Integrated Strategies for the use of Lipid Biomarkers in the Diagnosis of Ancient Mycobacterial Disease

Oona Y-C. Lee¹, Ian D. Bull², Erika Molnár³, Antonia Marcsik³, György Pálfi³, Helen D. Donoghue⁴, Gurdyal S. Besra¹ and David E. Minnikin*¹

¹School of Biosciences
University of Birmingham
Edgbaston, Birmingham B15 2TT, UK

²Bristol Biogeochemistry Research Centre
School of Chemistry
University of Bristol, Bristol BS8 1TS, UK

³Department of Anthropology
University of Szeged
H-6701 Szeged, Hungary

⁴Research Department of Infection
University College, London W1T 4JF, UK

*Email for correspondence: d.e.minnikin@bham.ac.uk

Abstract
The use of ancient DNA and lipid biomarkers has become established for the detection of tuberculosis and leprosy in archaeological material. Long-chain compounds are released by an efficient non-aqueous alkaline extraction and acidic components are converted to stable pentafluorobenzyl (PFB) esters, which can be preserved for immediate or future analysis. These long-chain components are fractionated into non-hydroxylated fatty acid esters and mycolic acid esters. Characteristic profiles of mycolates are recorded by fluorescence high-performance liquid chromatography (HPLC) of pyrenebutyric acid derivatives of PFB esters. Mycocerosate PFB esters are analysed by negative-ion chemical-ionisation gas chromatography-mass spectrometry (NICI-GCMS). In this study, ancient DNA analyses, along with mycolate and mycocerosate profiles, for a representative late 7th century AD skeleton from the Avar period in Hungary, are presented. It will be shown how mycolate and mycocerosate analyses of residual material from an ancient DNA analysis can be used to confirm a mixed leprosy-tuberculosis infection and provide an estimate of the relative amounts of the \textit{M. tuberculosis} and \textit{M. leprae} bacterial load obtained. In the present case, the bone pathology indicated only leprosy, but the biomarker analysis suggested predominance of tuberculosis over leprosy.

Keywords: Tuberculosis, Leprosy, Biomarkers, Lipids, Skeletons, Paleopathology

Introduction
The aim is to develop the use of specific lipid biomarkers, in addition to DNA, in the diagnosis of ancient tuberculosis and leprosy. An integrated approach is used to analyse the established lipid biomarkers, mycolic and mycocerosic acids. The mycolic acids (Fig. 1A) are high molecular weight, 70 to 90 carbon, fatty acids, which are linked into the cell envelope to provide a tough hydrophobic defensive permeability barrier (Minnikin 1982; Minnikin \textit{et al.} 2002). The mycocerosates (Fig. 1B) are multi-methyl branched fatty acids, components of waxes that also contribute to the hydrophobic barrier (Minnikin \textit{et al.} 2002). A brief “Glossary of Technical Terms” is provided later.

In pioneering studies, using fluorescence high performance liquid chromatography (HPLC) of methylanthryl derivatives, mycolic acid profiles complemented the detection of ancient DNA in confirming the presence of tuberculosis in calcified pleura (Donoghue \textit{et al.} 1998) and a medieval skeleton (Gernaey \textit{et al.} 2001). The presence of \textit{Mycobacterium tuberculosis} infection was also found to correlate well with burial records in a skeletal collection from Newcastle Infirmary (Gernaey \textit{et al.} 1998). The fluorescence HPLC
protocol was updated to use pyrenebutyric acid (PBA) derivatives of mycolic acid pentafluorobenzyl (PFB) esters (Fig. 1A). Using this method, the oldest proven case of tuberculosis was confirmed in a 9,000 year old female skeleton and an adjacent child (Herschkovitz et al. 2008). Similarly, tuberculosis was confirmed in the British Museum Granville Mummy (Donoghue et al. 2010) and a case of leprosy was identified from Uzbekistan (Taylor et al. 2009). The value of mycocerosate biomarkers for tuberculosis was demonstrated in the Coimbra Skeletal Collection, using negative ion chemical ionisation gas chromatography mass spectrometry (NICI-GCMS) (Redman et al. 2009).

This paper illustrates the integrated use of ancient DNA and mycolic and mycocerosic acid analysis in detecting ancient tuberculosis and leprosy, with particular reference to a mixed infection.

**Analytical Methods**

The archaeological material studied was a nasal scraping of skeleton KD 517 from the late 7th century site at Kiskundorozsma, Hungary (Molnár et al. 2006; Pálfi and Molnár 2009). This was a male, aged 35–40 years, with distinctive paleopathology indicative of leprosy; there was pronounced remodelling of the rhinomaxillary region, including abnormal porosity of the hard palate and slight changes to the tibiae and fibulae. In addition, bilateral porotic criba orbitalia was noted. There was palaeopathology for leprosy and/or tuberculosis in other skeletons from this site, so KD 517 was examined, for both *Mycobacterium leprae* and *M. tuberculosis* complex DNA, using protocols described previously (Donoghue et al. 2005).

To analyse lipid biomarkers, the residual extract from the DNA study on KD 517 (Molnár et al. 2006) was used along with standard *M. tuberculosis* and *M. leprae* biomass. Alkaline hydrolysis and
derivatisation, as described previously (Hershkovitz et al. 2008; Redman et al. 2009; Taylor et al. 2009; Donoghue et al. 2010), produced fractions containing pentafluorobenzyl esters of mycocerosic and mycolic acids. The PFB-mycocerosates were analysed by selected ion monitoring negative ion chemical ionisation gas chromatography mass spectrometry (SIM NICI-GCMS) (Redman et al. 2009). PFB-mycolates were converted to pyrenebutyric acid (PBA) esters (Fig. 1A), which were analysed by fluorescence HPLC (Hershkovitz et al. 2008; Donoghue et al. 2010). Profiles were recorded for standard samples of M. tuberculosis H37Rv and M. leprae, remaining from a previous study (Draper et al. 1983).

**Results**

In the initial ancient DNA analysis of KD 517, a specific region of the M. leprae RLEP region was targeted, as this is present at 32 copies/cell and is therefore very sensitive. A 129-bp amplicon was obtained and confirmed by nested PCR, which yielded a 99-bp product. In addition, primers for the single-copy 18kD antigen gene gave a 136-bp product, confirmed by nested PCR giving a 110-bp amplicon (Donoghue et al. 2005). M. tuberculosis complex-specific PCR primers for the multi-copy IS6110 region yielded a 123-bp amplicon that was confirmed by production of a 92-bp nested product.

The mycolic and mycocerosic acid profiles, obtained from extracts of skeleton KD 517 are shown in Figures 2 and 3, respectively. In the reverse phase HPLC profiles of total mycolates (Fig. 2A), the profile from KD 517 showed a closer superficial resemblance to M. tuberculosis than M. leprae, suggesting a predominance of the former. The total mycolates (Fig. 2A) were collected and analysed by normal phase HPLC, as shown in Figure 2B. The presence of a probable methoxymycolate component (Fig. 1) in the KD 517 profile was also indicative of M. tuberculosis, as
methoxymycolates are absent in \textit{M. leprae} (Fig. 2B) (Minnikin \textit{et al.} 1985). The fractions corresponding to α-mycolates were analysed by reverse phase HPLC (Fig. 2C), with a main C$_{80}$ component in KD 517 again suggesting a predominance of \textit{M. tuberculosis}. However, a relatively enhanced proportion of the C$_{76}$ α-mycolate introduced a possibility of leprosy (Fig. 2C). Tuberculosis infection was confirmed by a very clear methoxymycolate reverse phase HPLC profile, corresponding well with that from \textit{M. tuberculosis} (Fig. 2D). The ketomycolate reverse phase HPLC profile from KD 517 showed a main C$_{87}$ component, emphasising the predominance of \textit{M. tuberculosis}, but a significant C$_{83}$ ketomycolate strongly suggested the presence of \textit{M. leprae} (Fig. 2E). The ratio of α-, methoxy- and ketomycolates (100:44:33) in KD 517 (Fig. 2B) was similar to a ratio (100:47:34), calculated for a 3:2 approximate ratio of the \textit{M. tuberculosis} and \textit{M. leprae} standards. Similarly in KD 517, assuming this \textit{M. tuberculosis} to \textit{M. leprae} bacterial loading of 3:2, the total ketomycolates and its component peaks of C$_{83}$ and C$_{87}$ have a ratio of 100:14:30 (Fig. 2E), which corresponds to the calculated value of 100:17:29 for the 3:2 mixture.

The profiles recorded for mycocerosic acids are shown in Figure 3. The mycocerosates of KD 517 displayed four major C$_{29}$, C$_{30}$, C$_{32}$ and C$_{34}$ components (Fig. 3A). This suggests a mixed infection, as the mycocerosates of \textit{M. tuberculosis} lack C$_{34}$ components (Fig. 3B) and those from \textit{M. leprae} have minor amounts of C$_{30}$ mycocerosates (Fig. 3C). The C$_{34}$ component in the profiles from the \textit{M. tuberculosis} standard (Fig. 3B), with a retention time of 27.05 min, is not a mycocerosate, as authentic C$_{34}$ mycocerosates have retention times around 30 min (Fig. 3C). Again, assuming a 3:2 \textit{M. tuberculosis} loading, the 27:65:100:81 ratio of the C$_{29}$, C$_{30}$, C$_{32}$ and C$_{34}$ mycocerosates in KD517 (Fig. 3A) corresponded well for the C$_{29}$, C$_{30}$ and C$_{32}$ components in the ratio (28:67:100:42) calculated for the 3:2 mixture, but not so well for the C$_{34}$ mycocerosate.

\textbf{Discussion}

This communication illustrates how integrated analysis of ancient DNA and the principal lipid biomarkers, mycolic and mycocerosic acids, can provide a clear diagnosis of a mixed infection of ancient tuberculosis and leprosy. Since the DNA protocols use only aqueous extraction media, hydrophobic lipids remain in the residual materials. Strict protocols (Donoghue \textit{et al.} 2005; 2009) were followed in the DNA study to avoid any contamination of the residues, so that they could be examined for lipid biomarkers.

The ancient DNA results demonstrated the presence of clear amplicons for both leprosy and tuberculosis. There was clearly sufficient well-preserved \textit{M. leprae} DNA to enable amplification of a 136-bp target sequence from the single-copy 18 kD antigen locus. However, only the IS6110 locus was examined for the \textit{M. tuberculosis} complex. This is multi-copy and the copy number varies between strains. In addition, DNA is inherently less stable than lipid biomarkers and it is likely that its preservation is not consistent, either between
different sites on the bacterial genome, or between different microbial species. Therefore, no conclusions can be drawn from the DNA study regarding the relative quantities of microbial DNA from each organism.

The mycolic acid profiles, in Figure 2, show the overall strategy for the recognition of components characteristic of tuberculosis and leprosy. The initial “reverse-phase” mycolate HPLC analysis (Fig. 2A) involves a partition between a “non-polar” hydrophobic surface, with 18-carbon attached long chains, and a relatively “polar” solvent in which the mycolates are only partially soluble. This type of chromatography is essentially a size screen enabling the high molecular weight C\textsubscript{70} to C\textsubscript{90} mycolic acids to be distinguished from any mammalian lipids, which all have less than 60 carbons. Total mycolates are collected and subjected to “normal” phase HPLC (Fig. 2B), which separates different mycolate classes dissolved in a relatively “non-polar” solvent on a “polar” surface according to their overall polarity. This polarity is increased in mycolates by adding polar functions to the molecules. The α-mycolates (Fig. 1A) are the least polar but adding a methoxy group increases polarity, as the oxygen atom can bond strongly to the polar silica gel “normal” phase surface. The oxygen atom in the keto group bonds even more strongly to the silica gel so ketomycolates are eluted last (Fig. 2B). The standard normal phase HPLC profiles for \textit{M. tuberculosis} and \textit{M. leprae} are distinguished by the lack of methoxymycolates in the latter (Minnikin et al. 1985). The individual mycolate types are collected from the normal phase separation and each is subjected to a second reverse phase analysis to reveal the different homologous components present (Figs 2C-E). For \textit{M. tuberculosis} and \textit{M. leprae}, respectively, the main α-mycolates are C\textsubscript{80} and C\textsubscript{78} and the ketomycolates are C\textsubscript{87} and C\textsubscript{83}, the latter difference being the most discriminatory. The mycolate profiles from KD 517 show features indicating the presence of infection by both \textit{M. tuberculosis} and \textit{M. leprae}. Tuberculosis is confirmed by the presence of methoxymycolates in the normal phase profiles (Fig. 2B) and a highly characteristic methoxymycolate profile, comparing well with that given by the \textit{M. tuberculosis} standard (Fig. 2D).

The ketomycolate profile of KD 517 (Fig. 2E) is particularly informative, showing two distinct clusters of peaks centred on C\textsubscript{83}, for \textit{M. leprae} and C\textsubscript{87} for \textit{M. tuberculosis}. These two clusters correlate well with the profiles for standard \textit{M. leprae} and \textit{M. tuberculosis} (Fig. 2E). The ratio of the main C\textsubscript{87} and C\textsubscript{83} components in the ketomycolate profiles from KD 517 can provide an indication of the relative proportions of tuberculosis and leprosy infections. The profile of the α-mycolate of KD 517 (Fig. 2C) is less informative, with only an enhanced amount of the C\textsubscript{76} α-mycolate indicating that this may not be a homogeneous tuberculosis infection. Also, the fact that the main α-mycolate in KD 517 is C\textsubscript{80}, rather than C\textsubscript{78}, suggests a predominance of tuberculosis over leprosy. As shown in Results, a 3:2 ratio of tuberculosis to leprosy correlates with the HPLC proportions recorded in Figure 2.

The multimethyl-branched mycocerosic acids (Fig. 1B) are recognisable by their appearance as characteristic double peaks (Fig. 3), resulting from hydrolytic racemisation to give “diastereoisomers” which can be distinguished on gas chromatography (Minnikin et al. 1993; Redman et al. 2009). Another characteristic recognition feature, for mycocerosates, is the fact that the C\textsubscript{29} and C\textsubscript{30} components have almost identical gas chromatographic retention times of around 20 min, in the present study (Fig. 3). Since the C\textsubscript{29} acid is smaller by one carbon unit (Fig. 1B), it would be expected to be eluted before a C\textsubscript{30} acid, but mycocerosates with three methyl branches, rather than four, are relatively less volatile, so a characteristic chromatographic coincidence occurs (Fig. 3). The C\textsubscript{29} and C\textsubscript{30} mycocerosates are, of course, readily distinguished by their ions at m/z 437 and m/z 451 (Fig. 3); these ions correspond to negative carboxylate ions (-COO\textsuperscript{−}), which are characteristically produced from pentafluorobenzyl esters (Fig. 1B) on NICI mass spectrometry.

The recorded profiles of mycocerosic acids (Fig. 3) gave clear evidence for a tuberculosis-leprosy co-infection. Essentially, in the profile from KD 517 (Fig. 3A), the presence of C\textsubscript{34} mycocerosate suggests leprosy (Fig. 3C) (Draper et al. 1983; Minnikin et al. 1993) and enhanced C\textsubscript{29} is characteristic of tuberculosis (Fig. 3B). As noted above, the C\textsubscript{34} component, in the profiles from the \textit{M. tuberculosis} standard (Fig. 3B), is not a mycocerosate. The essential patterns for tuberculosis (Fig. 3B) and leprosy (Fig. 3D) are highlighted by an “L-shape” and a “crescent shape”, respectively. The profile for KD 517 can be considered as a superposition of these two patterns (Fig. 3A). Mass spectrometry is not intrinsically very quantitative, but the assumption that there is a 3:2 predominance of tuberculosis over leprosy is broadly supported by the mycocerosate data.
It must be emphasised that these estimates of disease prevalence have limited accuracy. The mycolate HPLC analysis has the potential to be very accurate, if higher resolution columns are employed; however, under the conditions used, the key peaks in Figure 2 are not all clearly separated. Also, the distribution of tuberculosis and leprosy is unlikely to be uniform throughout skeletons. It is interesting that, in the nasal scraping analysed, there was an apparent predominance of tuberculosis mycolic and mycocerosic acids over those characteristic of leprosy, as the nasal area is a favoured site for leprosy infection. This does not necessarily imply that there were more tubercle than leprosy bacilli in the entire body. Results may also vary over time, according to the relative progression of each individual disease in the infected subject. The proportions of individual mycolates in the \textit{M. tuberculosis} and \textit{M. leprae} standards will also naturally vary somewhat, again diluting the precision of such diagnoses. Overall, although the pathology indicates an early leprosy infection, this is a disease that takes several years to progress to the stage of visible lesions. However, tuberculosis is a more virulent infection that kills the majority of susceptible individuals before any skeletal palaeopathology can develop. Therefore, the finding of a greater proportion of \textit{M. tuberculosis} lipid biomarkers is consistent with the scenario of an individual infected with leprosy succumbing to tuberculosis (Donoghue \textit{et al.} 2005). However, the possibility of death from a totally unrelated cause cannot be discounted.

The present study illustrates the benefit of assembling a wide range of evidence to support, or otherwise, an initial suspicion of mycobacterial disease, based solely on pathology. Detailed pathological investigations (Molnár \textit{et al.} 2006; Pálfi and Molnár 2009) clearly suggested a diagnosis of leprosy in KD 517, with no obvious pathology suggestive of tuberculosis.

It is important when using specific biomarkers, such as DNA and lipids, to follow a strict and well-designed protocol. In recent years, it has been necessary to clearly define important criteria for the use of ancient DNA (Donoghue \textit{et al.} 2009) and mycolic acids (Minnikin \textit{et al.} 2010) in the diagnosis of ancient mycobacterial disease. A significant advantage of lipid biomarkers is that they are detected directly and, due to the great sensitivity of the methods, femtogram quantities can be detected without any amplification. The present study demonstrates the range and rigour of diagnostic data that can be achieved by an integrated, objective approach to biomarker use in the diagnosis of ancient mycobacterial disease. The detection of both DNA and lipid biomarkers in the same sample is a step forward in realising the maximum amount of information from valuable archaeological samples.

**Glossary of Technical Terms**

- **Biomarker:** a characteristic discriminatory cellular component.
- **DNA:** deoxyribonucleic acid, a key molecule carrying essential genetic information.
- **Gas chromatography:** a separation process where compounds are carried along in a gaseous “mobile phase” over a discriminatory “stationary phase” in a column.
- **Lipid:** a fatty, waxy cellular material insoluble in water.
- **Liquid chromatography:** a separation process where compounds are carried along in a liquid “mobile phase” over a discriminatory “stationary phase” in a column.
- **Mass spectrometry:** a technique for giving molecules an electric charge (ionisation) and determining their mass by focussing them with magnetic and electrical fields.
- **PCR:** polymerase chain reaction, a means for making multiple copies of DNA fragments.

**Acknowledgements**

Support was obtained from The Leverhulme Trust in an Emeritus Fellowship (DEM) and Project Grant F/00 094/BL (GSB, DEM, OY-CL). Provision of access to the UK Natural and Environmental Research Council Mass Spectrometry Unit at the University of Bristol is acknowledged (DEM, IDB).

**Literature Cited**


Biomarkers for Ancient Mycobacterial Disease


