Tuberculosis 95 (2015) S140–S144

Contents lists available at ScienceDirect

Tuberculosis

journal homepage: http://intl.elsevierhealth.com/journals/tube

Ancient DNA analysis – An established technique in charting the evolution of tuberculosis and leprosy

Helen D. Donoghue [a, b, *], Mark Spigelman [a, c], Justin O'Grady [a, 1], Ildikó Szikossy [d], Ildikó Pap [d], Oona Y.-C. Lee [e], Houdini H.T. Wu [e], Gurdyal S. Besra [e], David E. Minnikin [e]

[a] Centre for Clinical Microbiology, Division of Infection & Immunity, University College London, London, UK
[b] Centre for the History of Medicine, Division of Biosciences, University College London, UK
[c] Department of Anatomy and Anthropology, Sackler Faculty of Medicine, Tel Aviv University, Tel Aviv, Israel
[d] Department of Anthropology, Hungarian Natural Science Museum, Budapest, Hungary
[e] Institute of Microbiology and Infection, School of Biosciences, University of Birmingham, Edgbaston, Birmingham, UK

Keywords:
Ancient DNA
Evolution
Mycobacterium leprae
Mycobacterium tuberculosis
Molecular typing

S U M M A R Y

Many tuberculosis and leprosy infections are latent or paucibacillary, suggesting a long time-scale for host and pathogen co-existence. Palaeopathology enables recognition of archaeological cases and PCR detects pathogen ancient DNA (aDNA). *Mycobacterium tuberculosis* and *Mycobacterium leprae* cell wall lipids are more stable than aDNA and restrict permeability, thereby possibly aiding long-term persistence of pathogen aDNA. Amplification of aDNA, using specific PCR primers designed for short fragments and linked to fluorescent probes, gives good results, especially when designed to target multi-copy loci. Such studies have confirmed tuberculosis and leprosy, including co-infections. Many tuberculosis cases have non-specific or no visible skeletal pathology, consistent with the natural history of this disease. *M. tuberculosis* and *M. leprae* are obligate parasites, closely associated with their human host following recent clonal distribution. Therefore genotyping based on single nucleotide polymorphisms (SNPs) can indicate their origins, spread and phylogeny. Knowledge of extant genetic lineages at particular times in past human populations can be obtained from well-preserved specimens where molecular typing is possible, using deletion analysis, microsatellite analysis and whole genome sequencing. Such studies have identified non-bovine tuberculosis from a Pleistocene bison from 17,500 years BP, human tuberculosis from 9000 years ago and leprosy from over 2000 years ago.

1. Introduction

According to the World Health Organisation [1], one third of the global human population is infected with tuberculosis (TB) but most are latent infections. In people with no underlying risk factors, approximately 10% will develop an active infection during their lifetime [1]. However, underlying deficiencies in immunity caused by co-infections such as HIV, or co-morbidities such as cancer, greatly increase the chance of active infection. This high level of latent infection suggests a period of close co-evolution of *Mycobacterium tuberculosis* and its human host [2]. Phylogenetics indicate that the *M. tuberculosis* complex (MTBC) emerged via an evolutionary bottleneck and that existing lineages have emerged after a succession of unidirectional deletion events [3]. *M. tuberculosis* is an obligate pathogen and has no environmental reservoir. There appears to be an association between *M. tuberculosis* lineages with different human populations around the globe and this persists within modern cities with a population of diverse origins [4]. An association has been found between population density and the emergence of human infectious diseases [5]. This association is apparent in the early Neolithic period of human development and in the case of TB appears to be continuing with the emergence of highly transmissible and virulent

* Corresponding author. Centre for Clinical Microbiology, Division of Infection & Immunity, Royal Free Campus, University College London, London NW3 2PF, UK. Tel.: +44 (0)2077940500x31146.
E-mail addresses: h.donoghue@ucl.ac.uk (H.D. Donoghue), spigelman@btinternet.com (M. Spigelman), Justin.OGrady@uea.ac.uk (J. O'Grady), szikossy@nhmus.hu (I. Szikossy), papildi@hotmail.com (I. Pap), o.y.lee@bham.ac.uk (O.Y.-C. Lee), h.wu.2@bham.ac.uk (H.H.T. Wu), g.besra@bham.ac.uk (G.S. Besra), d.e.minnikin@bham.ac.uk (D.E. Minnikin).
[1] Current address: Norwich Medical School, University of East Anglia, Norwich, UK.

http://dx.doi.org/10.1016/j.tube.2015.02.020
1472-9792/

strains of *M. tuberculosis* in major cities that have a long record of continuous habitation [6].

Leprosy is a chronic human infection caused by *Mycobacterium leprae*. This has declined in recent years but caused approximately 219,000 new cases in 2011 [7], mainly in South East Asia, Africa and South America. It is a major cause of preventable disability and of social exclusion due to stigma. *M. leprae* is extremely slow growing and requires to be in an intracellular environment within a host, primarily human. *M. leprae* also appears to have experienced an evolutionary bottleneck and subsequent clonal expansion between pathogen and host [8]. Different strains of *M. leprae* can be distinguished by variable number tandem repeat (VNTR) and short tandem repeat typing. These can indicate short-term transmission via microsatellite analysis but are unstable due to poor DNA repair by *M. leprae* [9]. Stable long-term changes can be monitored by synonymous single nucleotide polymorphisms (SNPs) and these have identified lineages that are also associated with different human populations [10].

Much can be inferred by the study of modern isolates of both *M. tuberculosis* and *M. leprae*. However, the direct study of ancient specific biomarkers for these pathogens, such as ancient DNA (aDNA) and cell wall lipid biomarkers, has distinct advantages [11]. These biomarkers enable confirmation of infection in skeletal or soft tissue remains with non-specific or no palaeopathology, as only about 5% of TB cases are believed to result in bony changes. Ancient biomarkers may also answer historical questions, such as the nature of pre-Columbian TB or the role played by the slave trade across the Atlantic in the dispersal of TB and leprosy to the Americas. Mixed infections can also be detected. Analysis of aDNA may permit the determination of genetic lineages, genotypes or subgenotypes in specimens of known age, thus providing real time calibration of the date of their emergence. The association of ancient pathogen genotypes with different host populations may also pinpoint and date human migrations [10,12].

A useful approach in palaeomicrobiology is to obtain independent verification of findings by seeking different specific biomarkers in individual specimens. Our group has concentrated on initial examination for aDNA and subsequent independent analysis of mycobacterial specific cell wall lipids [13]. Until recently, our aDNA data have been obtained by DNA amplification using PCR, first conventional single-stage or nested PCR, and more recently using real-time PCR with specific fluorescent probes and primers for selected target regions. As aDNA is often highly fragmented, the use of specific probes has been very productive as they enable specific detection of selected target regions of as little as 60–80 base pairs (bp). The development of Next Generation Sequencing (NGS) and sophisticated bioinformatic analysis has enabled sequencing and analysis of non-amplified DNA using targeted enrichment approaches [14,15]. Rarely, in exceptionally well-preserved material, it is possible to perform shot-gun sequencing without target enrichment and to obtain an analysis of the entire DNA within a sample. This has been achieved for *M. leprae* found in mediaeval dental pulp [15] and in extremely well preserved lung tissue from a naturally mummified individual from 18th century Vác, Hungary [16].

2. Ancient DNA methodology

2.1. Extraction of M. tuberculosis and M. leprae aDNA

The following protocol gives sufficient time for samples to be disaggregated, but includes vigorous bead beading and snap freezing in dry ice to release aDNA from association with any residual lipids from the lipid-rich mycobacterial cell wall. Small samples (bone scrapings 20–80 mg; mummified tissue 10–40 mg) are taken from human remains, according to recommended protocols for aDNA, with separate rooms and equipment for different stages of the process [17]. Skeletal material is crushed in a sterile pestle and mortar and samples are added to 400 μl of Proteinase K/EDTA. The slurry is incubated at 56 °C and mixed on a bead beater daily until solubilised. An aliquot is treated with 40 μl of 0.1 mol^{-1} of *N*-phenacylthiozolium bromide (PTB), to cleave any covalent cross-links thus enabling DNA strand separation and amplification [17]. As PTB is inhibitory in the PCR reaction, an aliquot without PTB is processed in parallel, so that short DNA fragments can be precipitated from PTB-free silica supernatants (see below). Sample tube contents are transferred into lysis buffer containing 5 mol^{-1} guanidium thiocyanate and incubated for 1–3 days at 56 °C. To complete the disruption of bone and any mycobacterial remnants, samples are boiled, then snap-frozen in liquid nitrogen and thawed in a 65 °C water bath. This procedure is repeated twice. Sample tubes are centrifuged at 5000 g for 15 min at 5 °C and the supernates carefully removed into clean, sterile tubes. DNA is captured with silica suspension (NucliSens®) and mixed on a rotator wheel for 1 h. Tube contents are centrifuged and silica pellets washed once with wash buffer (NucliSens®), twice with 70% (v/v) ethanol (−20 °C) and once with acetone (−20 °C). After drying in a heating block, DNA is eluted using 60 μl elution buffer (NucliSens®), aliquoted and used immediately or stored at −20 °C. Silica supernates (500 μl) from PTB-negative samples are also taken from the lysis buffer, and 2.0 ml screw-capped Eppendorf tubes used to wash the silica. After chilling at 5 °C, supernates are mixed vigorously for 20 s with 200 μl of Protein Precipitation Solution (SLS Ltd, UK) and centrifuged for 3 min at 10,000 g. Any pellet is discarded and 600 μl isopropanol (−20 °C) added to 550 μl of each supernate. Tubes are mixed by inversion 50 times and spun 3 min. Supernates are discarded and tubes washed once with 500 μl 70% ethanol (−20 °C). After draining, tubes are dried in a heating block. Any precipitated DNA is re-hydrated with 60 μl elution buffer (NucliSens®), aliquoted and used immediately or stored at −20 °C. Negative extraction controls are processed in parallel with the test samples.

2.2. DNA amplification and detection

In the current protocol, two specific regions of each organism are targeted, using repetitive elements to increase the likelihood of detection of pathogen aDNA. For the *M. tuberculosis* complex, IS*6110* (1–25 copies/cell) and IS*1081* (6 copies/cell) are used [18]. For *M. leprae*, RLEP (37 copies/cell) and REPLEP (15 copies/cell) are used [19]. Initially, conventional PCR was used, with primers targeting DNA regions within the range 90-bp to 124-bp. PCR was performed in two stages, with 45 rounds of amplification followed, if necessary, by a nested reaction using internal primers, with a further 25 cycles of amplification. PCR products were detected by agarose gel amplification, gel slices were removed, the PCR products purified and sequenced. As aDNA is highly fragmented specific primers and fluorescent probes have since been designed to enable shorter DNA fragments to be specifically detected (Table 1). The Qiagen QuantiTect® Probe reaction mix is used with additional 2 mM BSA to reduce PCR inhibition and additional 2.0 mM $MgCl_2$ to facilitate primer binding. A hot-start *Taq* polymerase is used to minimise non-specific primer and template binding. Negative DNA extraction and PCR controls are processed alongside the test sample. Amplification is performed in a final volume of 25 μl using the Qiagen RotorGene® real-time platform. After enzyme activation for 15 min at 95 °C, amplification consists of 50–55 cycles of strand separation at 94 °C for 10 s, primer binding at 60 °C for 20 s and strand extension at 72 °C for 10 s. The probes enable direct observation of specific amplicons and the determination of cycle threshold (Ct) indicates relative concentration of template. Findings may be