RESEARCH ARTICLE

Recovery of a Medieval *Brucella melitensis* Genome Using Shotgun Metagenomics

Gemma L. Kay,[a] Martin J. Sergeant,[a] Valentina Giuffra,[b,c] Pasquale Bandiera,[c] Marco Milanese,[d] Barbara Bramanti,[e] Raffaella Bianucci,[e,f,g] Mark J. Pallen[a]

Division of Microbiology and Infection, Warwick Medical School, University of Warwick, Coventry, United Kingdom[a]; Division of Paleopathology, Department of Translational Research on New Technologies in Medicine and Surgery, University of Pisa, Pisa, Italy[b]; Department of Biomedical Sciences[c] and Department of History,[d] University of Sassari, Sassari, Italy; Department of Biosciences, Centre for Ecological and Evolutionary Synthesis, University of Oslo, Oslo, Norway[e]; Department of Public Health and Pediatric Sciences, Laboratory of Physical Anthropology, University of Turin, Turin, Italy[f]; Anthropologie Bioculturelle, Droit, Ethique et Santé, Aix Marseille Université, Marseille, France[g]

G.L.K. and M.J.S. and R.B. and M.J.P. contributed equally to this work.

ABSTRACT **Shotgun metagenomics provides a powerful assumption-free approach to the recovery of pathogen genomes from contemporary and historical material. We sequenced the metagenome of a calcified nodule from the skeleton of a 14th-century middle-aged male excavated from the medieval Sardinian settlement of Geridu. We obtained 6.5-fold coverage of a *Brucella melitensis* genome. Sequence reads from this genome showed signatures typical of ancient or aged DNA. Despite the relatively low coverage, we were able to use information from single-nucleotide polymorphisms to place the medieval pathogen genome within a clade of *B. melitensis* strains that included the well-studied Ether strain and two other recent Italian isolates. We confirmed this placement using information from deletions and IS*711* insertions. We conclude that metagenomics stands ready to document past and present infections, shedding light on the emergence, evolution, and spread of microbial pathogens.**

IMPORTANCE **Infectious diseases have shaped human populations and societies throughout history. The recovery of pathogen DNA sequences from human remains provides an opportunity to identify and characterize the causes of individual and epidemic infections. By sequencing DNA extracted from medieval human remains through shotgun metagenomics, without target-specific capture or amplification, we have obtained a draft genome sequence of an ~700-year-old *Brucella melitensis* strain. Using a variety of bioinformatic approaches, we have shown that this historical strain is most closely related to recent strains isolated from Italy, confirming the continuity of this zoonotic infection, and even a specific lineage, in the Mediterranean region over the centuries.**

Received 15 May 2014 **Accepted** 11 June 2014 **Published** 15 July 2014

Citation Kay GL, Sergeant MJ, Giuffra V, Bandiera P, Milanese M, Bramanti B, Bianucci R, Pallen MJ. 2014. Recovery of a medieval *Brucella melitensis* genome using shotgun metagenomics. mBio 5(4):e01337-14. doi:10.1128/mBio.01337-14.

Editor Paul Keim, Northern Arizona University

Address correspondence to Raffaella Bianucci, raffaella.bianucci@unito.it, or Mark J. Pallen, m.pallen@warwick.ac.uk.

Brucellosis is a widespread infection of livestock (sheep, goats, cattle, cows, and pigs) and remains one of the most common zoonotic infections, with more than 500,000 new human cases worldwide annually (1). Human brucellosis is most commonly caused by the species *Brucella melitensis* and is usually acquired through ingestion of unpasteurized dairy products or, less commonly, through ingestion of infected meat or direct occupational contact with animals (1). If left untreated, the infection usually follows a chronic course, spreading systemically to the organs of the reticuloendothelial system and often leading to osteoarticular disease (2).

Brucellosis is an ancient disease. Vertebral lesions consistent with brucellosis have been described in a >2-million-year-old male skeleton of *Australopithecus africanus* (3). Lesions consistent with brucellosis have been described in Bronze Age skeletons from the Levant and the Basque country, in adult skeletons from Herculaneum, and in medieval human remains (2).

A major ambition of paleopathology is to shed light on the influence of infectious disease on past populations. However, morphological analyses are limited in that few infections produce durable lesions and very different pathogens can produce similar pathologies (4). For example, similar lesions occur in tuberculosis and brucellosis, even though the causative organisms are quite different taxonomically and in cell structure. Amplification of pathogen DNA from human remains via PCR has provided an alternative source of information about a range of past infections (5). However, only a single study has reported success in using PCR amplification to confirm historical brucellosis (6).

In addition, there are problems with amplification-based approaches to the recovery of historical and ancient DNA. First, PCR is a competitive and highly sensitive reaction, prone to contamination even in dedicated facilities (7). Second, these approaches generally provide information on a single gene or gene fragment, affording little or no insight into pathogen biology, evolution, and epidemiology. Third, they require the onerous design and optimization of pathogen-specific primers; this limits the open-

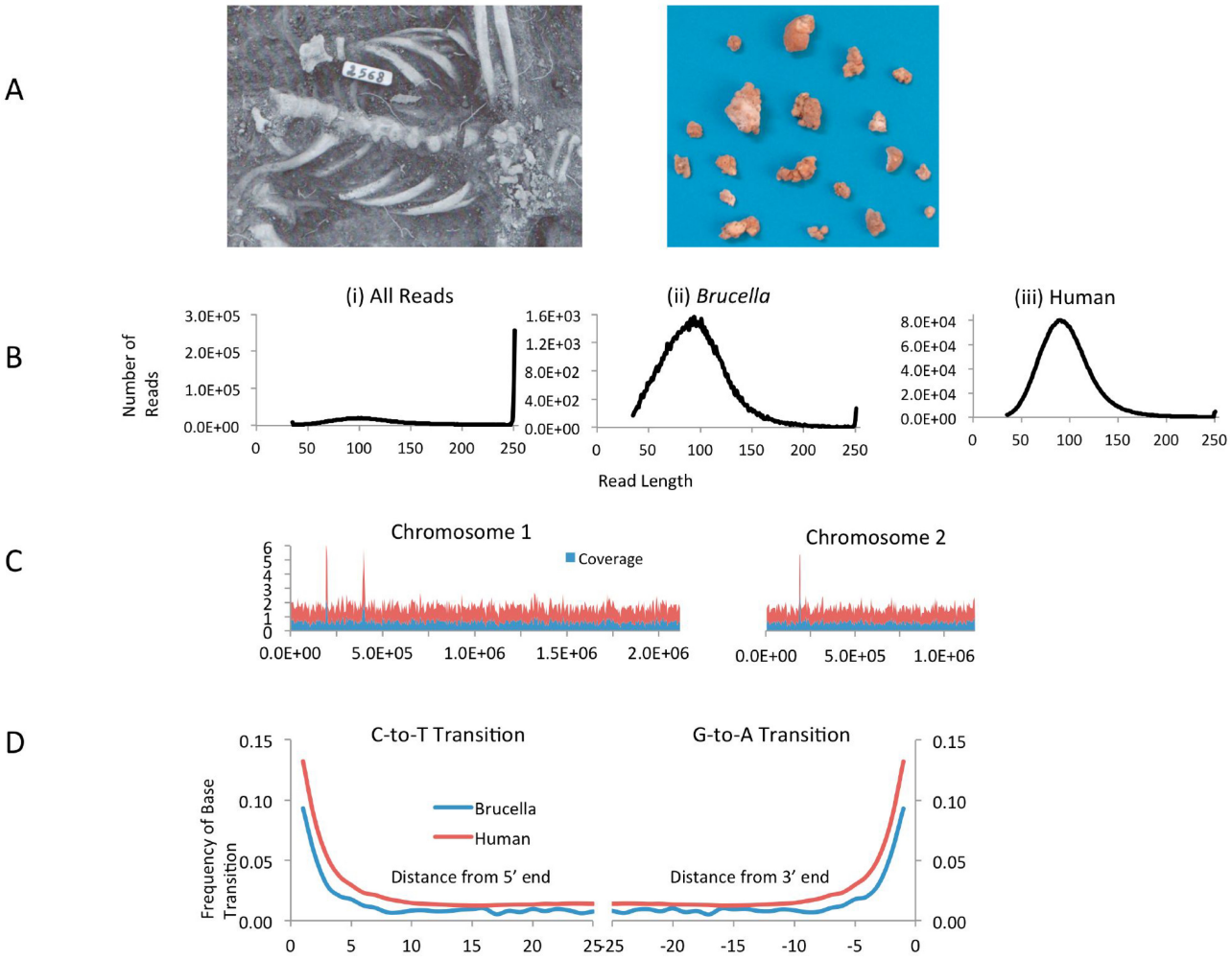

FIG 1 (A) Calcific nodules excavated from the pelvic girdle of skeleton 2658. (B) Size distribution of reads from sample 2568. (i) All reads from the initial MiSeq run (~2 m). (ii) Reads from the second Miseq run (>20 m) which aligned with the *B. melitensis* 16 M genome. (iii) Reads from the dedicated run which mapped to the human genome (hg19). (C) Coverage plot of reads from sample 2568 mapped against the two chromosomes from the *B. melitensis* 16 M genome. The plot shows the average coverage and standard deviation for each 5,000-bp region in the genome. (D) Evidence of sequence damage associated with aged DNA showing the frequency of C-to-T and G-to-A transitions at the 5′ and 3′ ends of DNA fragments, respectively.

endedness of the approach, so that one generally finds only what one is looking for. This last point also applies to genome capture approaches, which have proven successful in recovering *Yersinia pestis* genomes from samples from the Black Death and the Justinianic plague (8, 9).

Shotgun metagenomics—that is the unbiased sequencing *en masse* of DNA extracted from a sample without target-specific amplification or capture—provides an attractive alternative approach to the detection and characterization of pathogens in contemporary and historical human material. This approach has proven successful in obtaining genome-wide sequence data for *Borrelia burgdorferi*, *Mycobacterium tuberculosis,* and *Mycobacterium leprae* from long-dead human remains (10–12).

When confronted with calcified nodules from a 14th-century skeleton, we initially thought of tuberculosis. However, when we used shotgun metagenomics to identify potential pathogens in the sample, we were instead surprised to recover a medieval *Brucella melitensis* genome sequence.

RESULTS

Metagenomic recovery of *Brucella melitensis* sequences with signatures of medieval origin. The skeleton of a 50- to 60-year-old male (skeleton 2568) was excavated from the abandoned medieval village of Geridu (Sorso, Sassari, Italy) in northwest Sardinia in December 1997 (Fig. 1) (13). The skeleton showed features of diffuse idiopathic skeletal hyperostosis (DISH), including fusions between the fourth and tenth thoracic vertebrae, fusion of the fifth lumbar vertebra to the sacrum, and extraspinal enthesopathies (14). Thirty-two calcified nodules were found in the pelvic girdle, with diameters ranging from 0.6 by 0.7 cm to 2.2 by 1.6 cm (Fig. 1). A DNA extraction was performed on one of the nodules.

We obtained a DNA yield from the nodule of 30 ng, which was used to construct a TruSeq Nano Illumina library, which was run at low coverage on an Illumina MiSeq sequencer, alongside 10 other bar-coded libraries, 8 from other historical human tissue samples and 2 from blank controls. Just over two million se-

quences were obtained from sample 2568 on this run. An analysis of size distribution was performed, revealing a bimodal distribution with a broad peak running from 50 to 150 bp, with a second much taller peak running up to the maximum read length of 250 bp (Fig. 1). On the assumption that historical DNA fragments were restricted to the smaller peak, sequences over 150 bp in length were excluded from analysis of this run. Homology searches revealed sequences from sample 2568 that could be assigned with confidence to the genus *Brucella.*

We then attempted to map reads from all 11 samples against the genome of the *B. melitensis* reference strain 16 M. We obtained insignificant matches for 10 of the samples (<12 aligned reads per sample), whereas sample 2568 yielded >20,000 paired-end reads (equivalent to 10,000 sequences, as the fragments are shorter than the read length used in paired-end sequences) that mapped against the *B. melitensis* 16 M genome, providing approximately 0.7-fold coverage of a medieval *Brucella* genome from a strain that we have called Geridu-1. A coverage plot (Fig. 1) revealed even coverage across both chromosomes in the 16 M genome, ruling out spurious hits to conserved sequences from environmental bacteria.

To obtain additional sequencing reads, the sample 2568 library was sequenced at ~10-fold-higher coverage on a single dedicated MiSeq run, which yielded just over 20 million paired-end sequences. When these were mapped at high stringency, 23% of these reads mapped against the human genome and 0.48% against the *B. melitensis* 16 M genome (representing 6.5-fold coverage). Interestingly, when the reads that mapped to either the *B. melitensis* or human genomes were reanalyzed, they showed a much tighter size distribution than the library as a whole, with a peak centered on 100 bases (Fig. 1). In addition, the reads mapping to the *B. melitensis* or human genome showed abundant CT and GA base conversions at the 5′ and 3′ ends, which is indicative of the damage typical of ancient or aged DNA (Fig. 1). These findings are supportive of a medieval origin of these sequences (15).

SNP-based phylogenetic placement shows that the medieval *Brucella* genome is closely related to recent Italian isolates. Conventional phylogenetic methods based on identification of trusted single-nucleotide polymorphisms (SNPs) cannot be applied to low-coverage genome sequences. However, the technique of "phylogenetic placement" provides an alternative solution (16). Here, one draws on a fixed reference tree, computed from high-coverage genomes, and places the unknown query sequence on the tree using programs such as pplacer. We used a published set of phylogenetically informative SNPs for *Brucella* spp. (17) and analyzed reads from the initial MiSeq run that aligned to equivalent positions in the 16 M genome. Using this approach, despite the low coverage, we could show confidently (with a posterior probability of 1) that the Geridu-1 strain clustered most closely with the well-characterized Ether strain (ATCC 23458; the reference strain for *B. melitensis* biovar 3) and was nested within a clade of four *B. melitensis* strains (see Fig. S1 in the supplemental material). Indeed, with as few as 250 reads, we were able to accurately assign the Geridu-1 strain to the Ether clade (Fig. S2).

To refine the placement of the Geridu-1 genome, we constructed a broad-based phylogenetic tree from all available modern *B. melitensis* genomes. This showed that Geridu-1's close relative, the Ether strain, belonged in a distinctive clade along with four other strains, for which only draft genome sequences were available (see Fig. S3 in the supplemental material). We then compared all five strains from the Ether clade and the Geridu-1 genome recovered from the second MiSeq run by calling SNPs against the completed 16 M genome (Table S1) and using them to draw a phylogenetic tree for this clade (Fig. 2). The SNP table and tree showed that the Geridu-1 strain represented the earliest branching lineage within the Ether clade, separated by 450 to 500 SNPs from any other strain in the clade (Table S2).

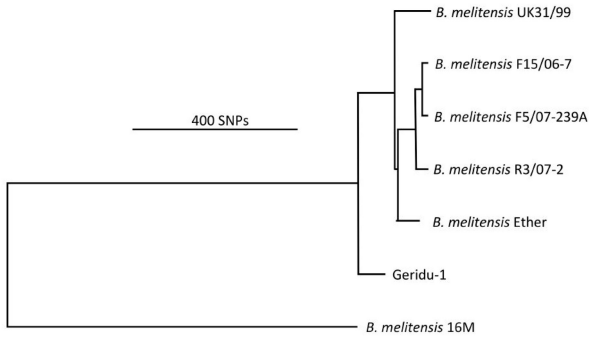

FIG 2 Phylogenetic tree showing the position of the medieval Geridu-1 strain within the Ether clade. Only those SNPs which correlated with sufficient coverage in the Geridu-1 alignment were included in the construction of the tree.

At least three of the five contemporary strains that belong to the Ether clade originate from the Italian peninsula or from Sicily: F15/06-7 is a 2006 human isolate from Sicily, F5/07-239A is a 2007 small-ruminant isolate from Italy, and the Ether strain itself was isolated from an Italian goat in 1961. There are no other Italian isolates in the set of currently available *B. melitensis* genome sequences, perhaps suggesting that the Ether clade originated in Italy and its associated islands, or at least in the western Mediterranean. We note that in a recent multiple-locus variable-number tandem-repeat analysis (MLVA) study (18), Italian strains clustered separately from most other European strains and we suspect that Geridu-1 belongs to the western Mediterranean cluster defined by MLVA, although additional genome sequencing of isolates from that cluster are required to confirm this.

Confirmation of phylogenetic placement using insertions and deletions. To confirm the placement of the Geridu-1 strain within the Ether clade of *B. melitensis*, we drew upon two additional sources of information: the distribution of deletions and the locations of insertion elements. First, we looked for deletions of >100 bases in length that occurred in the Geridu-1 genome in comparison to the 16 M strain. We identified 11 such deletions (see Table S3 in the supplemental material). We determined the distribution of these deletions in all available *B. melitensis* genomes. Nine deletions were found in only Geridu-1 and the five other strains in the Ether clade. Two occurred sporadically in other strains.

Next, we examined the distribution and location of the insertion element IS*711* (also called IS*6501*), which occurs widely in *Brucella* spp. (19). When we mapped reads from the Geridu-1 genome to the insertion element, we obtained 52.3-fold coverage. Dividing that by the average coverage for the Geridu-1 genome (6.5-fold) provided us with an estimate of eight IS*711* copies in the medieval strain. By analyzing reads that spanned the ends of IS*711* and the adjacent chromosome, we were able to confirm the existence of eight insertion points in Geridu-1, seven of which also occurred in *B. melitensis* 16 M and in all other available *B. meliten-*

sis genomes (Table S3). The IS*711* insertion present in Geridu-1 but absent from 16 M was located at position 517784 in chromosome 2, disrupting a gene encoding the hypothetical protein BMEII0494. This insertion point was found in no other *B. melitensis* strain genome, apart from the other five members of the Ether lineage. These patterns of deletion and IS*711* insertions confirm the placement of the Geridu-1 strain within the Ether clade of *B. melitensis*.

DISCUSSION

Here, we have shown that shotgun metagenomics can be used to obtain a *Brucella melitensis* genome sequence from a medieval sample without target-specific amplification or capture. The recently reported success of this approach with two mycobacterial diseases, leprosy and tuberculosis, has been ascribed to the unique properties of the mycobacterial cell wall in preserving bacterial DNA (11, 12). In contrast, this study confirms that whole-genome sequences from bacterial pathogens without resilient cell envelopes can be recovered from human remains by metagenomics hundreds or even thousands of years postmortem. However, unlike medieval mycobacterial DNA (12), medieval *Brucella* DNA does show signatures of damage associated with an ancient or historical provenance.

This observation complements several other lines of evidence that support the authenticity of data obtained from this historical sample. First, the sample was processed in facilities dedicated to ancient DNA research, where no work on *Brucella* cultures or DNA had ever taken place. Second, laboratory contamination was ruled out by the lack of significant hits for *Brucella* in libraries obtained from eight other historical specimens and two blanks sequenced in the same run. Third, we observed conclusive and extensive matches to a dedicated human and animal pathogen, obtaining even genomic coverage of a genome nested within the *B. melitensis* phylogeny. This rules out spurious and patchy hits for conserved genes from related environmental organisms as a source of error. Finally, we took care to sample the interior of the nodule, which eliminated the risk of contamination from soil.

Calcification of soft tissue is a recognized, albeit rare, complication of human brucellosis, so we assume that calcification of abdominal or pelvic tissues accounts for the appearance of pelvic nodules in this individual. The skeletal pathology provides no additional evidence of brucellosis. Instead, the widespread bony changes in skeleton 2568 are diagnostic of diffuse idiopathic skeletal hyperostosis (DISH), a systemic condition of unknown etiology that is characterized by the ossification or calcification of ligaments and enthuses, particularly of the thoracic spine, and which is seen most commonly in men over 50 years of age (14). DISH has been reported in numerous ancient and medieval skeletons and is thought to be associated with a privileged and/or sedentary lifestyle—for example, it was seen in the Medici family (20). Whether this historical case of brucellosis represents infection through direct contact with an infected animal (e.g., a shepherd exposed to birthing animals) or through ingestion of dairy products remains unclear, although the relatively old age at death and the coexistence of DISH as a potential disease of affluence might favor the latter.

Our findings fit the epidemiological, historical, and geographical contexts in that *B. melitensis* is most commonly acquired from sheep or goats (1) and sheep (and probably also goat) herding has a long history in Sardinia, as evidenced by distinctive genomic signatures in two local breeds of domesticated sheep and the presence of the Sardinian mouflon, which is thought to represent an archaic feral sheep population that reached Europe from the Neolithic domestication center (21).

By showing that our medieval *Brucella* strain is most closely related to recent Italian isolates, we have established the continuity of this zoonotic infection in the region over the centuries. Brucellosis still occurs on the island of Sardinia, although it is now relatively well controlled there compared to other regions of endemicity. Nonetheless, the struggle to control this infection, which has afflicted the peoples of the Mediterranean region since ancient times, is still ongoing. We conclude that metagenomic approaches stand ready to document past infections, shedding light on the emergence, evolution, and spread of *Brucella* and other microbial pathogens and to inform current and future infectious disease diagnosis and control.

MATERIALS AND METHODS

Source material. The source material consisted of 1 of 32 calcified nodules found in the pelvic girdle of an adult male skeleton (sample 2568) that was excavated from 1997 to 1999 from a cemetery in the medieval rural settlement of Geridu (Sorso, Sassari, Italy), in northwest Sardinia. Twenty-five single pit graves were identified in a well-organized part of the cemetery (sector 2500). Stratigraphic analysis identified two distinct burial phases. Phase I (9 single burials) dates to the first half of the 14th century CE; phase II (16 single burials) dates to the final period in which the cemetery was used (1350 to 1400 CE). Following the historical reports, the settlement of Geridu was definitely abandoned in 1426 CE. The sample 2568 remains were retrieved from 1 of the 16 burials dating to the second half of the 14th century CE (13). Sex determination was performed on the basis of the morphological features of the skull. Age at death was estimated on the basis of dental wear and sternal rib end modification (22). Lesions indicative of pathologies were recorded in accordance with the methods and standards set out in the Global History of Health Project (23).

DNA extraction. DNA extraction and library preparation were carried out in a dedicated ancient DNA laboratory in which no strains of *Brucella* had ever been cultured, no pathogen-specific PCRs had ever been performed, and in which the handler wore gloves, together with a mask, a gown, and a hood. The surface of the calcified nodule was removed using a drill bit that had been cleaned with bleach to eliminate potential surface contaminants. The undersurface was then removed by drilling at low speed to produce approximately 20 mg of powder. The sample was incubated at 37°C with shaking to demineralize it in 400 μl CTAB (cetyltrimethylammonium bromide) solution and 40 μl proteinase K for 1 week.

DNA was isolated with chloroform using the DNeasy plant minikit (Qiagen United Kingdom), with the following modifications to the manufacturer's protocol so the kit could be used for historical human tissues. In brief, the demineralized sample was centrifuged at 20,000 $\times$ *g* for 10 min and the supernatant was collected. To homogenize the sample, 400 μl of 0.2 M chloroform was added and the tube was inverted for 10 min. The sample was centrifuged at 6,000 $\times$ *g* for 3 min, and the supernatant was transferred to 2-ml tubes. The manufacturer's DNeasy plant minikit protocol was followed from step 6 with the following additional modifications. Three volumes of buffer AW1 was added to each sample and incubated at room temperature for 2 to 3 h. After the addition of buffer AW2, the samples were centrifuged for 3 min. DNA was eluted in two 50-μl aliquots (100-μl total). Extracted DNA was quantified in 5 μl of sample using the Qubit high-sensitivity double-stranded DNA (HS dsDNA) assay according to the manufacturer's instructions (Invitrogen Ltd., Paisley, United Kingdom) and then stored at −20°C until library preparation.

Sequencing. DNA extracted from sample 2568 was converted into a TruSeq Nano library for sequencing on an Illumina MiSeq sequencer according to the manufacturer's low-sample protocol (Illumina UK, Little

Chesterford, United Kingdom), with the following minor modifications. No fragmentation step was included, given the expectation that ancient DNA would already be heavily fragmented. DNA was end-repaired within the ancient-DNA laboratory, with the modification of incubation for 90 min at 30°C and size selection for <350-bp inserts. Analysis on an Agilent Bioanalyzer 2100 system provided an estimated size distribution of fragments with a peak length of 235 bp.

Once adapters (including bar codes) had been ligated, the sequencing library was moved to a PCR laboratory. In accordance with the manufacturer's instructions, the library fragments were PCR amplified; however, we used 10 PCR cycles instead of the usual 8. The library was quantified (2 μl per sample) using the Qubit HS dsDNA assay according to the manufacturer's instructions (Invitrogen Ltd., Paisley, United Kingdom) and then stored at −20°C. The sample 2568 library was diluted to 4 nM, as determined by analysis on an Agilent Bioanalyzer 2100 and using the Qubit HS dsDNA assay, and then pooled in equimolar amounts with 10 other bar-coded libraries (8 from other historical human tissue samples and 2 from blank controls). The entire library pool was then diluted to 12 pM and sequenced on the first MiSeq run using the Illumina MiSeq v2 2 × 250-bp paired-end protocol. In a second MiSeq run, dedicated entirely to the 2568 library, the 4 nM library was diluted to 12 pM without pooling.

Identification of *Brucella* sequences. Metagenomic sequence reads from both MiSeq runs using sample 2568 have been deposited in the European Nucleotide Archive (project accession number PRJEB6045). Sequences derived from sample 2568 in the first MiSeq run that had lengths of ≥150 bp were subjected to a BLASTN search against the NCBI NR database, and the results were analyzed with MEGAN (24). Reads from all samples on the initial MiSeq run were analyzed with Bowtie2 version 2.1.0 (25), allowing only 1 mismatch per 33 bases of the read (under the following settings: `--mp 1,1`; `--ignore-quals`; `--score-min L,0, −0.033`). The reads were mapped against the genomes of the following pathogens: *Plasmodium falciparum* 3D7, *Leishmania infantum* JPCM5, *Yersinia pestis* CO92, *Mycobacterium tuberculosis* H37Rv, and *Brucella melitensis* 16 M (GenBank accession numbers AL844501 to AL844509, AE014185 to AE014188, AE001362, FR796433 to FR796468, NC_003143, AL123456, NC_003317, and NC_003318).

Phylogenetic placement of *Brucella* sequences from Geridu at low coverage. Previously described lineage-defining SNPs (17) were used to construct a tree with FastTree 2.7.1 (26), using neighbor joining and generalized time-reversible models of nucleotide evolution. Reads from the metagenome were mapped against the reference strain *B. melitensis* 16 M using the default settings in Bowtie2 and the majority base called from each SNP position with no quality filtering. If no base was present at the position, a gap was used. The pplacer suite of programs (16) was used to assign the sequence to a position on the tree.

Phylogenetic analysis of *B. melitensis* strains. The phylogeny of all *B. melitensis* strains was analyzed, using information from the PATRIC database (27) and a *Brucella* SNP matrix from the Broad Institute (http://www.broadinstitute.org). To map these strains on the matrix, sequences of 160 bases surrounding each SNP were taken from the reference strain, *B. melitensis* 16 M, and this fragment was mapped (using Bowtie2) against each strain to ascertain the corresponding base in the genome. If no mapping occurred or the read mapped in two places, a gap was added instead. Using this extended matrix, concatenated SNPs for each strain were then used to construct a tree using FastTree 2.7.1.

Calling SNPs from strains in the Ether clade. Five modern draft genomes from strains in the Ether clade (UK31/99, F15/06-7, F5/07-239A, Ether, and R3/07-2) were aligned against *B. melitensis* 16 M using Progressive Mauve (28), and SNPs were called using the default settings. If any strain had a gap or N, the SNP was discarded. SNPs for the Geridu-1 strain were called from the mapped bam file using Samtools (29), with requirements that there was at least 6-fold coverage and that the mutant allele accounted for at least 80% of aligned sequences. The SNPs from Geridu-1 were combined with those from the Ether clade strains, excluding any SNP that occurred at a location where there was less than 6-fold coverage in the Geridu-1 sequence alignment. The remaining 2,332 SNPs (Table S1) were then used to construct a tree using FastTree 2.7.1 and to produce a pairwise dissimilarity matrix.

Analysis of breakpoints in Geridu-1 and other strains. From manual scrutiny of the coverage plot, we identified regions in the 16 M reference genome where for a span of >100 bp, no reads mapped from the Geridu-1 genome. To confirm the existence and refine the boundaries of these deletions, the Geridu-1 reads were remapped against 16 M using the `--local` option of Bowtie2, which allows the beginning and end of reads to be soft-clipped to obtain an improved alignment. The clipped regions of the reads at the edges of candidate deletions were used to identify the breakpoint created by the deletion. The distribution of the breakpoints in other strains was determined by retrieving the sequences flanking the deletion breakpoints in the Geridu-1 strain and performing a BLASTN search of available *B. melitensis* genomes. Similar approaches were used to determine the identity and distribution of insertion breakpoints associated with IS*11*.

Nucleotide sequence accession numbers. Metagenomic sequence reads from this study have been deposited in the European Nucleotide Archive (project accession number PRJEB6045). The following URL will pull all of the sequences that were used in our searches: http://www.ncbi.nlm.nih.gov/nuccore/AL844501,AL844502,AL844503,AL844504,AL844505,AL844506,AL844507,AL844508,AL844509,AE014185,AE014186,AE014186,AE014187,AE014188,AE001362,FR796433,FR796434,FR796435,FR796436,FR796437,FR796438,NC_003143,AL123456,NC_003317,NC_003318.

SUPPLEMENTAL MATERIAL

Supplemental material for this article may be found at http://mbio.asm.org/lookup/suppl/doi:10.1128/mBio.01337-14/-/DCSupplemental.

Figure S1, PDF file, 0.2 MB.
Figure S2, PDF file, 0.1 MB.
Figure S3, PDF file, 0.5 MB.
Table S1, PDF file, 0.1 MB.
Table S2, PDF file, 0.4 MB.
Table S3, PDF file, 0.1 MB.

ACKNOWLEDGMENTS

We acknowledge the RAS Legge Regionale 7 agosto 2007, (n) 7, bando 2010 Bando 2010-2011, for funding bioanthropological investigations. We thank Warwick Medical School for funding laboratory and bioinformatics work.

We thank Jackie Chan for storage of samples, Oliver Smith for supplying a drill and for advice on library preparation, Robin Allaby for providing access to the ancient-DNA laboratory, and Nick Loman for advice on phylogenetic placement. We thank the *Brucella* Group Sequencing Project, Broad Institute of Harvard and MIT, for making prepublication genomes available.

R.B., B.B., and M.J.P. designed the research; M.M., R.B., V.G., P.B., and G.L.K. performed the research; G.L.K., M.J.S., and M.J.P. analyzed the data; and M.J.P. and R.B. wrote the paper.

REFERENCES

1. **Franco MP, Mulder M, Gilman RH, Smits HL.** 2007. Human brucellosis. Lancet Infect. Dis. **7**:775–786. http://dx.doi.org/10.1016/S1473-3099(07)70286-4.
2. **D'Anastasio R, Staniscia T, Milia ML, Manzoli L, Capasso L.** 2011. Origin, evolution and paleoepidemiology of brucellosis. Epidemiol. Infect. **139**:149–156. http://dx.doi.org/10.1017/S095026881000097X.
3. **D'Anastasio R, Zipfel B, Moggi-Cecchi J, Stanyon R, Capasso L.** 2009. Possible brucellosis in an early hominin skeleton from Sterkfontein, South Africa. PLoS One **4**:e6439. http://dx.doi.org/10.1371/journal.pone.0006439.
4. **Ortner DJ.** 2011. Human skeletal paleopathology. Int. J. Paleopathol. **1**:-4–11. http://dx.doi.org/10.1016/j.ijpp.2011.01.002.
5. **Anastasiou E, Mitchell PD.** 2013. Palaeopathology and genes: investigat-

ing the genetics of infectious diseases in excavated human skeletal remains and mummies from past populations. Gene **528:**33–40. http://dx.doi.org/10.1016/j.gene.2013.06.017.

6. **Mutolo MJ, Jenny LL, Buszek AR, Fenton TW, Foran DR.** 2012. Osteological and molecular identification of brucellosis in ancient Butrint, Albania. Am. J. Phys. Anthropol. **147:**254–263. http://dx.doi.org/10.1002/ajpa.21643.
7. **Willerslev E, Cooper A.** 2005. Ancient DNA. Proc. Biol. Sci. **272:**3–16. http://dx.doi.org/10.1098/rspb.2004.2813.
8. **Bos KI, Schuenemann VJ, Golding GB, Burbano HA, Waglechner N, Coombes BK, McPhee JB, DeWitte SN, Meyer M, Schmedes S, Wood J, Earn DJ, Herring DA, Bauer P, Poinar HN, Krause J.** 2011. A draft genome of *Yersinia pestis* from victims of the Black Death. Nature **478:**506–510. http://dx.doi.org/10.1038/nature10549.
9. **Wagner DM, Klunk J, Harbeck M, Devault A, Waglechner N, Sahl JW, Enk J, Birdsell DN, Kuch M, Lumibao C, Poinar D, Pearson T, Fourment M, Golding B, Riehm JM, Earn DJ, Dewitte S, Rouillard JM, Grupe G, Wiechmann I, Bliska JB, Keim PS, Scholz HC, Holmes EC, Poinar H.** 2014. *Yersinia pestis* and the plague of Justinian 541-543 AD: a genomic analysis. Lancet Infect. Dis. **14:**319–326. http://dx.doi.org/10.1016/S1473-3099(13)70323-2.
10. **Keller A, Graefen A, Ball M, Matzas M, Boisguerin V, Maixner F, Leidinger P, Backes C, Khairat R, Forster M, Stade B, Franke A, Mayer J, Spangler J, McLaughlin S, Shah M, Lee C, Harkins TT, Sartori A, Moreno-Estrada A, Henn B, Sikora M, Semino O, Chiaroni J, Rootsi S, Myres NM, Cabrera VM, Underhill PA, Bustamante CD, Vigl EE, Samadelli M, Cipollini G, Haas J, Katus H, O'Connor BD, Carlson MR, Meder B, Blin N, Meese E, Pusch CM, Zink A.** 2012. New insights into the Tyrolean Iceman's origin and phenotype as inferred by whole-genome sequencing. Nat. Commun. **3:**698. http://dx.doi.org/10.1038/ncomms1701.
11. **Chan JZ, Sergeant MJ, Lee OY, Minnikin DE, Besra GS, Pap I, Spigelman M, Donoghue HD, Pallen MJ.** 2013. Metagenomic analysis of tuberculosis in a mummy. N. Engl. J. Med. **369:**289–290. http://dx.doi.org/10.1056/NEJMc1302295.
12. **Schuenemann VJ, Singh P, Mendum TA, Krause-Kyora B, Jäger G, Bos KI, Herbig A, Economou C, Benjak A, Busso P, Nebel A, Boldsen JL, Kjellström A, Wu H, Stewart GR, Taylor GM, Bauer P, Lee OY, Wu HH, Minnikin DE, Besra GS, Tucker K, Roffey S, Sow SO, Cole ST, Nieselt K, Krause J.** 2013. Genome-wide comparison of medieval and modern *Mycobacterium leprae*. Science **341:**179–183. http://dx.doi.org/10.1126/science.1238286.
13. **Milanese M.** 2006. Vita e morte dei villaggi rurali tra Medioevo ed Età Moderna. Dallo scavo della villa de Geriti ad una pianificazione della tutela e della conoscenza dei villaggi abbandonati della Sardegna. Archeologia e storia degli insediamenti rurali abbandonati della Sardegna, p 50–53. *In* Milanese M (ed), Quaderni del Centro di Documentazione dei Villaggi Abbandonati della Sardegna 2, Florence. All'Insegna Edizioni del Giglio, Florence, Italy.
14. **Mader R, Verlaan JJ, Buskila D.** 2013. Diffuse idiopathic skeletal hyperostosis: clinical features and pathogenic mechanisms. Nat. Rev. Rheumatol. **9:**741–750. http://dx.doi.org/10.1038/nrrheum.2013.165.
15. **Briggs AW, Stenzel U, Johnson PL, Green RE, Kelso J, Prüfer K, Meyer M, Krause J, Ronan MT, Lachmann M, Pääbo S.** 2007. Patterns of damage in genomic DNA sequences from a Neandertal. Proc. Natl. Acad. Sci. U. S. A. **104:**14616–14621. http://dx.doi.org/10.1073/pnas.0704665104.
16. **Matsen FA, Kodner RB, Armbrust EV.** 2010. pplacer: linear time maximum-likelihood and Bayesian phylogenetic placement of sequences onto a fixed reference tree. BMC Bioinformatics **11:**538. http://dx.doi.org/10.1186/1471-2105-11-538.
17. **Foster JT, Beckstrom-Sternberg SM, Pearson T, Beckstrom-Sternberg JS, Chain PS, Roberto FF, Hnath J, Brettin T, Keim P.** 2009. Whole-genome-based phylogeny and divergence of the genus *Brucella*. J. Bacteriol. **191:**2864–2870. http://dx.doi.org/10.1128/JB.01581-08.
18. **Al Dahouk S, Flèche PL, Nöckler K, Jacques I, Grayon M, Scholz HC, Tomaso H, Vergnaud G, Neubauer H.** 2007. Evaluation of *Brucella* MLVA typing for human brucellosis. J. Microbiol. Methods **69:**137–145. http://dx.doi.org/10.1016/j.mimet.2006.12.015.
19. **Mancilla M, Ulloa M, López-Goñi I, Moriyón I, Maria Zarraga A.** 2011. Identification of new IS*711* insertion sites in *Brucella abortus* field isolates. BMC Microbiol. **11:**176. http://dx.doi.org/10.1186/1471-2180-11-176.
20. **Giuffra V, Giusiani S, Fornaciari A, Villari N, Vitiello A, Fornaciari G.** 2010. Diffuse idiopathic skeletal hyperostosis in the Medici, Grand Dukes of Florence (XVI century). Eur. Spine J. **19**(Suppl 2):S103–S107. http://dx.doi.org/10.1007/s00586-009-1125-3.
21. **Ciani E, Crepaldi P, Nicoloso L, Lasagna E, Sarti FM, Moioli B, Napolitano F, Carta A, Usai G, D'Andrea M, Marletta D, Ciampolini R, Riggio V, Occidente M, Matassino D, Kompan D, Modesto P, Macciotta N, Ajmone-Marsan P, Pilla F.** 2014. Genome-wide analysis of Italian sheep diversity reveals a strong geographic pattern and cryptic relationships between breeds. Anim. Genet. **45:**256–266. http://dx.doi.org/10.1111/age.12106.
22. **Loth SR, Iscan MY.** 1989. Morphological assessment of age in the adult: the thoracic region, p 105–136. *In* Iscan MY (ed), Age markers in the human skeleton. Charles C Thomas, Springfield, IL.
23. **Steckel RH, Larsen CS, Sciulli PW, Walker PL.** 2005. The Global History of Health Project: data collection codebook. Ohio State University, Columbus, OH. http://global.sbs.ohio-state.edu/new_docs/Codebook-01-24-11-em.pdf.
24. **Huson DH, Auch AF, Qi J, Schuster SC.** 2007. MEGAN analysis of metagenomic data. Genome Res. **17:**377–386. http://dx.doi.org/10.1101/gr.5969107.
25. **Langmead B, Salzberg SL.** 2012. Fast gapped-read alignment with Bowtie 2. Nat. Methods **9:**357–359. http://dx.doi.org/10.1038/nmeth.1923.
26. **Price MN, Dehal PS, Arkin AP.** 2010. FastTree 2—approximately maximum-likelihood trees for large alignments. PLoS One **5:**e9490. http://dx.doi.org/10.1371/journal.pone.0009490.
27. **Wattam AR, Abraham D, Dalay O, Disz TL, Driscoll T, Gabbard JL, Gillespie JJ, Gough R, Hix D, Kenyon R, Machi D, Mao C, Nordberg EK, Olson R, Overbeek R, Pusch GD, Shukla M, Schulman J, Stevens RL, Sullivan DE, Vonstein V, Warren A, Will R, Wilson MJ, Yoo HS, Zhang C, Zhang Y, Sobral BW.** 2014. PATRIC, the bacterial bioinformatics database and analysis resource. Nucleic Acids Res. **42:**D581–D591. http://dx.doi.org/10.1093/nar/gkt1099.
28. **Darling AE, Mau B, Perna NT.** 2010. progressiveMauve: multiple genome alignment with gene gain, loss and rearrangement. PLoS One **5:**e11147. http://dx.doi.org/10.1371/journal.pone.0011147.
29. **Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R, 1000 Genome Project Data Processing Subgroup.** 2009. The Sequence Alignment/Map format and SAMtools. Bioinformatics **25:**2078–2079. http://dx.doi.org/10.1093/bioinformatics/btp352.