Accepted Manuscript

Evolutionary changes in the genome of Mycobacterium tuberculosis and the human genome from 9000 years BP until modern times

Dr Mark Spigelman, Helen D. Donoghue, Ziad Abdeeb, Suheir Ereqat, Issa Sarie, Charles, L. Greenblatt, Ildikó Pap, Ildikó Szikossy, Israel Hershkovitz, Gila Kahila Bar-Gal, Carney Matheson

PII:          S1472-9792(15)00023-2
DOI:         10.1016/j.tube.2015.02.022
Reference:   YTUBE 1283

To appear in:  Tuberculosis


This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.
Evolutionary changes in the genome of *Mycobacterium tuberculosis* and the human genome from 9000 years BP until modern times

Mark Spigelman\(^1,2,3,\*\), Helen D Donoghue\(^1,4,\*\), Ziad Abdeeb\(^5\), Suheir Ereqat\(^5\), Issa Sarie\(^5\), Charles, L. Greenblatt\(^3\), Ildikó Pap\(^6\), Ildikó Szikossy\(^6\), Israel Hershkovitz\(^2\), Gila Kahila Bar-Gal\(^7\), Carney Matheson\(^8\)

\(^1\)Centre for Clinical Microbiology, Division of Infection & Immunity, University College London, London, UK

\(^2\)Department of Anatomy and Anthropology, Sackler Faculty of Medicine, Tel Aviv University, Tel Aviv, Israel

\(^3\)Kuvin Center for the Study of Infectious & Tropical Diseases and Ancient DNA, Hadassah Medical School, The Hebrew University, Jerusalem, Israel

\(^4\)Centre for the History of Medicine, Division of Biosciences, University College London, London, UK

\(^5\)Al-Quds Nutrition and Health Research Institute, Faculty of Medicine, Al-Quds University, Abu-Deis, P.O. Box 201760, West Bank, Palestine

\(^6\)Department of Anthropology, Hungarian Natural Science Museum, Budapest, Hungary

\(^7\)Koret School of Veterinary Medicine, The Hebrew University of Jerusalem, Rehovot, Israel

\(^8\)Paleo-DNA Laboratory, Departments of Anthropology and Biology, Lakehead University, Thunder Bay, Ontario, Canada

+ Equal First authors

Email addresses:

spigelman@btinternet.com; h.donoghue@ucl.ac.uk; zabdeen13@gmail.com;
nereqat@med.alquds.edu; isarie63@gmail.com; charlesg@ekmd.huji.ac.il; papildi@hotmail.com;
anatom2@post.tau.ac.il; gila.kahila@mail.huji.ac.il; cmatheso@lakehead.ca;

\*Corresponding author: Dr Mark Spigelman, 2 Clarence Terrace, Regents Park, London NW1 4RD

Telephone: 44(0)2072249095 e-mail: spigelman@btinternet.com

Word count summary: 193

Word count (excluding summary and references): 2910
Summary

The demonstration of *Mycobacterium tuberculosis* DNA in ancient skeletons gives researchers an insight into its evolution. Findings of the last two decades sketched the biological relationships between the various species of tubercle bacilli, the time scale involved, their possible origin and dispersal. This paper includes the available evidence and on-going research. In the submerged Eastern Mediterranean Neolithic village of Atlit Yam (9000 BP), a human lineage of *M. tuberculosis*, defined by the TbD1 deletion in its genome, was demonstrated. An infected infant at the site provides an example of active tuberculosis in a human with a naïve immune system. Over 4000 years later tuberculosis was found in Jericho. Urbanization increases population density encouraging *M. tuberculosis*/human co-evolution. As susceptible humans die of tuberculosis, survivors develop genetic resistance to disease. Thus in 18th century Hungarian mummies from Vác, 65% were positive for tuberculosis yet a 95-year-old woman had clearly survived a childhood Ghon lesion.

Whole genome studies are in progress, to detect changes over the millennia both in bacterial virulence and also host susceptibility/resistance genes that determine the NRAMP protein and Killer Cell Immunoglobulin-like Receptors (KIRs). This paper surveys present evidence and includes initial findings.

**Key words:** Ancient DNA; evolution; KIR historical specimens; *Mycobacterium tuberculosis*; *SLC11A1* gene; Solute Carrier family genes
1. Introduction

Microbial infections played a key role in shaping life on earth and have been a major selector for the evolution of all present species. Evidence exists that demonstrate infectious diseases were already present in our remote ancestors.\(^1,2\) Considering the impact of *Mycobacterium tuberculosis* (MTB), in all probability it has had a greater influence on the genetic selection of the *Homo sapiens* population than any other infectious agent.

The molecular identification of human pathogens in ancient human remains has recently opened new scientific fields that provide considerable insight into the history and evolution of host, pathogen and their interaction. This allows us to track changes in the ancestral tubercle bacillus as it became more and more exposed to the internal environment and immune system of its human host. Conversely, it is possible to track changes in the genes of the human population that confer resistance or susceptibility to disease over time.

TB is related to population density,\(^3\) transmitted from human to human living in close contact. However, the origin of the disease, the earliest hosts of MTB and its evolution remain unclear. The evolution of the bacteria cannot be considered in isolation. It is important to realise how TB has influenced the human development over the millennia, particularly our resistance/susceptibility genes. MTB experienced an evolutionary bottleneck when it became an obligate pathogen and has a clonal relationship with different human lineages.\(^4\) Subsequent co-evolution has resulted in the majority of TB infections being latent. In past eras of low human population density, MTB adapted over time in response to host-adaptive changes and *vice versa*. This process, which can be defined as mutualism, is a biological interaction between individuals of two different species where both individuals derive a fitness benefit. As the host becomes more resistant, strains better able to colonise the resistant host will predominate, thus starting off another cycle. More virulent MTB strains will attack their human host, killing the most susceptible and leaving the more resistant as survivors. However, when human populations were sparse, this could break the chain of transmission of the pathogen. The development of antibiotics has shortened the mutualistic cycle significantly, but the combination of HIV co-infection, antimicrobial therapy and increased global human population density is leading to the emergence of some MTB strains that are both more transmissible but also more virulent.\(^5\)

2. The impact of palaeomicrobiological investigations of archaeological human material

2.1 Questions to be addressed

Archaeologists should seek to correlate research questions with historical events. For example, did past invasions introduce new pathogens, or more virulent strains of pathogens into susceptible populations? Thousands of indigenous peoples in the Americas died from exposure to European strains of MTB, measles and smallpox.\(^6\) Another possible scenario is that invaders may have
brought new pathogens with them on return to their place of origin. A good example of this is the
introduction by European colonialists of venereal syphilis from South America.

A further question one has to ask is what was the genetic status of Homo erectus or
predecessor species regarding the underlying genetic basis of host resistance and susceptibility to
tuberculosis. Did ancestral hominids have the precursors of modern host susceptibility/resistance
genes or were these acquired late? Is the ‘Out of Africa’ theory of the origin of human TB proposed
by Gutierrez et al? capable of being verified by a study of human remains, or will these show that
TB developed in several areas and that this is the explanation for the variability of the organism in
different geographical areas?

The majority of TB patients in the world today never progress to active disease. The World
Health Organisation estimates that approximately one-third of the global population is infected but
only 10% of immunocompetent progress to active disease during their lifetime. Our current
immunity may be the result of Darwinian selection only, or may depend upon whether particular
genres are switched on or off – a mechanism that can result in rapid adaptation. It must be
remembered that other non-genetic factors influence human susceptibility to infection such as
dietary deficiencies, stress and trauma. Long-term climatic changes have an impact on vegetation
and agriculture whereas local variations in climate may influence transmission of MTB by
infectious aerosols. Temperature changes will determine whether humans spend more time in the
open air or enclosed spaces, for example.

2.2 Significant findings

With the first reported finding of MTB DNA in ancient skeletons based on amplification of a small
(123 bp) DNA target that was specific for the MTB-complex a new era of research into microbial
pathogen evolution became possible. In addition to skeletal remains, calcified and mummified
tissues also proved to be good sources of MTB ancient DNA (aDNA). Our knowledge was
enhanced with the finding of MTB in a 17000-year-old Pleistocene bison from Natural Trap Cave,
Wyoming. Spoligotyping revealed that the Pleistocene bison lesions contained aDNA from the
M. tuberculosis complex, possibly MTB or Mycobacterium africanum, but distinct from
Mycobacterium bovis. The consensus bison spoligotyping pattern was compared with the
combined database collated by the National Institute of Public Health and Environment (RIVM),
Utrecht, The Netherlands and the Veterinary Science Division, Department of Agriculture and
Rural Development, Belfast, N. Ireland. No exact matches were found on the database. However,
in a computer analysis comparing a library of defined species, the highest similarity was from M.
africanum (82.3%), then M. tuberculosis - MTB (76.6%), with M. bovis having only 72.7% similarity.

The original aDNA findings in the Pleistocene bison were confirmed ten years later by finding
species-specific MTB cell wall lipid biomarkers. We have used this method of independent
confirmation of our MTB aDNA findings since 1998 because lipid analysis uses methods based
on the direct detection of femtogram quantities of target molecules, with no need for any
amplification. This is a more rigorous method of independent confirmation than sending part of the specimen to another laboratory for analysis.

The Pleistocene bison contained MTB-complex aDNA but the particular lineage has not yet been identified. The earliest known human MTB was detected and characterised in samples from the submerged Neolithic site of Atlit Yam, a 9000-year-old settlement submerged in the sea off the coast of Haifa in Israel. The findings were confirmed by lipid analysis and the preservation was sufficiently good that it was possible to confirm that the MTB had experienced the TbD1 deletion, found only in human lineages. This is of particular significance as this was a Pre-Pottery site with the earliest evidence of animal domestication in the Levant.

We were fortunate as a group to secure samples from two large collections of natural mummies – one from the 18th to early 19th century from Vác, Hungary and the second from early Christian Nubia dated to 500-1400 CE at Kulubnarti in Northern Sudan. The importance of these collections was that the DNA preservation is well above average as in both locations the bodies were naturally mummified with no chemicals used. Indeed, the Kulubnarti material demonstrated co-infections of MTB with Leishmania spp, and using the Hungarian material, it was possible to determine the main MTB genetic lineages and perform molecular typing. Our work on the Pleistocene bison together with the Hungarian Vác mummies was cited and assisted in developing the hypothesis proposed in an excellent early paper on MTB evolution by Brosch et al.

To fill the time gap between the Nubian Kulubnarti mummies and the Atlit Yam skeletal remains, specimens from the Bronze Age township of Jericho have been examined. Initially bones from early excavations from the 1950’s were studied, in a collaboration involving colleagues from Munich, Al Quds University and Jerusalem. Unfortunately, although these specimens yielded possible MTB aDNA, this could not be confirmed independently. Material from the excavation of Ain es-Sultan refugee camp area, where ancient Jericho (Tel es-sultan) ~4000 BC has yielded MTB aDNA, which has been confirmed by lipid analysis. The infecting pathogen was from a TbD1-deleted MTB lineage. At present a metagenomic study on this specimen is in progress at McMasters University.

The Hungarian mummy project based on 265 bodies, most wholly or partially mummified, from a sealed crypt, is unique as there is contemporaneous archival information about many of the individuals. This enabled the identification of some family groups and also made it possible to study TB in a large population from a fixed period and single location. It was possible to type the MTB aDNA within a family and to show that each member was infected with a slightly different strain. Recently, lung tissue from the older daughter in this family group has been shown by non-enriched whole genome sequencing, to contain two different strains of MTB, with apparent sequential deletions, that appear to be ancestral to a modern outbreak strain in Germany. In contrast, MTB aDNA was found in a calcified lymph node from the mediastinum of a 95-year-old mummy, where initially all tissues were negative but an X-ray showed the calcified node. This
demonstrates that in this well-preserved group of mummies it is possible to identify cases of active and of latent infection. It was these findings that led to our interest in host susceptibility and resistance genes.

3. Host susceptibility and resistance

In addition to the retrieval of the pathogen DNA, a pilot study is investigating the genes believed to be responsible for susceptibility or resistance to the disease to determine if these genes differ in any way between those who were infected and those who appear immune. The study of the host susceptibility/resistance factors in the mummies and their descendants will give information on the role of host genetics in the pathogenesis of infectious disease, and contribute to the design of new therapeutic strategies. The study involves two host targets, the \textit{SLC11A1} gene (previously named \textit{NRAMP}) and Killer Cell Immunoglobulin-like Receptor genes (KIRs). The plan is to seek any correlation between presence and absence of tuberculosis, with the presence of certain alleles in these resistance genes. Already, our initial research on material from Hungarian and Sudanese mummies has revealed some interesting genetic patterns.

KIRs are members of a group of regulatory molecules found on subsets of lymphoid cells, first identified by their ability to impair some specificity on natural killer (NK) cytolysis. The \textit{KIR} locus, which maps to chromosome 19q(13.4) within the 1 Mb \textit{Leukocyte Receptor Complex (LRC)}, contains a family of polymorphic and highly homologous genes. KIR genes are tandemly arrayed over a physical distance of about 150 Kb, displaying the remarkable feature of gene content variation among haplotypes. The KIR molecules recognize the Human Leukocyte Antigen (HLA) class I molecules, which are encoded by genes within the Major Histocompatibility Complex (MHC) chromosome 6. Interactions between KIR isotypes that inhibit natural killer (NK) cell activity and specific HLA class I allotypes protect healthy cells from spontaneous destruction by NK cell mediated cytolysis. Other KIR isotypes stimulate the activity of NK cells demonstrating that KIR play a significant role in the control of the innate immune response. Recent studies report a greater repertoire of inhibitory KIR genes among TB patients than controls and a direct association of certain KIR and HLA-C genes with resistance to pulmonary TB. Different KIR genes have a role in inhibiting or increasing susceptibility towards TB and the complimentary MHC ligands need to be tested for the functional relevance of the associated genes.

A contemporaneous study of the \textit{SLC11A1} gene is in progress at Lake Head University. The promoter region has been studied in modern populations and been linked to a number of infections and autoimmune diseases, caused by \textit{M. tuberculosis}, \textit{M. bovis}, \textit{Mycobacterium leprae}, \textit{Mycobacterium lepraemurium}, \textit{Salmonella typhimurium} and \textit{Leishmania donovani}. The identification of sequence variants has prompted research into the evolution of nuclear genes, inheritance patterns, selective pressures, and changes in both allele frequencies and disease linkages over time. Linkage studies can help ascertain the resistance and susceptibility factors of diseases and can assist modern medicine by providing a better understanding of the infectious
processes themselves.\textsuperscript{25,26} The Allele 2 variation of the promoter region was found to be present in every patient infected with tuberculosis, indicating that this level of allelic expression may well be related to the resistance or susceptibility of an individual to infectious diseases. Allele 3 seems to produce the highest level of $SLC11A1$ expression, which confers a resistance to microbial infection to the individual, but increases susceptibility to autoimmune diseases. Conversely, Allele 2 produces the lowest level of $SLC11A1$ expression, conferring individual resistance to autoimmune diseases, but also a greater susceptibility to microbial infections. It is possible that this contradiction in allelic expressions may have resulted from inverse selective pressures, serving to maintain both alleles within the human population. Allelic variants of $SLC11A1$ have been identified as risk factors for paediatric TB.\textsuperscript{27} Other studies of host susceptibility and resistance genes have indicated that different human lineages may exhibit differing susceptibilities to TB infection.\textsuperscript{28} There is also limited evidence that genetic expression may vary according to sex and age.\textsuperscript{29} An intriguing finding is that human genetic susceptibility varies according to the differing clinical forms of TB.\textsuperscript{30} Limited data are now available on amplified aDNA (Tables 1 and 2) from 18 individuals from 18\textsuperscript{th} century Vác, Hungary and early Christian Nubia (Table 2).\textsuperscript{25} The promoter microsatellite polymorphisms of the $SLC11A1$ gene look encouraging as patterns are emerging (Table 2). Both the KIR and $SLC11A1$ studies are on-going and results will be disclosed on completion.

4. Conclusions

This study seeks to show the progress that has been achieved in paleomicrobiological research over the last two decades and indicates its contribution to the study of human pathogen co-evolution. Understanding the adaptations that the host and the pathogen have undergone through history, together with the resistance/susceptibility adaptations, may shed light on future interactions of humans with MTB. It is highly important to understand the process of mutualism – the biological interaction between individuals of two different species, where each derives a fitness benefit – in the present era of personalized medicine.

Acknowledgements

The contribution made by our many collaborators, researchers and students is gratefully acknowledged. Special acknowledgement is due to Dr Angela Gernaey (deceased) who helped pioneer the early mycolic acid work on the bison bone.

Ethical approval

Not required

Funding

Recent research has been partly supported by Hungarian National Scientific Research Foundation (OTKA) Grant No. 61166 (IP, IS). The German National Science Foundation (Deutsche
Forschungsgemeinschaft)- DFG grant number NE575/4-1 supported the early Jericho work.

**Author contributions**

MS conceived the original aDNA studies and HD, GKB-G, SE and CM performed experiments. HD, GKBG, CG and CM analysed ancient DNA data. MS, IH, ZA, IP and IS provided data and supplied specimens. Is S is head archaeologist of the Jericho excavations MS wrote the first and final drafts, HD prepared revised drafts and all authors approved the final version.

**Competing interests**

None declared

**Table legends**

**Table 1.** The SLC11A1 gene promoter microsatellite primer set

**Table 2.** The repeats identifying each SLC11A1 allele

**Table 2.** Genotypes of the SLC11A1 gene found in Hungarian and Nubian Mammies.
References


Table 1. The SLC11A1 gene promoter microsatellite primer set

<table>
<thead>
<tr>
<th>Primer</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>C1</td>
<td>ACT CGC ATT AGG CCA ACG AG</td>
</tr>
<tr>
<td>C2(FAM)* (6FAM)</td>
<td>TTC TGT GCC TCC CAA GTT AGC</td>
</tr>
</tbody>
</table>

- The antisense primer marked with florescence dye
- The primer was published by Bellamy et al., 1998

Table 2. Genotypes of the SLC11A1 gene found in Hungarian and Nubian Mammies

<table>
<thead>
<tr>
<th>Sample</th>
<th>Allele</th>
<th>Genotype</th>
<th>M. tuberculosis infection</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2/3</td>
<td>Heterozygote</td>
<td>Positive chest</td>
</tr>
<tr>
<td>2</td>
<td>2/3</td>
<td>Heterozygote</td>
<td>Positive chest</td>
</tr>
<tr>
<td>3</td>
<td>2</td>
<td>Homozygote</td>
<td>Positive chest and abdomen</td>
</tr>
<tr>
<td>4</td>
<td>2</td>
<td>Homozygote</td>
<td>Positive chest and abdomen</td>
</tr>
<tr>
<td>5</td>
<td>2</td>
<td>Homozygote</td>
<td>Positive chest and abdomen</td>
</tr>
<tr>
<td>6</td>
<td>2/3</td>
<td>Heterozygote</td>
<td>Positive right lung and abdomen</td>
</tr>
<tr>
<td>7</td>
<td>2</td>
<td>Homozygote</td>
<td>Positive chest</td>
</tr>
<tr>
<td>8</td>
<td>2/3</td>
<td>Heterozygote</td>
<td>Positive chest, abdomen and pluera</td>
</tr>
<tr>
<td>9</td>
<td>3</td>
<td>Homozygote</td>
<td>Positive left chest, left lung, left pelvis and abdominal wall</td>
</tr>
<tr>
<td>10</td>
<td>#</td>
<td></td>
<td>Positive soft tissue, pleura, rib</td>
</tr>
<tr>
<td>11</td>
<td>3</td>
<td>Homozygote</td>
<td>Not Infected</td>
</tr>
<tr>
<td>12</td>
<td>3</td>
<td>Homozygote</td>
<td>Not Infected</td>
</tr>
<tr>
<td>13</td>
<td>3</td>
<td>Homozygote</td>
<td>Not Infected</td>
</tr>
<tr>
<td>14</td>
<td>2/4</td>
<td>Heterozygote</td>
<td>Unknown</td>
</tr>
<tr>
<td>15</td>
<td>3/4</td>
<td>Heterozygote</td>
<td>Unknown</td>
</tr>
</tbody>
</table>

# Mutation present – to be confirmed

Allele 1(201bp) = A(CA)₅TG(CA)₅TG(CA)₁₅C; Allele 2(199bp) = A(CA)₅TG(CA)₅TG(CA)₁₅C; Allele 3 (197bp) = A(CA)₅TG(CA)₅TG(CA)₁₅C; Allele 4(199bp) = A(CA)₅TG(CA)₅C
Figure 1: Co-evolution between human and pathogens
Evolution of one species in response to characteristics of another