess intra-regional strain diversity, we investigated skeletons from multiple sites of the same region (Cambridgeshire, England) and the remains of people buried in two leprosaria in Portugal and Spain. We were able to reconstruct 19 *M. leprae* genomes which were also suitable for phylogenetic analysis, doubling the number of published ancient genomes, and to identify a new SNP subtype, named 3Q, in Belarus. The analysis of these new genomes supports the previous findings of high genetic diversity in medieval leprosy strains across Europe and indicates that this was also the case in at least one other medieval leprosarium site, the Hospital of Sant Llàtzer (Barcelona).

**Results**

**Sample information and dating**

Samples from a total of 41 individuals were investigated in this study (Additional file 1: Supplementary Note 1, Table S1 and S2) [1, 21, 30, 31, 35, 50, 52–88]. All but two individuals (IDS097 and BEL024) were previously associated with Hansen’s disease due to either their archeological context or pathological lesions compatible with a diagnosis of Hansen’s disease (Additional file 1: Supplementary Note 1) [1, 21, 30, 31, 35, 50, 52–88], and archeologically dated from the Early Medieval period to the twentieth century (Table 1). Fifteen individuals were radiocarbon dated (Table 1, Additional file 1: Supplementary Note 2, Fig. S1, Table S1) [89–95], seven to the late medieval period (eleventh to fourteenth century), and two to the modern era (fifteenth to twentieth century). The Edix Hill skeleton (Cambridgeshire, eastern England), the earliest investigated here, is archeologically dated to the sixth–seventh centuries, confirmed by radiocarbon dating.

We were able to detect ancient *M. leprae* DNA in libraries from 20 (~48.8%) individuals from ten archeological sites, spanning the period of the sixth–seventh centuries CE to the twentieth century (Additional file 1: Supplementary Note 1, Table S1 and S2) [1, 21, 30, 31, 35, 50, 52–88]; 19 of these yielded sufficient *M. leprae* DNA for genome reconstruction. The individuals positive for *M. leprae* represent a variety of geographical regions located in six modern day European countries (Fig. 1) including two leprosaria: Lagos (southern Portugal) and Barcelona (northeast Spain). From these two leprosaria, we were able to reconstruct one and nine genomes, respectively (Fig. 1). Direct dating of 14 individuals (Table 1) confirmed the archeological age estimations, except for sample PAV’09 I.5 (Portugal) (Additional file 1: Supplementary Note 1) [1, 21, 30, 31, 35, 50, 52–88].

**Genome reconstruction**

For 19 individuals, we were able to reconstruct *M. leprae* genomes with a 1-fold coverage for at least 60% of the genome sequence and a mean coverage of 1.46–110.61× (Fig. 1, Table 1). To confirm the ancient nature of DNA, we examined it for characteristic damage patterns. These are an increased frequency of cytosine to thymine base exchanges at the fragment end, which result from the deamination of cytosine to uracil during the DNA degradation process [98]. In addition, ancient DNA is highly fragmented, resulting in a short fragment length [98]. The DNA fragments used for genome reconstruction have a mean fragment length of 51–86 bp (Additional file 1: Table S2) [99–106]. The frequency of C to T base misincorporation of all non-UDG libraries results in 2–20% (Additional file 1: Supplementary Note 3, Fig. S2) [106, 107]. These genomes were sufficiently covered for a reconstruction of a Maximum Parsimony and Maximum Likelihood tree and for SNP typing, and SNP annotation (Table 1, Fig. 2A, Additional file 1: Supplementary Note 3, Fig. S3A, S3B, S4A, S4B, Table S3, Additional file 2: Table S4), [42, 45–47, 108–111]. Sixteen of these 19 reconstructed genomes have a minimum of 3-fold coverage for at least 60% of the genome sites. These 16 high-coverage genomes were used further for molecular dating by BEAST (Fig. 2B, Fig. 3, Additional file 1: Fig. S5, S6, S7) [111–115].

**Phylogenetic analysis**

We combined the 19 new genomes with 177 published modern and ancient *M. leprae* genomes [44, 45, 47, 49, 50, 96, 97, 108, 116–121] to investigate the genetic affinities of the newly reconstructed genomes. All genomes from our study are placed in one of the previously defined eight branches of the *M. leprae* phylogeny [45]. Four Iberian genomes are placed in the most basal branch, branch 0 (Table 1, Fig. 2A, B, Additional file 1: Fig. S3A and S3B) [109–111]. This includes UF25, UF703, and UF803 from the leprosarium in Barcelona (Spain) and COR XVIII (Cordiñanes, León, northwest Spain). They cluster with two medieval *M. leprae* genomes from Hungary (SK11) and Denmark (Jorgen507) presented in a previous study [45]. None of the newly reconstructed genomes is located in branch 5, but the medieval Belarusian genome BEL024 is placed in branch 4. This genome, as well as the ancient genomes Abu Sinai1630 (Egypt) [49] and Body-188 (Czech Republic) [45], are diverged basally to most of the modern branch 4 genomes (Table 1, Fig. 2A, B, Additional file 1: Supplementary Note 3, Fig. S3A and S3B) [109–111], except for S15.
In total, 12 of our 19 (63.2%) genomes position on branch 3 (Table 1, Fig. 2A, B, Additional file 1: Fig. S3A and S3B) [109–111]: the non-leprosarium genomes from Bergen (Norway), CHRY023, CHRY044, EDI066, and JDS097 (all eastern England), and Kirk Hill (Scotland); leprosarium samples UF700 and UF101 (Barcelona, northeast Spain); and three additional leprosarium samples from the same site (UF8, UF11, and UF21), and PAVd’09_I.5 (southern Portugal).

The Bergen sample forms a sister branch to the previously published Jorgen625 (Denmark) strain [44]. Two *M. leprae* strains from Barcelona, UF101 and UF700, are closely related (SNP distance $d = 23$) and branch off basally to strains isolated from modern British red squirrels (Table 1, Fig. 2A, B, Additional file 1: Supplementary Note 3, Fig. S3A, S3B, S4A and S4B) [50, 109–111]. The samples from Kirk Hill (Scotland), and CHRY023 and JDS097 (both eastern England) are phylogenetically placed close to the previously published SK2 genome [44] from Winchester, southern England. Together with PAVd’09_I.5 (Portugal), and the three Barcelona leprosarium samples, UF8, UF11, and UF21 (fifteenth-eighteenth centuries), they form a separate cluster consecutively diverging from the evolutionary line leading to the American branch 3 cluster.

The samples R7546-671 (Russia) and UF800 (Barcelona, Spain) are placed in branch 2F (Table 1, Fig. 2A, Additional file 1: Fig. S3A and S3B) [109–111]. The Russian genome R7546-671 takes a basal position to the two modern Ethiopian genomes placed in branch 2F. The genome UF800 is located basal to the previously published ancient genomes SK8 and SK14 (Winchester, southern England) [44, 96]. The branches 2H, 2E, and 1 are so far defined by modern *M. leprae* genomes exclusively (Fig. 2A, B, Additional file 1: Fig. S3A and S3B) [109–111].

In summary, 12 of our ancient *M. leprae* genomes are placed in branch 3, four in branch 0, two in branch 2F, and one in branch 4. The two genomes located in branch 2F are an ancient one from the leprosarium in Barcelona, clustering with medieval European *M. leprae* sequences, and the historic sample from Russia, falling

---

**Table 1** Overview of all newly sequenced *Mycobacterium leprae* genomes. The age of the samples is either given in archeological dates (italic) or radiocarbon dates. The age of all directly dated samples is provided in calibrated CE. The listed SNP types are determined according to the new SNP typing system [46]. Following this new system, a new SNP was determined for the BEL024 sample (labeled with an asterisk), but according to the SNP typing system by Monot and colleagues [42] the sample would be classified as 3L.

| Sample       | Sample age (14C dates: non-italics; archeological ages: italics) | Location                              | Mean coverage | Coverage $\geq 1x$ in % | Coverage $\geq 3x$ in % | Genotype (new) | Branch |
|--------------|------------------------------------------------------------------|---------------------------------------|---------------|-------------------------|-------------------------|----------------|--------|
| R7546-671    | 19th–20th century CE                                            | St. Petersburg, Russia                | 16.51         | 97.16                   | 94.96                   | 2F             | 2F     |
| UF11         | 18th century CE                                                 | Sant Llàtzer, Barcelona, Spain       | 6.71          | 85.67                   | 61.80                   | 3I-1           | 3      |
| UF8          | 16th century CE                                                 | Sant Llàtzer, Barcelona, Spain       | 1.46          | 67.52                   | 18.99                   | 3I-1           | 3      |
| UF21         | 1431–1611 cal CE                                                | Sant Llàtzer, Barcelona, Spain       | 4.11          | 92.14                   | 67.70                   | 3I-1           | 3      |
| UF25         | 1423–1466 cal CE                                                | Sant Llàtzer, Barcelona, Spain       | 33.09         | 97.40                   | 95.73                   | 3I-1           | 3      |
| JDS097       | 1231–1384 cal CE                                                | Hospital of St. John, Cambridge, UK  | 12.81         | 96.89                   | 94.27                   | 3I-1           | 3      |
| PAVd09_I.5   | 1283–1396 cal CE                                                | Valle da Gafaria, Lagos, Portugal    | 96.82         | 97.40                   | 97.44                   | 3I-1           | 3      |
| Bergen       | 1268–1388 cal CE                                                | Nonneseter, Bergen, Norway           | 110.61        | 97.45                   | 97.44                   | 3I-1           | 3      |
| UF700        | 1035–1165 cal CE                                                | Sant Llàtzer, Barcelona, Spain       | 19.45         | 97.53                   | 96.91                   | 3I-1           | 3      |
| UF101        | 1027–1157 cal CE                                                | Sant Llàtzer, Barcelona, Spain       | 21.28         | 13.70                   | 97.39                   | 3I-1           | 3      |
| UF800        | 12th–early 13th century CE                                      | Sant Llàtzer, Barcelona, Spain       | 3.34          | 86.27                   | 52.63                   | 2F             | 2F     |
| COR_XVIII    | 12th–early 13th century CE                                      | Cordiñanes de Valdeón, León, Spain  | 2.49          | 67.28                   | 32.39                   | 3K-0           | 0      |
| UF703        | 1040–1208 cal CE                                                | Sant Llàtzer, Barcelona, Spain       | 26.94         | 97.44                   | 96.19                   | 3K-0           | 0      |
| KirkHill     | 1030–1155 cal CE                                                | Kirk Hill, St Andrews, Scotland      | 6.86          | 94.85                   | 81.01                   | 3I-1           | 3      |
| UF803        | 1023–1157 cal CE                                                | Sant Llàtzer, Barcelona, Spain       | 6.18          | 91.01                   | 69.77                   | 3K-0           | 0      |
| CHRY044      | 1034–1175 cal CE                                                | Cherry Hinton, Cambridgeshire, UK    | 18.09         | 96.31                   | 92.35                   | 3I-1           | 3      |
| BEL024       | 1035–1203 cal CE                                                | Byhau, Magileu, Belarus             | 43.86         | 97.71                   | 97.51                   | 3Q (New*)      | 4      |
| CHRY023      | 1034–1162 cal CE                                                | Cherry Hinton, Cambridgeshire, UK    | 7.01          | 96.53                   | 89.75                   | 3I-1           | 3      |
| EDI006       | 575–650 cal CE                                                  | Edix Hill, Cambridgeshire, UK        | 23.71         | 97.64                   | 97.43                   | 3I-1           | 3      |
basal to modern Ethiopian *M. leprae* genomes. Most of the newly reconstructed medieval *M. leprae* genomes are placed in branch 3 (Table 1) and cluster with medieval European genomes from previous studies [44, 45] all located basal to modern *M. leprae* genome clades (Fig. 2A, B). Most intriguing, the phylogenetic tree uncovered a high genetic relatedness of two medieval genomes from Barcelona (UF101 and UF 700) and the modern *M. leprae* genomes isolated from red squirrels [50] (Fig. 2A, B). One of the genomes reconstructed (from the individual BEL024) falls in branch 4, basal to most of the other branch 4 genomes (Fig. 2A, B). Lastly, four of our Iberian *M. leprae* genomes are placed in branch 0 clustering with two previously published medieval genomes [45]. This cluster is basal to modern human leprosy sequences and forms a sister clade to the genomes reconstructed from modern non-human primates.

**Genotyping and SNP effect analysis of the new strains**

We also performed a more detailed analysis of the SNPs identified in our genomes including genotyping and SNP effect analysis to increase the resolution of our analysis. For the genotyping of all 19 newly reconstructed genomes, we used the method developed by Monot and colleagues [42] (Fig. 2A, Additional file 1: Table S3) [42, 45–47, 108] to allow comparability with previously published data. Briefly, there are 84 informative markers (78 SNPs and six indels in homopolymeric tracts) used for classification into the 16 SNP subtypes of *M. leprae* [42]: 1A-D, 2E-H, 3I-M, and 4 N-P. For a more straightforward application, the SNP types (SNP type 1–4) and the SNP subtypes (A–P) can be determined using a combination of three and 16 loci, respectively [42]. Deeper resolution in SNP subtyping was recently published in the SNP subtypes 3I (3I-1, 3I-2) and 3K, and the corresponding specific markers were also applied in our analysis [46]. All newly sequenced ancient genomes from branch 2F, 3, and 0 belong to the SNP subtype 2F, 3I-1, and 3 K-0 respectively (Fig. 2A). The genome (BEL024) from medieval Belarus was identified as genotype 3L according to the Monot classification [41, 42], but, phylogenetically it forms a lineage separate from the canonical 3L strains and diverge basally to the SNP subtype 3M (strain Body-188 from the Czech Republic). We propose to label this lineage as 3Q (Fig. 2A).
Using SnpEff [122], a range of 49 to 167 SNPs with potential effects were identified in the newly reconstructed genomes presented here (Additional file 2: Table S4) [44, 45, 47, 49, 50, 96, 97, 117–121, 123, 124]. Additionally, the samples have between 28 and 72 non-synonymous SNPs in coding regions. Eighteen samples have unique SNPs located within coding genes: BEL024 (n = 4), Bergen (n = 2), CHRY023 (n = 3), CHRY044 (n = 6), EDI006 (n = 1), Kirk Hill (n = 5), PAV’d09.15 (n = 2), COR_XVIII (n = 5), R7546-671 (n = 4), UF101 (n = 2), UF21 (n = 2), UF25 (n = 6), UF700 (n = 5), UF703 (n = 2), UF800 (n = 7), UF803 (n = 1), and UF8 (n = 3); details of these unique SNPs can be found in Additional file 2: Table S4 [44, 45, 47, 49, 50, 96, 97, 117–121, 123, 124]. Six SNPs were located within coding genes that are related to virulence factors, affecting amino acid and purine metabolism (leucine synthesis), mammalian cell entry (mce) operons, secretion system, and cell surface components (Additional file 2: Table S4, Additional file 1: Table S6) [44, 45, 47, 49, 50, 96, 97, 108, 116, 118–124]. In parallel, only a few SNPs were found in modern strains compared to ancient strains from branches 3 and 4, for which ancient strains are basal in the branch (exception of Br2016-45 and the red squirrel strains for branch 3). Only one SNP (t954663c, ml0805, pseudogene) was acquired by all modern strains (Fig. 2A, from Br2016-46 to NHDP-55). A total of nine SNPs were acquired by the modern strains from branch 4 but only one led to a non-synonymous mutation in a gene coding for a phospho-
N-acetylmuramoyl-pentapeptide transferase (a1079902g, \textit{murX}). Indel analysis was not performed due to limited genome quality.

Finally, we analyzed the specific SNPs to branch 3—the branch in which most of our genomes fall—in comparison to the other branches. We found 16 SNPs including 50\% of missense mutations (Additional file 2: Table S4) \[44, 45, 47, 49, 50, 96, 97, 117–121, 123, 124\]. Within branch 3 we also detected a close genetic relationship between modern \textit{M. leprae} strains isolated from red squirrels [50] and medieval \textit{M. leprae} genomes reconstructed from two medieval Barcelona individuals (UF101 and UF700). Therefore, we investigated further into details of SNP differences and similarities of these \textit{M. leprae} genomes. The modern \textit{M. leprae} genomes from red squirrels differ from each other by zero to three nucleotides (0–0.11\%). The two medieval \textit{M. leprae} genomes from the individuals UF101 and UF700 from Barcelona differ in 23 nucleotides (0.82\%). When comparing either the \textit{M. leprae} genome from UF101 or UF700 to the \textit{M. leprae} genomes from red squirrels, they are different at 41–45 positions (1.47 to 1.61\%). Furthermore, the two medieval \textit{M. leprae} genomes isolated human individuals and the five modern \textit{M. leprae} strains from red squirrels share three specific SNPs: a synonymous SNP at position 1348383, and two non-synonymous SNPs at the positions 2271752 and 2495453. The T-to-G SNP at position 2271752 is coding for “rpoC” gene and calling glycine instead of valine, the G-to-T SNP at position is coding for the gene “mntH” and here the amino acid serine is called instead of alanine.

\textbf{Estimation of divergence times (BEAST analysis)}

Bayesian time-aware phylogenetic reconstruction (Additional file 1: Supplementary Note 3) \[111–115\] was performed using a relaxed molecular clock model, because the strict clock was rejected based on the coefficient of variation distribution among the branch rates [125]. The inferred phylogeny (> 0.98 posterior support for all but a few nodes; Fig. 2B, Additional file 1: Fig. S5) \[111–115\] supports the topology of the Maximum Parsimony tree (Fig. 2A, Additional file 1: Fig. S3A, S4A) \[109–111\]. The most recent common ancestor of all the sequences included in the analysis was estimated to ca. 1900 BCE (2910–1110 BCE 95\% HPD, see Fig. 2B, Additional file 1: Supplementary Note 3, Fig. S5) \[111–115\].

Bayesian skyline estimation of demographic changes through time shows two sudden increases in \textit{M. leprae}'s
effective population size ($N_e$) starting around 250 CE ($\pm$ 250 years) and 1600 CE ($\pm$ ca. 100 years), which potentially tie in with major changes in human connectivity (Fig. 3). In the first instance, this could include Roman expansion, and the second coincides with the arrival of Europeans in the Americas.

**Discussion**

In this study, we reconstructed 19 *M. leprae* genomes from ancient individuals from across Europe (Table 1), doubling the number of available ancient genomes and providing insights into the distribution of *M. leprae* lineages in understudied regions. Using 11 of these genomes, we can evaluate *M. leprae*’s diversity within the Iberian Peninsula. The remaining eight genomes can further inform our understanding of the distribution of the pathogen within Eastern Europe, the British Isles, and Scandinavia. We now have an improved understanding of *M. leprae*’s diversity across medieval Europe; we show that the long-term predominance of branch 3 genomes seen in North-West Europe applies also to South-West Europe and that there are strains from three major branches (0, 2F, and 3) that circulate within single locations, in this case medieval leprosaria (Fig. 2A, Additional file 1: Fig. S3A, S3B, S4A and S4B). We also discovered another branch 4 strain, defined as SNP subtype 3Q, in Eastern Europe, recently most frequent in Africa [47], allowing for a better resolution of this branch’s prevalence in that region (Fig. 2A, Additional file 1: Fig. S3A, S3B, S4A and S4B).

We were interested in understanding whether other leprosarium sites showed similar levels of strain diversity as St Jørgen (Odense, Denmark), where three different strains were found in individuals from a similar time frame [45]. Therefore, we attempted to add data from two other leprosaria and were successful in obtaining multiple genomes from individuals buried at one of them, a leprosarium in Barcelona. As demonstrated by the inferred phylogeny (maximum parsimony and maximum likelihood; Fig. 2A, B, Additional file 1: Fig. S3A, S3B) [109–111], the genomes from the Barcelona leprosarium are placed in three branches: branch 3, branch 2F, and branch 0, similarly to the genomes reconstructed from the St Jørgen cemetery. Based on radiocarbon dating, at least three Barcelona strains date to the same time period (eleventh–thirteenth centuries) and thus show that the high strain diversity identified at St Jørgen in Denmark [45] is not unique. Strains from two different branches (2 and 3) were also identified at the Winchester leprosarium [44]. This suggests that by the thirteenth century, the presence of multiple, phylogenetically distant strains in leprosaria was a common phenomenon.

As not all individuals with Hansen’s disease were buried in leprosarium contexts, and to assess circulating strains in one location over a long period of time, we also investigated the region of Cambridge. Here, we have the highest number of genomes from one geographical area from non-leprosarium contexts; two early Anglo-Saxon individuals that date prior to the widespread foundation of leprosaria in England (from the fifth–seventh centuries: Great Chesterford; Edix Hill, ED1006), and two from the tenth–twelfth century (Cherry Hinton, CHRY023, CHRY044) and one of thirteenth century date (Hospital of St John, JDS097) when leprosaria are known to exist in this region of England [1]. Interestingly, all belong to branch 3 and thus show low diversity (Fig. 2A, Additional file 1: Fig. S3A, S3B, S4A and S4B), even though branch 2 strains were present in England by at least the thirteenth century [44] and possibly as early as the eleventh century, based on SNP typing [126].

Although it is necessary to study a greater number of individuals from non-leprosarium contexts from the same location, the high strain diversity found in the leprosarium at Barcelona raises interesting questions about the nature of leprosarium sites and whether individuals entering some of these institutions originated from diverse locations. It is thought that most leprosaria were founded to serve local people but some possibly also admitted the “wandering leprous,” meaning individuals with no accommodation alternatives [23, 127], which could have included pilgrims [128]. Furthermore, high-status leprosaria may have attracted individuals from more diverse backgrounds, who perhaps had more opportunities to travel during their lifetime (e.g., pilgrimage, military encounters, trade), before and while they resided in those specialized hospitals. The surviving documentation from the leprosarium at Barcelona is particularly revealing; although initially founded for people with “leprosy” in towns, the account books from the end of the fourteenth century show that many non-local people who had travelled or undertook pilgrimage, were present there [67]. In addition, there are also a number of accounts and tales of people undertaking a pilgrimage specifically in the hope of obtaining a cure for “leprosy” [2].

Questions surrounding the importance of pilgrimage in the spread of the disease have already been raised through the identification of the “Pilgrim Burial” at Winchester and the mobility of people with the disease might have been underestimated in the past [128]. Barcelona was likely important for those travelling to Santiago de Compostela, a major pilgrimage destination in medieval Christian Europe [128]. Further work in this area might reveal the important role of different types of mobility in the spread of the disease, as has been
attempted for other past infections [129, 130]. To investigate the potential role of leprosaria as “diversity pools,” the questions of whether there is low diversity in surrounding populations from Odense and Barcelona and if there is high diversity in leprosarium sites in Cambridge’s region of East Anglia in eastern England need to be assessed. Assessing evidence in Norwich, an important East Anglian port town around 100 km northeast of Cambridge, where there is evidence for Hansen’s disease in skeletons from a cemetery possibly associated with a leprosarium [131] might be revealing [132]. There are also other (non-leprosarium) sites in East Anglia that have revealed skeletons with Hansen’s disease [1]. It may also be beneficial to assess individuals from a newly excavated leprosarium, St Leonard’s in Peterborough, approximately 65 km northwest of Cambridge. The use of stable isotope analysis to explore the origin and mobility histories of those buried there at a population level in a leprosarium context would also be beneficial. Further work on historical sources from leprosaria and other documentation might also be able to reveal more about the distances and routes that people travelled.

While we identified strains from multiple branches, branch 3 genomes were most common, a trend also seen in our previous work [45]. Although future research needs to assess more individuals with Hansen’s disease from regions where we have limited or no genomes, it is becoming increasingly evident that branch 3 strains were both widespread and predominant in medieval Europe, especially in Western regions. For England and East Anglia in particular, we can confirm that this branch was likely dominant throughout the disease’s known 900-year history (approximately 500 CE–1400 CE) as suggested by Inskip and colleagues [133]. If we consider the presence of branch 3 genomes at the leprosarium in Barcelona, we can also observe a long history of branch 3 strains from around 1100 CE to the eighteenth century. Combined, the Cambridge and Barcelona data show that branch 3 strains have been circulating in Western Europe for over 1200 years, with little genetic variation; they differ in only around 1100 positions of the circa 3.2 Mbp M. leprae genome (Additional file 2: Table S4) [44, 45, 47, 49, 50, 96, 97, 117–121, 123, 124]. Modern branch 3 strains closely related to them can still be found in the USA and South America [44, 45, 47] as well as in two animal reservoirs, further reflecting the complex history of this branch as well as the previously observed slow-evolving nature of the pathogen [44, 45, 134]. Overall, the 17 medieval and two early modern European genomes of branch 3, including 12 of our new strains, provide a so far unique resolution for the history and past dissemination of this branch. This research has refined our information about the relevance of link to population dynamics and the spread of M. leprae (Fig. 2A, B, Additional file 1: Fig. S3A, S3B and S4A, S4B).

For Eastern Europe, the identification of a novel SNP subtype, here named 3Q, in an individual from Belarus (9th to 11th CE) which sits in branch 4 basal to the Body188 genome from the Czech Republic [45], might reflect a long-term presence of strains from this branch in the region. Currently, branch 4 is composed of six SNP subtypes: 3L [49], 3M [44, 45], 4N/O [135], 4N, 4O, and 4P [47]. While the three SNP subtypes 4N, 4O, and 4P are well described in modern samples, they have yet to be identified in ancient samples. Conversely, SNP subtypes 3L and 3M strains are very rare in modern M. leprae samples, being mainly identified on islands (such as Martinique or New Caledonia [42]) and mostly described from ancient remains from Eastern Europe and North Africa [42–45, 49]. In comparison to strains of other branches circulating in medieval Europe, the SNP subtypes 3M and 3L seem to be less successful in surviving in the modern world. This discovery of a new SNP subtype in ancient remains points to a potential loss of diversity over time. For example, it is possible that modern hygienic practices or cross-immunity/competition with other diseases such as tuberculosis have had an impact on diversity [1, 26, 38]. This is also reflected in the poor resolution of the basal structures of branch 4 in the phylogenetic tree. However, the apparent loss of diversity might also be an artifact of our limited knowledge of modern diversity. Overall, it highlights the importance of future investigations in medieval Eastern Europe and North Africa in order to trace back the evolutionary history of branch 4 and capture its past and present diversity.

Interestingly, we observe a close phylogenetic relationship between the nineteenth and twentieth century strain from Russia to those in modern Ethiopia [47] (Fig. 2A, B, Additional file 1: Fig. S3A, S3B, S4A and S4B). A potential explanation for this relationship may come from a series of historic events, pointing to direct contacts between Russia and Ethiopia at the end of the nineteenth century and the beginning of the twentieth century: In the late nineteenth century, Russian settlers arrived in Ethiopia in order to establish “New Moscow” in the region of modern Djibouti [136, 137]. Although the Russian settlers were forced to leave by the French army in 1889, Russian-Ethiopian relations continued and resulted in Russian support of the Ethiopians in the Ethiopian-Italian war [137]. After the victory of Adwa in 1896, which secured the independence of Ethiopia, formal political relations between Russia and Ethiopia started. These historic interactions may have caused the exchange of pathogens between the regions, including the M. leprae strain presented here. Historic reports confirm the prevalence of Hansen’s disease in Russia in
the late nineteenth to early twentieth century [138, 139] (Additional file 1: Supplementary Note 1) [1, 21, 30, 31, 35, 50, 52–88], further supporting potential transmissions.

In concordance with previous studies focusing on Northwest Europe [44, 45, 96], our results confirm that a high M. leprae diversity is also present in other parts of medieval Europe. Despite our new information on strain diversity and increased resolution regarding the reconstruction of the history of Hansen’s disease, we still cannot resolve its origin. The lack of ancient samples from potential source areas and older time periods prevents us from favoring one of the two models proposed by Schuenemann and colleagues [45]. Even the genome Abusir1630 from Ancient Egypt [49], so far the only ancient genome from outside of Europe, is located basal to branch 4 and therefore cannot contribute to discussions of the origin of the other branches present in medieval Europe. In comparison, regions with a significant prevalence of Hansen’s disease today have different but consistently lower levels of strain diversity than discovered for medieval European leprosaria as well as the strain diversity estimated for the entire medieval European continent. However, we observe different levels of modern strain diversity in distinct geographic regions: higher levels are present in some endemic countries such as Brazil (branches 1, 3, and 4) and India (branches 1 and 2) while lower levels exist on small islands such as Madagascar (only branch 1), as well as in Ethiopia (branch 2), or West African countries (branch 4) [44, 45, 47, 50, 119]. This modern diversity variation may be linked to past migrations resulting in new introductions of strains, potentially European ones.

While it is always difficult to relate specific events to increased diversity due to the wide time ranges associated with archeological dating and their use in molecular dating approaches, we see two significant episodes of M. leprae population expansion in the Bayesian skyline plot (Fig. 3) that coincide with important shifts in human connectivity. The first date range 250 CE ± 250 years broadly covers that of Roman conquest and expansion, while the later date range, 1600 CE ± 150 years, is consistent with rapid advances in knowledge and technology in the late Medieval period which culminated with the arrival of Europeans in the Americas and the beginning of regular long distance (transatlantic) trade. Historical and archeological data have already highlighted a link between the increasing expansion of the Roman Empire and the spread of Hansen’s disease [1, 43, 140] and our results support this hypothesis. Coincidently, it is from this period that we start to see some of the earliest evidence in the osteological record, especially in Europe [1]. For the Late Medieval period, genetic data show the link between European M. leprae branch 3 strains and those currently circulating in the Americas [44, 45, 134, 141]. Modern armadillos harbor the same branch 3 strains as those identified in ancient European samples, showing a direct link between the two regions [108]. In both situations, increased global connectivity may have introduced new strains to regions that already had an endemic strain, but the introduced strains outcompeted them or found new hosts. This is a situation common today with rapid and frequent global travel being a key part of the global economy [142].

While individuals in populations that have never experienced a particular infectious disease may be more vulnerable to new incoming diseases [143], radical changes and disruptions in a social organization often associated with colonization events, like those explored here, are also key in increasing indigenous populations’ vulnerability. Important factors can include malnutrition, conflict, breakdown of social networks, forced labor/slavery, or similar conditions of stress [144, 145]. As Hansen’s disease outcomes are strongly dependent on the immune response of potential hosts [146], its arrival into immunologically compromised populations, perhaps in poor communities, may be significant here and these factors may have provided greater opportunities for the bacterium to spread and multiply, explaining the increase in effective population size.

Bayesian molecular clock inference provided us with estimates of the age of the most recent common ancestors and divergence of branches of M. leprae [114]. Our evolutionary timescale estimates are concordant with some previously published [45, 47], and a discrepancy with the latest published estimate [49] calls for cautious interpretations and further investigation in the future. However, the 95% credibility intervals for the age of the most recent common ancestor estimates do overlap (so they do not differ significantly). Nonetheless, with the higher resolution time-aware phylogeny, we can now refine the potential estimates for the time of the interspecies transmission between red squirrels and humans, and contribute to “One Health” that explores links between humans, other animals, and the environment. In the future, this multidisciplinary and multimethod-driven approach will be key to investigating leprosy’s evolutionary history to understand past and recent spread and transmission of Hansen’s disease. This important approach combines evidence from archeology and modern genetics and can help us understand the importance and relevance of red squirrels in spreading the pathogen in the past, and what this means for the present. Work by Avanzi and colleagues [50] and others [45, 47] showed that the closest sequenced genome to that in modern squirrels was retrieved from a fifth to sixth century male individual (SK2) from Great Chesterford, Essex (eastern England and close to London). However, with our data we identified additional ancient
Conclusion
In conclusion, our new ancient genomes from Europe, including an eastern European one with a new SNP type from Belarus, support previous observations of high diversity of *M. leprae* in the past by finding a similar phylogeography in Europe. In addition, we observe a high diversity at a leprosarium, indicating that leprosaria received infected individuals with diverse strains and from various geographic backgrounds. New estimates on the past population diversity of this pathogen further allow insights into its global history in relation to major historic events. Although we were able to refine our understanding of the interspecies transmission of *M. leprae* between red squirrels and humans, this important One Health aspect cannot be resolved by studying only ancient human strains. With the inclusion of ancient animal samples, the picture of *M. leprae*’s evolutionary history enables new avenues of research for approaching this aspect in future studies.

Methods

Material: sample information
In our study, we analyzed 81 samples from 41 individuals from the early medieval to the early modern period (Additional file 1: Table S1) [1, 21, 30, 31, 35, 50, 52–88], to address questions about intra- and interregional genetic diversity of ancient *M. leprae*. Predominantly, we included samples from regions where no genome-wide data were as yet available. As part of early screening work, we also assessed one Scottish Bronze Age sample (Additional file 1: Supplementary Note 1) [1, 21, 30, 31, 35, 50, 52–88]. Unfortunately, it had a very low DNA preservation, but because of the destructive nature of the process, and the temporal importance of the sample, ethically speaking it is important to mention this negative result. It is therefore fully excluded from all work presented here. However, it should be noted that *M. tuberculosis* complex DNA was identified by GM Taylor, but could not be replicated [15]. All other individuals studied here, except one (JDS097), were previously associated with Hansen’s disease due to their archeological context or due to pathological lesions compatible with a diagnosis of Hansen’s disease (Additional file 1: Supplementary Note 1) [1, 21, 30, 31, 35, 50, 52–88]. We investigated bone and/or tooth samples of these 41 individuals to confirm their archeological association with leprosy at the genetic level through the reconstruction of ancient *M. leprae* genomes.

DNA extraction
DNA extraction and pre-amplification steps of all leprosy samples analyzed in this study were undertaken either in the cleanroom facilities at the University of Tübingen, the University of Zurich, the University of Cambridge, or the University of Tartu. Post-amplification steps were performed in separate DNA laboratories at the University of Tübingen, the University of Zurich, University of Cambridge, the Max Planck Institute for the Science of Human History (MPI-SHH) in Jena, and the University of Tartu. All laboratories fulfill the requirements for ancient DNA research [149, 150]. To minimize the risk of potential contamination, all samples were specially pre-treated before DNA extraction (Additional file 1: Supplementary Note 3) [151]. For DNA extractions, we applied a well-established guanidine-silica-based extraction protocol developed for ancient DNA work [151] and used 30–120 mg of bone powder. For the DNA-extraction step, positive and negative controls were produced; positive controls to determine whether the DNA was successful or not, negative controls to identify potential contamination. The negative controls were carried along with all laboratory experiments and were also sequenced, the positive control till the first step of library preparation.
**Library preparation**

In this study, we produced double-stranded non-UDG and UDG-treated, as well as single-stranded DNA libraries (see Additional file 1: Supplementary Note 3) [98, 152–155]. The double-stranded DNA libraries produced in Tübingen and Zürich were used for a screening capture, an in-solution capture approach, used to enrich the DNA for three specific leprosy genes and direct sequencing (Additional file 1: Supplementary Note 3) [156, 157]. The samples processed in Tübingen were enriched additionally for human mitochondrial DNA [44, 156, 157] (Additional file 1: Supplementary Note 3) [156, 157]. The samples processed in Tübingen and Zürich were used for a screening capture, an in-solution capture approach, used to enrich the DNA for three specific leprosy genes and direct sequencing (Additional file 1: Supplementary Note 3) [156, 157]. The samples processed in Tübingen were enriched additionally for human mitochondrial DNA [44, 156, 157] (Additional file 1: Supplementary Note 3) [156, 157]. The samples processed in Tübingen were enriched additionally for human mitochondrial DNA [44, 156, 157] (Additional file 1: Supplementary Note 3) [156, 157]. The samples processed in Tübingen were enriched additionally for human mitochondrial DNA [44, 156, 157] (Additional file 1: Supplementary Note 3) [156, 157]. The samples processed in Tübingen were enriched additionally for human mitochondrial DNA [44, 156, 157] (Additional file 1: Supplementary Note 3) [156, 157]. The samples processed in Tübingen were enriched additionally for human mitochondrial DNA [44, 156, 157] (Additional file 1: Supplementary Note 3) [156, 157]. The samples processed in Tübingen were enriched additionally for human mitochondrial DNA [44, 156, 157] (Additional file 1: Supplementary Note 3) [156, 157]. The samples processed in Tübingen were enriched additionally for human mitochondrial DNA [44, 156, 157] (Additional file 1: Supplementary Note 3) [156, 157]. The samples processed in Tübingen were enriched additionally for human mitochondrial DNA [44, 156, 157] (Additional file 1: Supplementary Note 3) [156, 157]. The samples processed in Tübingen were enriched additionally for human mitochondrial DNA [44, 156, 157] (Additional file 1: Supplementary Note 3) [156, 157]. The samples processed in Tübingen were enriched additionally for human mitochondrial DNA [44, 156, 157] (Additional file 1: Supplementary Note 3) [156, 157]. The samples processed in Tübingen were enriched additionally for human mitochondrial DNA [44, 156, 157] (Additional file 1: Supplementary Note 3) [156, 157]. The samples processed in Tübingen were enriched additionally for human mitochondrial DNA [44, 156, 157] (Additional file 1: Supplementary Note 3) [156, 157]. The samples processed in Tübingen were enriched additionally for human mitochondrial DNA [44, 156, 157] (Additional file 1: Supplementary Note 3) [156, 157]. The samples processed in Tübingen were enriched additionally for human mitochondrial DNA [44, 156, 157] (Additional file 1: Supplementary Note 3) [156, 157]. The samples processed in Tübingen were enriched additionally for human mitochondrial DNA [44, 156, 157] (Additional file 1: Supplementary Note 3) [156, 157]. The samples processed in Tübingen were enriched additionally for human mitochondrial DNA [44, 156, 157] (Additional file 1: Supplementary Note 3) [156, 157]. The samples processed in Tübingen were enriched additionally for human mitochondrial DNA [44, 156, 157] (Additional file 1: Supplementary Note 3) [156, 157]. The samples processed in Tübingen were enriched additionally for human mitochondrial DNA [44, 156, 157] (Additional file 1: Supplementary Note 3) [156, 157]. The samples processed in Tübingen were enriched additionally for human mitochondrial DNA [44, 156, 157] (Additional file 1: Supplementary Note 3) [156, 157]. The samples processed in Tübingen were enriched additionally for human mitochondrial DNA [44, 156, 157] (Additional file 1: Supplementary Note 3) [156, 157]. The samples processed in Tübingen were enriched additionally for human mitochondrial DNA [44, 156, 157] (Additional file 1: Supplementary Note 3) [156, 157]. The samples processed in Tübingen were enriched additionally for human mitochondrial DNA [44, 156, 157] (Additional file 1: Supplementary Note 3) [156, 157]. The samples processed in Tübingen were enriched additionally for human mitochondrial DNA [44, 156, 157] (Additional file 1: Supplementary Note 3) [156, 157]. The samples processed in Tübingen were enriched additionally for human mitochondrial DNA [44, 156, 157] (Additional file 1: Supplementary Note 3) [156, 157]. The samples processed in Tübingen were enriched additionally for human mitochondrial DNA [44, 156, 157].

**DNA sequencing**

DNA was sequenced either at the MPI-SHH Jena, the Functional Genomics Center at the University of Zurich (FGCZ), or the Institute of Genomics Core Facility at the University of Tartu (UTIG).

**Downstream analysis of sequencing data**

For a detailed description of the downstream analysis, see Additional file 1: Supplementary Note 3 [44, 49, 96, 97, 99–108, 116–118, 120, 121, 160]. Briefly, the sequenced DNA was screened for positive M. leprae reads using the EAGER pipeline [99]. For those samples containing sufficient authentic DNA reads mapping against the M. leprae genome, the complete ancient genome was reconstructed using software integrated into the EAGER pipeline [99] (Additional file 1: Supplementary Note 3) [100–107]. Sequencing reads of previously published samples included in our study were processed identically (Supplementary Note 3) [44, 45, 47, 49, 50, 96, 97, 100–108, 116–121, 160]. In addition mitochondrial haplogroups were determined for those libraries included in mitochondrial DNA capture and the data of the directly sequenced libraries were used for molecular sex determination (Additional file 1: Supplementary Note 3, Table S1) [99, 161–166].

**SNP typing, SNP alignment, and SNP effective analysis**

We performed SNP typing [41, 42], SNP alignment [123], and SNP effect analysis [122] for those samples in which we were able to reconstruct an ancient genome with a 1-fold coverage of at least 60% of the genome. These genomes were also included in the reconstruction of a Maximum Parsimony tree (Fig. 2A, Additional file 1: Fig. S3A, S4A and Supplementary Note 3) [109–111] and Maximum Likelihood tree (Additional file 1: Fig. S3B and S4B) [109–111]. For SNP typing, the genomes were first aligned to the reference genome Mycobacterium leprae TN (NC_002677.1). In a second step, branch-specific SNPs were detected by comparing all genomes, and third step, specific SNPs for ancient genomes were determined by comparing the modern and ancient genomes of particular branches.

**Estimation of divergence time (BEAST analysis)**

Only the high-quality samples with a minimum 3-fold coverage of at least 60% of the genomic sites were used for Bayesian time-aware phylogeny and past population dynamics inference with BEAST [114] (Additional file 1: Supplementary Note 3, Fig. S5, S6 and S7) [111–115]. The resulting Skyline plot shows changes in the effective population size ($N_e$) through time (Fig. 3). Although $N_e$ is not an estimate of the actual census size of the population, it should well reflect the relative changes in the number of transmitted bacteria.

**Supplementary Information**

The online version contains supplementary material available at https://doi.org/10.1186/s12915-021-01120-2.

**Acknowledgements**

We thank the Cambridgeshire County Council including Quinton Carroll, Kasia Gadaniec, and Ben Donnelly-Symes for access to the Cherry Hinton and Edix Hill material. We thank Corinne Duhig for information on the Edix Hill skeletal remains, and Adrián Maldonado for comments on the manuscript. We thank Cambridge Archaeological Unit for access to the skeletons from the Hospital of St John the Evangelist, Cambridge. We also thank the Barcelona History Museum (Museu d’Història de Barcelona, MUHBA) for providing us with access to the skeletal remains from Barcelona.

**Authors’ contributions**

V.J.S., S.A.I., C.L.S., S.P., and J.K. conceived and designed the study. S.A.I., A.K., A.A., N.M., R.I.T, N.Y.B., A.P.B., D.S.K., S.S.H, V.J.M.M, M.T.F., N.A.F., A.L.S, A.M.S., V.D., L.M., L.G.G., S.N.W., C.L., N.S., S.T.C., C.A.R., A.S., and C.C. provided samples and archeological context. V.J.S., S.A.I., and C.L.S. supervised the work. S.P., M.G., G.J.N., E.R., R.I.T, P.B., L.S, K.T, and C.L.S. performed the experimental work. J.N., M.G., M.K., M.M., C.A., and S.P. analyzed the sequenced data. V.J.S., C.L.S., J.K., JR, and S.T.C. provided funding. S.P., S.A.I., V.J.S., and J.N. wrote the manuscript with input from all authors. All authors read and approved the final manuscript.
Funding
This work is supported by the University of Zurich’s University Research Priority Program “Evolution in Action: From Genomes to Ecosystems” (V.J.S., J.N.), the Senckenberg Centre for Human Evolution and Palaeoenvironment (S-HEP) at the University of Tübingen (V.J.S., J.K.), the Max Planck Society (I.K.), the European Union through the European Regional Development Fund [Project No. 2014-2020/04.1.06-0030] (C.L.S., M.G., M.K.), the Estonian Research Council personal research grants PUT1339 (A.K.), PIRG243 (C.L.S., L.S.), and PRG1027 (A.K., L.S., K.T.), by the Wellcome Trust [Award no. 200036/Z/15/Z] and St John’s College, Cambridge (C.C., J.E.R., S.A.I., C.L.S.), National Science Centre in Poland [Project No. 2018/31/B/HS3/01464] (M.M.), the European Union’s Horizon 2020 Research and Innovation Program under the Marie Skłodowska-Curie Grant No. 845479 (CA), and the Fondation Raoul Follereau (S.T.C.). A.M.S., A.L.S., L.M., V.D., S.N.W., and V.M.J.M. were carried out at the R&D Unit Research Centre for Anthropology and Health, University of Coimbra (financed by Fundação para a Ciência e Tecnologia - FCT/MCTES, project reference UIDB/00281/2020). Additionally, V.M.J.M. was supported by FCT/MCTES research project IF/001862/2014 and LM was financed by the FCT/MCTES doctoral grant SFRH/BD/130165/2017. M.T.F.’s work was carried out at the R&D Unit Center for Functional Ecology - Science for People and the Planet (CEF), with reference UIDB/04004/2020, financed by FCT/MCTES through national funds (PIDDAC). C.L. was financed by FCT/MCTES, under the project with the reference SFRH/BPD/117128/2016.

Availability of data and materials
Raw sequencing data are available on NCBI (BioProject ID: PRJNA721828).

Declarations
Ethics approval and consent to participate
All human samples are more than 70 years old and anonymous, and therefore they do not require ethical approval for the genetic analysis under current Swiss law (https://www.admin.ch/chpal Zot/consultation/old/0061313/index.html). In addition, for the samples processed in Cambridge (UK), all are over 100 years of age and therefore are not subject to the Human Tissue Act (2004). The samples were provided under agreements on scientific collaboration with relevant people linked to the following skeletal collections, who all fully support this study and are listed in the acknowledgments: Edix Hill (Barrington A.), Cambridgeshire; Church End, Cherry Hinton, Cambridgeshire; The Hospital of St John the Evangelist, Cambridge; The Nonnieseter site, Bergen, Norway; All Saints Church, Kirk Hill, Scotland; Valle da Gafaria, Lagos, Portugal leprosarium; Cordiñanes de Valdeón, León, Spain; Studenka necropolis, Mahileu region, Belarus; The Hospital of Sant Llàtzer, Barcelona, Spain; Blokhuzien, The Netherlands; Santarém, Portugal; Beja leprosarium, Beja, Ermida de Santo André; D.G. Rokhlin Collection, Moscow, Russia. We also followed guidance for the ethical treatment of human remains following BABA0 (https://www.baba0.org.uk/assets/Uploads/BABA0-Code-of-Ethics-2019.pdf).

Consent for publication
Not applicable.

Competing interests
The authors declare no competing interests.

Author details
1Institute of Evolutionary Medicine, University of Zurich, Winterthurerstrasse 190, 8057 Zurich, Switzerland. 2Institute for Archaeological Sciences, University of Tübingen, Rümelinstrasse 19-23, 72070 Tübingen, Germany. 3Institute for Bioinformatics and Medical Informatics, University of Tübingen, Sand 14, 72076 Tübingen, Germany. 4Estonian Biocentre, Institute of Biology, Kazan Federal University, Kazan, Russian Federation. 5Centre of New Biological Sciences, University of Coimbra, Research Centre for Anthropology and Health, Calçada Marim de Freitas, 3000-456 Coimbra, Portugal. 6Laboratory of Forensic Anthropology, Department of Life Sciences, University of Coimbra, Centre for Functional Ecology, Calçada Marim de Freitas, 3000-456 Coimbra, Portugal. 7Area of Antropofisiica, Departamento de Biodiversidade y Gestión Ambiental, Universidad de León, Campus de Vegazana, 24071 León, Spain. 8Laboratory of Biological Anthropology, Department of Biology, School of Science and Technology, University of Evora, Evora, Portugal. 9Laboratório de Ciências Forenses e Psicopatológicas E. G. Moniz (LCFPEM), Centro de Investigação Interdisciplinar E. G. Moniz (CIEM). 10Department of Archaeology, History, Cultural studies and religion, University of Bergen, 5020 Bergen, Norway. 11Department of Life Sciences, University of Coimbra, Research Centre for Anthropology and Health, Calçada Marim de Freitas, 3000-456 Coimbra, Portugal. 12The Institute of Archaeology of the Russian Academy of Sciences, 117292, Dn. Uljanova str. 19, Moscow, Russian Federation. 13Department of Archaeology, History, Cultural studies and religion, University of Bergen, 5020 Bergen, Norway. 14Department of Life Sciences, University of Coimbra, Research Centre for Anthropology and Health, Calçada Marim de Freitas, 3000-456 Coimbra, Portugal. 15Laboratory of Forensic Anthropology, Department of Life Sciences, University of Coimbra, Centre for Functional Ecology, Calçada Marim de Freitas, 3000-456 Coimbra, Portugal. 16Área de Antropologia Física, Departamento de Biodiversidade y Gestión Ambiental, Universidad de León, Campus de Vegazana, 24071 León, Spain. 17Institute of Biomedicine (BIOMED), Universidad de León, Campus de Vegazana, 24071 León, Spain. 18Laboratory of Biological Anthropology, Department of Biology, School of Science and Technology, University of Evora, Evora, Portugal. 19European Union through the European Regional Development Fund – Evropo21, Lisboa, Portugal. 20Laboratory of Biological Anthropology and Human Osteology (LABOH), CRIA/FCSH, Universidade NOVA de Lisboa, Lisboa, Portugal. 21UNIARQ – University of Lisbon, Lisbon, Portugal. 22OSTEO Research, Camino de la Iglesia 1, Barrio de ma, Santistube De Pedraza, 40171 Segovia, Spain. 23Global Health Institute, Ecole Polytechnique Fédérale de Lausanne, Lausanne, Switzerland. 24Institut Pasteur, 25-28, rue du Docteur Roux, 75724 Paris Cedex 15, France. 25Department of Archeology, History of Belarus and Special Historical Disciplines, Mogilev State A. Kuleshov University, St Kosmonavtov 1, Mogilev 212022, Republic of Belarus. 26Department of Archaeology, Durham University, South Road, Durham DH1 3 LE, UK. 27Department of Scottish History and Archaeology, National Museums Scotland, Chambers Street, Edinburgh EH1 1JF, UK. 28Department of Archaeology, University of Cambridge, Downing Street, Cambridge CB2 3R, UK. 29Senckenberg Centre for Human Evolution and Palaeoenvironments, University of Tübingen, Rümelinstrasse 19-23, 72070 Tübingen, Germany. 30St John’s College, University of Cambridge, Cambridge CB2 1TP, UK. 31School of Archaeology and Ancient History, University of Leicester, Leicester LE1 7RH, UK.

Received: 28 May 2021 Accepted: 7 August 2021
Published online: 05 October 2021

References
1. Roberts CA. Leprosy: Past and Present: University of Florida Press; 2020. https://doi.org/10.2307/ltv.16g2zm.
2. Rawcliffe C. Learning to Love the Leper: Aspects of Institutional Charity in Anglo Norman England. Anglo Norman Studies. 2003;xxiii:231-50.
3. Dematre L. Leprosy in Premodem Medicine: A Malady of the Whole Body. JHU Press; 2007.
4. Roberts CA, Manchester K. The archaeology of disease: Cornell University Press; 2007.
5. Kalisch PA. An overview of research on the history of leprosy. Part 1. From Celsus to Simpson, Circa. 1 A.D. Part 2. From Vishov to Møller-Christense, 1845-1973. Int J Lepr Other Mycobact Dis. 1975;43(2):129-44.
6. Miller TS, Nesbitt JW. Walking corpses: Leprosy in Byzantium and the medieval West. Cornell University Press; 2014.
7. Longfield-Jones GM, Mirko D. Grmek, Diseases in the ancient Greek world, trans. Tailrace Mueller and Leonard Muellier, Baltimore and London, Johns Hopkins University Press, 1989, 8vo, pp. xii, 458, £29.00. Med Hist. 1989;33(4): 500.
8. Robbins G, Tripathy VM, Misra VN, Mohanty RK, Shinde VS, Gray KM, et al. Ancient skeletal evidence for leprosy in India (2000 BC). PLoS One. 2009;4(5): e5669 https://doi.org/10.1371/journal.pone.0005669.
9. Mark S. Alexander the Great, seafaring, and the spread of leprosy. J Hist Med Allied Sci. 2002;57(3):285–311 https://doi.org/10.1093/jhmas/57.3.285.
10. Mariotti V, Dutour O, Belcastro MG, Facchini F, Brasili P. Probable early presence of leprosy in Europe in a Celtic skeleton of the 4th-3rd century BC (Casaletto di Reno, Bologna, Italy). Int J Osteoarchaeol. 2005;15(5):311–25 https://doi.org/10.1002/oa.775.
11. Dzierzkzy-Rogalski T. Rytmy i antyrhythmy biologiczne w życiu człowieka: Wiedza Powszechna; 1980.
12. Gibson S, Greenblatt C, Spigelman M, Gorski A, Donoghue HD, Vemon K, et al. The Shroud Cave-a unique case study linking a closed loculus, a shroud and ancient mycobacteria. Ancient Biomolecules; 2002. p. 134.
155. Kircher M, Sawyer S, Meyer M. Double indexing overcomes inaccuracies in multiplex sequencing on the Illumina platform. Nucleic Acids Res. 2012; 40(1):e3 https://doi.org/10.1093/nar/gkr771.

156. Maricic T, Whittem M, Paabo S. Multiplexed DNA sequence capture of mitochondrial genomes using PCR products. PLoS One. 2010;5(11):e14004 https://doi.org/10.1371/journal.pone.0014004.

157. Furtwängler A, Reiter E, Neumann GU, Siebke I, Steuri N, Hafner A, et al. Ratio of mitochondrial to nuclear DNA affects contamination estimates in ancient DNA analysis. Sci Rep. 2018;8(1):14075 https://doi.org/10.1038/s41598-018-32083-0.

158. Hodges E, Rooks M, Xuan Z, Bhattacharjee A, Benjamin Gordon D, Brizuela L, et al. Hybrid selection of discrete genomic intervals on custom-designed microarrays for massively parallel sequencing. Nat Protoc. 2009;4(6):960–74 https://doi.org/10.1038/nprot.2009.68.

159. Tió-Coma M, Avanzi C, Verhard EM, Pierneef L, van Hooij A, Benjak A, et al. Genomic characterization of Mycobacterium leprae to explore transmission patterns identifies new subtype in Bangladesh. Front Microbiol. 2020;11:1220 https://doi.org/10.3389/fmicb.2020.01220.

160. HybridYeastHiC: Analysis of Hi-C in hybrid yeast and related ChIP-seq and RNA-seq data. Github. https://github.com/shendurelab/HybridYeastHiC. Accessed 4 Apr 2021.

161. Renaud G, Slon V, Duggan AT, Kelso J. Schmutzi: estimation of contamination and endogenous mitochondrial consensus calling for ancient DNA. Genome Biol. 2015;16(1):224 https://doi.org/10.1186/s13059-015-0776-0.

162. Vianello D, Sevini F, Castellani G, Lomartire L, Capi M, Franceschi C. HAPLOFIND: a new method for high-throughput mtDNA haplogroup assignment. Hum Mutat. 2013;34(9):1189–94 https://doi.org/10.1002/humu.22356.

163. Weissensteiner H, Pacher D, Kloss-Brandstätter A, Forer L, Specht G, Bandelt H-J, et al. HaploGrep 2: mitochondrial haplogroup classification in the era of high-throughput sequencing. Nucleic Acids Res. 2016;44(W1):W56–63 https://doi.org/10.1093/nar/gkw233.

164. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics. 2009;25(14):1754–60 https://doi.org/10.1093/bioinformatics/btp324.

165. Mittnik A, Wang C-C, Pfengle S, Daubaras M, Zarnia G, Hallgren F, et al. The genetic prehistory of the Baltic Sea region. Nat Commun. 2018;9(1):442 https://doi.org/10.1038/s41467-018-02825-9.

166. Skoglund P, Storl J, Götherström A, Jakobsson M. Accurate sex identification of ancient human remains using DNA shotgun sequencing. J Archaeol Sci. 2013;40(12):4477–82 https://doi.org/10.1016/j.jas.2013.07.004.

Publisher’s Note
Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.