Paleopathological analysis of disease in the pre-colonial Americas

Dissertation
der Mathematisch Naturwissenschaftlichen Fakultät
der Eberhard Karls Universität Tübingen
zur Erlangung des Grades eines
Doktors der Naturwissenschaften
(Dr. rer. nat.)

vorgelegt von
Elizabeth Angel Nelson
aus Phoenix, Arizona, USA

Tübingen
2020

Gedruckt mit Genehmigung der Mathematisch-Naturwissenschaftlichen Fakultät der Eberhard Karls Universität Tübingen.

Tag der mündlichen Qualifikation:	22.12.2020
Dekan:	Prof. Dr. József Fortágh
1. Berichterstatter:	Prof. Dr. Johannes Krause
2. Berichterstatter:	Prof. Dr. Nicholas Conard

Acknowledgements

I am immensely grateful to have received support, encouragement and advice from several people throughout this journey. Here, I offer my thanks to those who have contributed to the success of this dissertation project:

I would first like to express my thanks to my advisor Prof. Johannes Krause and my supervisor Dr. Kirsten I. Bos for bringing me into the world of Archaeogenetics and Molecular Paleopathology. Thank you for this opportunity, for all you have taught me, for valuing my anthropological background and for guiding me through these compelling projects.

Many thanks to the additional members of my committee: To Prof. Nick Conard, for agreeing to supervise me and for being a source of support and encouragement especially during the final stages of my Ph.D. To Dr. Cosimo Posth for agreeing to serve on my doctoral committee and for being a source of knowledge, encouragement, and enthusiasm throughout my Ph.D. To Prof. Jane E. Buikstra for continually providing mentorship, advice, inspiration, support and encouragement throughout my academic career for which I am immeasurably grateful – Abrazos!

To Dr. Alexander Herbig, who with generosity of time and energy, provided a wealth of knowledge with invaluable guidance on problem solving and understanding of computational analysis.

To Dr. Maria Spyrou who trained me in laboratory methods and analysis, provided me with feedback on presentations and papers, and whose advice, support and patience have helped me reach this goal. Thank you for teaching and thank you for being my friend. I remain incredibly grateful for you and your friendship.

To Ashild Vågene, who guided me through analysis of ancient TB genomes, provided me valuable feedback and insightful discussions (and a beautiful orchid). Thank you for being a wonderful office mate and friend and for continuing to support me throughout this process to help me 'make it so'.

A special thanks to the laboratory staff and technicians who supported this work through sequencing, in-solution capture and their availability to provide advice and guidance. Specifically, thank you to Guido Brandt and Antje Wissgot for their support in these projects.

To collaborators who have taught me, inspired me and supported me throughout this experience: Prof. Dr. Tiffiny A. Tung, and Prof. Dr. J. Marla Toyne, Prof. Anne Stone. To my fellow Ph.D. collaborators Kelly Blevins and Evelyn Guevara who through their collaborative spirits, insights, and contributions made this thesis possible. I'm grateful for the opportunity to work with each of you. Thank you to each of you for contributing your knowledge to our projects.

To those at institutes and universities across the world, who have inspired me through their research and scholarly work and encouraged me through their mentorship and/or friendship: Prof. Verena Schuenemann, Prof. Dr. Rick Smith, Dr. Adam Powell, Dr. Felix Key, Prof. Dr. Courtney Hofman, Dr. Tanvi Honap, Prof. Dr. Deborah Bolnick, Prof. Jeffrey Frost, Prof. Keolu Fox and Prof. Naomi Cleghorn.

To Prof. Dr. Haagen Klaus and who has been incredible mentor and friend, thank you for teaching me, talking with me, and being endlessly supportive in this experience. Thank you for sharing your knowledge on Andean archaeology, bioarchaeology, theoretical frameworks and Star Trek.

Thank you to my Jena friends and colleagues in the Department of Archaeogenetics, MPI-SHH. To Dr. Ainash Childebeyeva, Richard Hagan, Thiseas Lamnidis, Tiago Ferraz da Silva, Karen Giffin, and Gunnar Neumann, who (professionally) offered insightful discussions on research and (personally) are some of the kindest people I know. To Rodrigo Barquera, thank you for your friendship and for always having a smile, a hug, a great idea and of course, chocolate.

Thank you to Aditya Kumar Lankapalli, James Fellows Yates, and Aida Andrades Valtueña for answering my seemingly endless questions, for providing friendship, advice, support and laughter throughout my Ph.D.

Thank you to Franzie Aron and Cody Parker for our friendship retreats to the veranda, for keeping me company on the long days and late nights in the lab, for the laughter we shared and for your support. Cody, thank you for all you did to make Ph.D. life a little easier for me and my son.

To Kathrin Nägele, thank you for being my guide through Germany, for being so incredibly generous and caring, for helping me (including reviewing this thesis, the translation of my thesis summary and the map in this thesis), for baking for me and for buying me plants, and, most importantly, for always being there for me and my son.

To Dr. Allison Mann, thank you for being a good friend throughout my Ph.D. and for always making me smile, sharing so many laughs, and being one of the most humble and gifted people I have ever met.

Thank you to my friends in the Archaeology Department, particularly Nils Vanwezer, Blair Jobe, Ayushi Nayak and Sam Brown for creating a community of buddies with such good memories. Thank you being such a source of support and encouragement especially when I needed it the most. To Sam Brown, thank you for being such a positive force in my Ph.D. experience and for always encouraging me, providing me with motivational post-its, making the best martini's and always making me smile. I am so thankful for your friendship.

To Dr. Marcel Keller, thank you for assisting in the German translation of my summary. Thank you for the many talks over coffee and wine, for the post cards, for your support and encouragement. "Thank you for your friendship."

To Dr. Susanna Sabin, my officemate, my roommate, my colleague and my friend, thank you for reviewing this thesis, for putting up with my constant singing and giggling, for dancing with me at the Hauptbahnhoff, for being my confidant and commiserative person, and for supporting me in some of the most challenging times. I'm forever thankful for you.

Thank you to Dr. Shevan Wilkin for making this Ph.D. experience full of laughter and adventures. Thank you for being there at difficult times, for encouraging and (sometimes aggressively) motivating me to reach this goal. Thank you for somehow thinking I'm cool enough to be a spy. Thank you for being a caring friend and for helping me retain my sanity through this process.

Finally, I thank my family who have shown me unconditional love and support. To my parents, Dr. Thomas Nelson and Shirley Nelson, thank you for your love and support throughout my studies. To my partner, Prof. Michael S. Allen, thank you for believing in me more than I do myself, for your love and patience, and for supporting me through this experience and encouraging me to reach my goals. To my son, Isaac Weston, thank you for believing in me, for being adaptable, for understanding the long days and nights working, and for being willing to go on this adventure with me. Without your love and support I would not have made it this far.

Table of Contents

Abbreviations

BCE	Before Common Era
aDNA	Ancient Deoxyribonucleic Acid
bp	Base pair(s)
C	Cytosine
CE	Common Era
DNA	Deoxyribonucleic acid
HOPS	Heuristic Operations for Pathogen Screening
HTS	High-throughput sequencing
IS	Insertion sequence
LH	Late Horizon
LIP	Late Intermediate Period
LIV	Lower Illinois River Valley
MH	Middle Horizon
MALT	MEGAN Alignment Tool
MEGAN	Metagenomic Alignment Viewer
MTBC	*Mycobacterium tuberculosis* complex
NGS	Next Generation Sequencing
ORV	Osmore River Valley
PCR	Polymerase Chain Reaction
SNP	Single Nucleotide Polymorphism
T	Thymine
TB	Tuberculosis
U	Uracil
UDG	Uracil-DNA-glycosylase
yBP	Years Before Present

Summary

Paleopathology has evolved into an interdisciplinary field with contributions from anatomical, biomedical, bioarchaeological, histological, microbiological and biomolecular specialties. Traditionally founded in the practice of morphological skeletal observations, today there is an increasing reliance on Next Generation Sequencing (NGS) methods leading to ancient pathogen genomes providing the primary evidence for infectious diseases that circulated amongst human populations in the past. The increase in ancient pathogen genomes provides valuable insight to human-pathogen migration, evolution and ecology for infectious diseases which still burden the world's population today. In this dissertation, I apply skeletal and molecular methods to evaluate the presence of disease in Indigenous groups of North and South America. Here, I employ recent molecular and computational developments to detect, reconstruct, and characterize ancient tuberculosis (TB) genomes from the pre-colonial Americas.

In the first paper I investigate the presence of the metabolic bone disease, skeletal fluorosis, at the Ray Site, Illinois, USA. I evaluated 117 individuals of the archaeological community of this Middle/early-Late Woodland site (BCE 50 – 400 CE) constructing biological profiles and recorded paleopathological data of skeletal and dental material. I identified eight individuals displaying a shared set of osseous changes including osteosclerosis, ossification of tendinous and ligamentous attachments, high frequency of fractures, and dental defects. The development of a differential diagnosis in conjunction with evaluation of environmental factors led to the conclusion that the observed condition was most likely skeletal fluorosis. This is the first description of skeletal fluorosis from an archaeological community in North America.

In the second paper, I applied both skeletal and molecular methods to investigate the presence of infectious disease in communities from Huari, Peru, the former capital of the Wari Empire located in the Ayacucho Basin of the Andes. Groups included in this study were recovered from commingled skeletal deposits of three sectors of the site dating to terminal Wari and post Wari eras of the Late Intermediate Period (LIP, 1000 – 1400 CE). I selected 34 skeletal samples for molecular analysis including skeletal elements presenting indicators of bacterial infection, likely tuberculosis or brucellosis, and those with little to no change. Due to the commingled nature of this collection, it is unknown if the skeletal elements selected for this study that presented no

pathological changes represented individuals who did not have a skeletal response to the disease. The generation of NGS data and application of metagenomic pathogen screening revealed the presence of DNA fragments belonging to *Mycobacterium tuberculosis* complex (MTBC) members, the causative agent of tuberculosis, in 8 of the vertebral samples, including those which show osseous lesions characteristic of TB and those presenting little to no pathological change. The history of paleopathological investigations of TB in the Andes is discussed with the demonstration of a taxonomic binning tool for broad pathogen screening of shotgun sequenced data, and its value when analyzing commingled skeletal elements including those which display non-specific or slight pathological changes.

The third paper investigates MTBC evolution and ecology on a local scale in Huari populations who experienced dramatic socio-political and climatic shifts during the LIP. Here, I expand the previous detection of MTBC strains in LIP populations at Huari, to incorporate 69 more skeletal samples that represent a spectrum of osseous changes. MALT analysis allowed for identification MTBC strains in an additional 6 vertebrae. Subsequently, I reconstructed 7 MTBC genomes belonging to the *M. pinnipedii* strain group, associated with modern seals and sea lions. Analysis shows that although these strains are closely related to the coastal genomes, the *M. pinnipedii* strains recovered from the three contemporaneous sectors of Huari demonstrate diversity within a discrete geographical area and time period.

My fourth manuscript presents a reconstructed ancient MTBC genome from the eastern Andean slopes recovered from the Chachapoya *chullpa* of Diablo Huasi located in Amazonas, Peru. This skeleton displayed pathological changes that were non-specific in lesion distribution and morphology. Although I employed a broad pathogen screening method with limited ascertainment bias, TB was the only pathogen detected. Its preservation permitted the subsequent reconstruction of a 15-fold TB genome. Phylogenetic analysis reveals that the Diablo Huasi TB strain is closely related to those from the neighboring ancient coastal and highland populations. I discuss the utility of molecular methods for paleopathological analyses resulting in atypical pathological presentation and highlight the value of incorporating skeletal and molecular data to paleopathology. This work expands the known geographic range of ancient *M. pinnipedii* strains

to the eastern piedmont of the Northern Peruvian Andes, thus providing insight into their local ecology and evolution.

For the fifth manuscript, I present preliminary data for a collaborative project investigating tuberculosis from archaeological human remains recovered from pre-colonial contexts across the Americas including Peru, Chile, Mexico, and the USA. A total of 28 samples found to be putatively positive in PCR detection were sequenced and subsequently analyzed. Here, I present challenges in ancient TB DNA analysis and the laboratory methods and computational analyses we employ to overcome these challenges. I discuss the generation of new data for downstream analysis.

Zusammenfasung

Die Paläopathologie hat sich zu einem interdisziplinären Feld entwickelt, welches Beiträge anatomischer, biomedizinischer, bioarchäologischer, histologischer, mikrobiologischer und biomolekularer Fachgebiete integriert. In der Tradition der Praxis morphologischer Skelettbeobachtungen begründet, wächst die Bedeutung von NGS-Methoden. Die so rekonstruierten alten Pathogengenome liefern den Hauptbeweis für, in der Vergangenheit verbreitete, Infektionskrankheiten. Die Zunahme von verfügbaren Genomen alter Krankheitserreger liefert wertvolle Einblicke in die Migration, Evolution und Ökologie der Menschen und ihrer infektiösen Krankheitserreger wie TB, welche die Weltbevölkerung bis heute belasten. In dieser Dissertation wende ich paläoanthropologische und molekulare Methoden an, um Krankheiten in indigenen Gruppen Nord- und Südamerikas nachzuweisen. Ich nutze neue molekulare und bioinformatische Techniken um alte TB-Genome aus dem vorkolonialen Amerika zu erkennen, zu rekonstruieren und zu charakterisieren.

In der ersten Arbeit untersuchte ich das Auftreten einer metabolischen Knochenerkrankung, der Skelettfluorose, am Ray-Standort in Illinois, USA. Anhand von 117 Personen aus der archäologischen Gemeinschaft dieser mittel- / früh- und spätzeitlichen Woodland Stätte (BCE 50-400 CE), erstellte ich biologische Profile und zeichnete paläopathologische Daten zu Skelett- und Zahnmaterial auf. Ich konnte acht Personen mit häufigen Knochenveränderungen indentifizieren. Die Veränderungen schlossen Osteosklerose, Ossifikation von Sehnen- und Bandansätzen und

eine hohe Häufigkeit von Frakturen und Zahndefekten ein. Die Entwicklung einer Differentialdiagnose, in Verbindung mit der Bewertung von Umweltfaktoren, führte zu dem Schluss, dass der beobachtete Zustand höchstwahrscheinlich dem Befund einer Skelettfluorose entsprach. Dies ist die erste Beschreibung einer Skelettfluorose aus einem archäologischen Kontext in Nordamerika.

In der zweiten Arbeit wendete ich paläoanthropologische und molekulare Methoden an, um das Vorhandensein von Infektionskrankheiten in Gemeinden aus der ehemaligen Hauptstadt des Wari-Reiches, Huari, im Ayacucho-Becken der Anden,, zu untersuchen. Die in diese Studie einbezogenen Gruppen wurden aus einer Mischung aus Skelettfunden dreier Sektoren des Gebiets gewonnen. Die Sektoren datieren auf die späte und letzte Wari- Periode und die daran anschließende späten und mittlere Periode (LIP, 1000 - 1400 CE). Ich wählte 34 Proben für die molekulare Analyse aus, einschließlich solcher, die Zeichen einer vermuteten bakteriellen Infektion durch Tuberkulose oder Brucellose aufwiesen, aber auch Skelette mit wenigen oder keinen skelettalen Veränderungen. Mittels MALT wurde das Vorhandensein verschiedener Pathogene analysiert und Vertreter des *Mycobacterium tuberculosis* complex (MTBC), dem Erreger der Tuberkulose, in 8 der Proben nachgewiesen. Die positiven Fälle schlossen Skelette mit typischen TB-ähnliche Läsionen ein, aber auch Skelette die nur geringe bis keine pathologischen Veränderungen aufweisen. Der erfolgreiche Nachweis der Krankheitserreger in Skelett Elementen unterschiedlicher Herkunft und unterschiedlichem Ausprägung der Krankheiten zeigt den Wert der hier angewandten Kombination aus Shotgun-Sequenzierung und breitem Pathogen-screening.

In der dritten Publikation untersuche ich die die MTBC-Evolution und -Ökologie auf lokaler Ebene in Huari-Populationen, welche während des LIP dramatische gesellschaftspolitische und klimatische Veränderungen erfuhren. Durch die Untersuchung von weiteren 69 Skelettproben verschiedener Verknöcherungsgrade durch MALT-screening erweiterte ich den Nachweis auf von MTBC-Stämmen auf 6 weitere Individuen. In allen Individuen erfolgte der Nachweis in Wirbeln. Anschließend rekonstruierte ich 7 MTBC Genome der *M. pinnipedii*-Stammgruppe, die heute Robben und Seelöwen befällt. Die Analyse zeigt eine enge Verwandschaft zu den an der Küste vorkommenden Stämmen, jedoch eine hohe Vielfalt innerhalb der Stämme aus verschiedenen Sektoren von Huari.

Mein viertes Manuskript zeigt ein rekonstruiertes altes MTBC-Genom aus den östlichen Andenhängen, gewonnen aus dem Chachapoya chullpa von Diablo Huasi im Amazonasgebiet Perus. Dieses Skelett zeigte pathologische Veränderungen, die in Bezug auf Läsionsverteilung und Morphologie nicht spezifisch waren. Obwohl ich eine breite Pathogen-Screening-Methode einsetzte, war TB das einzige nachgewiesene Pathogen. Die gute Erhaltung ermöglichte die anschließende Rekonstruktion eines 15 fach abgedecken TB Genomes. Die phylogenetische Analyse zeigt, dass der Diablo Huasi TB Stamm eng mit denen der benachbarten alten Küsten- und Hochlandpopulationen verwandt ist. Ich diskutiere den Nutzen molekularer Methoden für paläopathologische Analysen und hebe den Wert der Einbeziehung von Skelett- und molekularen Daten in die Paläopathologie hervor. Diese Arbeit erweitert das bekannte geografische Verbreitungsgebiet der alten *M. pinnipedii* Stämme auf das östliche Piemont der nordperuanischen Anden und gibt so Einblick in ihre lokale Ökologie und Entwicklung.

Für das fünfte Manuskript präsentiere ich vorläufige Daten für ein größeres Projekt zur Untersuchung der Tuberkulose aus archäologischen menschlichen Überresten. Diese wurden aus vorkolumbianischen Kontexten in ganz Amerika, einschließlich Peru, Chile, Mexiko und den USA, geborgen. Insgesamt 28 Proben wurden beim PCR Nachweis als mutmaßlich positiv befunden und anschließend sequenziert und analysiert. Hier stelle ich Herausforderungen in der Analyse alter TB Dann, sowie die Labormethoden und Computeranalysen vor, die wir zur Überwindung dieser Herausforderungen anwenden. Ich diskutiere die Generierung neuer Daten für die anschließende Analyse.

List of publications

The following five papers are included and discussed in this dissertation:

I. **Elizabeth A. Nelson,** Christine L. Halling, Jane E. Buikstra. Evidence of Skeletal Fluorosis at the Ray Site, Illinois, USA: a pathological assessment, discussion of environmental factors. International Journal of Paleopathology, 26, 48-60.

II. **Elizabeth A. Nelson**, Jane E. Buikstra, Alexander Herbig, Tiffiny A. Tung, Kirsten I. Bos. Advances in Paleopathology of Tuberculosis in Pre-contact Andean South America. International Journal of Paleopathology, 29, 128-140.

III. **Elizabeth A. Nelson**, Aditya Kumar Lankapalli, Maria Spyrou, Susanna Sabin, Åshild Vågene, Ainash Childebayeva, Ben Rohrlach, James A. Fellows Yates, Martha Cabrera, Jose Ochotama, Tiffiny A. Tung, Alexander Herbig, Kirsten I. Bos. Tuberculosis in the Wake of the Wari Empire. *Manuscript*.

IV. **Elizabeth A. Nelson**, Evelyn Guevara, J. Marla Toyne, Alexander Herbig, Johannes, Krause, Kirsten I. Bos. Tuberculosis in the pre-colonial Chachapoya of Peru. *Manuscript*.

V. **Elizabeth A. Nelson***, Kelly E. Blevins*, Jane E. Buikstra, Alexander Herbig, Johannes Krause, Anne C. Stone, Kirsten I. Bos. Preliminary Work: Identifying and Characterizing TB Strains Across the pre-colonial Americas. *Manuscript*.

*contributed equally to this work

Additionally, I am a co-author on the following articles published during pursuit of my doctorate:

Liliana M Davalos, Rita M. Austin, Mairin A Balisi, Rene L Begay, Courtney A. Hoffman, Melissa E. Kemp, Justin R. Lund, Cara Monroe, Alexis M. Mychajliw, **Elizabeth A. Nelson**, Maria A. Nieves-Colón, Sergio A. Redondo, Susanna Sabin, Krystal S. Tsosie, Joseph M. Yracheta, (2020). Pandemics' historical role in creating inequality. Science, 368(6497), 1322-1323.

Rodrigo Barquera, Thiseas C. Lamnidis, Aditya Kumar Lankapalli, Arthur Kocher, Diana I. Hernández-Zaragoza, **Elizabeth A. Nelson**, Adriana C. Zamora-Herrera, Patxi Ramallo, Natalia Bernal-Felipe, Alexander Immel, Kirsten I. Bos, Víctor Acuña-Alonzo, Chiara Barbieri, Patrick Roberts, Alexander Herbig, Denise Kühnert, Lourdes Márquez-Morfín, Johannes Krause, (2020). Origin and health status of first-generation Africans from early Colonial Mexico. Current Biology.

Maria A. Spyrou, Marcel Keller, Rezeda Tukhbatova, **Elizabeth A. Nelson**, Aida Andrades Valtueña, Don Walker, Amelie Alterauge, Niamh Carty, Hermann Fetz, Michaël Gourvennec, Robert Hartle, Michael Henderson, Kristin von Heyking, Sacha Kacki, Elizabeth L. Knox, Christian Later, Joris Peters, Jürgen Schreiber, Dominique Castex, Sandra Lösch, Michaela Harbeck, Alexander Herbig, Kirsten I. Bos and Johannes Krause. A phylogeography of the second plague pandemic revealed through the analysis of historical Y. pestis genomes, (2019). Nature Communications, 10(1), 1-13.

Kirsten I. Bos, Denise Kühnert, Alexander Herbig, Luis Roger Esquivel-Gomez, Aida Andrades Valtueña, Rodrigo Barquera, Karen Giffin, Aditya Kumar-Lankapalli, **Elizabeth A. Nelson**, Susanna Sabin, Maria Spyrou, Johannes Krause, (2019). Paleomicrobiology: Diagnosis and Evolution of Ancient Pathogens. Annual review of microbiology, 73.

Cosimo Posth, Nathan Nakatsuka, Iosif Lazaridis, Pontus Skoglund, Swapan Mallick, Thiseas C. Lamnidis, Nadin Rohland, Kathrin Nägele, Nicole Adamski, Emilie Bertolini, Nasreen Broomandkhoshbacht, Alan Cooper, Brendan J. Culleton, Tiago Ferraz, Matthew Ferry, Anja Furtwängler, Wolfgang Haak, Kelly Harkins, Thomas K. Harper, Tábita Hünemeier, Ann Marie Lawson, Bastien Llamas, Megan Michel, **Elizabeth A. Nelson**…& Johannes Krause, and David Reich, (2018). Reconstructing the Deep Population History of Central and South America. Cell, 175(5), 1185-1197.

Verena J. Schuenemann, Aditya Kumar Lankapalli, Rodrigo Barquera, **Elizabeth A. Nelson**, Diana Iraíz Hernández, Víctor Acuña Alonzo, Kirsten I. Bos, Lourdes Márquez Morfín, Alexander Herbig, Johannes Krause, (2018). Historic Treponema pallidum genomes from Colonial Mexico. Plos Neglected Tropical Diseases, 12(6): e0006447.

Own contributions

I. I performed skeletal and dental analysis and data collection of 117 individuals together with Christine L. Halling. I analysed all collected data and developed the differential diagnosis with advisement from Dr. Jane E. Buikstra. I wrote all sections of the paper together with Dr. Jane E. Buikstra.

II. I performed all skeletal analysis and sample collection under the supervision of Dr. Tiffiny A. Tung. I performed all molecular laboratory work for data generation up to shotgun sequencing (performed by MPI-SHH laboratory technicians) under the supervision of Dr. Kirsten I. Bos and Dr. Maria Spyrou. I performed all analysis of sequencing data including processing, metagenomic pathogen screening and evaluation of detected MTBC reads under the supervision of Dr. Maria Spyrou, Dr. Alexander Herbig and Kirsten I. Bos. I wrote all sections of the paper together with all authors, i.e. Dr. Jane E. Buikstra, Dr. Tiffiny A. Tung, Dr. Alexander Herbig and Dr. Kirsten I. Bos.

III. I performed all laboratory work for this study up to sequencing and TB capture of samples (performed by MPI-SHH laboratory technicians) with laboratory training from Dr. Maria Spyrou during the first portion. I performed the following data analysis under the supervison of Dr. Alexander Herbig and Dr. Kirsten I. Bos: metagenomic pathogen screening, phylogenetic analysis, evaluation of environmental background, SNP evaluation, functional analysis of SNPs and evaluation of heterozygous positions. I performed filtering of non-confident SNPs together with Aditya K. Lankalipalli. I wrote all sections of this paper with contributions from Aditya K. Lankapalli, Dr. Ainash Childebeyeva, Dr. Tiffiny A. Tung, Dr. Alexander Herbig and Dr. Kirsten I. Bos.

IV. I performed all aspects of laboratory work for samples DIA002.A and DIA002.C for this project, including screening and preparation for target enrichment (sequencing and targeted capture performed by MPI-SHH laboratory technicians). Laboratory work for sample DIA002.B was prepared by Evelyn Guevara who I supervised through this process. I

performed all aspects of data analysis with support from Dr. Alexander Herbig and Dr. Kirsten I. Bos. I wrote all portions of this manuscript under the supervision of Dr. Kirsten I. Bos with additional feedback from Dr. Alexander Herbig.

V. I carried out amplifications for the libraries prepared at ASU and prepared them for targeted capture and sequencing. I performed all analysis of preliminary data together with Kelly E. Blevins, who I was training through this process, under the supervision of Dr. Alexander Herbig and Dr. Kirsten I. Bos. I prepared the second generation of data discussed in shortly in the manuscript from extraction to preparation for sequencing. I wrote all portions of this manuscript under the supervision of Dr. Kirsten I. Bos with additional feedback from Dr. Alexander Herbig.

1 Introduction

1.1 Paleopathology: The past informs the present

Mortui viventes docent: "the dead teach the living" – Paleopathology Association

Throughout time, diseases have influenced the global human population, shaping societies and cultures, political frameworks, economic structures, and human biology. Archaeological data, paleopathological data, historical records and molecular analysis have contributed to our understanding of diseases of the past and their social impacts. Infectious disease outbreaks have had a massive effect on the human population and societies, leading to significant population decline and, in some cases, changing the course of history. European scholars note the example of the fall of feudalism following the Black Death, demonstrating the power of infectious disease on socio-political structures (Herlihy, 1997). For scholars researching Indigenous populations of the Americas, we consider the introduction of pathogens to the Americas during European colonization as contributing to one of the most significant demographic and population shifts ever recorded. Although, disease, be it infectious, genetic, or metabolic, was not absent from Indigenous American populations prior to European arrival, the introduction of infectious microbes and viruses through colonization effected a population decline with estimates up to 90% of the Indigenous population (Kaplan et al., 2011; Koch et al., 2019) and contributed to the fall of the Aztec and the Inka empires (De Rojas, 2012).

The devastating social and demographic impacts of disease appear to have been recognized and weaponized very early in history. The employment of infectious disease in political conflicts is exemplified in accounts of the Spanish attempts to provide their French adversaries wine with the blood of those afflicted Hansen's Disease (leprosy) (Frischknecht, 2003) and, the most infamous and controversial event of the Golden Horde's launching of plague ridden human corpses over the walls of Kaffa to drive out the Genoese merchants (Derbes, 1966; Wheelis, 2002). In the Americas, throughout history and under successive political officials, infectious agents have been used and disease outbreaks have been capitalized on to promote colonial expansion (Dávalos et al., 2020). It is believed smallpox was the most frequently used biological weapon in colonial histories of the

Americas from Spanish Conquest to the 'Conquest of the West' with the infamous so-called smallpox blankets given to Indigenous groups (Selgelid, 2003).

In contrast, historic infectious outbreaks have also led to examples of human perseverance and ingenuity with the development of outbreak management strategies and therapies, including the precaution of quarantine, the development of health councils and increase sanitation measures, many still used today (Gall et al., 2016; Kinzelbach, 2013; Tognotti, 2013). Studies of historic and ancient outbreaks demonstrate the precautions to be taken in the midst of infectious outbreaks and the consequences for indifference (Giles-Vernick et al., 2010). In fact, in 2013, David Morens (NIH) with Anthony Fauci (NIH), now leader of the US coronavirus task force, drew upon historic outbreaks and outlined environmental and social influences of disease emergence, including shifts in climate and ecosystems, as well as, socio-cultural factors such as animal-human interactions, political tensions, malnutrition, poverty and social inequality. They warned that the major factors underlying infectious disease emergence and reemergence as observed in historic records were once again at play and the world was primed for another pandemic (Morens and Fauci, 2013). In the first two decades of the 21st century alone we have seen the epidemics and pandemics of infectious diseases including Ebola (2014) and SARS-CoV2 (2019). During these modern outbreaks, *some* political and social bodies demonstrate the application of the lessons learned in from past pandemics in effort to control and combat the spread of disease.

It is not only historic records which provide valuable information about diseases of the past but also the field of paleopathology. Morphologically based paleopathology provides for the identification of disease, evaluation of pathological changes of individual case studies, understanding the geographic distribution of diseases and permits the detection of shifts in population health and disease over time. When viewed within the environmental and cultural contexts, anatomical morphological analysis has the ability to signal environmental risks of disease in a population which may have long been overlooked. One study demonstrating the value of such studies was led by Pierpaolo Petrone and colleagues who reported the metabolic disease skeletal of fluorosis in a population recovered from Herculaneum dating to 79 CE (Petrone et al., 2011). The paleopathological analysis and description of fluorosis was followed by the publication of additional supporting data including radiography, chemical tests and the inclusion of modern

epidemiological data of children from Naples, Italy (Petrone et al., 2013). The inclusion of modern data revealed that fluoride toxicity was endemic in the area at levels greater than what was previously assumed. Thus, drawing attention to the extreme human health hazard of the drinking water in modern day Naples.

Likewise, the synergism of morphological and molecular methods has provided valuable information on relevant infectious diseases. The specialization of molecular paleopathology, which is the application of ancient DNA (aDNA) methods to archaeological and historic human remains, has produced a number of unexpected and valuable results helping to provide solutions and providing insight to pathogen evolution. The reconstruction of ancient genomes permits phylogenetic and evolutionary analyses through the evaluation of genetic change over time, contributing information on strain diversity, adaptation and virulence. The accumulation of beneficial mutations in response to a new host or environment allows a pathogen to emerge or re-emerge (Siddle and Quintana-Murci, 2014). Therefore, to develop effective preventative and management strategies for infectious disease outbreaks it is necessary to have an in-depth understanding of their evolutionary histories. One such example is the pioneering study involving molecular paleopathological research on the great influenza pandemic of 1918-1919 that led to the characterization of the causative influenza strain (Reid et al., 1999). This finding laid the foundation for the development of antiviral drugs, therapies and vaccines (Drancourt and Rauolt, 2005;); thus, demonstrating the value of aDNA pathogen research to help combat modern outbreaks. Furthermore, advances in aDNA research have allowed for the recovery of certain ancient pathogens in unexpected regions of the world (Majander et al., 2020; Vågene et al., 2018), the identification of possible changes in transmission mode in ancient pathogens (Rasmussen et al., 2015; Spyrou et al., 2018) and for the genomic investigation of factors influencing the rise and decline of historic outbreaks through genome wide comparisons of ancient and modern strains (Schuenemann et al., 2013). Additionally, molecular paleopathology has revealed previously unknown evolutionary histories concerning one of the world's greatest disease burdens today, tuberculosis (TB) (Bos et al., 2014). This study contributed insight to ancient zoonotic events which are currently considered the most significant infectious disease threat for human populations (Jones et al., 2008; Morens and Fauci, 2013).

Historic and modern studies have established that the success of an infectious disease to emerge and disseminate is dependent on multiple factors of socio-cultural responses, the pathogen's ability to reproduce and adapt, environmental changes and the susceptibility of the population (Bliven and Maurelli, 2016; Cohen, 2000; Morens and Fauci, 2013). To develop effective preventative and management strategies for infectious disease outbreaks it is necessary to have an in-depth understanding of their evolutionary histories. Together, archeological, historical and aDNA data can provide a comprehensive view of the disease-scape of the past to prepare us for the future.

1.2 The interdisciplinary study of ancient disease

Paleopathology is defined as the study of ancient disease, however, this definition may give a false sense of simplicity. The dynamic relationship between disease and the factors (social, biological, environmental, etc.) through which diseases emerge and spread are complex and incompletely understood (Maurens and Fauci, 2013). Therefore, it is understandable that nature of paleopathology would necessitate an interdisciplinary approach to adequately address questions of population health and disease (Buikstra, 2010).

The genesis of the field dates to the mid-19th century, where interest in ancient disease moves beyond the so-called *Antecedent Phase* (16th – 19th century) involving excavation and collection of pathological cases toward the first systematic studies of ancient disease (Aufderheide and Rodríguez-Martin, 1998; Buikstra, 2010). Since conceptualization of the field it has developed from interested physicians (e.g. Rudolf Virchow, Grafton Elliot Smith, Frederic Wood Jones) producing uncontextualized clinical case reports to dedicated physical anthropologists and paleopathologists developing rigorous differential diagnoses (Buikstra, 2010; Grauer, 2012; Grauer, 2018; Klaus, 2017). During the 1960's, paleopathology emerged as a tool to understand past populations and cultural dimensions between health and society (Møller-Christiansen, 1961). This *New Paleopathology* contributes broader population studies (Toyne et al., 2020), deeper individual cases and relevant interpretations and discussions of disease (Buikstra and Williams, 1991a, b).

Paleopathology has evolved into an interdisciplinary field combining the humanities, biomedical science, and social studies (Buikstra, 2010; Grauer, 2012). The increased interaction with multiple

disciplines, including archaeologists, bioarchaeologists and historians, has permitted comprehensive and holistic studies of disease in past populations. The inclusion of historical records, environmental data, oral histories and artistic depictions from archaeological and historic sources provides a broader understanding of disease presentation (symptoms and signs), the dissemination of disease, the socio-political and cultural interpretations of disease and population responses to disease. Modern paleopathology applies an interdisciplinary and multi-method approach producing multiple lines of data for individual cases, as well as, population-based studies (Roberts, 2016). Yet, the field of paleopathology is still coming to terms with its nature as an interdisciplinary endeavor where actors in this research share overlapping research agendas with sometimes drastically different end-goals. However, within paleopathology, each research foci (within broader projects and without) produces a variety of transdisciplinary contributions, all of equal value towards understanding infectious disease. Moreover, the synergy of molecular and morphological methods has encouraged method-driven approaches, leading to significant advances in the field and yielding an abundance of relevant information concerning population health and disease.

1.2.1 Morphology based methods and the study of ancient disease

Traditionally, the field of paleopathology has been based on morphological analysis of skeletal and soft tissue remains of past peoples in a laboratory setting (Ortner, 2003) with little interaction with field-based archaeology. Thus, research has focused on diseases which manifest in the remaining tissues found and recovered from archaeological excavations or maintained in historical anatomical collections. Primary data in the field of paleopathology is generated through a multi-dimensional process of observation, description and diagnosis. This begins with morphological anatomical/skeletal analysis in which the construction of biological profile is then followed by the recording and describing of abnormal or pathological anatomical and skeletal changes (Buikstra and Ubelaker, 1994). Finally, the development of a differential diagnosis, which is a complex probabilistic exercise, involves three general steps: 1) the description of the observed pathological condition, 2) the identification of possible pathological etiologies and 3) the comparison to known disease patterns and removal of contrasting forms of disease from consideration. This process has led to the identification of genetic, metabolic, degenerative, developmental, neoplastic and infectious diseases (Buikstra, 2010; Klaus, 2014; Suby, 2020; Titelbaum, 2020).

Interpretation of morphological data based on abnormal skeletal changes requires knowledge of bone anatomy, physiology, etiology and various causative factors of pathological bone formation (Klaus 2014; Ragsdale and Lehmer 2012). Paleopathological texts (Aufderheide and Rodriguez Martin, 1998; Buikstra, 2018; Ortner et al., 2003) typically showcase extreme or obvious cases of skeletal manifestation of disease which can lead to a bias in morphologically based paleopathological research. Therefore, detecting slight to moderate skeletal pathology requires specialized training. This is even more necessary for identifying variations in pathological morphology and distribution (Toyne et al., 2020). Through interdisciplinary research we gain a deeper understanding of lesion formation, thus enabling recognition of early and later stages of pathological changes associated with specific conditions. Clinical data, incorporated with anatomical knowledge and pathophysiological research, illuminates the spectrum of changes expected in a condition, which is frequently broad and varied; thus, providing insight to demographic patterns in population disease.

Among the skeletal and soft tissue signs of disease there are some that are considered diagnostic for a specific disease based on morphology and/or skeletal distribution. However, a fundamental challenge to this field is that many skeletal pathologies are non-specific and cannot confidently be attributed to a disease. Pathological processes can result in the manifestation of indistinguishable patterns of abnormal bone formation or destruction (Ortner, 2003). Likewise, many diseases and infectious agents have varying rates of osseous response or more commonly leave no skeletal evidence of their presence (Ortner, 2003; Roberts, 2016). A further confounding problem is introduced by taphonomic changes resulting from postmortem manipulation in which environmental causes, looting and various burial programs/practices are to blame for visible changes on skeletal remains. These postmortem changes complicate observations as they can be difficult to distinguish from skeletal pathologies. Damage the bone and soft tissue surfaces can obscure observations and/or result in incomplete skeletal individuals, thus prohibiting the observation of disease distribution (Aufderheide, 2011; Grauer, 2018). Likewise, paleopathologists must recognize a key assumption inherent to the field: we assume conditions and diseases manifest on the skeleton similarly in the past as they do today (Buikstra, 2010). Acknowledgement of this assumption and awareness of variation in human response to pathogens,

in addition to the aforementioned challenges, encourages interdisciplinary approaches and the application of ancient DNA methods.

Paleopathological studies based on morphological data have revealed the presence of diseases in unexpected populations (Buikstra, 1976; Vlok et al., 2020), raising questions to the origin and dissemination of the disease (Alison et al., 1981). Additionally, they have provided a broader understanding of population health and disease (Buikstra and Williams, 1981) and developed complex interpretations of disease in the past (Scaffidi et al., 2020). These accomplishments were not without a history of critique and subsequent development. Most notable would be the criticisms by Medical Historian Saul Jarcho (1966), who called out the field for shortcomings in communication, data sharing and development, and the failure to develop relevant contributions. Although many issues still persist, paleopathology has since established methods and standards for recording data (e.g. Buikstra and Ubelaker, 1994; Brickley and McKinley, 2004; Ragsdale, 1992), in addition to the development of shared databases (WORD, *Osteoware*) (Bekvalac, 2017; Wilczak and Dudar, 2011), the increase in interdisciplinary work and, in doing so, has seen great advances in invasive and non-invasive methods.

1.2.2 Ancient DNA methods and the study of ancient disease

Ancient DNA research involves the analysis of DNA molecules recovered from a variety of sample and tissue types originating from museums, anatomical collections, environmental and archaeological contexts (e.g. Bos et al., 2011; Mascher et al., 2016; Pääbo et al., 2004; Rasmussen et al., 2010; Sabin et al., 2020; Warinner et al., 2014). As observed with archaeological remains of skeletons, textiles and architecture, DNA is subject to the effects of time with progressive degradation and decomposition by enzymatic, microbial, oxidative and hydrolytic processes (Lindahl, 1993; Pääbo et al., 2004). Without active cellular DNA repair mechanisms to counteract these processes, aDNA becomes damaged and fragmented, typically less than 100 base pairs (bp) (Dabney et al., 2013a; Sawyer et al., 2012). This damage is usually the hydrolytic damage of cytosine (C) to uracil (U) and can result in nucleotide misincorporations across the sequence fragment (Briggs et al., 2007; Dabney et al. 2013a; Pääbo et al., 2004). Over time, ancient DNA molecules sustain multiple types of damage which can vary across samples in the degree and pace

of damage accumulation due to influence of environmental factors (Adler et al., 2011; Allentoft et al., 2012). Yet, many aDNA fragments still retain valuable information of the organism, providing unique biological data to our understanding of the past.

Early studies in molecular paleopathology were based on Polymerase Chain Reaction (PCR) methods (Saiki et al., 1985). For the first time, genetic evidence of ancient infectious disease through the amplification DNA elements deriving from ancient pathogens could be contributed to paleopathological analyses (Arriaza et al., 1995; Drancourt et al., 1998, 2004; Hershkovitz et al., 2008; Salo et al., 1994; Zink et al., 2001). This method, applied to paleopathology, involves the targeted amplification of DNA sequences believed to be specific to pathogenic organisms for detection of a pathogen. Although PCR methods offer high sensitivity, the level specificity is highly variable depending on the organism, the target regions and the length of primers used in amplification. Additionally, this method introduces a bias towards well preserved, long fragments which are most commonly deriving from modern sources and contaminants (Pääbo and Wilson, 1988). This early research on ancient pathogens did not yet have quality measures or standardized methods, nor was there any ability to authenticate the antiquity of the pathogen understudy (Cooper and Poinar, 2000). Furthermore, findings were often contested due the inability to independently validate results (Gilbert et al., 2005; Shapiro et al., 2006). Additionally, ancient pathogen DNA research using PCR methods has been largely limited to detection-based analyses. At this time, arguments against the use of molecular methods were put forth stating that destructive methods can only ethically be applied if they can provide results morphological assessments cannot, and many diseases of interest were well documented in historical records or manifested in diagnostic skeletal signs challenging the ethical application of the destructive PCR methods (Roberts et al., 2016).

The introduction of high-throughput Next Generation sequencing (NGS) methods transformed the field. This method allowed large-scale parallel sequencing of untargeted DNA molecules for the first time (Bentley et al., 2008; Mardis et al., 2008; Margulies et al., 2005), enabling the detection of ancient pathogens, the recovery of metagenomic samples and the reconstruction of whole genomes. This provided a new level of sensitivity and resolution which permitted confident genome wide analysis for evolutionary studies and thus, enabled the development of authentication

measures of ancient DNA. Since the first ancient *Yersinia pestis* genome was reconstructed in 2011 (Bos et al.), application of this technology has become the gold-standard with studies increasing at a rapid pace. With high-throughput sequencing (HTS), the application of aDNA methods moved beyond detection only to the goal of understanding ancient pathogen evolution, migration and ecology through characterizing ancient bacterial pathogens at a genomic level.

The early challenges faced by aDNA, are now being addressed by methodological advances accompanying the NGS revolution. With NGS, developments have been made to overcome issues of contamination with measures of data quality and authentication of ancient DNA (Key et al., 2017). Where DNA damage once posed limitations to PCR based research, NGS permits not only the recovery of fragments shorter than 100 bp but the observation of DNA degradation for the authentication of its antiquity. Methods for recovery of ancient DNA have been optimized for short fragments to accommodate this feature of highly fragmented ancient DNA (Dabney et al., 2013b). Likewise, research and developments exploring optimal sampling strategies for both human and pathogen DNA recovery are now guiding aDNA research design (Pinhasi et al., 2015).

Of particular benefit to molecular paleopathology is that NGS methods allow for the generation of metagenomic data through large scale paralleled sequencing of non-target reads. In contrast to the targeting of specific regions of DNA in PCR based methods, NGS methods produce an abundance of data representing the host and organisms that constitute an extracted DNA sample. This results in the recovery of host and microbial DNA. However, unlike targeted approaches, this method limits ascertainment bias and permits the detection of multiple pathogens or multiple strains from a single sample (Barquera et al., 2020; Giffin et al., 2020; Kay et al., 2014). As a result, this has allowed for the detection and genomic reconstruction of ancient pathogens from various samples including skeletal remains from contexts that challenge diagnosis including skeletal remains presenting non-specific pathological changes or do not present skeletal pathology (Bos et al., 2014; Spyrou et al., 2018; Vågene et al., 2018; papers II, III, and IV of this thesis). Owing to these technological advances, molecular paleopathology has resulted in the identification and evolutionary analysis of multiple pathogens from a variety of archaeological human remains (Bos et al., 2014; Kay et al., 2014; Schuenemann et al., 2018; Spyrou et al., 2018; Warinner et al., 2014).

The value of ancient DNA is well demonstrated; however, similar to morphological paleopathological studies, it is not without limits and challenges. Some archaeological samples that present pathological lesions may not yield detectable pathogen DNA. This could be due to poor sampling location, preservation issues where no detectable level of pathogen DNA remains, or perhaps the current lack or scarcity of available genetic data for the pathogen. Likewise, diseases which originate from metabolic syndromes or genetic influences may not be well characterized in the biomedical world, and therefore, the application of aDNA methods is of limited benefit. In such cases, morphological skeletal analysis still provides insight to the health and disease of the individual or population under study (Buikstra, 2010; Roberts, 2016). This is the case with skeletal fluorosis in which genetic markers associated with increased susceptibility to the condition have been identified in mice, however, such genetic markers have not yet been identified in human populations. Morphology and distribution of skeletal changes in fluorosis are well documented in modern clinical populations as are the environmental and dietary influences leading to the disease. Therefore, a well-developed differential diagnosis should include an interdisciplinary approach with multiple lines of evidence and be accompanied by dietary and environmental data (Nelson et al., 2019).

1.2.3 The promise of molecular and morphological paleopathology

The complementary nature of morphological and molecular analyses in paleopathological research has allowed for the identification of diseases across broad temporal and geographic spaces. Constructing a confident differential diagnosis is highly challenging in the absence of diagnostic skeletal lesions or in cases that originate from problematic burial programs which obscure observations. Ancient DNA provides the possibility for pathogen identification, genomic characterization and evolutionary analyses, increasing the biomedical relevance of the field. The synergy of molecular and morphological paleopathology holds the possibility for greater understanding of ancient and historic infectious disease outbreaks involving highly virulent pathogens and acute infections which likely leave no skeletal signs yet may be reflected in mortuary patterns (Buikstra, 2010).

The inclusion of ancient DNA methods to paleopathology, specifically NGS methods, has encouraged and facilitated the already aborning paradigm shift away from one-dimensional

identification and description of skeletal pathologies toward evolutionary thought and investigations of the emergence of infectious disease and evolution of pathogens. This interdisciplinary approach in paleopathology has allowed for more comprehensive interpretations in a biocultural approach: exploring infectious and non-infectious disease and the dynamic interaction of humans and environmental and socio-cultural influences on disease (Armelagos, 2008; Zuckerman et al., 2012). Comprehensive studies including skeletal analyses, archeological data and aDNA of both humans and pathogens have begun to explore the interplay of environment, social histories, human biology and infectious disease (Barquera et al., 2020). While this thesis focuses on pathogen DNA to explore ancient disease, human DNA research presents an exciting opportunity to explore one of the key aspects of *The Osteological Paradox* (Wood et al., 1992): hidden heterogeneity in susceptibility to disease. Through continual advances for overcoming challenges in the field, molecular paleopathology has become a bridge for sub-disciplines and related disciplines to address a multitude of questions pertaining to a variety of aspects concerning disease including socio-cultural and environmental influences of disease, as well as, evolutionary studies. The reconstruction of ancient genomes has the potential to add molecular insights to view the human population and disease dynamics in epidemiological transitions throughout human history (Harper and Armelagos, 2010) to help us better prepare for future transitions in this ever-changing world.

1.3 Molecular Paleopathology and pre-colonial tuberculosis

With the advent of NGS, results of the first studies incorporating whole genome reconstruction and genomic analysis to paleopathology were unprecedented, providing deep time information on the evolution and ecology of infectious organisms. This was most certainly the case for the first reconstructed ancient tuberculosis genomes (Bos et al., 2014). Tuberculosis is caused by a closely related group of mycobacteria, the *Mycobacterium tuberculosis* complex (MTBC). These bacillary bacteria spread most commonly through respiratory transmission and are known to cause disease in both humans and other mammals. The MTBC is comprised of 9 lineages that are phylogeographically distributed and have some level of association with geographic populations (Gagneux et al., 2006; Ngabonziza et al., 2020). Today, lineage 4 dominates Europe and the Americas, thus suggesting introduction through European colonization. However, skeletal evidence of TB in the Americas prior to European contact is abundant, particularly in Chile and

Peru, with the earliest case appearing in Chile, South America ca. 290 CE (Allison et al., 1981; Roberts and Buikstra, 2003). Although an abundance of skeletal data existed and TB had been identified through PCR methods (Salo et al., 1994), existence of pre-colonial TB was still contested (Wilbur and Buikstra, 2006). Because tuberculosis was believed to be an "Old World" disease, introduced to the Americas during European colonization, some argued the skeletal changes were likely caused by another infectious agent and that the PCR results were sequences of closely related environmental organisms causing false positive detection. To support this idea, critics called attention to the susceptibility of Indigenous Americans to TB during European colonization, hypothesizing this susceptibility came from the novelty of the pathogen to the naïve Indigenous immune system (Stead et al., 1995; Stead 1997).

In 2014, the application of NGS methods led to the reconstruction of the *Mycobacterium tuberculosis* complex genomes from pre-colonial South American archaeological remains (Bos et al., 2014). This study provided new insights to MTBC evolution, ecology and strain distribution through the characterization of an unanticipated strain infecting ancient human populations, *M. pinnipedii*, adapted for modern seals and sea lions. This study provided evidence suggesting possible zoonotic transmission of ancient TB to pre-colonial populations of the Americas (Bos et al., 2014). Additionally, Bayesian molecular dating of the MTBC, including these ancient strains, estimated the most recent common ancestor to be as young as approximately 6,000 yBP. This contradicted an earlier study using modern strains which suggested scenarios of out-of-Africa dispersals of the MTBC (Comas et al., 2013). The evidence presented in this work debunked previously established ideas of initial introduction of MTBC during European colonization by presenting the *pinniped hypothesis* whereby authors posit a scenario of seals introducing MTBC to the Americas (Bos et al., 2014). These findings overturned old arguments contesting pre-colonial TB despite the existence of an abundance of paleopathological skeletal evidence from archaeological contexts across the Americas (Stead et al., 1995; Stead 1997). They currently remain the only published genomes of pre-colonial MTBC from the Americas.

Tuberculosis led to 1.5 million deaths in 2019 (WHO, 2019) and it continues to be one of the leading causes of death in the world. The spread and success of TB is largely due to features of the organism such as latency and a broad range of host species, leading to multiple zoonotic events

(Brites and Gagneux et al., 2015). Considered one of the most successful pathogens, the reemergence of TB is closely linked to increased virulence and pathogenicity through genetic changes leading to multi-drug resistant (MDR) and extensively drug resistant (XDR) strains (Morse, 2004). In order to develop successful countermeasures for this global burden the evolutionary history of MTBC species across multiple populations, ecologies and hosts must be well understood (Karesh et al., 2004). Although TB is one of the most devastating diseases and has a long history of human infection, there remains no broadly effective vaccine (Orr et al., 2013). The need for infectious disease research on both past and present populations has never been more relevant or urgent. To that end, the molecular portions of my thesis (papers II-V) focus on adding to the known strain diversity of pre-colonial TB in the Americas in order to better understand the evolution and ecology of the MTBC.

2 Thesis objectives

In this thesis, my primary goals are to demonstrate the relevance and value of paleopathology as an interdisciplinary study to our understanding of ancient disease, specifically tuberculosis. While highlighting the synergistic relationship between the traditional morphology-based paleopathology and the more recently introduced molecular paleopathology, I add to the current understanding of ancient pre-colonial MTBC strains.

The complementary nature of aDNA methods to paleopathological research has been demonstrated in a number of projects. However, both methods are not without their limits and each specialization is experiencing continual development. Morphology based paleopathology involves challenges rooted in the archaeological context, the environment, cultural practices, the absence of skeletal manifestation of disease and the occurrence of non-specific skeletal changes shared across multiple diseases. Molecular paleopathology is limited by incomplete databases, methodological or environmental issues in pathogen DNA recovery, the need for improved authentication measures and the abundance of microbes existing in interment matrices which challenge confident reconstruction of ancient pathogen genomes (Spyrou et al., 2019; Warinner et al., 2017). However, in recent years, collaboration between bioarchaeologists and aDNA specialists has led to the development of advances in both methodologies and research design. In this thesis, I present papers which illustrate the strengths and limitations of both fields concerning ancient disease research. This is accomplished through the presentation of a metabolic disease affecting the skeleton known as *skeletal fluorosis*, which at this time cannot be investigated using genetic means and is therefore reliant on the development of a rigorous and comprehensive differential diagnosis. In contrast, I present three papers (II-IV) in which the sample sets resulted in a broad differential diagnosis due to either the commingled nature of the archaeological remains or the presentation of unusual skeletal pathology. Complementing these studies with the application of molecular methods permitted the identification of infectious agents and provided multiple lines of evidence for a confident diagnosis of TB, thus permitting the reconstruction of ancient MTBC genomes and the evolutionary analysis of the identified strains.

Tuberculosis has a long history of human infection yet remains one of the leading causes of death today, killing more people than any other infectious agent (WHO, 2019). Archaeological evidence

of tuberculosis from pre-colonial contexts in the Americas suggests a significant increase in TB during the Late Intermediate Period (1000 – 1400 CE), a time period characterized by climatic, cultural and political shifts in Peru. Understanding environmental, socio-political, cultural and biological influences that may have promoted this increase in TB holds great relevance for today as these factors remain at play. Skeletal and molecular methods must come together to produce data that will provide environmentally and culturally contextualized genomic information on strain diversity. In this thesis, I present the contextualization of ancient MTBC genomes from ancient populations across geographically dispersed locations and ecologies. Furthermore, I work to demonstrate the value and relevance of interdisciplinary approaches in paleopathology, not only for a comprehensive view of ancient disease, but also to help address infectious diseases which hold relevance for today such as TB.

3 Results

3.1 Skeletal fluorosis at the Ray Site, Illinois, USA (paper I)

Title: "Evidence of skeletal fluorosis at the Ray Site, Illinois, USA: a pathological assessment and discussion of environmental factors"

Elizabeth A. Nelson, Christine L. Halling, and Jane. E. Buikstra

Published in the *International Journal of Paleopathology*, 2019, 26, 48-60.

Study synopsis

Fluorosis is a metabolic disease that occurs when toxic levels of fluoride are consumed. The pathological effect of fluoride on the skeleton, called *skeletal fluorosis*, is an endemic problem in populations at risk of chronic ingestion of excess fluoride. Today, fluorosis is most frequently associated with anthropogenic factors such as industrial pollutants and socio-cultural practices. However, many regions of the world have natural environments conducive to the development of skeletal fluorosis. Archaeological evidence of skeletal fluorosis has been studied from such environments in Bahrain, Naples and the United Arab Emirates. However, despite American regions of high fluoride content in soils and water, this condition was absent from paleopathological studies focused in the Americas. In this study, I present a paleopathological survey of a cemetery population from an area of naturally high fluoride content, the Illinois River Valley in western Illinois, USA. The main results of this paper are the following:

1. Skeletal pathological analysis of 117 individuals resulted in the identification of a shared suite of pathological changes in eight individuals.

2. Shared pathological changes were largely non-specific and include osteosclerosis, enthesopathies, periosteal bone formation, dental pitting and dental mottling. Analysis of skeletal distribution of pathological changes revealed the condition was bilateral/systemic and that the axial skeleton was more affected than the appendicular skeleton. Morphological estimations of biological sex and age-at-death estimations revealed middle aged (~40-50 years of age) males were more frequently affected with this skeletal syndrome.

3. The result of the analysis was a broad differential diagnosis, however, the environmental context of high levels of fluoride in soil and water suggests the observed pathology is most likely skeletal fluorosis. This result is the first description of skeletal fluorosis from an archaeological context in the Americas and demonstrates the development of a rigorous differential diagnosis and inclusion of environmental data for a comprehensive assessment.

Although the observed pathological changes were not considered severe and present a minimal to moderate skeletal response to excess fluoride, the inclusion of environmental data supports this finding. This work should support further paleopathological research on skeletal fluorosis in the Americas and encourage the development of comprehensive analyses incorporating multiple lines of data.

3.2 Advances in molecular detection of tuberculosis in Andean paleopathology (paper II)

Title: "Advances in the molecular detection of tuberculosis in pre-contact Andean South America"
Elizabeth A. Nelson, Jane E. Buikstra, Alexander Herbig, Tiffiny A. Tung, and Kirsten I. Bos
Published in the *International Journal of Paleopathology*, 2020, 29, 128 -140

Study synopsis

The Andean region of South America has played a central role in the field of paleopathology. Archaeological investigations in this region have made significant contributions to tuberculosis research, where a number of pioneering paleopathological studies that adopted different analytical approaches have permitted in-depth investigation of the disease during the pre-colonial period. This paper presents a discussion of advances in the field of paleopathology and highlights the strengths and limitations of current approaches for the molecular detection of ancient pathogens, with focus on tuberculosis. As an example of valuable advances to the field of paleopathology, this paper demonstrates and describes a molecular screening approach for the detection of ancient tuberculosis in individuals from Late Intermediate Period (1000 – 1400 CE) contexts at the site of Huari, Peru. The main results of this paper include:

1. The identification of skeletal pathology in commingled vertebrae from Late Intermediate Period contexts (1000 – 1400 CE) of Huari, Peru, in the Andean highlands. Vertebrae displayed lytic lesions consistent with bacterial infection, namely tuberculosis or brucellosis, in addition to non-specific changes. A broad differential diagnosis was developed which includes tuberculosis, brucellosis and mycotic infections.

2. Through the application of high-throughput sequencing and broad pathogen screening methods, *Mycobacterium tuberculosis* complex DNA was identified in eight out of 34 vertebrae. These samples included two vertebrae displaying well developed tuberculosis-like lesions, however, the remaining six vertebrae showed either slight, non-specific reactive bone, hypervascularity which may be considered within normal variation or no macroscopically observable skeletal pathology.

3. Molecular analysis of the composition of each extracted sample revealed a dominance of environmental bacteria DNA in the samples. This evaluation will assist in designing appropriate measures for DNA recovery and genome reconstruction.

This paper highlights the value of molecular methods as a complement to morphological skeletal analyses for the identification of past disease. This is demonstrated through the application of this method to commingled skeletal remains which, by their nature, impede the development of confident differential diagnoses. Furthermore, despite the abundance of environmental bacteria, this method enabled the confident detection and evaluation of molecular signatures of tuberculosis infections despite the complex environmental signal derived from the depositional context.

3.3 Seven Pre-colonial MTBC genomes from Huari, Peru (paper III)

Title: "Tuberculosis in the wake of Wari Imperial Decline"

Elizabeth A. Nelson, Aditya K. Lankipalli, Maria Spyrou, Susanna Sabin, Åshild Vågene, Ainash Childebeyeva, Ben Rohrlach, James A. Fellows Yates, Martha Cabrera, Jose Ochotama, Tiffiny A. Tung, Alexander Herbig, and Kirsten I. Bos

Manuscript ready for submission (winter 2020)

Study synopsis

Demographic patterns of tuberculosis are structured by environmental and socio-political factors, including malnutrition, socio-political distress and climatic variability. Such trends may also be observed in past archaeological contexts such as the Late Intermediate Period (LIP, 1000–1400 CE) of the Andes, a period of climatic and socio-political transitions. This time was characterized by the decline of the first Andean empire, the Wari, which was followed by a shift to more dispersed settlements, dietary changes, an increase in lethal and sub-lethal violence and the appearance of skeletal changes suggestive of tuberculosis. Late Intermediate Period archaeological contexts present the largest number of paleopathological reports of Andean tuberculosis. Application of molecular methods to LIP skeletal evidence of tuberculosis has resulted in the reconstruction of *Mycobacterium tuberculosis* complex (MTBC) members and subsequently revealed the strain infecting coastal pre-colonial populations of the Osmore River Valley, Peru, as closely related to modern *M. pinnipedii*, a strain known to affect modern seals and sea lions. However, the full geographic expanse of this strain, and likewise, the full diversity of MTBC strains circulating in the Andes has remained largely unknown. This study presents evidence of MTBC strains closely related to the previously identified ancient, coastal *M. pinnipedii* strains circulating in populations experiencing the fall and wake of the Wari empire located in the central Andes over 500 km away from the coast. The key findings of this paper are as follows:

1. Through molecular analysis of 103 skeletal samples, I have reconstructed seven ancient MTBC genomes from three sectors of Huari, the former capital of the Wari empire dating to the Terminal Wari (950-1050 CE) and post-Wari (1275-1400 CE) eras. Phylogenetic analysis reveals these strains as closely related to the previously characterized ancient *M. pinnipedii* genomes recovered from the coastal Osmore River Valley (ORV) sites. Furthermore, analysis shows unexpected strain diversity circulating within a narrow geographic and temporal context with strains from one sector, Vegachayoq Moqo, being more closely related to the coastal ORV strains than to the neighboring contemporary sector, Monqachayoq, located within less than a kilometer away.

2. Using Bayesian molecular dating methods, the divergence time of this subclade of ancient MTBC strains with the modern *M. pinnipedii* strains is estimated to be less than 1,000 years ago. Suggesting skeletal evidence of tuberculosis from contexts earlier to this time period may come from an ancestral species contributing similar pathological skeletal changes.

This work presents the first molecular evaluation of ancient MTBC strains on a local scale in a population experiencing the wake of imperial decline and additional environmental stressors. Likewise, this may reflect that the Late Intermediate Period increase in skeletal tuberculosis was caused by the broad circulation of similar MTBC strains closely related to *M. pinnipedii*.

3.4 A pre-colonial MTBC genome from the Chachapoya site of Diablo Huasi (paper IV)

Title: "Tuberculosis in a pre-colonial Chachapoya individual from the northeastern Peruvian Andes"

Elizabeth A. Nelson, Evelyn Guevara, J. Marla Toyne, Alexander Herbig, Johannes Krause, Kirsten I. Bos

Manuscript ready for submission (Spring 2021).

Study synopsis

Although ancient *Mycobacterium tuberculosis* complex strains closely related to *M. pinnipedii* have been identified and described from contexts of the coastal and highland regions of Late Intermediate Period (LIP, 1000 – 1400 CE) Peru, tuberculosis has yet to be molecularly characterized in the eastern Andean subtropical forest. However, skeletal evidence of tuberculosis from this region is described in reports of the Chachapoyas cultural region. A population-based study of Kuelap in the Chachapoyas region of the eastern Andean slopes revealed that approximately 6% of the population exhibited skeletal signs of tuberculosis (Toyne et al., 2020). However, these individuals displayed skeletal changes of tuberculosis that appear as a variant in both morphology and lesion distribution compared to populations from the western regions of the Andes. The study included this thesis discusses one such individual from the Chachapoyas cliff side tomb of Diablo Huasi located in the subtropical forest of the Amazonas region of Peru. The individual under study displayed pathological skeletal lesions that are described as having some similarities with tuberculosis skeletal manifestation yet are atypical of tuberculosis and therefore

may be caused by other pathological processes. Skeletal pathological analysis resulted in a broad differential diagnosis and therefore a metagenomic approach was used to perform broad pathogen screening to detect infectious agents which may be contributing to the unusual skeletal pathology. Here, I present the first LIP MTBC genome from the eastern piedmont of the northern Peruvian Andes. Analysis of this genome has led to the following findings:

1. The reconstruction of a 15-fold ancient MTBC genome from Diablo Huasi, Peru. This finding supports skeletal evidence of the presence of TB in the subtropical cloud forests of the Chachapoyas region.

2. Phylogenetic analysis of the Diablo Huasi MTBC genome reveals this strain is closely related to the previously described contemporaneous MTBC strains from coastal and highland populations of the Andean cultural region.

3. Although the phylogenetic placement and radiocarbon dating of the sample offer some support of the sample's antiquity, further authentication of the antiquity of this genome was not possible through observations of DNA damage typical of ancient DNA. Furthermore, it was not possible to attribute this low DNA damage to contamination by a modern pathogen or closely related environmental mycobacteria as analysis of the capture product showed little evidence of incorporated environmental mycobacteria in the reconstruction of this genome.

In conclusion, this paper demonstrates the value of high throughput sequencing and metagenomic methods applied to archaeological cases that show unusual presentation of skeletal pathologies. Although radiocarbon dating of the individual and the phylogenetic placement of the recovered genome support the antiquity of the pathogen DNA, this genome presents low DNA damage. The recovery of this genome provides a greater understanding of geographic and ecological diversity of ancient *M. pinnipedii* strains adding MTBC representation from the eastern cloud forest of Amazonas, Peru.

3.5 Preliminary steps in identifying, authenticating and reconstructing ancient MTBC genomes (paper V)

Title: "Preliminary investigations of MTBC across the pre-colonial Americas"

Elizabeth A. Nelson, Kelly E. Blevins, Jane E. Buikstra, Josefina Mansilla Lory, Alexander Herbig, Johannes Krause, Anne C. Stone, Kirsten I. Bos

Manuscript ready for submission (2021)

Study synopsis

Skeletal evidence of tuberculosis is found throughout American continents, however, our understanding of *Mycobacterium tuberculosis* complex (MTBC) strain diversity is currently limited to western South America. Likewise, we have only identified ancient MTBC strains in the Americas which are closely related to the modern *M. pinnipedii*. This paper includes samples from various pre-colonial site locations across North and South America to further explore MTBC strain diversity and ecology in the Americas and gain a better understanding on the geographic and temporal presence of *M. pinnipedii*. In this study, we present the challenges that are often in present in ancient DNA research including many that are unique to ancient MTBC studies. Most ancient samples are recovered from environmental matrices teeming with closely related environmental mycobacteria. While environmental contamination is not a concern unique to ancient MTBC studies, this body of research is fraught with challenges due to the novel nature of ancient MTBC research and the abundance of environmental mycobacteria closely related to the MTBC. The measures taken in the construction of ancient MTBC genomes for confident identification, authentication and evaluation of data are presented in this work and highlighted as results show evidence of a new strain of MTBC recovered from the Mesoamerican region of North America. Key findings in this work are as follows:

1. Four ancient MTBC genomes were reconstructed from archaeological samples recovered from Chile and central Mexico. Phylogenetic analysis reveals a strain circulating in central Mexican populations prior to European arrival which has not yet been identified. This strain is outside of the current known subclade of ancient South American MTBC and modern *M. pinnipedii* strains.

2. Inspection of the sequencing data reveals these genomes do not exhibit DNA damage typical of ancient DNA. This will be of particular importance for the ancient strains in Mexico, as they show a unique phylogenetic position and will need further authentication to support this finding.

3. One ancient MTBC genome was recovered from a tooth. This will be the first identification of ancient MTBC from a dental sample and is an unexpected finding due to the known behavior and pathophysiology of MTBC members which have an affinity for oxygen rich environments. However, this finding is not impossible as in very rare cases, tuberculosis may become septicemic.

4. Metagenomic evaluation of the microbial composition of the extract from the tooth yielding the MTBC genome found no oral microbial signature. This is unusual for a dental sample even in the presence of decontamination measures.

In summary, this work had many promising results, but due to problems that arose in initial data preparation, transport of the samples and inconsistencies in the quality of data, new extractions and data generation were performed. This sample set provides an excellent example of the importance of authentication measures that should be performed. To present these new findings it is important to have support from authentication measures and quality assessments.

Figure 1. Map showing the site locations of pre-colonial genomes presented in this thesis. This thesis presents 12 newly reconstructed genomes from five archaeological sites across North and South America. The number of genomes from each site are as follows: Tlatelolco (n=1), Tenochtitlan (n=1), Diablo Huasi (n=1), Huari (n=7), Pica 8 (n=1). The previously published pre-colonial South American genomes (Bos et al., 2014) are shown as blue triangles. This figure was created with assistance from Kathrin Nägele.

4 Discussion

4.1 Morphological analysis of skeletal pathology (papers I, II, III)

For this thesis, a total of 117 complete skeletal individuals (Ray Site) (paper I), 2,226 commingled vertebrae and more than 11 commingled ribs (sectors of Huari) (papers II and III) were analyzed by E.A.N. (doctoral candidate) for pathology using morphology-based methods. These sample sets demonstrate the value of preservation and skeletal completeness (Ray Site population, paper I) with the contrast of challenges introduced by burial programs such as commingling and disturbed burials (Huari populations, papers II and III). Likewise, the skeletal manifestation of the diseases identified in these populations differ drastically in their specificity, and subsequently, the ease at which etiological identification is possible. For example, in some cases tuberculosis will cause diagnostic skeletal changes, however, skeletal fluorosis results in non-specific changes with characteristic distribution across the skeleton. Diseases associated with dietary toxicity and metabolic syndromes such as skeletal fluorosis often have gradual effects that accumulate through chronic ingestion or exposure of toxic elements. Therefore, clinical and paleopathological investigations of skeletal fluorosis must take into account the full spectrum of progressive changes and inferred symptoms (Petrone et al. 2013; Littleton 1999; Teotia and Teotia, 1988). As demonstrated in paper I of this thesis, early signs of skeletal fluorosis result in slight to moderate skeletal expressions of the disease which are shared across multiple diseases (Brickley et al., 2020; Izuora et al., 2013; Littleton, 1990). Because the observed pathological changes were non-diagnostic and may result in multiple conditions, analysis of the Ray Site remains resulted in a broad differential diagnosis. Fortunately, due to the skeletal completeness of these individuals, the distribution of skeletal changes could be taken into account and the shared suite of pathological changes across the skeleton and across the affected population were observed. This allowed for a narrowing of the possibilities and increased the confidence of the differential diagnosis. Skeletal fluorosis is a key example for the influence environment may directly have on population health. Therefore, it is paramount that in the development of a differential diagnosis the environmental context is considered.

In contrast, TB may manifest on the skeleton with diagnostic pathological lesions and lesion distribution. As discussed in paper II, distribution of tuberculosis infection varies between adult

and subadults, whereby subadults present more frequent cranial involvement with lytic lesion development on the frontal bone (Ortner, 2003; Roberts and Buikstra, 2019; Straus, 1933). Typically, adult skeletal tuberculosis results in spinal destruction by the development of lytic lesions, usually on the thoracic and lumbar vertebrae, as first described by Percival Pott in 1779. Although the now termed "Pott's disease", or destruction of the intervertebral disc space and adjacent vertebrae, is the most common manifestation of skeletal tuberculosis, extraspinal lesions are possible (Tuli, 2016) as described in paper II. However, archaeological contexts often present varied skeletal completeness due to weathering, looting, bioturbation or burial programs of the cultures under study (papers II and III). Skeletal assessments to construct a confident differential diagnosis of pathology are highly dependent on skeletal completeness and the quality of preservation. For example, when assessments include a limited number of skeletal elements of the individual the evaluation of lesion distribution and patterning across the skeleton is not possible and the resulting differential diagnosis will include many possible conditions (papers II, III, and IV). These observations of commingled burials and atypical lesion distribution represent opportunities for the incorporation of molecular methods.

4.1.1 The value of morphological analysis to molecular research (papers II, III, IV)

As demonstrated in papers II, III and IV of this thesis, sample selection is an important aspect of successful recovery of pathogens and is directly related to understanding the pathophysiology of an infectious disease, such as tuberculosis. Pathogens have differing behavior and varied responses on the human skeleton. Paleopathological and pathophysiological knowledge contribute to sample selection in two ways, both involved in identifying the sample most likely to yield pathogen DNA based on 1) pathogen behavior, and 2) the stage of disease (active or healing).

Knowledge of pathogen behavior and physiology influences the anatomical location in which a pathogen is most likely to be recovered (paper II). For example, MTBC bacilli disseminate through the body and commonly concentrate in areas of arterial blood supply and high oxygen tension which promotes mycobacterial growth. Therefore, well vascularized, marrow-rich bones, such as vertebrae, will likely yield the richest recovery of TB DNA (paper II). Furthermore, direction of pathogen spread is reflected in vertebral lesions where the central type and skipped lesions,

paradiscal lesions and anterior type lesions are each representative of the vessels involved in the hematogenous delivery of the bacilli (Tuli, 2016). Septicemic tuberculosis is rare and therefore, skeletal signs such as these should indicate presence of TB bacilli in these elements and by extension the possible DNA. This is in contrast to bloodborne infections such as *Yersinia pestis*, in which teeth serve as the most appropriate sample for ancient pathogen recovery (Bos et al., 2011; Margaryan et al., 2015; Rasmussen et al., 2015; Spyrou et al., 2018). The study in paper IV may illustrate this point as MTBC was recovered from the vertebra of individual DIA002, yet no trace of MTBC DNA was detected in the tooth or petrous. In fact, all positive detections of MTBC DNA included in this thesis were recovered from vertebral samples, except for sample reported to come from a tooth (paper V). Although this tooth sample was deemed as questionable due to the lack of any oral signature in the sample, it should be noted that TB septicemia, although rare, could lead to detection of MTBC DNA in the dental pulp chamber.

Understanding the stage of disease and the ability to identify active or healing stages of disease in osteological samples is of particular importance for ancient TB research. Tuberculosis-like lesions develop in the bone generally 2 to 3 years after the primary infection of bone by way of vascular channels (Girling et al., 1988). Vertebrae are the most frequently affected where lesions typically form in the intervertebral joint space (the anterior or inferior surfaces of adjacent vertebrae) or on the anterior surface of the vertebral body. As discussed in paper II, lesions are produced by the intercellular infection of multinuclear osteoclasts, leading to the dysregulation of the cell and pathological activation of an osteolytic response (Hoshino et al., 2014). This osteolytic response continues throughout active TB infection, however, during latency the disease regresses and bone may heal resulting in extensive new bone development (Mousa, 1998). Therefore, it is believed skeletal samples which show bone destruction in the form of resorptive lesions with a loss of the extracellular matrix and present little to no bone formation serve as the most likely candidates for MTBC DNA recovery.

For population-based studies it should also be considered that skeletal changes in response to MTBC infection occur in only a small percentage of the infected population and varies widely across populations (1-20%) (Davidson & Horowitz, 1970; Davies, 1994; Jaffe, 1972; Resnick & Niwayama, 1988). This suggests the observation of TB skeletal lesions in a single individual likely

represents a much larger number of infections. In such cases, slight and non-specific skeletal changes should be included in consideration for molecular testing as presented in papers II and III of this thesis. In these papers (II and III), the Huari vertebrae displayed a spectrum of skeletal presentation ranging from characteristic lytic TB lesions to non-specific changes to no detectable pathology.

Molecular paleopathology uses inherently destructive methods on human remains to generate data and these remains are often limited and of significant scientific and cultural value (Margaryan et al., 2018; Roberts, 2016). Therefore, the need to optimize sample selection to limit destruction to collections of human remains and increase experiment success should be a priority. Ethical research design includes the selection of skeletal samples based on knowledge of pathogen behavior and skeletal disease process to increase the likelihood of successful retrieval of ancient pathogen DNA and justify destruction of the sample.

4.1.2 The value of molecular analysis to TB paleopathological research (papers II-V)

The value of NGS molecular methods for puzzling cases of commingled burials, incomplete skeletons, non-specific skeletal changes or the absence of skeletal manifestation is evident throughout this thesis (papers II, III, and IV). The ability to indiscriminately sequence the microbial composition of an extraction allowed for the detection of MTBC through a metagenomic screening approach with the Megan Alignment Tool (MALT) (Vågene et al., 2018). Many of the vertebrae from Huari which yielded MTBC DNA showed only slight skeletal changes or no macroscopically detectable skeletal changes (papers II and IV). Additionally, the application of metagenomic methods allowed for the evaluation of environmental background included in an extraction (papers II, III, IV, V). As demonstrated in this thesis, this enables efficient research design to not only detect pathogens but also for the selection of samples for efficient in-solution capture and genome reconstruction. The recovery of genome wide data permits phylogenetic analysis and evaluation of pathogen evolution. The application of NGS methods to paleopathology provides an opportunity for the evaluation of pathognomonic data to understand evolutionary histories of ancient diseases at an unprecedented resolution increasing the biomedical relevance and value of the field.

In analyses of MTBC strain evolution and ecology using aDNA methods, the quality and confidence in genome reconstruction are of utmost importance for all evolutionary analyses that will follow. The method employed in this thesis for targeted enrichment and subsequent genome reconstruction was an in-solution hybridization capture (Fu et al., 2013) in which amplified indexed libraries were added to a solution of single stranded DNA probes complementary to DNA of MTBC members, saturating the sample and binding to the target MTBC DNA. These probes were based on a computationally constructed ancestor for MTBC (Comas et al., 2010) hypothetically allowing for the recovery of ancient and modern MTBC members. Although these probes are specific for MTBC in great measure, environmental mycobacteria are ubiquitous and their genetic similarity may result in the unintentional inclusion in capture product (Warinner et al., 2017). Therefore, in papers III – IV, following genome reconstruction and preparation of the sequenced data through quality filtering and removal of duplicate reads, authentication of the observed genetic variation was performed using SNP Evaluation (https://github.com/andreasKroepelin/SNP_evaluation) based on parameters outlined by Keller et al., 2019 but augmented for MTBC DNA. The evaluation of variant positions allowed for the detection and removal of non-confident SNPs which may have derived from environmental microbes and permitted the production of a reconstructed MTBC genome with high confidence and quality. The application of these methods allowed for the reconstruction of 12 ancient MTBC genomes from various skeletal samples across multiple environments in the pre-colonial Americas (papers II – V) (Figure 2). Although four of the genomes presented in this thesis are of a preliminary nature, the remaining eight could be used for confident phylogenetic and evolutionary analysis to evaluate MTBC strains circulating in the Late Intermediate Period (~1000 – 1400 CE) of the South American Andean cultural region.

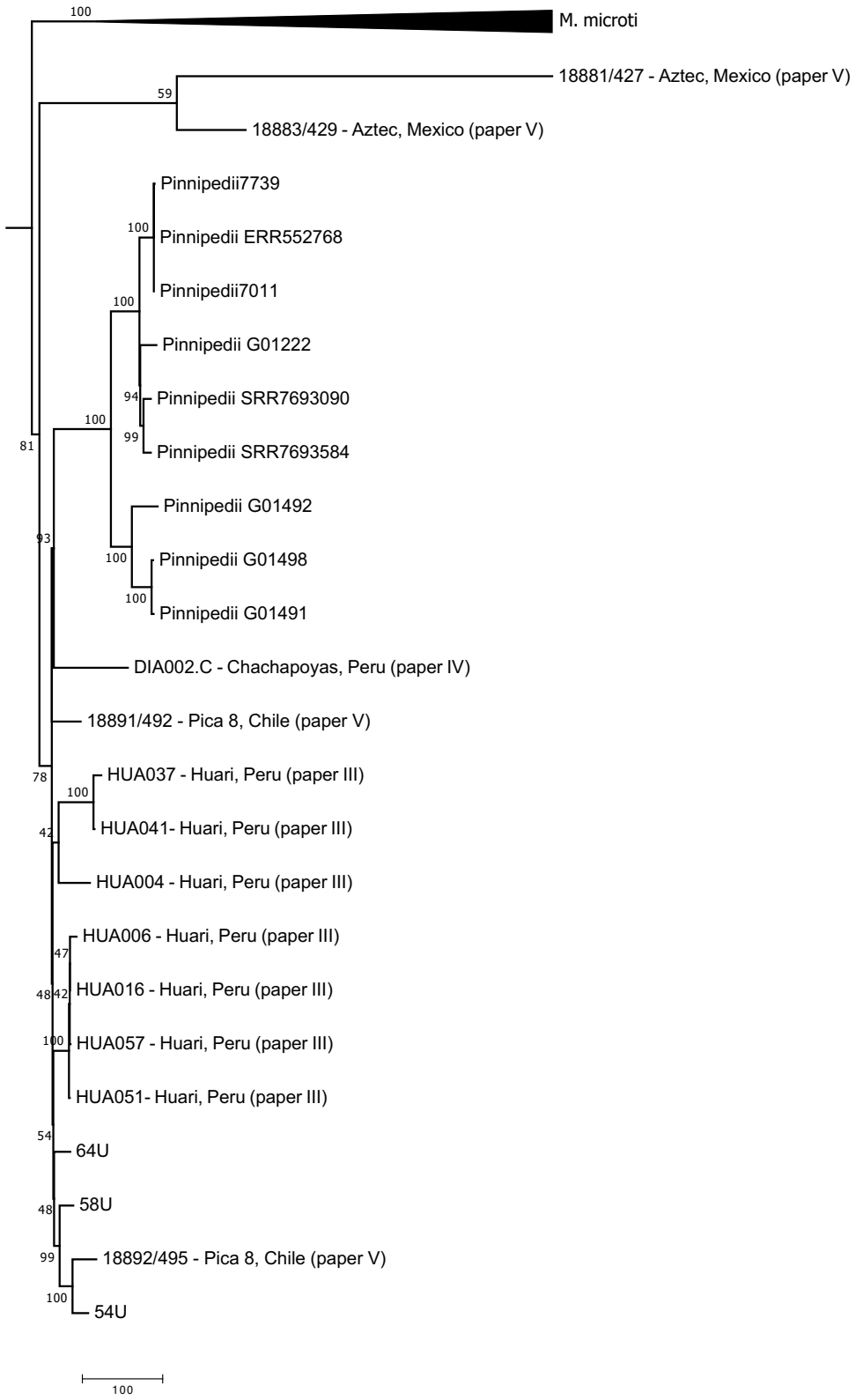


Figure 2. Phylogeny of the reconstructed MTBC genomes presented in this thesis. This Maximum Parsimony phylogeny was constructed with a 98% partial deletion of data thereby using 36,684 sites out of

a total of 56441. The tree was constructed using a Subtree-Pruning-Regrafting model. The values on the nodes are bootstraps generated with 500 iterations. The Huari, Chachapoya and Bos et al., 2014 genomes were filtered to remove any non-confident positions which may be deriving from environmental bacteria. Due to the low coverage of sample 18881/427 a 3-fold coverage minimum for SNP calling was used. The alignment was prepared using MultiVCFAnalyzer v0.85 (Bos et al., 2014). This phylogeny was prepared using MEGA7 (Kumar et al., 2016).

4.2 Tuberculosis in the Late Intermediate Period Andean region (papers II – V)

The Late Intermediate Period (LIP, ~1000 – 1400 CE) of Andean South America was characterized by significant cultural transitions and climatic variability. The early LIP was ushered in by the decline of the Wari empire and the disintegration of the Tiwanaku state (Isbell, 2008), contemporaneous with a severe drought in the Andean highlands and a movement to more dispersed settlements (Fehren-Schmitz et al., 2014; Seltzer and Hastorf, 1990; Tung, 2012). Concurrent with these climatic and socio-political shifts, there is a rise in cases of skeletal tuberculosis from North and South American archaeological contexts (Buikstra and Roberts, 2003). Archaeological evidence shows the highest density of cases in South America existing in Peru and Chile giving rise to the hypothesis that this was the region where MTBC first developed in the Americas. The largest skeletal series of ancient TB from the Andean cultural region during the LIP comes from the Estuquiña in the Moquegua Valley where 37 individuals were identified (Buikstra and Williams, 1991). Additional sites across the Andean cultural region provide skeletal evidence of TB during this time period (discussed in paper II). However, pre-colonial cases reach a peak during the LIP and TB seems to decline during Inka rule in the Late Horizon (1475 – 1534 CE).

This thesis contributes genomic evidence to support further investigations of how the LIP climatic and socio-political shifts, and their associated consequences, promoted the spread of TB across the Andean region (papers II and III). Specifically, I present analyses of reconstructed LIP genomes from multiple sites which reveal closely related ancient MTBC strains from across the Andean cultural region. Phylogenetic analysis of these genomes reveals the ancient MTBC strains are clustering together with some diversity and are closely related to modern *M. pinnipedii* strains (Figure 2). The identification of ancient strains, closely related to the modern *M. pinnipedii*, from

coastal, highland and high-altitude subtropical forests of the northeastern Andean slopes reveals a previously unknown geographic and ecological presence of this strain. Although this landscape may appear challenging and seem impassible, Andean cultures have a long history of interaction, trade and movement across multiple ecological zones (Covey, 2008; Guengerich, 2015; Pomeroy, 2013; Toohey, 2016). For LIP populations, such activities certainly would have facilitated the movement of infectious disease. However, due to the broad host tropism of MTBC members (Brites et al., 2018), it is not possible speak to human-to-human transmission as several alternative mammalian hosts are found this region, including llamas, alpacas and guinea pigs, all of which are known to contract and transmit TB (Grange and Yates, 1994).

The conditions of the LIP and the rise of skeletal TB reflects the dynamics of socio-political disruption, climatic fluctuations, dietary shifts and migration (movement to dispersed settlements), all of which are known influence the spread of TB (Lönnroth et al., 2009). Additional reconstructed genomes from archaeologically and environmentally contextualized pre-colonial Andean TB cases will provide greater resolution to understand factors related to this increase in TB and their possible influence on MTBC evolution and spread. There currently exists only three published ancient South American MTBC genomes, however this thesis adds 12 genomes to the known diversity of MTBC in the pre-colonial Americas. Thus, further developing a database to address broader questions of biological, climatic, environmental and socio-cultural interplay influencing tuberculosis.

4.2.1 Open questions on ancient South American MTBC research (papers II-V)

Although morphological reports of skeletal tuberculosis are abundant across South America, only a few ancient genomes exist (Bos et al., 2014; papers II-V) providing limited temporal and geographic representation. Therefore, the breadth of diversity and geographic presence of MTBC strains remain not well understood. With the development of more data crossing a larger time transect we may begin to understand not only strain diversity but the rise of TB during the LIP and its subsequent decline during the Late Horizon (1400 – 1532 CE). Likewise, because human *M. pinnipedii* is no longer known to regularly circulate in human populations, this may shed light on the evolutionary fate of the ancient *M. pinnipedii* strains. For South American populations *M.*

pinnipedii is believed to have been replaced by European associated strains of TB (L4), however, there remains a lack of understanding of when (if ever) and why this lineage ceased to infect humans. Furthermore, despite the proposed "pinniped hypothesis" of initial MTBC introduction to the Americas by way of seals or sea lions, the phylogenies presented in this thesis have yet to offer strong support for this idea. Likewise, much earlier skeletal evidence of TB exists throughout North and South America which has not yet been molecularly characterized (Allison et al., 1973; Allison et al., 1981; Ritchie, 1952; Roberts and Buikstra, 2003). Therefore, the origin of MTBC to the Americas remains unconfirmed. To address this, further faunal and human archaeological samples will need to be analyzed.

Currently, global ancient MTBC research is growing and methods are developing at a rapid pace. Due to the limited number of ancient TB genomes in existence (Bos et al., 2014; Kay et al., 2015; Sabin et al., 2020), certain aspects of this research remain unstandardized and uncharacterized. One challenge faced in this dissertation is the authentication of ancient TB genomes which has been complicated by unusual patterns of DNA damage in some select samples (papers IV and V). Reconstructed genomes from papers IV and V demonstrate DNA damage below expectations for ancient DNA. A similar pattern was observed previously for another mycobacterial disease, *Mycobacterium leprae,* whereby the pathogen DNA had little to no damage (Schuenemann et al., 2013). This was attributed to the pathogen's waxy cell wall composed of lipids and mycolic acids, believed to help protect the DNA from environmental degradation. The same argument could be made for *Mycobacterium tuberculosis* complex strains as they are of the same genus and possess the same cellular wall structure (Brennan, 2003). However, this assessment is complicated by the variability in TB DNA damage across the currently existing reconstructed ancient TB genomes (Figure 3).

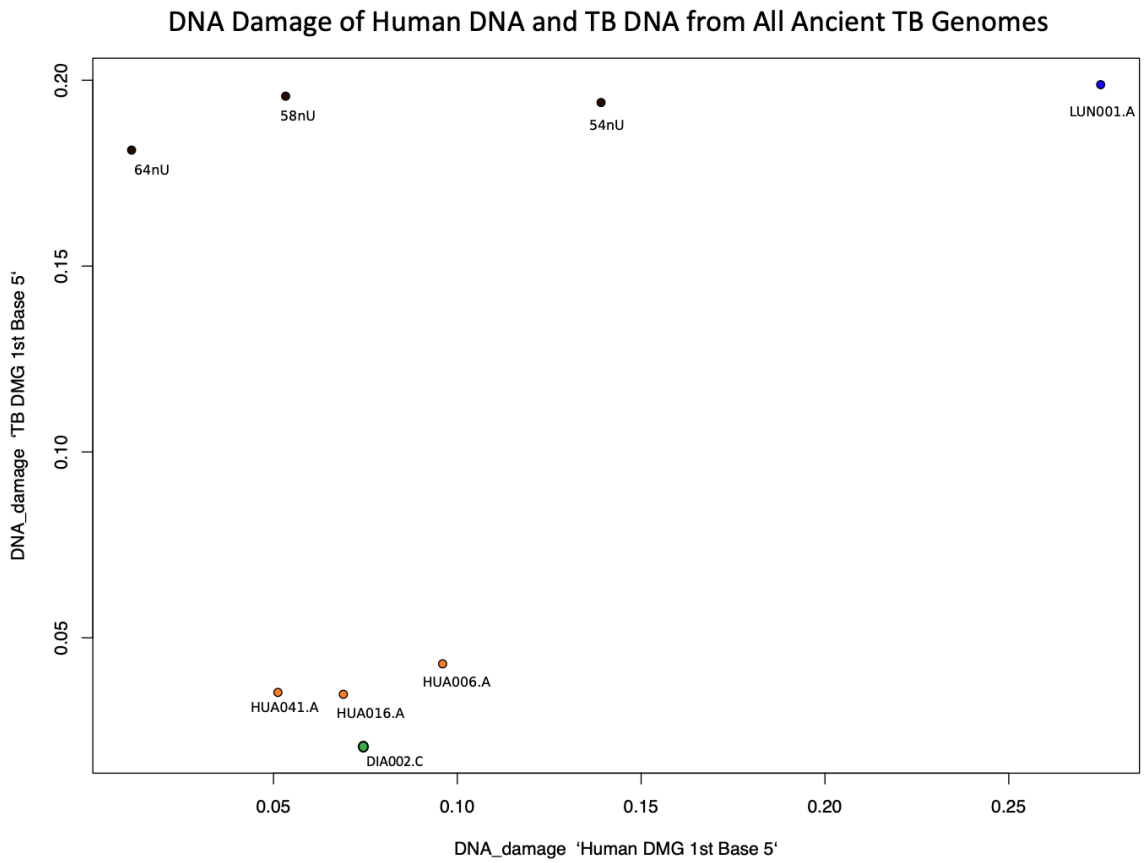


Figure 3. Scatterplot of DNA damage patterns of selected ancient TB genomes illustrating variability in damage patterns. The orange circles indicate genomes from Huari, Peru (paper III). The green circle indicates the genome reconstructed from the Chachapoyas region of Peru (paper IV). The black circles indicate ancient TB genomes from Peru (Bos et al., 2014). The blue circle indicates an ancient TB genome from Sweden (Sabin et al., 2020). Shotgun data was mapped to the human reference (Hg19) and the reconstructed MTBC ancestor using BWA-aln (Li and Durban, 2009) and damage was calculated using mapDamage (Jónsson et al., 2013). All samples are vertebrae from the LIP (1000-1400 CE) except for LUN001.A which is a 17th century lung nodule. Figure 3 shows the genomes prepared in this thesis have a noticeably lower TB DNA damage. Samples from paper V were not shown as those samples were not shotgun sequenced and only TB capture data was available.

The field of paleopathology currently faces questions on best practices and best methods for ancient pathogen research. Sampling strategies have been explored and established for ancient human DNA (Pinhasi et al., 2015). In contrast, best sampling strategies vary across pathogens and samples. Additionally, ancient pathogen research requires a variable quantity of samples for successful discovery depending on the project design and the disease being researched. For example, an ancient MTBC genome was recovered and reconstructed from one lung nodule (Sabin

et al., 2020), whereas the seven Huari MTBC genomes presented in paper III were recovered after analysis of 103 skeletal elements. Analysis of MTBC DNA recovery and anatomical location of the sampling site of the Huari vertebrae showed no relationship between MTBC DNA and sampling location (paper III). How do we create better sampling strategies and better research design? This problem is confounded by the many ancient pathogens that are of interest and the variability in samples likely to yield their DNA (Nelson et al., 2020). For all the achievements that have been made in molecular paleopathology there is still a need for principal methodological developments.

4.3 Molecular paleopathology: evolutionary thought and the biocultural approach

My final point of discussion involves the role of anthropological theory in molecular paleopathology. In the past decade alone, the study of ancient disease has seen continuous rapid and innovative growth with the development of new analytical methods and technologies. High-throughput sequencing platforms revolutionized bioarchaeological studies of past populations and ancient disease. Application of molecular methods to paleopathology answers the call for more technological advancements in methods of diagnosis (Ortner, 1991; Ubelaker, 2003). While this thesis has addressed the utility of molecular methods in identifying diseases that may have gone unrecognized, the merging of molecular and morphological methods should be viewed as an opportunity to explore the context of disease and develop more complex interpretations involving human biology, microbial threats, environmental factors and the cultural contexts (Grauer, 2012; Wright and Yoder, 2003). The abundance of data from high-throughput sequencing permits investigations of evolution, phylogeny, adaptation and ecology of both well-known and relatively under researched organisms in paleopathology (Zuckerman et al., 2012). These analyses have great value to understanding pathogens and disease today. However, the nature of these data can lead to a form of evolutionary thought at the risk of leaving the biocultural approach behind. When research begins to focus on the biological aspects of pathogen presence and evolution without considering socio-cultural factors, the study of ancient disease becomes isolated from the multi-factorial, dynamic event of disease (Zuckerman and Armelagos, 2011). Social relations and status, cultural practices, political frameworks, economic and environmental factors are a few of those which influence the emergence, transmission and persistence of disease, and by extension its

success. Because of this morphological and molecular paleopathology, in the study of ancient population health and disease, should contextualize evolutionary analyses to be relevant to modern populations and gain full scientific and ethical value (Armelagos and Van Gerven, 2003).

Disease is a powerful and complex phenomenon involving interplay of the environment, culture, society and biology. It has been involved in the restructuring of society (Herlihy, 1997), used as a tool in the intentional structured violence against marginalized groups (Selgelid, 2003; Davalos et al., 2020), and affected great influence on our own biology (Krause-Kyora et al., 2018; Lindo et al., 2016). Disease arises and propagates in these intersectional landscapes and without recognizing it within the appropriate theoretical frameworks, a great opportunity is lost to make meaningful reason of infectious disease emergence and success. Synthesizing information and contextualizing results require interdisciplinary contributions providing valuable insight on the pathogen and the context. Each specialization of paleopathology and by extension each actor in this research shares unique insight to contribute valuable insight to actively shape and reshape our understanding of disease. Paleopathology as a whole has never been better positioned to contribute to broad, comprehensive understandings of disease and provide transdisciplinary and widely relevant insights. The subfields of paleopathology still have challenges, yet through increased communication and collaboration we have the opportunity to extraordinarily impact population health and disease research across time and space.

5 Conclusion

In this thesis, I have used both morphological and molecular methods to conduct comprehensive paleopathological investigations of metabolic and infectious diseases in the pre-colonial Americas. I demonstrate the use of morphologically based paleopathological methods on a skeletal population from an environment rich in natural fluoride to present the first description of skeletal fluorosis from an American archaeological population. In addition, I present morphological assessments of TB and molecular analyses of ancient MTBC genomes from a variety of contexts across the Americas. In their respective manuscripts these results are discussed through a contextual framework including environmental, cultural, geographic and socio-political factors.

The value of comprehensive paleopathological analyses has been confirmed in this thesis which includes studies of skeletal samples presenting a variety of pathological changes, both non-specific and diagnostic. In this thesis, the incorporation of environmental and molecular data provides a greater understanding of disease in the Americas prior to European colonization. Although skeletal fluorosis is now considered a rare disease in North America, the study presented in this thesis contributes new knowledge to the field of North American paleopathology. This study argues for the consideration of fluorosis in studies presenting paleopathological differential diagnoses as it is likely this disease was far more common in populations from this region than previously recognized. Future bioarchaeological and paleopathological efforts may reveal more cases in time.

Alternatively, the investigation of MTBC strains presented in this thesis provides evolutionary information on an infectious organism that currently considered one of the world's greatest disease burdens (WHO, 2019). The insights provided in this thesis on MTBC strain diversity in the Americas and the geographic and ecological expanse of ancient *M. pinnipedii* strains suggest more connections will be made through future research efforts including ancient data sets from across the Americas. As there currently exist only three published ancient MTBC genomes from the Americas, in addition to the 12 presented in this thesis, additional sampling efforts will need to be made in order to draw confident conclusions on pre-colonial tuberculosis. As molecular data increase so should our incorporation of the biocultural context. As demonstrated in this thesis we gain insight on the success and dissemination of TB when we view our genomic results through the biocultural and environmental contexts. Through interdisciplinary endeavors we will gain insight to the introduction and evolution of MTBC in ancient South America and beyond.

6 References

Adler, C. J., Haak, W., Donlon, D., Cooper, A., & Consortium, G. (2011). Survival and recovery of DNA from ancient teeth and bones. Journal of Archaeological Science, 38(5), 956–964.

Allentoft, M. E., Collins, M., Harker, D., Haile, J., Oskam, C. L., Hale, M. L., Campos, P. F., Samaniego, J. A., Gilbert, M. T. P., & Willerslev, E. (2012). The half-life of DNA in bone: Measuring decay kinetics in 158 dated fossils. Proceedings of the Royal Society B: Biological Sciences, 279(1748), 4724–4733.

Allison, M. J., Gerszten, E., Munizaga, J., Santoro, C., & Mendoza, D. (1981). Tuberculosis in pre-Columbian Andean populations. Prehistoric Tuberculosis in the Americas, 49–61.

Allison, M. J., Mendoza, D., & Pezzia, A. (1973). Documentation of a case of tuberculosis in pre-Columbian America. American Review of Respiratory Disease, 107(6), 985–991.

Armelagos, G. J. (2008). Biocultural anthropology at its origins: Transformation of the new physical anthropology in the 1950s. The Tao of Anthropology, 269–282.

Armelagos, G. J., & van Gerven, D. P. (2003). A century of skeletal biology and paleopathology: Contrasts, contradictions, and conflicts. American Anthropologist, 105(1), 53–64.

Arriaza, B. T., Salo, W., Aufderheide, A. C., & Holcomb, T. A. (1995). Pre-Columbian tuberculosis in northern Chile: Molecular and skeletal evidence. American Journal of Physical Anthropology, 98(1), 37–45.

Aufderheide, A. C. (2011). Soft tissue taphonomy: A paleopathology perspective. International Journal of Paleopathology, 1(2), 75–80.

Aufderheide, A. C., Rodríguez-Martín, C., & Langsjoen, O. (1998). The Cambridge encyclopedia of human paleopathology (Vol. 478). Cambridge University Press Cambridge.

Barquera, R., Lamnidis, T. C., Lankapalli, A. K., Kocher, A., Hernández-Zaragoza, D. I., Nelson, E. A., Zamora-Herrera, A. C., Ramallo, P., Bernal-Felipe, N., Immel, A., Bos, K., Acuña-Alonzo, V., Barbieri, C., Roberts, P., Herbig, A., Kühnert, D., Márquez-Morfín, L., & Krause, J. (2020). Origin and Health Status of First-Generation Africans from Early Colonial Mexico. Current Biology, 30(11), 2078-2091.e11. https://doi.org/10.1016/j.cub.2020.04.002

Barras, V., & Greub, G. (2014). History of biological warfare and bioterrorism. Clinical Microbiology and Infection, 20(6), 497–502. https://doi.org/10.1111/1469-0691.12706

Bekvalac, J. J. (2017). Standardised osteological recording of archaeological skeletal material using an Oracle platform database: The Wellcome Osteological Research Database (WORD). American Journal of Physical Anthropology, 162, 116–116.

Bentley, G. E., Ubuka, T., McGuire, N. L., Chowdhury, V. S., Morita, Y., Yano, T., Hasunuma, I., Binns, M., Wingfield, J. C., & Tsutsui, K. (2008). Gonadotropin-inhibitory hormone and its receptor in the avian reproductive system. General and Comparative Endocrinology, 156(1), 34–43. https://doi.org/10.1016/j.ygcen.2007.10.003

Bliven, K. A., & Maurelli, A. T. (2016). Evolution of bacterial pathogens within the human host. Virulence Mechanisms of Bacterial Pathogens, 1–13.

Bos, K. I., Harkins, K. M., Herbig, A., Coscolla, M., Weber, N., Comas, I., Forrest, S. A., Bryant, J. M., Harris, S. R., & Schuenemann, V. J. (2014). Pre-Columbian mycobacterial genomes reveal seals as a source of New World human tuberculosis. Nature, 514(7523), 494–497.

Bos, K. I., Schuenemann, V. J., Golding, G. B., Burbano, H. A., Waglechner, N., Coombes, B. K., McPhee, J. B., DeWitte, S. N., Meyer, M., & Schmedes, S. (2011). A draft genome of Yersinia pestis from victims of the Black Death. Nature, 478(7370), 506–510.

Brennan, P. J. (2003). Structure, function, and biogenesis of the cell wall of Mycobacterium tuberculosis. Tuberculosis, 83(1–3), 91–97.

Brickley, M. B., Ives, R., & Mays, S. (2020). The bioarchaeology of metabolic bone disease. Academic Press.

Brickley, M., & McKinley, J. I. (2017). Guidelines to the standards for recording human remains. Institute of Field Archaeologists paper no. 7.

Briggs, A. W., Stenzel, U., Johnson, P. L., Green, R. E., Kelso, J., Prüfer, K., Meyer, M., Krause, J., Ronan, M. T., & Lachmann, M. (2007). Patterns of damage in genomic DNA sequences from a Neandertal. Proceedings of the National Academy of Sciences, 104(37), 14616–14621.

Brites, D., & Gagneux, S. (2015). Co-evolution of Mycobacterium tuberculosis and Homo sapiens. Immunological Reviews, 264(1), 6–24. https://doi.org/10.1111/imr.12264

Brites, D., Loiseau, C., Menardo, F., Borrell, S., Boniotti, M. B., Warren, R., Dippenaar, A., Parsons, S. D. C., Beisel, C., Behr, M. A., Fyfe, J. A., Coscolla, M., & Gagneux, S. (2018). A New Phylogenetic Framework for the Animal-Adapted Mycobacterium tuberculosis Complex. Frontiers in Microbiology, 9. https://doi.org/10.3389/fmicb.2018.02820

Buikstra, J. E. (1976). The Caribou Eskimo: General and specific disease. American Journal of Physical Anthropology, 45(3), 351–367. https://doi.org/10.1002/ajpa.1330450303

Buikstra, J. E. (2010). Paleopathology: A contemporary perspective. A Companion to Biological Anthropology, 395.

Buikstra, J. E. (2018). Ortner's identification of pathological conditions in human skeletal remains. Academic Press.

Buikstra, J. E., & Ubelaker, D. H. (1994). Standards for data collection from human skeletal remains. Arkansas archaeological survey. Research Series, 44.

Buikstra, J. E., & Williams, S. (1991a). Tuberculosis in the Americas: Current perspectives. Human Paleopathology, Current Syntheses and Future Options, Edited by D. Ortner and A. Aufderheide, Págs, 161–172.

Buikstra, J. E., & Williams, S. (1991b). Tuberculosis in the Americas: Current Perspectives. In Human Paleopathology: Current Syntheses and Future Options edited by D. Ortner and A. Aufderheide.

Buikstra, J., Williams, S., Ortner, D., & Aufderheide, A. (1991). Human Paleopathology: Current Syntheses and Future Options.

Cohen, M. L. (2000). Changing patterns of infectious disease. Nature, 406(6797), 762–767.

Comas, I., Chakravartti, J., Small, P. M., Galagan, J., Niemann, S., Kremer, K., Ernst, J. D., & Gagneux, S. (2010). Human T cell epitopes of Mycobacterium tuberculosis are evolutionarily hyperconserved. Nature Genetics, 42(6), 498–503. https://doi.org/10.1038/ng.590

Comas, I., Coscolla, M., Luo, T., Borrell, S., Holt, K. E., Kato-Maeda, M., Parkhill, J., Malla, B., Berg, S., & Thwaites, G. (2013). Out-of-Africa migration and Neolithic coexpansion of Mycobacterium tuberculosis with modern humans. Nature Genetics, 45(10), 1176–1182.

Cooper, A., & Poinar, H. N. (2000). Ancient DNA: do it right or not at all. Science, 289(5482), 1139–1139.

Covey, R. A. (2008). Multiregional perspectives on the archaeology of the Andes during the Late Intermediate Period (c. AD 1000–1400). Journal of Archaeological Research, 16(3), 287–338.

Dabney, J., Knapp, M., Glocke, I., Gansauge, M.-T., Weihmann, A., Nickel, B., Valdiosera, C., García, N., Pääbo, S., Arsuaga, J.-L., & Meyer, M. (2013). Complete mitochondrial genome sequence of a Middle Pleistocene cave bear reconstructed from ultrashort DNA fragments. Proceedings of the National Academy of Sciences, 110(39), 15758–15763. https://doi.org/10.1073/pnas.1314445110

Dabney, J., Meyer, M., & Pääbo, S. (2013). Ancient DNA damage. Cold Spring Harbor Perspectives in Biology, 5(7), a012567.

Dávalos, L. M., Austin, R. M., Balisi, M. A., Begay, R. L., Hofman, C. A., Kemp, M. E., Lund, J. R., Monroe, C., Mychajliw, A. M., & Nelson, E. A. (2020). Pandemics' historical role in creating inequality. Proceedings of the National Academy of Sciences of the United States of America, 117(25).

Davidson, P. T., & Horowitz, I. (1970). Skeletal tuberculosis: A review with patient presentations and discussion. The American Journal of Medicine, 48(1), 77–84.

Davies, P. D. O. (1994). Tuberculosis in the elderly. Journal of Antimicrobial Chemotherapy, 34(suppl_A), 93–100.

De Rojas, J. L. (2012). Tenochtitlan: Capital of the Aztec empire. University Press of Florida.

Derbes, V. J. (1966). De Mussis and the great plague of 1348. JAMA, 196(1), 59–62.

Drancourt, M., Aboudharam, G., Signoli, M., Dutour, O., & Raoult, D. (1998). Detection of 400-year-old Yersinia pestis DNA in human dental pulp: An approach to the diagnosis of

ancient septicemia. Proceedings of the National Academy of Sciences, 95(21), 12637–12640.

Drancourt, M., & Raoult, D. (2005). Palaeomicrobiology: Current issues and perspectives. Nature Reviews Microbiology, 3(1), 23–35.

Drancourt, M., Roux, V., Dang, L. V., Tran-Hung, L., Castex, D., Chenal-Francisque, V., Ogata, H., Fournier, P.-E., Crubézy, E., & Raoult, D. (2004). Genotyping, Orientalis-like Yersinia pestis, and Plague Pandemics. Emerging Infectious Diseases, 10(9), 1585–1592. https://doi.org/10.3201/eid1009.030933

Fehren-Schmitz, L., Haak, W., Mächtle, B., Masch, F., Llamas, B., Cagigao, E. T., Sossna, V., Schittek, K., Cuadrado, J. I., & Eitel, B. (2014). Climate change underlies global demographic, genetic, and cultural transitions in pre-Columbian southern Peru. Proceedings of the National Academy of Sciences, 111(26), 9443–9448.

Frischknecht, F. (2003). The history of biological warfare. EMBO Reports, 4(Suppl 1), S47–S52. https://doi.org/10.1038/sj.embor.embor849

Fu, Q., Meyer, M., Gao, X., Stenzel, U., Burbano, H. A., Kelso, J., & Pääbo, S. (2013). DNA analysis of an early modern human from Tianyuan Cave, China. Proceedings of the National Academy of Sciences, 110(6), 2223. https://doi.org/10.1073/pnas.1221359110

Gagneux, S., DeRiemer, K., Van, T., Kato-Maeda, M., De Jong, B. C., Narayanan, S., Nicol, M., Niemann, S., Kremer, K., & Gutierrez, M. C. (2006). Variable host–pathogen compatibility in Mycobacterium tuberculosis. Proceedings of the National Academy of Sciences, 103(8), 2869–2873.

Gall, G. E. C., Lautenschlager, S., & Bagheri, H. C. (2016). Quarantine as a public health measure against an emerging infectious disease: Syphilis in Zurich at the dawn of the modern era (1496–1585). GMS Hygiene and Infection Control, 11.

Giffin, K., Lankapalli, A. K., Sabin, S., Spyrou, M. A., Posth, C., Kozakaitė, J., Friedrich, R., Miliauskienė, Ž., Jankauskas, R., & Herbig, A. (2020). A treponemal genome from an historic plague victim supports a recent emergence of yaws and its presence in 15 th century Europe. Scientific Reports, 10(1), 1–13.

Gilbert, M. T. P., Bandelt, H.-J., Hofreiter, M., & Barnes, I. (2005). Assessing ancient DNA studies. Trends in Ecology & Evolution, 20(10), 541–544.

Giles-Vernick, T., Craddock, S., & Gunn, J. (2010). Influenza and public health: Learning from past pandemics. Earthscan Publications Ltd.

Girling, D. J., Darbyshire, J. H., Humphries, M. J., & O'Mahoney, G. (1988). Extra-pulmonary tuberculosis. British Medical Bulletin, 44(3), 738–756.

Grange, J. M., & Yates, M. D. (1994). Zoonotic aspects of Mycobacterium bovis infection. Veterinary Microbiology, 40(1–2), 137–151.

Grauer, A. L. (2012). Introduction: The scope of paleopathology. A Companion to Paleopathology, 1–14.

Grauer, A. L. (2018). A century of paleopathology. American Journal of Physical Anthropology, 165(4), 904–914.

Guengerich, A. (2015). Settlement organization and architecture in late intermediate period Chachapoyas, northeastern Peru. Latin American Antiquity, 362–381.

Hardie, R. M., & Watson, J. M. (1992). Mycobacterium bovis in England and Wales: Past, present and future. Epidemiology and Infection, 109(1), 23.

Harper, K., & Armelagos, G. (2010). The changing disease-scape in the third epidemiological transition. International Journal of Environmental Research and Public Health, 7(2), 675–697.

Herlihy, D. (1997). The Black Death and the transformation of the West. Harvard University Press.

Hershkovitz, I., Donoghue, H. D., Minnikin, D. E., Besra, G. S., Lee, O. Y.-C., Gernaey, A. M., Galili, E., Eshed, V., Greenblatt, C. L., Lemma, E., Bar-Gal, G. K., & Spigelman, M. (2008). Detection and Molecular Characterization of 9000-Year-Old Mycobacterium tuberculosis from a Neolithic Settlement in the Eastern Mediterranean. PLOS ONE, 3(10), e3426. https://doi.org/10.1371/journal.pone.0003426

Hoshino, A., Hanada, S., Yamada, H., Mii, S., Takahashi, M., Mitarai, S., Yamamoto, K., & Manome, Y. (2014). Mycobacterium tuberculosis escapes from the phagosomes of infected human osteoclasts reprograms osteoclast development via dysregulation of cytokines and chemokines. Pathogens and Disease, 70(1), 28–39.

Isbell, W. H. (2008). Wari and Tiwanaku: International identities in the central Andean Middle Horizon. In The handbook of South American archaeology (pp. 731–759). Springer.

Izuora, K., Twombly, J. G., Whitford, G. M., Demertzis, J., Pacifici, R., & Whyte, M. P. (2011). Skeletal fluorosis from brewed tea. The Journal of Clinical Endocrinology & Metabolism, 96(8), 2318–2324.

Jaffe, H. L. (1972). Skeletal lesions caused by certain other infectious agents. Metabolic, Degenerative and Inflammatory Diseases of Bone and Joints. Lea & Febiger, Philadelphia, Pa, 1015–1031.

Jarcho, S. (1966). The development and present condition of human palaeopathology in the United States. In Human palaeopathology (pp. 3–30). Yale University Press New Haven.

Jones, K. E., Patel, N. G., Levy, M. A., Storeygard, A., Balk, D., Gittleman, J. L., & Daszak, P. (2008). Global trends in emerging infectious diseases. Nature, 451(7181), 990–993.

Jónsson, H., Ginolhac, A., Schubert, M., Johnson, P. L., & Orlando, L. (2013). MapDamage2. 0: Fast approximate Bayesian estimates of ancient DNA damage parameters. Bioinformatics, 29(13), 1682–1684.

Kaplan, J. O., Krumhardt, K. M., Ellis, E. C., Ruddiman, W. F., Lemmen, C., & Goldewijk, K. K. (2011). Holocene carbon emissions as a result of anthropogenic land cover change. The Holocene, 21(5), 775–791.

Karesh, W. B., Dobson, A., Lloyd-Smith, J. O., Lubroth, J., Dixon, M. A., Bennett, M., Aldrich, S., Harrington, T., Formenty, P., Loh, E. H., Machalaba, C. C., Thomas, M. J., & Heymann, D. L. (2012). Ecology of zoonoses: Natural and unnatural histories. The Lancet, 380(9857), 1936–1945. https://doi.org/10.1016/S0140-6736(12)61678-X

Kay, G. L., Sergeant, M. J., Zhou, Z., Chan, J. Z.-M., Millard, A., Quick, J., Szikossy, I., Pap, I., Spigelman, M., Loman, N. J., Achtman, M., Donoghue, H. D., & Pallen, M. J. (2015). Eighteenth-century genomes show that mixed infections were common at time of peak tuberculosis in Europe. Nature Communications, 6(1), 6717. https://doi.org/10.1038/ncomms7717

Key, F. M., Posth, C., Krause, J., Herbig, A., & Bos, K. I. (2017). Mining Metagenomic Data Sets for Ancient DNA: Recommended Protocols for Authentication. Trends in Genetics, 33(8), 508–520. https://doi.org/10.1016/j.tig.2017.05.005

Kinzelbach, A. (2006). Infection, contagion, and public health in late medieval and early modern German imperial towns. Journal of the History of Medicine and Allied Sciences, 61(3), 369–389.

Klaus, H. D. (2017). Paleopathological rigor and differential diagnosis: Case studies involving terminology, description, and diagnostic frameworks for scurvy in skeletal remains. International Journal of Paleopathology, 19, 96–110. https://doi.org/10.1016/j.ijpp.2015.10.002

Klaus, H. D. (2020). Metabolic diseases in Andean paleopathology: Retrospect and prospect. International Journal of Paleopathology, 29, 54–64. https://doi.org/10.1016/j.ijpp.2019.06.008

Klaus, H. D., Wilbur, A. K., Temple, D. H., Buikstra, J. E., Stone, A. C., Fernandez, M., Wester, C., & Tam, M. E. (2010). Tuberculosis on the north coast of Peru: Skeletal and molecular paleopathology of late pre-Hispanic and postcontact mycobacterial disease. Journal of Archaeological Science, 37(10), 2587–2597.

Koboldt, D. C., Steinberg, K. M., Larson, D. E., Wilson, R. K., & Mardis, E. R. (2013). The next-generation sequencing revolution and its impact on genomics. Cell, 155(1), 27–38.

Koch, A., Brierley, C., Maslin, M. M., & Lewis, S. L. (2019). Earth system impacts of the European arrival and Great Dying in the Americas after 1492. Quaternary Science Reviews, 207, 13–36.

Koontz Scaffidi, B. (2020). Spatial paleopathology: A geographic approach to the etiology of cribrotic lesions in the prehistoric Andes. International Journal of Paleopathology, 29, 102–116. https://doi.org/10.1016/j.ijpp.2019.07.002

Krause-Kyora, B., Nutsua, M., Boehme, L., Pierini, F., Pedersen, D. D., Kornell, S.-C., Drichel, D., Bonazzi, M., Möbus, L., & Tarp, P. (2018). Ancient DNA study reveals HLA susceptibility locus for leprosy in medieval Europeans. Nature Communications, 9(1), 1–11.

Lindahl, T. (1993). Instability and decay of the primary structure of DNA. Nature, 362(6422), 709–715.

Lindo, J., Huerta-Sánchez, E., Nakagome, S., Rasmussen, M., Petzelt, B., Mitchell, J., Cybulski, J. S., Willerslev, E., DeGiorgio, M., & Malhi, R. S. (2016). A time transect of exomes from a Native American population before and after European contact. Nature Communications, 7(1), 1–11.

Littleton, J. (1999). Paleopathology of skeletal fluorosis. American Journal of Physical Anthropology: The Official Publication of the American Association of Physical Anthropologists, 109(4), 465–483.

Lönnroth, K., Jaramillo, E., Williams, B. G., Dye, C., & Raviglione, M. (2009). Drivers of tuberculosis epidemics: The role of risk factors and social determinants. Social Science & Medicine, 68(12), 2240–2246.

Majander, K., Pfrengle, S., Kocher, A., Neukamm, J., du Plessis, L., Pla-Díaz, M., Arora, N., Akgül, G., Salo, K., & Schats, R. (2020). Ancient Bacterial Genomes Reveal a High Diversity of Treponema pallidum Strains in Early Modern Europe. Current Biology.

Mardis, E. R. (2008). Next-generation DNA sequencing methods. Annu. Rev. Genomics Hum. Genet., 9, 387–402.

Margaryan, A., Hansen, H. B., Rasmussen, S., Sikora, M., Moiseyev, V., Khoklov, A., Epimakhov, A., Yepiskoposyan, L., Kriiska, A., Varul, L., Saag, L., Lynnerup, N.,

Willerslev, E., & Allentoft, M. E. (2018). Ancient pathogen DNA in human teeth and petrous bones. Ecology and Evolution, 8(6), 3534–3542. https://doi.org/10.1002/ece3.3924

Margulies, M., Egholm, M., Altman, W. E., Attiya, S., Bader, J. S., Bemben, L. A., Berka, J., Braverman, M. S., Chen, Y.-J., & Chen, Z. (2005). Genome sequencing in microfabricated high-density picolitre reactors. Nature, 437(7057), 376–380.

Mascher, M., Schuenemann, V. J., Davidovich, U., Marom, N., Himmelbach, A., Hübner, S., Korol, A., David, M., Reiter, E., & Riehl, S. (2016). Genomic analysis of 6,000-year-old cultivated grain illuminates the domestication history of barley. Nature Genetics, 48(9), 1089–1093.

Møller-Christiansen, V. (1961). Bone lesions in leprosy. Copenhagen: Munksgaard.

Morens, D. M., & Fauci, A. S. (2013). Emerging infectious diseases: Threats to human health and global stability. PLoS Pathog, 9(7), e1003467.

Morse, D. (1961). Prehistoric tuberculosis in America. American Review of Respiratory Disease, 83(4), 489–504.

Morse, S. (2004). Factors and determinants of disease emergence. Revue Scientifique et Technique-Office International Des Épizooties, 23(2), 443–452.

Morse, S. S., Mazet, J. A., Woolhouse, M., Parrish, C. R., Carroll, D., Karesh, W. B., Zambrana-Torrelio, C., Lipkin, W. I., & Daszak, P. (2012). Prediction and prevention of the next pandemic zoonosis. The Lancet, 380(9857), 1956–1965.

Mousa, H.-L. (1998). Tuberculosis of bones and joints: Diagnostic approaches. International Orthopaedics, 22(4), 245–246.

Ngabonziza, J. C. S., Loiseau, C., Marceau, M., Jouet, A., Menardo, F., Tzfadia, O., Antoine, R., Niyigena, E. B., Mulders, W., Fissette, K., Diels, M., Gaudin, C., Duthoy, S., Ssengooba,

W., André, E., Kaswa, M. K., Habimana, Y. M., Brites, D., Affolabi, D., … Supply, P. (2020). A sister lineage of the Mycobacterium tuberculosis complex discovered in the African Great Lakes region. Nature Communications, 11(1), 2917. https://doi.org/10.1038/s41467-020-16626-6

Orr, M. T., Fox, C. B., Baldwin, S. L., Sivananthan, S. J., Lucas, E., Lin, S., Phan, T., Moon, J. J., Vedvick, T. S., & Reed, S. G. (2013). Adjuvant formulation structure and composition are critical for the development of an effective vaccine against tuberculosis. Journal of Controlled Release, 172(1), 190–200.

Ortner, D. J. (1991). Theoretical and methodological issues in paleopathology. Human Paleopathology: Current Syntheses and Future Options, 5–11.

Ortner, D. J. (2003). Identification of pathological conditions in human skeletal remains. Academic Press.

Pääbo, S., Poinar, H., Serre, D., Jaenicke-Després, V., Hebler, J., Rohland, N., Kuch, M., Krause, J., Vigilant, L., & Hofreiter, M. (2004). Genetic analyses from ancient DNA. Annu Rev Genet, 38(1), 645–679.

Pääbo, S., & Wilson, A. C. (1988). Polymerase chain reaction reveals cloning artefacts. Nature, 334(6181), 387–388.

Petrone, P., Giordano, M., Giustino, S., & Guarino, F. M. (2011). Enduring fluoride health hazard for the Vesuvius area population: The case of AD 79 Herculaneum. PloS One, 6(6), e21085.

Petrone, P., Guarino, F. M., Giustino, S., & Gombos, F. (2013). Ancient and recent evidence of endemic fluorosis in the Naples area. Journal of Geochemical Exploration, 131, 14–27.

Pinhasi, R., Fernandes, D., Sirak, K., Novak, M., Connell, S., Alpaslan-Roodenberg, S., Gerritsen, F., Moiseyev, V., Gromov, A., & Raczky, P. (2015). Optimal ancient DNA yields from the inner ear part of the human petrous bone. PloS One, 10(6), e0129102.

Pomeroy, E. (2013). Biomechanical insights into activity and long distance trade in the south-central Andes (AD 500–1450). Journal of Archaeological Science, 40(8), 3129–3140.

Ragsdale, B. D., & Lehmer, L. M. (2012). A knowledge of bone at the cellular (histological) level is essential to paleopathology. A Companion to Paleopathology, 227–249.

Ragsdale, B. L. (1992). Task force on terminology: Provisional word list. Paleopathology Newsletter, 78, 7–8.

Rasmussen, M., Li, Y., Lindgreen, S., Pedersen, J. S., Albrechtsen, A., Moltke, I., Metspalu, M., Metspalu, E., Kivisild, T., & Gupta, R. (2010). Ancient human genome sequence of an extinct Palaeo-Eskimo. Nature, 463(7282), 757–762.

Rasmussen, S., Allentoft, M. E., Nielsen, K., Orlando, L., Sikora, M., Sjögren, K.-G., Pedersen, A. G., Schubert, M., Van Dam, A., Kapel, C. M. O., Nielsen, H. B., Brunak, S., Avetisyan, P., Epimakhov, A., Khalyapin, M. V., Gnuni, A., Kriiska, A., Lasak, I., Metspalu, M., … Willerslev, E. (2015). Early Divergent Strains of Yersinia pestis in Eurasia 5,000 Years Ago. Cell, 163(3), 571–582. https://doi.org/10.1016/j.cell.2015.10.009

Reid, A. H., Fanning, T. G., Hultin, J. V., & Taubenberger, J. K. (1999). Origin and evolution of the 1918 “Spanish” influenza virus hemagglutinin gene. Proceedings of the National Academy of Sciences, 96(4), 1651–1656.

Ritchie, W. A. (1952). Paleopathological evidence suggesting pre-Columbian tuberculosis in New York State. American Journal of Physical Anthropology, 10(3), 305–318.

Roberts, Charlotte A., & Buikstra, J. E. (2003). The bioarchaeology of tuberculosis: A global perspective on a re-emerging disease. University Press of Florida.

Roberts, Charlotte A., & Buikstra, J. E. (2019). Bacterial infections. In Ortner's Identification of pathological conditions in human skeletal remains (pp. 321–439). Academic Press.

Roberts, Charlotte A. (2016). Palaeopathology and its relevance to understanding health and disease today: The impact of the environment on health, past and present. Anthropological Review, 79(1), 1–16.

Sabin, S., Herbig, A., Vågene, Å. J., Ahlström, T., Bozovic, G., Arcini, C., Kühnert, D., & Bos, K. I. (2020). A seventeenth-century Mycobacterium tuberculosis genome supports a Neolithic emergence of the Mycobacterium tuberculosis complex. Genome Biology, 21(1), 1–24.

Saiki, R., Scharf, S., Faloona, F., Mullis, K., Horn, G., Erlich, H., & Arnheim, N. (1985). Enzymatic amplification of beta-globin genomic sequences and restriction site analysis for diagnosis of sickle cell anemia. Science, 230(4732), 1350. https://doi.org/10.1126/science.2999980

Salo, W. L., Aufderheide, A. C., Buikstra, J., & Holcomb, T. A. (1994). Identification of Mycobacterium tuberculosis DNA in a pre-Columbian Peruvian mummy. Proceedings of the National Academy of Sciences, 91(6), 2091–2094.

Sawyer, S., Krause, J., Guschanski, K., Savolainen, V., & Pääbo, S. (2012). Temporal patterns of nucleotide misincorporations and DNA fragmentation in ancient DNA. PloS One, 7(3), e34131.

Scaffidi, B. K. (2020). Spatial paleopathology: A geographic approach to the etiology of cribrotic lesions in the prehistoric Andes. International Journal of Paleopathology, 29, 102–116.

Schuenemann, V. J., Singh, P., Mendum, T. A., Krause-Kyora, B., Jäger, G., Bos, K. I., Herbig, A., Economou, C., Benjak, A., Busso, P., Nebel, A., Boldsen, J. L., Kjellström, A., Wu, H., Stewart, G. R., Taylor, G. M., Bauer, P., Lee, O. Y.-C., Wu, H. H. T., … Krause, J. (2013). Genome-Wide Comparison of Medieval and Modern Mycobacterium leprae. Science, 341(6142), 179–183. https://doi.org/10.1126/science.1238286

Selgelid, M. J. (2003). Smallpox revisited? American Journal of Bioethics, 3(1), 5–11.

Seltzer, G. O., & Hastorf, C. A. (1990). Climatic change and its effect on prehispanic agriculture in the central Peruvian Andes. Journal of Field Archaeology, 17(4), 397–414.

Shapiro, B., Rambaut, A., & Gilbert, M. T. P. (2006). No proof that typhoid caused the Plague of Athens (a reply to Papagrigorakis et al.). International Journal of Infectious Diseases, 10(4), 334–335.

Siddle, K. J., & Quintana-Murci, L. (2014). The Red Queen's long race: Human adaptation to pathogen pressure. Current Opinion in Genetics & Development, 29, 31–38.

Spyrou, M. A., Bos, K. I., Herbig, A., & Krause, J. (2019). Ancient pathogen genomics as an emerging tool for infectious disease research. Nature Reviews Genetics, 20(6), 323–340. https://doi.org/10.1038/s41576-019-0119-1

Spyrou, M. A., Tukhbatova, R. I., Wang, C.-C., Valtueña, A. A., Lankapalli, A. K., Kondrashin, V. V., Tsybin, V. A., Khokhlov, A., Kühnert, D., & Herbig, A. (2018). Analysis of 3800-year-old Yersinia pestis genomes suggests Bronze Age origin for bubonic plague. Nature Communications, 9(1), 1–10.

Stead, W. W. (1997). The origin and erratic global spread of tuberculosis: How the past explains the present and is the key to the future. Clinics in Chest Medicine, 18(1), 65–77.

Stead, W. W., Eisenach, K. D., Cave, M. D., Beggs, M. L., Templeton, G. L., Thoen, C. O., & Bates, J. H. (1995). When did Mycobacterium tuberculosis infection first occur in the New World? An important question with public health implications. American Journal of Respiratory and Critical Care Medicine, 151(4), 1267–1268.

Strauss, D. C. (1933). Tuberculosis of the flat bones of the vault of the skull. Surg Gynaecol Obstet, 57, 384–398.

Suby, J. A. (2020). Paleopathological Research in Southern Patagonia: An Approach to Understanding Stress and Disease in Hunter-Gatherer Populations. Latin American Antiquity, 31(2), 392–408.

Teotia, S. P. S., & Teotia, M. (1988). Endemic skeletal fluorosis: Clinical and radiological variants. Fluoride, 21(1), 39–44.

Titelbaum, A. R. (2020). Developmental anomalies and South American paleopathology: A comparison of block vertebrae and co-occurring axial anomalies among three skeletal samples from the El Brujo archaeological complex of northern coastal Peru. International Journal of Paleopathology, 29, 76–93.

Tognotti, E. (2013). Lessons from the history of quarantine, from plague to influenza A. Emerging Infectious Diseases, 19(2), 254.

Toohey, J. L. (2016). Reconciling archaeological and ethnohistoric data for coast-highland interaction in the Cajamarca region. Ñawpa Pacha, 36(2), 185–208.

Toyne, J. M., Esplin, N., & Buikstra, J. E. (2020). Examining variation in skeletal tuberculosis in a late pre-contact population from the eastern mountains of Peru. International Journal of Paleopathology, 30, 22–34.

Tuli, S. M. (2016). Tuberculosis of the skeletal system. JP Medical Ltd.

Tung, T. A. (2012). Violence, ritual, and the Wari Empire: A social bioarchaeology of imperialism in the ancient Andes. University Press of Florida.

Ubelaker, D. H. (2003). Anthropological perspectives on the study of ancient disease. Archaeology, Ecology and Evolution of Infectious Disease, 93–102.

Ubelaker, D. H., & Buikstra, J. E. (1994). Standards for data collection from human skeletal remains. Arkansas Archaeological Survey Research, 44, 206.

Vågene, Å. J., Herbig, A., Campana, M. G., García, N. M. R., Warinner, C., Sabin, S., Spyrou, M. A., Valtueña, A. A., Huson, D., & Tuross, N. (2018). Salmonella enterica genomes from victims of a major sixteenth-century epidemic in Mexico. Nature Ecology & Evolution, 2(3), 520–528.

Vlok, M., Oxenham, M. F., Domett, K., Tran Thi, M., Nguyen Thi Mai, H., Matsumura, H., Trinh, H. H., Higham, T., Higham, C., & Nghia, T. H. (2020). Two probable cases of infection with Treponema pallidum during the Neolithic period in Northern Vietnam (~ 2000-1500B. C.). Bioarchaeology International, 4, 15–36.

Warinner, C., Herbig, A., Mann, A., Fellows Yates, J. A., Weiß, C. L., Burbano, H. A., Orlando, L., & Krause, J. (2017). A robust framework for microbial archaeology. Annual Review of Genomics and Human Genetics, 18, 321–356.

Warinner, C., Rodrigues, J. F. M., Vyas, R., Trachsel, C., Shved, N., Grossmann, J., Radini, A., Hancock, Y., Tito, R. Y., & Fiddyment, S. (2014). Pathogens and host immunity in the ancient human oral cavity. Nature Genetics, 46(4), 336–344.

Wheelis, M. (2002). Biological Warfare at the 1346 Siege of Caffa. Emerging Infectious Diseases, 8(9), 971–975. https://doi.org/10.3201/eid0809.010536

Wilbur, A. K., & Buikstra, J. E. (2006). Patterns of tuberculosis in the Americas: How can modern biomedicine inform the ancient past? Memórias Do Instituto Oswaldo Cruz, 101, 59–66.

Wilczak, C. A., & Dudar, J. C. (2011). Osteoware software manual. Washington DC: Smithsonian Institution.

Wood, J. W., Milner, G. R., Harpending, H. C., Weiss, K. M., Cohen, M. N., Eisenberg, L. E., Hutchinson, D. L., Jankauskas, R., Cesnys, G., & Česnys, G. (1992). The osteological paradox: Problems of inferring prehistoric health from skeletal samples [and comments and reply]. *Current Anthropology*, *33*(4), 343–370.

Wright, L. E., & Yoder, C. J. (2003). Recent progress in bioarchaeology: approaches to the osteological paradox. *Journal of Archaeological Research*, *11*(1), 43-70.

World Health Organization. (2019). WHO Global Tuberculosis Report. WHO. https://www.who.int/tb/publications/global_report/en/

Zink, A., Haas, C. J., Reischl, U., Szeimies, U., & Nerlich, A. G. (2001). Molecular analysis of skeletal tuberculosis in an ancient Egyptian population. Journal of Medical Microbiology, 50(4), 355–366.

Zuckerman, M. K., & Armelagos, G. J. (2011). The origins of biocultural dimensions in bioarchaeology. Social Bioarchaeology, 15–43.

Zuckerman, M., Turner, B., & Armelagos, G. (2012). Evolutionary thought and the rise of the biocultural approach in paleopathology. A Companion to Paleopathology, A Grauer (Ed.). Wiley-Blackwell: New York, 34–58

7 Figures

8 Appendix

The following five papers are included and discussed in this dissertation:

I. Elizabeth A. Nelson, Christine L. Halling, Jane E. Buikstra. Evidence of Skeletal Fluorosis at the Ray Site, Illinois, USA: a pathological assessment, discussion of environmental factors. *International Journal of Paleopathology*, 26, 48-60.

II. Elizabeth A. Nelson, Jane E. Buikstra, Alexander Herbig, Tiffiny A. Tung, Kirsten I. Bos. Advances in Paleopathology of Tuberculosis in Pre-contact Andean South America. *International Journal of Paleopathology*, 29, 128-140.

III. Elizabeth A. Nelson, Aditya Kumar Lankapalli, Maria Spyrou, Susanna Sabin, Åshild Vågene, Ainash Childebayeva, Ben Rohrlach, James A. Fellows Yates, Martha Cabrera, Jose Ochotama, Tiffiny A. Tung, Alexander Herbig, Kirsten I. Bos. Tuberculosis in the wake of the Wari empire. *Manuscript*.

IV. Elizabeth A. Nelson, Evelyn Guevara, J. Marla Toyne, Alexander Herbig, Johannes, Krause, Kirsten I. Bos. Tuberculosis in the pre-colonial Chachapoya of Peru. *Mansucript*.

V. Elizabeth A. Nelson*, Kelly E. Blevins*, Jane E. Buikstra, Alexander Herbig, Johannes Krause, Anne C. Stone, Kirsten I. Bos. Preliminary Work: Identifying and Characterizing TB Strains Across the pre-colonial Americas. *Manuscript*.

*contributed equally to this work

Paper I

E. A. Nelson, C. L. Halling, J. E. Buikstra (2019).

Evidence of Skeletal Fluorosis at the Ray Site, Illinois, USA: a pathological assessment, discussion of environmental factors

International Journal of Paleopathology, 26, 48-60.

International Journal of Paleopathology 26 (2019) 48–60

Contents lists available at ScienceDirect

International Journal of Paleopathology

journal homepage: www.elsevier.com/locate/ijpp

Evidence of Skeletal Fluorosis at the Ray Site, Illinois, USA: a pathological assessment and discussion of environmental factors

Elizabeth A. Nelson[a,*], Christine L. Halling[b], Jane E. Buikstra[c]

[a] Department of Archaeogenetics, Max Planck Institute for the Science of Human History, Jena, Germany
[b] Department of Justice, State of Louisiana, United States
[c] Center for Bioarchaeological Research, Arizona State University, United States

ARTICLE INFO

Keywords:
Metabolic Bone Disease
Fluoride toxicity
North America
Differential Diagnosis

ABSTRACT

Objective: To carefully assess skeletal lesions in close environment context in order to evaluate whether skeletal fluorosis was present in individuals living in the prehistoric Midwest, USA.
Materials: Skeletal remains from minimally 117 individuals recovered from the Ray Site, located in western Illinois (USA) and dated to the Middle/early Late Woodland periods (50 BC-AD 400).
Methods: Macroscopic evaluation of all recovered skeletal elements.
Results: Eight individuals display a constellation of abnormal bony changes, including osteosclerosis, a high frequency of fractures, and dental abnormalities.
Conclusions: The osteosclerotic changes along with the naturally high fluoride content of west central Illinois soil and water suggests the presence of skeletal fluorosis.
Significance: This is the first report of skeletal fluorosis from archaeologically recovered human remains from North America.
Limitations: The ambiguous nature of the skeletal changes associated with fluorosis, especially in the less severe stages of the disease, renders determination of the etiology difficult.
Suggestions for Further Research: The continuation of paleopathological investigations of fluoride toxicity within archaeological communities recovered from this region with emphasis on the incorporation of biomedical and environmental data. Furthermore, complementary analyses of the chemical composition and the histological presentation of the skeletons could provide support for this diagnosis.

1. Introduction

Fluorosis is a pathological condition that occurs after chronic ingestion of high levels of fluoride. Although fluoride is generally considered an essential trace element and is beneficial to bone and dental health, there is a growing debate on the role of fluoride for human health, as adverse effects may occur with chronic consumption of high concentrations of fluoride (Littleton, 1999; Petrone et al., 2011; Barbier et al., 2010; Cerklewski, 1997). Excess fluoride in the human body is known to damage the musculoskeletal, endocrine, gastrointestinal, renal, reproductive, and neurological systems, and to interfere with various cellular processes (Barbier et al., 2010; Petrone et al., 2013). Symptoms of headache, nausea, constipation, and body aches occur within the first two weeks of continuous ingestion of toxic levels (Waldbott, 1998). Many symptoms subside or reduce in severity over time; however, continued ingestion of excess fluoride will cause increased rigidity of vascular structures, backache, muscular stiffness, blindness, and neurological complications, such as spasticity or lack of response (Jha et al., 2011; Reddy, 2009; Waldbott, 1998).

Prolonged consumption of elevated levels of fluoride results in bone and dental changes, such as excessive ossification of soft and skeletal tissue, producing enthesopathies and an increased density of bone and periostosis, in addition to dental mottling and pitting (Jolly et al., 1969; Kilborn et al., 1950; Møller and Gudjonsson, 1932; Rich and Ensinck, 1961). Compositional change of the mineral matrix of bone occurs as fluoride replaces the hydroxyl component of bone. The fluoroapatite mineral compromises the structure of bone by decreasing compliance and subsequently increasing the incidence of fractures (Littleton, 1999; Petrone et al., 2013). Ultimately, with continued fluoride consumption, immobilization will result from joint fusion (Teotia et al., 1971). It is unclear how long ingestion of high levels must occur before signs of skeletal fluorosis appear. Skeletal expression of the disease has been

* Corresponding author.
E-mail address: nelson@shh.mpg.de (E.A. Nelson).

https://doi.org/10.1016/j.ijpp.2019.05.003
Received 1 April 2019; Received in revised form 28 April 2019; Accepted 13 May 2019
Available online 20 June 2019
1879-9817/

reported in exposure levels as low as 0.7 ppm (Jolly et al., 1973). Though symptoms may occur in acute exposure, skeletal fluorosis is chronic, involving skeletal changes with intake in as few as 1-4 years (Siddiqui, 1955).

In spite of the known skeletal changes associated with fluorosis, particularly in regions of naturally high fluoride (Teotia and Teotia, 1988; Singh et al., 2018; Wiener et al., 2018), the disease remains understudied in archaeological populations. Few paleopathological reports exist in regions with water and soil fluoride levels reaching those known to cause skeletal changes (Lukacs, 1984; Littleton, 1999; Petrone et al., 2013; Yoshimura et al., 2006). For example, despite the existence of natural environments in many regions of North America conducive to fluorosis, this condition has not yet been reported in North American archaeological contexts. While this may suggest an absence of fluorosis on this continent, we argue that the etiology and history of this metabolic bone disease is so poorly understood that it may often be misdiagnosed or overlooked archaeologically. In this paper, we investigate the presence of skeletal signs suggestive of fluorosis within a community from a region rich in natural fluorite deposits.

2. Materials and Methods

2.1. The Ray Site: archaeological context

The Ray Site (11Br-104) is located at the confluence of the Illinois and LaMoine Rivers in Brown County, IL, (US) (Fig. 1), and includes a series of pit burials grouped in four clusters that extend across a ridge that rises above the floodplain. Threatened by agricultural activity, erosion, and looting, the site was excavated by avocational archaeologists, led by Mary and Glen Hanning, between 1975 and 1980, with assistance from the Center for American Archeology (Flotow 1983, 2006). Excavation of the site yielded a minimum of 117 individuals aligned northwest to southeast along the ridge crest. Most had been interred in a linear pattern, with four apparent clusters of three or more graves (Fig. 2). There was also an isolated concentration of nine burials in the southeastern area of the site (Flotow 1983, 2006). The site dates to the Middle Woodland period (50 BC-AD 400, Bullion and King, 2015), with one burial group (Group 3) dating to the earliest portion of the subsequent Late Woodland Period (AD 400-1000), based upon radiometric dating and diagnostic artifacts, including pipes. A 95% probability range for median calibrated radiocarbon dates is 55 cal BC - cal AD 495, which suggests a long Middle Woodland occupation extending into the more recent Late Woodland period.

The elaborate Middle Woodland ramp and tomb structures, common to lower Illinois valley (LIV) sites, are not present at the Ray Site, but there is evidence of extended burial rituals that contributed to archaeological recovery of incomplete skeletons and isolated elements. Grave accompaniments typical of LIV Middle Woodland sites are reported, including Hopewell and Havana pottery, mica, lamellar blades, hematite, bone awls or pins, and copper (Flotow, 1983, 2006).

The presence of diagnostic nonlocal "Hopewell" items made of copper and mica, which were introduced into the region either by trade or resource quests, along with shared Middle Woodland ceramic and lithic artifact styles and lack of genetic isolation (Bullion and King, 2015), suggests that the community who created the Ray Site cemetery shared heritage and cultural attributes with their more heavily studied neighbors in the LIV and the Central Illinois River Valley (CIV). Most commonly located in the floodplain, on low terraces, or talus slopes,

Fig. 1. Location of the Ray Site. The Ray Site is located in Illinois, USA, illustrated by the inset map of the continental United States of America, wherein the red star indicates the location of the site. The larger map marks the site's location with a red oval on the ridge crest at the confluence of the La Moine River to the north and the Illinois River to the east.

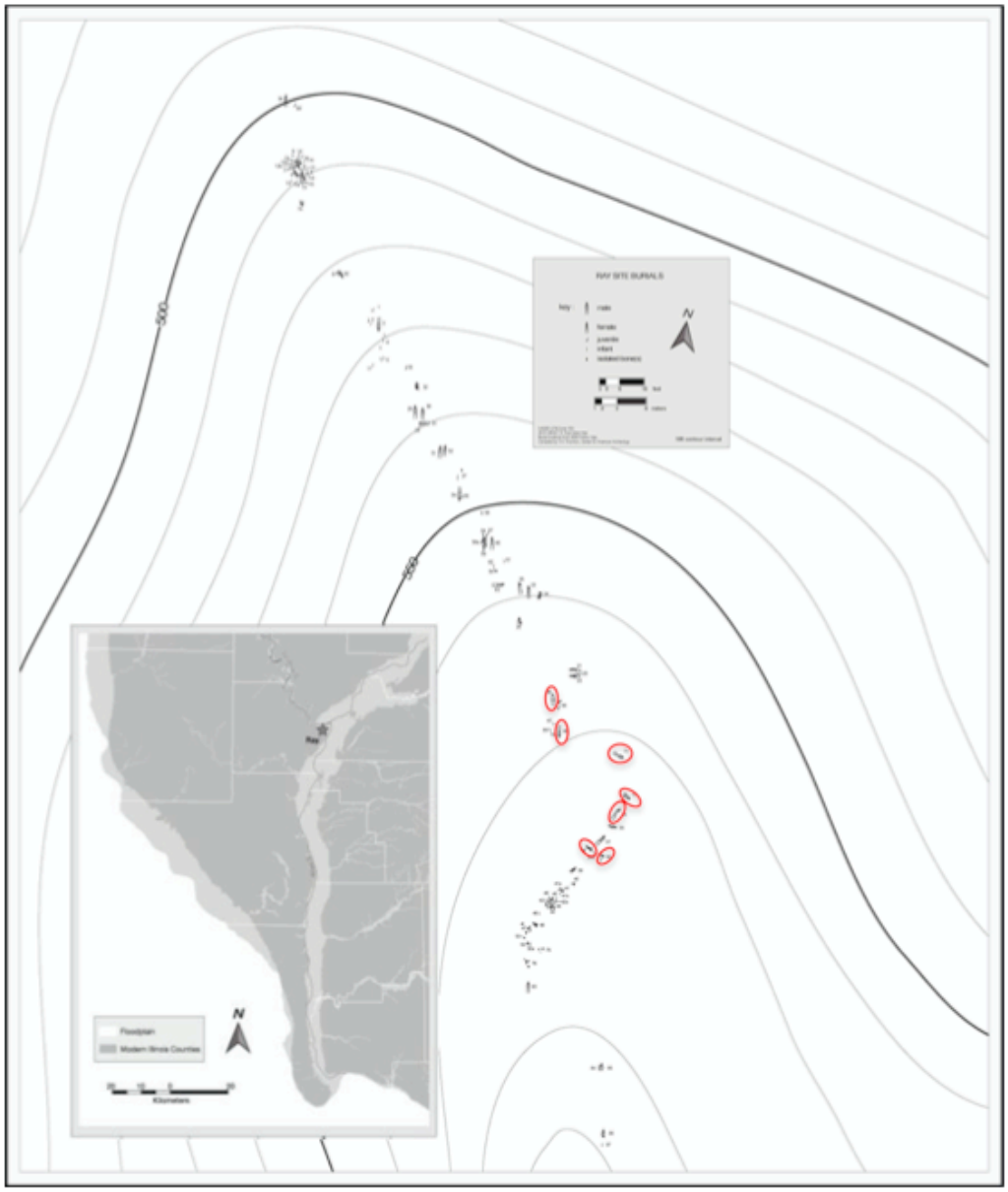


Fig. 2. Map of the Ray Site burials. The inset map shows the Ray Site's location within west Illinois (grey star). The red ovals indicate seven of the eight individuals identified as displaying shared skeletal changes. The eighth individual, Ray 2-77, was recovered during a later excavation from a group burial to the southeast corner of the site and is not shown in the map.

Middle Woodland villages are characterized archaeologically as settled communities of 30-50 people (Asch, 1976; Buikstra, 1976) whose diet included animals harvested from the nearby river, seasonal flyways, and oak-hickory forests, along with a range of cultivated plants known as the Eastern North American agricultural complex (Smith, 1992, 2006; Smith and Yarnell, 2009). Maize, however, would not enter the local diet for several centuries. While bluff crest cemeteries served local Middle Woodland villages, multi-community gathering places are recognized at floodplain sites (Buikstra and Charles, 1999; Buikstra et al., 1998).

2.2. Environmental Context

Illinois is an archaeologically rich region, with a varied landscape. Portions of this landscape are naturally abundant in fluoride, which places both ancient and contemporary people at risk for excessive fluoride ingestion leading to fluorosis. The most recent publically available fluvial analysis of the region near the Ray Site is reported to have fluoride levels of 2.2 – 3.0 mg/L (ppm) (US Dept of Health, Education and Welfare, 1969). The water tested for this analysis was not subjected to modern fluoridation or filtration. Surface analysis of Illinois soils in the area yielded fluoride content ranging from 275 mg/kg to 303 mg/kg (Omueti and Jones, 1977a, b). While these modern values of surface soil fluoride content may be influenced by anthropogenic factors, such as farming practices, it is not unreasonable to assert that they reflect values during the Woodland period; this is a rural area of Illinois with little to no industrial and municipal development to influence levels of fluoride content.

High levels of environmental fluoride do not necessarily result in fluorosis, as a variety of factors influence the presence and

bioavailability of fluoride. Environmental factors, such as weathering, temperature, and the natural geological composition of rocks and soil, affect the presence of available fluoride (Barbier et al., 2010; Petrone et al., 2011). Fluoride-binding minerals in the groundwater facilitate movement of the ionic form from the surrounding rocks. Consequently, fluoride is found in soil, water, plants, and marine resources (Littleton, 1999; Omueti and Jones, 1977a; Petrone et al., 2011). Fluoride inhalation from anthropogenic sources such as industrial atmospheric particulate is possible; this pathway has been linked to respiratory disease in coal burning nations, notably China (Barbier et al., 2010; Watanabe et al., 2000). Furthermore, water storage methods involving clay and soil may increase the amount of fluoride in water molecules through evaporation or leaching from clay vessels used for storage (Littleton, 1999).

Due to the high concentration of fluoride in the ground water and soil, it is possible that archaeological populations who resided in this region experienced fluoride toxicity. We, therefore, evaluated 117 remains from the Ray Site, located in a region known to harbor high levels of fluoride in the soil and groundwater.

2.3. Data Collection

We estimated age-at-death and biological sex using standard data collection protocols (Buikstra and Ubelaker, 1994). For adults, these methods emphasize morphological observation techniques of age-associated skeletal changes and sexually dimorphic features of the os coxa and skull (Brooks and Suchey, 1990; Meindl et al., 1990; Phenice, 1969; Walker, 2005). Dental development and eruption were used to estimate age-at-death for non-adults (Ubelaker, 1989), supplemented by evaluating the development and fusion of epiphyses (Scheuer and Black, 2004).

Abnormal bone changes and dental defects were observed macroscopically using a 10x hand-held loupe. We recorded abnormal bone formation and bone destruction according to comprehensive protocols outlined in Buikstra and Ubelaker (1994) and evaluated lesions using paleopathological descriptions and clinical literature (Buikstra and Ubelaker, 1994; Brickley and Ives, 2010; Ortner, 2003; Resnick and Niwayama, 1988). We recorded and described lesion appearance, location, and activity at time of death (active or healing), and distribution. Likewise, we documented the location and appearance of dental defects including dental pitting and mottling, which were recorded using the Dean's Index (Dean, 1936).

Referencing previously published work on skeletal and dental fluorosis (Littleton, 1999; Petrone et al., 2013; Brickley and Ives, 2010; Dean, 1936), we compiled a list of pathological changes known to occur in skeletal fluorosis (Table 1). Osseous sclerotic activity affecting the cranium includes ossification of ligamentous and tendinous attachments on the cranium, often leading to a narrowing of the foramen magnum, as well as the densification and eventual obliteration of the diploë (Littleton, 1999; Resnick and Niwayama, 1983). Rib and vertebral enlargement are the result of calcification of the associated soft tissue, giving the surface of the ribs a roughened appearance and vertebral bodies an expansive, enlarged appearance with occasional fusion of the vertebrae through excessive osteophyte development (Petrone et al., 2013; Krishnamachari, 1986; Resnick and Niwayama, 1983). Fractures are a common secondary symptom of skeletal fluorosis due to the change of the mineral composition of bone from hydroxyapatite to fluoroapatite, reducing the compliance of the tissue (Fabiani et al., 1999). Only individuals presenting five or more skeletal changes were considered possible cases of skeletal fluorosis.

3. Results

3.1. Age and Sex

The Ray Site assemblage includes a minimum of 117 individuals: 30 adult males (> 20 years of age), 29 adult females (> 20 years of age), 37 juveniles (fetal to 19 years of age), and 21 individuals for whom age and sex could not be estimated (Fig. 3). The adult skeletal remains of undetermined sex and age partially results from lengthy Woodland funerary rites that reduced body completeness through corpse manipulation. Plowing during the historic period (> 1830) also exacerbated deterioration (Flotow, 1983, 2006).

3.2. Pathology

In our survey of the Ray Site skeletal remains, we discovered eight individuals (Table 2) with similar pathological changes that appear

Table 1
Pathological changes consistent with skeletal fluorosis.

Skeletal change	Description	References
Osteosclerotic activity of the cranium	Ossification of ligamentous and tendinous attachment sites; cranial bones become thick and heavy with eventual obliteration of the diploë.	Littleton, 1999; Resnick and Niwayama, 1983; Singh et al., 1962
Enthesophyte formation	Ossification of ligamentous and tendinous attachments of affected bones.	Littleton, 1999; Petrone et al., 2013; Resnick and Niwayama; Petrone et al., 2013; Rogers and Waldron, 1995
Periosteal bone formation	Periosteal bone deposition producing both woven and sclerotic expansion of bone.	Resnick and Niwayama, 1983; Petrone et al., 2013; Littleton, 1999; Roholm, 1937
Rib enlargement	Dense, sclerotic bone formation on the ventral and visceral surfaces leading to irregular appearance; Thickening and sclerosis of trabecular bone; Irregular margins of ribs due to ossification of associated ligaments and attachments.	Petrone et al., 2013; Tamer et al., 2007; Littleton, 1999; Resnick and Niwayama, 1983
Sclerotic activity: Vertebrae/Os coxae	Dense, sclerotic bone formation leading to irregular surface appearance; Thickening and sclerosis of trabecular bone; ossification of associated ligaments and attachments.	Krishnamachari, 1986; Littleton, 1999; Petrone et al., 2013
Enamel defects	Dental pitting, opacity, and mottling.	Jolly et al., 1969; Dean, 1936; Everett, 2011; Fejerskov et al., 1990, 1996
Osteoarthritis	Secondary response of osteoarthritic changes in joints of affected bones.	Littleton 1999, Ozsvath, 2009; Petrone et al., 2013; Horowitz et al., 1984
Fractures	Increase incidence of fractures due to the compromised structure resulting from a change in the mineral composition of bones.	Fabiani et al., 1999; Petrone et al., 2011; Boivin et al., 1989; Agarwal, 1973; Haettich et al., 1991

Adapted from Brickley and Ives (2010: 246-247).

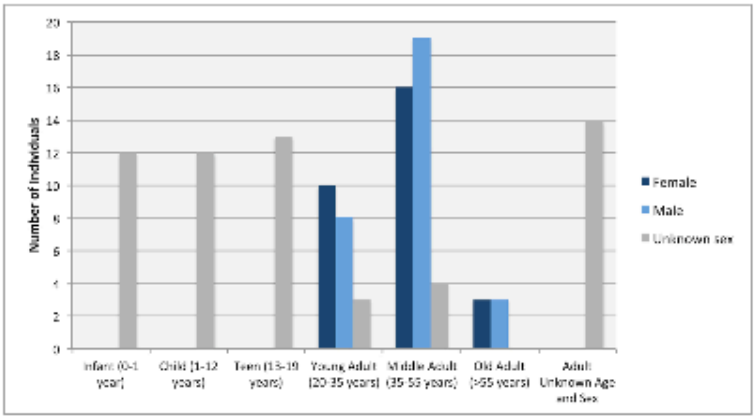


Fig. 3. Demography of the Ray Site (30 adult males, 29 adult females, 37 juvenile, and 21 indeterminable individuals).

Table 2
Demographic profile of individuals displaying similar skeletal pathology.

Individual	*Biological Sex*	Age at Death
Ray 2–69	Male	40–49 years
Ray 2–77	Male	40–50+ years
Ray 3–18	Female	40–50+ years
Ray 3–19	Male	40–49 years
Ray 3–36	Female	50+ years
Ray 3–38	Female	35–45 years
Ray 3–71	Male	35–50+ years
Ray 4–66	Male	25–35 years

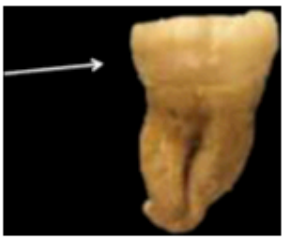


Fig. 4. Ray 3-19, right mandibular third molar displaying brown staining and mottling of enamel (indicated by arrow).

unusual for archaeological remains from west-central Illinois. These individuals presented changes such as excessive bone formation, enthesophyte formation, periosteal deposition, fractures, increased bone density, and dental mottling and pitting. These abnormal skeletal changes are present across the eight individuals (Table 3); however, the degree of development of abnormal changes and the skeletal elements involved varies between individuals.

Dental defects, such as mottling and pitting, were observed in five of the eight individuals; however, dental assessments were limited due to ante-mortem (Ray 2-77) and post-mortem tooth loss and severe tooth wear (Ray 3-18 and Ray 3-71), as well as soil deposits adhering to the tooth and obscuring the assessment (Ray 4-66)(SI Figs. 1–6). None of the individuals presented a full dental arcade and, at most, presented half of the adult dentition for assessment. Five individuals displayed dental defects, including opacities, brown staining (Fig. 4) and small enamel pits. These changes are present on the labial surface of the incisors, as well as on buccal and occlusal surfaces of the molars of the individuals (Ray 2-69, Ray 3-19, Ray 3-38, Ray 4-66, Ray 3-36).

In general, the axial skeleton displays excessive abnormal bone formation more frequently than the appendicular skeleton (Figs. 5 and 6,). As an example, Fig. 6 illustrates the enthesophyte development present on the os coxae, femora, ribs, tibiae, vertebrae, and patellae from the afflicted individuals. A thoracic vertebra from Ray 2-69 displays ossification of the anterior longitudinal ligament with some osteophytic development surrounding the zygapophyses (Fig. 6a). The left patella of the same individual shows ossification of the quadriceps tendon on the most superior and inferior margins of the bone (Fig. 6b). Likewise, the ribs of Ray 2-77 display excessive dense bone development on the ventral and visceral surfaces with accompanied ossification of the superior and inferior intercostal muscles and obliterated trabeculae (Fig. 6c). The right os coxa of Ray 3-39 shows enthesophyte formation along the iliac crest and ischiopubic ramus, along with bony spicules on the ischial tuberosity and the superior aspect of acetabulum (Fig. 6d). Ray 3-19 also displays ossification of the superior and inferior margins of the ribs, with sclerotic bone deposition on the ventral and visceral surfaces of the ribs (Fig. 6e).

Table 3
Pathological features present on the eight individuals from the Ray Site. (M = male, F = female).

Individual	Sclerotic activity: Cranium	Enthesophyte formation: Long bones	Periosteal bone formation	Rib enlargement	Sclerotic activity: Vertebra/Os coxa	Vertebral enlargement	Enamel defects	Osteoarthritis	Fractures
Ray 2–69 (M) (40–49 yrs)	♦	♦	♦	♦	♦		♦	♦	♦
Ray 2–77 (M) (35–50 yrs)	♦	♦	♦	♦	♦	♦		♦	♦
Ray 3–18 (F) (40–50+)		♦	♦		♦	♦		♦	
Ray 3–19 (M) (40–49 yrs)	♦	♦	♦	♦	♦	♦	♦	♦	♦
Ray 3–36 (F) (50+)	♦	♦	♦		♦	♦	♦	♦	
Ray 3–38 (F) (35–45 yrs)	♦	♦	♦		♦	♦	♦		
Ray 3–71 (M) (35–50+ yrs)	♦	♦	♦	♦	♦	♦		♦	♦
Ray 4-66 (M) (20–35 yrs)		♦	♦	♦	♦	♦	♦	♦	♦

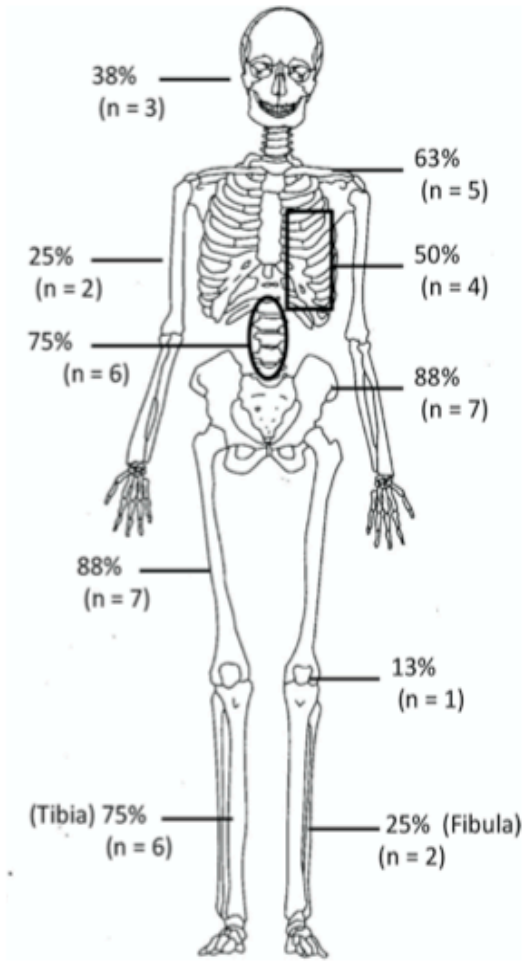


Fig. 5. Frequency and distribution of pathological skeletal involvement for the eight individuals with altered bone. Adapted from Buikstra and Ubelaker, 1994.

Fig. 6. Bony changes on elements of individuals Ray 3-18 (a), 2-69 (b), 3-36 (d), and 3-19 (c,e).

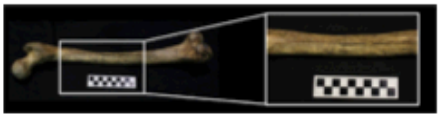


Fig. 7. Ray 3-71, enthesophytes present on linea aspera of right femur.

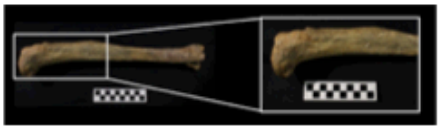


Fig. 8. Ray 2-69, right tibia displays dense sclerotic deposition with ossification of the patellar tendon at the site insertion.

Enthesophytes are most frequently observed on the linea aspera of the femur, the superior and inferior margins of the ribs (Fig. 6e), the tibial tuberosity (Fig. 7), the vertebrae (Fig. 6a), and the linea aspera of the femora (Fig. 8). Individuals Ray 2-77, Ray 2-69, Ray 3-19, Ray 3-71, and Ray 4-66 display abnormal sclerotic ossification of the rib surfaces (visceral and external). In addition, Ray 2-77 and Ray 3-19 also display the ossification of the intercostal muscles which enlarged the ribs superior-inferiorly (Fig. 6e). The iliac crest and ischiopubic rami of the os coxae frequently display enthesophytes (Fig. 6d). Bony changes affecting the vertebral column include osteophyte and enthesophyte development on vertebral bodies and transverse processes. In seven of the individuals (Table 3), ossification of spinal ligaments on some thoracic and lumbar vertebrae produced bony spicules at the superior margin of the laminae, the superior margin of the body, and the most posterior aspect of the spinous process. These enthesophytes are clearly the ossification of soft tissues, forming bony spicules and producing abnormal margins. These changes were not uniform throughout the vertebral column, as not every vertebra displayed abnormal changes.

Non-specific pathological skeletal changes are also present, including periosteal deposition and fractures. Periosteal deposition is most commonly observed on the bones of the leg, with bilateral involvement that includes both healed and actively remodeling lesions. The femora, clavicles, ribs, and malar are in some cases affected, though are far less commonly involved. Males present a more extreme expression that females, with an increased presence of enthesophytes and osteophytes. Two makes display bilateral expansive, dense, irregular bony surfaces at the distal end of the femoral diaphysis. The enlarged ribs previously described were observed only in males (Table 3). As illustrated previously in Fig. 2, seven of the eight individuals were buried in a cluster on the southeast section of the site. Associated artifacts and burial treatments did not distinguish these from the other interments (Flotow 1983, 2006).

Five individuals present a total of eight fractures, with two skeletons displaying multiple fractures (Ray 3-71 and Ray 3-19). Ray 3-71 presents a collapsed thoracic vertebra resulting from a compression fracture, which fused to the superior vertebra, and a well-healed fractured rib. Ray 3-19 suffered a fractured rib and right fibula (distal portion of diaphysis), as well as a fractured right humerus at the distal 1/3 of the diaphysis, which remained misaligned during healing. The right radius of individual Ray 2-69 displays a fracture at the distal portion of the

diaphysis, while the right ulna displays no sign of trauma. Ray 4-66 presents a broken but well healed rib. There are no indications of disuse atrophy in any of the affected limbs. Of 117, only seven individuals display fractures and 5 of those individuals are from our group of the eight individuals displaying shared pathological attributes, which we discuss in our differential diagnosis. The five individuals displaying fractures are all estimated to be male.

4. Discussion

4.1. Differential Diagnosis

Although the changes observed in these eight individuals are consistent with fluorosis, the nonspecific nature of the skeletal changes emphasizes the importance of constructing a differential diagnosis that includes other possible conditions. Most paleopathological descriptions of disease (i.e., Aufderheide and Rodríguez-Martín, 1998; Ortner, 2003) derive from examples of advanced stages of disease and therefore report extreme bony changes. In the individuals reported here, abnormal bony changes are less severe. This leads us to develop a broad differential diagnosis since many of the diagnoses can produce similar lesions in early stages of development. The shared pathological changes that characterize these remains are associated with nine present-day conditions. These include ankylosing spondylitis, diffuse idiopathic skeletal hyperostosis (DISH), hematogenous osteomyelitis, hypoparathyroidism, myositis ossificans (*fibrodysplasia ossificans progressiva*), osteopetrosis, treponemal infection, Paget's disease of bone, and skeletal fluorosis (Tables 3 and 4). Here, we briefly discuss each of these conditions.

4.1.1. Ankylosing spondylitis

Ankylosing spondylitis (AS) is a heritable, inflammatory arthritis that is genetically influenced by an HLA-B27 variation and progresses in severity throughout life. AS involves excessive ossification and begins with osteophytic development that ultimately results in joint fusion, particularly in the lower limbs, and most frequently, the pelvic girdle and vertebral column (Braun and Sieper, 2007; Evans et al., 2011; Hamersma et al., 2001; Lories and Luyten, 2013). This condition first presents in the lower region of the body, particularly the sacroiliac joint, and progresses cranially (Ortner, 2003: 571; Olivieri et al., 2009). It is known to affect males more often than females (Meirelles et al., 1999; Ortner, 2003; Van der Linden et al., 1984). Affected individuals usually begin to show signs during the third decade of life; manifestation of the condition within the second decade is not uncommon (Jadon et al., 2013). Consistent with our observations, excessive ossification occurs in middle-aged males more frequently than females of the same age cohort. AS often presents as osteophytic development on the joint margins of the lower limb; however, these changes are localized to joint surfaces. Furthermore, diagnostic characteristics of AS are not observed in the Ray Site individuals, including more frequent involvement of the axial skeleton, the remodeled square shape of vertebral bodies, and changes in joint surfaces involving of expansion and continuity of trabeculae (Ortner, 2003). Therefore, the presentation of bony changes including muscle and ligamentous attachment sites in the Ray Site sample, as well as distribution of changes, namely the frequent involvement of the tibiae, suggest AS is an unlikely diagnosis (Olivieri et al., 2009; Ortner, 2003).

4.1.2. Diffuse idiopathic skeletal hyperostosis (DISH)

Diffuse idiopathic skeletal hyperostosis (DISH) is characterized by the ossification of ligaments and tendons associated with joints (Mader et al., 2013; Ortner, 2003: 558). DISH is most commonly known to affect the thoracic spine, with some lumbar and cervical involvement (Ortner, 2003: 558). A few researchers have noted extraspinal involvement; however, the clinical consensus is that this is not characteristic of DISH (Mader et al., 2009; Mader et al., 2013). Although skeletal signs may be severe, the condition may also be asymptomatic (Mader et al., 2013; Olivieri et al., 2009). Though the criteria for diagnosing DISH are continuously evolving, the condition itself has a long history and is prevalent in most developed countries (Mader et al., 2013). DISH occurs in men more frequently than in women and is most common in older individuals (Mader et al., 2013; Olivieri et al., 2009; Ortner, 2003). The pathological development of bone and resultant compromised structure increases the risk and occurrence of fractures (Belanger and Rowe, 2001; Ortner, 2003). The increased risk of fractures and predilection for older male individuals is consistent with observations from the Ray Site. However, fractures associated with DISH occur most frequently on the spine (Belanger and Rowe, 2001); only one of the individuals included in our differential diagnosis displays a spinal fracture. Furthermore, the individuals from the Ray Site present pathological osseous development such as enthesophyte development along with periosteal deposition, primarily on extraspinal elements, which leads us to conclude that DISH is not an appropriate diagnosis.

4.1.3. Hematogenous osteomyelitis

Hematogenous osteomyelitis is a globally occurring infectious condition caused by several species of bacteria (Aufderheide and Rodríguez-Martín, 1998; Conrad, 2010; Flensborg et al., 2013; Lew and Waldvogel, 2004; Ortner, 2003: 181-182). This infection is associated with abnormal bone formation and resorption, often resulting in the formation of involucra (Flensborg et al., 2013; Ortner, 2003: 181-1821Ortner, 2003Ortner, 2003Ortner, 2003: 181-1821Ortner, 2003Ortner, 2003Ortner, 2003: 181-1821Ortner, 2003Ortner, 2003: 181-1821). In juveniles, the appendicular skeleton is the most frequent site of infection, specifically the femoral metaphyses (Conrad, 2010). Although this condition is rare in adults, skeletal involvement most frequently occurs in the vertebrae, but it also includes sternoclavicular joints, and sacroiliac joints, along with metaphyseal and diaphyseal regions of long bones (Aufderheide and Rodríguez-Martín, 1998; Flensborg et al., 2013; Lew and Waldvogel, 2004). Vertebral involvement is most significant in the lumbar vertebrae, less so in the thoracic and cervical segments (Kumar, 2013). The absence of involucra and the lack of vertebral involvement in the Ray Site individuals are inconsistent with hematogenous osteomyelitis in adults, and thus, it is an unlikely diagnosis.

4.1.4. Hypoparathyroidism

Hypoparathyroidism is a rare disorder characterized by deficiency of parathyroid hormone (PTH). The deficiency disrupts the balance of bone resorption and bone formation, ultimately resulting in decreased osteoclastic activity and subsequent increased calcium concentration leading to increased bone density (Bilezikian et al., 2011; Resnick and Niwayama, 1988). Individuals with hypoparathyroidism are reported to have ligamentous calcifications, particularly in the spine (Chaykin et al., 1969). Fractures are seldom discussed, although some researchers report fractures of hand and foot bones in 24% of people diagnosed with hypoparathyroidism (Rubin et al., 2008). Dental defects include weak enamel and thin misshapen tooth roots (Chaykin et al., 1969). This condition occurs in adults and adolescents alike; there appears to be no sex-based predilection (Bilezikian et al., 2011). Although the dental changes associated with hypoparathyroidism make it an unlikely diagnosis for our target sample, the limited clinical information does not allow us to rule it out. Thus, it remains a possibility.

4.1.5. Myositis ossificans progressiva

Myositis ossificans progressiva, also known as *fibrodysplasia ossificans progressiva*, is characterized by extensive ossification of tendons, ligaments, and muscles (Culbert et al., 2013). The early stage is characterized by enthesophyte formation. In advanced stages, the condition progresses to soft tissue involvement, including muscles, joint capsules, and skin. Those with the genetically influenced condition appear normal at birth, apart from congenital malformation of the hallux,

Table 4
Differential diagnosis with skeletal description and epidemiology.

Pathology	Skeletal description	Epidemiology
Ankylosing spondylitis (Chaykin et al., 1969, Lin et al., 2012; Lories and Luyten, 2013; Meirelles et al., 1999,; Olivieri et al., 2009)	Primarily affects sacroiliac joint and vertebrae. Peripheral enthesophyte and osteophyte formation observed and in advanced stages ossification/ankylosis of elements occurs. Ankylosis proceeds from sacroiliac joint superiorly to lumbar, thoracic, and the cervical region in severe cases. Though extra-vertebral involvement is not common, enthesophyte development has been noted at ankles, hips, knees, shoulders, and sternoclavicular joints. Some associated osteopenia is common.	A highly heritable disease affecting .55% of European populations and .23% of Chinese but is rare in African and Japanese populations. Two to three times more prevalent in males than females. Symptoms usually begin in second or third decade of life. Pathological changes increase with age.
DISH (Diffuse Idiopathic Skeletal Hyperostosis) (Belanger et al., 2001, Mader et al., 2009, Olivieri et al., 2009).	Spine is primary site of involvement with thoracic vertebrae most commonly affected. Bony changes include large syndesmophytes forming a "flowing" appearance. In thoracic region, changes are most often observed on the right side of the anterior aspect of thoracic vertebral bodies. Lumbar spine shows similar involvement displayed equally on both sides of anterior aspect. Cervical involvement is rare, as is posterior vertebral involvement. Increased incidence of fractures.	Unknown etiology. Said to occur in 28% of individuals 45-85 years old. Occurs in males more than females. More common in individuals of European descent; also found in Asian (Korean) individuals.
Fluorosis (Barbier et al., 2010; Brickley and Ives, 2010; Jolly et al., 1969; Littleton, 1999; Ortner, 2003; Petrone et al., 2011; Petrone et al., 2013; Reddy, 2009; Teotia et al., 1971; Teotia et al., 1988).	Excessive osteosclerotic deposition. Periosteal deposition, enthesophyte formation, and ultimately fusion of joints. Sclerotic activity most significant on lower limbs, os coxae, vertebrae, ribs, and crania. Dental defects include pitting and mottling. Increased incidence of fractures.	Affects men more than women and commonly older adults. Cases reported globally: Middle East, Africa, Europe, North America, and Asia. Disease conditional to ingestion of fluoride, diet, and genetic factors
Hematogenous Osteomyelitis (Aufderheide and Rodríguez-Martín, 1998; Conrad, 2010; Flensborg et al., 2013; Kumar, 2013; Lew and Waldvogel, 2004; Ortner, 2003)	May involve multiple parts of bone including medullary cavity, periosteum, and cortical bone. Long bones most commonly affected. Vertebral involvement in adults. Abnormal periosteal bone formation. Pathology involves abscesses, involucra, and sequestra. Abnormal bone formation near joints and surrounding joint capsules.	One in 5,000 children affected; one in 450,000 adults affected. Occurs globally and is the most common type of osteomyelitis. Skeletal involvement is more marked in children than adults, but it can occur in both. Males are more commonly affected than females.
Hypoparathyroidism (Bilezikian et al., 2011; Chaykin et al., 1969; Resnick and Niwayama, 1988)	Osseous build up does not display a distinct diagnostic pattern. Excessive endosteal and periosteal deposition. Lesions include calcification of ligaments, particularly those of the spine. Cortical and trabecular increases in density. Dental defects include weak enamel and misshapen roots.	One person per million is affected. Most prevalent in Finns, Sardinians, and Iranian Jews. Occurs in adults and children alike. No predilection for biological sex.
Myositis ossificans progressive (Auderheide and Rodríguez-Martín, 1998; DiMaio and Francis, 2001; Pignolo et al., 2011; Culbert et al., 2013)	Characterized by excessive ossification of tendons, ligaments, and muscles. Early stages display bony spicule and enthesophytes, while advanced stages present significant soft tissue involvement (muscles, joint capsules, and skin). May lead to fusion of joints. Individuals display congenital malformation of the first toes appearing abnormally small. Progresses posterior to anterior and axial to appendicular.	One in two million are affected. Considered a childhood syndrome, may continue into adulthood often in response to trauma or stress. No geographic or ethnic predispositions observed though the condition is genetically hereditary. Fatal near the second decade of life.
Osteopetrosis (Albers-Schonberg disease) (Agarwal, 2013; Bollerslev et al., 2013; Henriksen et al., 2010; Reddy, 2011; Sobacchi et al., 2013)	Osteosclerotic activity affecting skull, spine, pelvis, and appendicular bones. Defects are most significant on metaphyses of long bones causing an "Erlenmeyer flask" appearance. Increased incidence of fractures along with increased density of bones. Increased bone mass leads to altered craniofacial morphology involving increased frontal bossing, macrocephaly, loss of mandibular angle, and prognathism. Dentally this disease results in tooth eruption defects and dental caries. Longitudinal growth may be impaired resulting in short stature. The skeletal defects often lead to osteomyelitis and osteoarthritis.	Several variants of this condition exist, all displaying similar conditions but originating from failure of osteoclastic development or mutations on at least one of 10 genes. Genetically recognized as two types: autosomal recessive (AR) and autosomal dominant (AD). Incidence varies between types with AR occurring in 1 of every 250,000 and AD 5 of every 100,000. Onset usually takes place in childhood though it may be perinatal or late adolescence. AR is usually fatal in infancy but AD has a normal life expectancy. Affects both men and women equally.
Paget's disease of bone (Aufderheide and Rodríguez-Martín, 1998; Brickley and Ives, 2010; Mays, 2008; Ortner, 2003; Ralston, 2013; Ralston et al., 2008; Roberts ad Manchester, 2007; Tan and Ralston, 2014)	Compromised mineralization of bone leading to dense, sclerotic, irregular deposition of bone. Affects the cranium and post-crania. Skeletal elements most affected are the cranium, sacrum, spine and femora. Affected crania often exhibit expansive, dense diploe. Characteristic coarsening of trabeculae and mosaic pattern assist in histological identification.	Known to occur primarily in Europe, North America, Australia, and New Zealand with few accounts in Africa and Asia. More common in males than females. Prevalence increases with age.
Treponematosis (Cook, 2005; Cook and Powell, 2005; Hackett, 1975; Kent and Mosher et al., 2013; Ortner, 2003)	Resorptive, lytic, "gummatous" lesions followed by bone regeneration. Periosteal deposition may be localized and may present some bony exostoses at sites of major overlying muscle. Cortical thickening and expansion of the endosteal surface may occur. Affects the cranial vault and appendicular skeleton. Characteristic lesions include *caries sicca* of the cranial vault and saber tibia.	Globally occurring. 4 diseases caused by two species of *Treponema:* Pinta located in Central America, yaws is found in tropical environments near the equator, bejel is found in northern and southern hemispheres, and venereal syphilis is currently global.

which is described as short and small (Aufderheide and Rodríguez-Martín, 1998; Pignolo et al., 2011). The disease progresses in a specific anatomical pattern, first affecting the posterior region of the axial skeleton with later involvement of the anterior region and appendages (Pignolo et al., 2011). Identification of myositis ossificans progressiva involves soft tissue ossification and hallux malformity of both the metatarsal and proximal phalanx (Aufderheide and Rodríguez-Martín, 1998: 26, 367; Pignolo et al., 2011). The Ray sample does not display the characteristic symmetrical congenital malformations of the toes, nor is the characteristic pattern of progressive ossification of soft tissue from the posterior region of axial skeleton to anterior appendages present. Consequently, we do not consider myositis ossificans an appropriate diagnosis.

4.1.6. Osteopetrosis

Osteopetrosis, also known as Albers-Schönberg disease, includes a group of conditions, varying in symptom severity and genetic causative factors (Agarwal, 2013; Sobacchi et al., 2013). The two primary groups are divided into autosomal recessive (ARO) and autosomal dominant (ADO) (Sobacchi et al., 2013). ARO is considered to be "malignant" osteopetrosis and is characterized by increased bone density, increased incidence of fractures and impaired longitudinal bone growth (Stark and Savarirayan, 2009). Affected individuals typically present short stature, an abnormally small thorax, frontal bone bossing, and macrocephaly (Sobacchi et al., 2013; Stark and Savarirayan, 2009). ADO, classically considered Albers-Schönberg disease, results in osteosclerosis and increased fracture incidence; however, ADO is considered to be less severe in comparison to ARO. Characteristic disease presentation includes scoliosis, osteoarthritis of the hip, osteomyelitis of the jaw, and increased incidence of abscesses and dental caries (Bollerslev et al., 2013; Stark and Savarirayan, 2009). The Ray Site sample does not display macrocephaly, small thoraces, or shorter statures in comparison to the other members of the Ray Site population. Therefore, osteopetrosis is rejected as an appropriate diagnosis.

4.1.7. Non-venereal treponemal infection

Non-venereal treponemal infection has a substantial presence in the North American archaeological record and has been observed at sites located in Illinois (Cook, 2005; Cook and Powell, 2005; Mosher et al., 2013). Skeletal involvement in modern treponemal infections most commonly occurs in yaws, as compared to endemic syphilis or bejel; pinta does not affect skeletal tissue (Ortner, 2003: 290). Pathological skeletal changes associated with treponemal infection are typically present on the appendicular skeleton and cranium (Ortner, 2003: 118; 2008). Although there are slight differences in the changes observed across treponemal infections, typical skeletal presentation of the infection includes periostitis and osteomyelitis of the long bones (including the digits), cortical thickening and medullary cavity new bone formation, lytic cranial involvement (i.e., superficial cavitations, radial scars, stellate lesions, etc.), and saber-shin, bowing and pseudo-bowing of the tibia (Ortner, 2003; Hackett, 1975). The woven and well-integrated periosteal deposition displayed on the individuals from the Ray Site may be suggestive of treponemal infection, as this is consistent with "periosteal plaques" of new woven bone and sclerotic bone (Hackett, 1975). However, even within treponemal disease, periosteal reaction is considered non-specific (Ortner, 2003: 304-319; Powell and Cook, 2005). Other diagnostic features of treponemal infection such as *caries sicca* and gummatous lesions are absent in our sample of eight individuals (Cook and Powell, 2005; Hackett, 1975; Ortner, 2003). The pathological lesions observed in the eight individuals from the Ray Site include prominent osseous development on sites of ligamentous and tendinous attachments and sclerotic bone formation of the cranium. The periosteal deposition is dense and sclerotic, while lytic lesions are absent. There is no involvement of the digits, nor is there any appearance of bowing of the tibiae. In light of these pathological differences, we conclude that treponemal infection is an unlikely explanation for this suite of skeletal lesions.

4.1.8. Paget's disease of bone

Paget's disease of bone is of unknown etiology. Although some have argued a viral origin (Abe et al., 1995; Mills et al., 1984), this theory is controversial and there appears to be a combination of environmental and genetic factors at work (Hocking et al., 2002; Gennari et al., 2010; Roodman and Windle, 2005). Genetic factors clearly play a role in both rare and classic Paget's disease of bone (Aufderheide and Rodríguez-Martín, 1998; Mays, 2008; Ralston, 2013; Ralston et al., 2008). Paget's is considered a "European disease", occurring mostly in persons of European heritage, including those in North America, Australia and New Zealand, with few accounts in Africa and Asia (Brickley and Ives, 2010: 226; Ralston, 2013; Roberts and Manchester, 2007). The disease is reported to affect males more often than females and commonly occurs in individuals over 55 years old (Ortner, 2003: 435; Ralston, 2013). The presence of Paget's disease in the archaeological record is limited and significantly localized to Europe (Ortner, 2003: 38).

Paget's disease is characterized by abnormal bone formation resulting in deformed morphology throughout much of the skeleton. Dense, sclerotic, irregular bone deposition occurs on both the cranium and post-cranial skeleton (Ortner, 2003; Ralston 2013; Ralston et al., 2008). Compromised structure results in an increased incidence of vertebral fractures and "bowing" of the lower limbs, particularly the femora (Brickley and Ives, 2010; Roberts and Manchester, 2007). Skeletal elements most commonly affected are the pelvis, femur, lumbar spine, bones of the skull, and tibia (Ralston, 2013). There is progressive development from the sacrum to the cervical spine and cranium. Cranial changes begin with an osteoporotic reaction on the calvarium but then lead to a laminated thickening of the endocranial surface and eventual expansion of the diploë and present areas of sclerotic and porotic processes creating a pumice bone appearance (Ortner, 2003: 436-440; Ralston et al., 2008). The non-uniform expansion of the diploë results in a visibly deformed skull (Ortner, 2003: 437-439); this in conjunction with the abnormal posture produced by irregular bone formation results in a distinct pathological appearance of the individual.

The Ray sample displays sclerotic bone deposition throughout the skeleton. However, the diploë do not appear to have been actively expanding at time of death, as expected in Paget's disease. There are also no skeletal signs presented in any of the eight individuals suggesting progressive involvement extending cranially from the sacrum to the cervical spine. The facial deformity characteristic of Paget's disease of bone is also absent. This suggests that Paget's disease is an unlikely diagnosis.

4.1.9. Skeletal fluorosis

Skeletal fluorosis is a bone forming pathological condition caused by ingestion of high levels of fluoride (Largent, 1961; Littleton, 1999; Møller and Gudjonsson, 1932; Petrone et al., 2011; Roholm, 1937). Chronic ingestion of toxic levels of fluoride causes changes in the skeleton and dentition (Littleton, 1999). Characteristic signs of skeletal fluorosis include the presence of enthesophytes and osteophytes, along with dense periosteal deposition, enlargement of ribs, increased incidence of fractures, and osteosclerotic development on vertebral bodies (Ayoob and Gupta 2006; Brickley and Ives, 2010; Littleton, 1999; Petrone et al., 2013). Early signs of skeletal involvement include trabecular coarsening, along with ossification of muscle attachments and irregular bony outgrowths around joint surfaces (Izuora et al., 2011; Littleton, 1999). With continued ingestion of toxic levels of fluoride, individuals display excessive production of new bone, resulting in ankylosis of joints, and subsequent immobility (Izuora et al., 2011; Littleton, 1999; Ortner, 2003: 406; Petrone et al., 2011). Cranial changes include increased density of diploë, along with general osteosclerotic activity, primarily at the base of the cranium (Brickley and Ives 2010: 246; Littleton 1999).

Table 5
Differential diagnoses of pathological signs present on eight individuals from the Ray Site.

Pathological feature:	Ankylosing spondylitis	DISH	Hematogenous osteomyelitis	Hypopara-thyroid	Myositis ossificans	Osteopetrosis	Paget's disease of bone	Skeletal fluorosis	Non-venereal treponematosis
Osteosclerosis of cranium	-	♦	-	♦	♦	♦	♦	♦	♦
Enthesophyte formation	♦	♦	-	-	♦	-	-	♦	-
Periosteal deposition	♦		♦	♦	-	-	-	♦	♦
Increased incidence of fractures	-	♦	-	-	-	♦	♦	♦	-
Rib:									
Enlargement	-	-	-	♦	♦	♦	♦	♦	-
Expanded/irregular inferior margins	-	-	-	-	♦	♦	-	♦	-
Vertebra:	♦	♦	♦		♦	♦			
Enlargement				♦			♦	♦	-
Osteophyte/Enthesophyte formation				♦			♦	♦	♦
Dental enamel defects	-	-	-	♦	-	♦	-	♦	-
Associated osteoarthritis	♦	♦	-	-	-	♦	♦	♦	♦
Characteristics:									
Affecting primarily older individuals	-	♦	-	-	-	-	♦	♦	-
Affects men more than women	♦	♦	♦	-	-	-	♦	♦	-
Bilateral/Systemic	♦	♦		♦		♦	♦	♦	♦
Displayed most commonly in lower limbs	-	-	♦	-	-	♦	-	♦	-
Abnormally thick/heavy bones	-	-	♦	♦	♦	♦	♦	♦	♦

Key: (♦ = present), (- = absent), (¯ = unknown/insufficient clinical data) citations listed on Tables 1 and 4.

The use of dental defects as diagnostic criteria for fluorosis is problematic. Although fluoride has cariostatic properties, this only occurs within a certain range of fluoride ingestion (.5 ppm – 1.0 ppm), while toxic fluoride consumption leads to dental defects such as opacity, along with enamel staining and pitting (Dean, 1936; Petrone et al., 2013; Freni, 1994). Linear enamel hypoplasia (LEH) was not included in our differential diagnosis because biological and chemical research shows that fluorotic dental pitting develops via a pathway distinctive from LEH. It ultimately derives from post-eruptive damage resulting from poorly mineralized outer enamel (Aoba and Fejerskov, 2002; Fejerskov et al., 1977, 1990, 1996; Horowitz et al., 1984; Thylstrup and Fejerskov, 1979). Alvarez et al. (2009) also underscore that difficulties in diagnosing fluorosis arise when fluoride and non-fluoride-induced dental defects become confused. More work must be done to establish differences between dental changes due to fluorosis and those that may derive from a variety of systemic or local factors. It is now widely accepted by dental researchers that dental opacity observed in fluorosis is not *true* linear enamel hypoplasia, and therefore, linear enamel hypoplasia should not be considered a diagnostic feature of fluorosis (Aoba and Fejerskov, 2002; Fejerskov et al., 1977, 1996; Horowitz et al., 1984; Thylstrup and Fejerskov, 1979). Therefore, we did not include dental linear enamel hypoplasia as a diagnostic feature, however, in the Ray Site population we did observe and record dental defects including opacities, mottling, and dental pitting, which are pathological features associated with dental fluorosis.

Multiple factors influence an individual's response to fluorosis. Skeletal fluorosis appears to occur more frequently in males than females, and in older individuals, though this varies between populations (Ortner, 2003; Littleton, 1999). Factors unique to the individual, such as activity patterns and genetic predisposition, can result in variation across a population (Littleton, 1999; Mousny et al., 2008). Individuals with relatively high levels of physical activity tend to show increased expression of fluorotic skeletal changes (Brickley and Ives, 2010; Littleton, 1999). This results from both increased fluoride intake due to hydration demands of physical activity and bone remodeling in response to stress and loading (Littleton, 1999). Furthermore, epidemiological research reveals there are human "non-responder" populations (Dequeker and Declerck, 1993; Duursma et al., 1987), as well as populations that are sensitive to increased fluoride ingestion (Russell, 1962; Butler et al., 1985; Yoder et al., 1998; Choubisa et al., 2001). A number of studies demonstrate that sex, age, activity, and ancestry all play a role in the response to fluoride (Littleton, 1999; Petrone et al., 2013; Everett, 2011; Duursma et al., 1987; Rich et al., 1964).

We contend that the pathological lesions noted in eight of the 117 Ray Site individuals are etiologically linked to the presence of fluorosis. Skeletal changes displayed by the eight affected Ray Site individuals are consistent with skeletal fluorosis. The presence of dental pitting and mottling, dense periosteal deposition, calcification of ligamentous and tendinous attachments, and fractures supports this diagnosis. There are no individuals who display fused joints or complete ankylosis of the vertebra, typical of advanced stages. Although these severe changes are absent, the presence of bony changes consistent with skeletal fluorosis, in conjunction with the natural environment of the Illinois River Valley, support the hypothesis of fluoride toxicity. This differential diagnosis is illustrated in an abbreviated form in Table 5.

4.2. Environmental Context

The proposed diagnosis of skeletal fluorosis is further supported by the environmental context in which these individuals lived. The most recent publically available fluvial analysis of the Ray Site location, near the La Moine and Illinois rivers, is reported to have fluoride levels of 2.2 – 3.0 mg/L (ppm) (U.S. Dept. of Health, Education, and Welfare, 1969). The water tested for this analysis was not subjected to modern fluoridation or filtration. Surface analysis of Illinois soils in the area yielded fluoride content ranging from 275 mg/kg to 303 mg/kg (Omueti and Jones, 1977a, b). While these modern values of surface soil fluoride content may be influenced by anthropogenic factors, such as farming practices, it is not unreasonable to assert that they reflect values during the Woodland period; this is a rural area of Illinois with little to no industrial and municipal development that would have influenced levels of fluoride content.

High levels of environmental fluoride do not necessarily result in the pathological condition of fluorosis, as a variety of factors influence the environmental presence and bioavailability of fluoride. Environmental factors, such as weathering, temperature, and the natural geological composition of rocks and soil, affect the presence of available fluoride (Barbier et al., 2010; Petrone et al., 2011). Fluoride-binding minerals in the groundwater facilitate movement of the ionic form from the surrounding rocks. Consequently, fluoride is found in soil, water, plants, and marine resources (Littleton, 1999; Omueti and Jones, 1977a; Petrone et al., 2011). Fluoride inhalation from anthropogenic sources such as industrial atmospheric particulate is possible; this pathway has been linked to respiratory disease in coal burning nations, notably

China (Barbier et al., 2010; Watanabe et al., 2000). Furthermore, water storage methods involving clay and soil may increase the amount of fluoride in water molecules through evaporation or leaching from the material used for storage (Littleton, 1999).

It is probable that due to the naturally high level of environmental fluoride, the individuals of the Ray Site community experienced fluoride toxicity. The low percentage of individuals presenting skeletal signs of fluorosis may be attributed to underlying and still unknown genetic factors, as well as individual variation in physical activity, diet, and access to fresh water resources. Communities exploiting the resources of the Illinois River Valley area would have been in continuous contact with fluoride through plant foods, fresh-water resources such as mollusks and fish, and most significantly, groundwater. This interaction with the environment would have likely facilitated fluoride ingestion and encouraged a response to the high level of fluoride.

5. Conclusion and Directions for Future Research

In this paper, we have reported eight individuals, from the Ray Site, who display a shared suite of pathological skeletal changes. Considered together, these shared abnormal skeletal changes suggest that skeletal fluorosis was present at the Ray site. Although these individuals do not display severe signs of the condition, the observed suite of pathological changes, along with the environmental context, which places the community at risk for the development of skeletal fluorosis, leads us to conclude the Ray Site community most likely experienced fluorosis.

Although minimal to moderate fluoride toxicity still occurs in North America due to natural, geologic factors (Goal, 2006; Freni, 1994), this condition has not yet been reported in remains from American archaeological contexts. We argue that this condition may have been more common in archaeological populations than previously recognized, particularly in groups where only minimal to moderate expression is present. This study emphasizes the importance of considering environmental influences upon health, in addition to clinically derived assessments of pathological lesions. A holistic examination of possible influences of environmental and lifestyle factors, along with a suite of indicators of pathological processes, provides the best reference for differentially diagnosing the observed changes and identifying the most likely causes.

Acknowledgments

We would like to express our appreciation the Center for American Archeology for granting access to the remains of the Ray Site, as well as, providing workspace. Our gratitude also goes to the Department of Integrative Physiology and Anatomy, University of North Texas Health Science Center for loaning equipment for data collection. We also wish to thank the following individuals: Dr. Claire Kirchhoff (University of North Texas Health Science Center), Jason King (Director of Research for the Center of American Archeology) for his help in providing resources and access to the collection as well as providing maps of the burial distribution and Ray Site location, and Taylor Thornton, Research Associate of the Center for American Archeology, for the burial map provided for this publication. Also, thank you to Dr. Anne Grauer and the reviewers of the International Journal of Paleopathology for their insightful comments to the improvement of this paper.

Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at doi:https://doi.org/10.1016/j.ijpp.2019.05.003.

References

Abe, S., Ohno, T., Park, P., Higaki, S., Unno, K., Tateishi, A., 1995. Viral behavior of paracrystalline inclusions in osteoclasts of Paget's disease of bone. Ultrastructural pathology 19 (6), 455–461.

Agarwal, S., 2013. Skeletal dysplasia with increased bone density: Evolution of molecular pathogenesis in the last century. Gene 528 (1), 41–45.

Alvarez, J.A., Rezende, K.M.P., Marocho, S.M.S., Alves, F.B.T., Celiberti, P., Ciamponi, A.L., 2009. Dental fluorosis: exposure, prevention and management. Journal of Clinical and Experimental Dentistry 1 (1), 14–18.

Aoba, T., Fejerskov, O., 2002. Dental fluorosis: chemistry and biology. Critical Reviews in Oral Biology & Medicine 13 (2), 155–170.

Asch, D.L., 1976. The middle woodland population of the lower Illinois valley: a study in paleo-demographic methods. Scientific Papers(1).

Aufderheide, A.C., Rodríguez-Martín, C., 1998. The Cambridge Encyclopedia of Human Paleopathology. Cambridge University Press, Cambridge, UK.

Ayoob, S., Gupta, A.K., 2006. Fluoride in drinking water: a review on the status and stress effects. Critical Reviews in Environmental Science and Technology 36 (6), 433–487.

Barbier, O., Arreola-Mendoza, L., Del Razo, L.M., 2010. Molecular mechanisms of fluoride toxicity. Chemico-Biological Interactions 188 (2), 319–333.

Belanger, T.A., Rowe, D.E., 2001. Diffuse idiopathic skeletal hyperostosis: musculoskeletal manifestations. Journal of the American Academy of Orthopaedic Surgeons 9 (4), 258–267.

Bilezikian, J.P., Khan, A., Potts, J.T., Brandi, M.L., Clarke, B.L., Shoback, D., Sanders, J., 2011. Hypoparathyroidism in the adult: Epidemiology, diagnosis, pathophysiology, target-organ involvement, treatment, and challenges for future research. Journal of Bone and Mineral Research 26 (10), 2317–2337.

Bollerslev, J., Henriksen, K., Nielsen, M.F., Brixen, K., Van Hul, W., 2013. Genetics in Endocrinology: Autosomal dominant osteopetrosis revisited: lessons from recent studies. European Journal of Endocrinology 169 (2), R39–R57.

Braun, J., Sieper, J., 2007. Ankylosing spondylitis. The Lancet 369 (9570), 1379–1390.

Brickley, M., Ives, R., 2010. The bioarchaeology of metabolic bone disease. Academic Press, London.

Brooks, S., Suchey, J.M., 1990. Skeletal age determination based on the os pubis: a comparison of the Acsádi-Nemeskéri and Suchey-Brooks methods. Human Evolution 5 (3), 227–238.

Buikstra, J.E., Ubelaker, D.H., 1994. Standards for Data Collection from Human Skeletal Remains. Arkansas Archeological Survey Research Series 44. Arkansas Archaeological Survey, Fayetteville.

Buikstra, J. E., 1976. Hopewell in the Lower Illinois River Valley: a regional approach to the study of human biological variability and prehistoric behavior, Scientific Papers 2. Evanston, IL: Northwestern University Archaeological Program.

Buikstra, J.E., Charles, D.K., 1999. Centering the ancestors: Cemeteries, mounds, and sacred landscapes of the ancient North American midcontinent. Archaeologies of landscape: Contemporary perspectives, pp. 201–228.

Buikstra, J.E., Charles, D.K., Rakita, G.F., 1998. Staging Ritual: Hopewell Ceremonialism at the Mound House Site, Greene County, Illinois. Center for American Archeology, Kampsville Studies in Archaeology and History 1, Kampsville, Ill.

Bullion, E.A., King, J.L., 2015. Relatedness and Social Organization at the Ray Site (11BR104): A Biological Distance Analysis of a Middle Woodland Ridge Top Cemetery in the Illinois Valley. American Journal of Physical Anthropology 156, 95.

Butler, W.J., Segreto, V., Collins, E., 1985. Prevalence of dental mottling in school-aged lifetime residents of 16 Texas communities. American Journal of Public Health 75 (12), 1408–1412.

Cerklewski, F.L., 1997. Fluoride bioavailability—nutritional and clinical aspects. Nutrition research 17 (5), 907–929.

Chaykin, L.B., Frame, B., Sigler, J.W., 1969. Spondylitis: a clue to hypoparathyroidism. Annals of Internal Medicine 70 (5), 995–1000.

Choubisa, S.L., Choubisa, L., Choubisa, D.K., 2001. Endemic fluorosis in Rajasthan. Indian Journal of Environmental Health 43 (4), 177–189.

Conrad, D.A., 2010. Acute hematogenous osteomyelitis. Pediatrics in Review 31 (11), 464–471.

Cook, D.C., 2005. Syphilis? Not quite: paleoepidemiology in an evolutionary context in the Midwest. In: Powell, M.L., Cook, D.C. (Eds.), The Myth of Syphilis: The Natural History of Treponematosis in North America. Xxxx, University of Florida Press, Gainesville, pp. 177–199.

Cook, D.C., Powell, M.L., 2005. Piecing the puzzle together: North American treponematosis in overview. In: Powell, M.L., Cook, D.C. (Eds.), The Myth of Syphilis: The Natural History of Treponematosis in North America. Xxxx, University of Florida Press, Gainesville, pp. 442–479.

Culbert, A.L., Chakkalakal, S.A., Convente, M.R., Lounev, V.Y., Kaplan, F.S., Shore, E.M., 2013. Fibrodysplasia (myositis) ossificans progressiva. In: Thakker, R.V., Whyte, M.P., Eisman, J., Igarashi, T. (Eds.), Genetics of Bone Biology and Skeletal Disease. Academic Press, New York, NY, pp. 375–393.

Dean, H.T., 1936. Chronic Endemic Dental Fluorosis: (Mottled Enamel). Journal of the American Medical Association 107 (16), 1269–1273.

Department of Health, Education, and Welfare (US), 1969. Public Health Service survey of natural fluoride content of community water supplies. Public Health Service (US), Washington.

Dequeker, J., Declerck, K., 1993. Fluor in the treatment of osteoporosis. An overview of thirty years clinical research. Schweizerische Medizinische Wochenschrift 123 (47), 2228–2234.

Duursma, S.A., Glerum, J.H., Van Dijk, A., Bosch, R., Kerkhoff, H., Van Putten, J., Raymakers, J.A., 1987. Responders and non-responders after fluoride therapy in osteoporosis. Bone 8 (3), 131–136.

Evans, D.M., Spencer, C.C., Pointon, J.J., Su, Z., Harvey, D., Kochan, G., Oppermann, U., Dilthey, A., Pirinen, M., Stone, M.A., Appleton, L., 2011. Interaction between ERAP1 and HLA-B27 in ankylosing spondylitis implicates peptide handling in the mechanism for HLA-B27 in disease susceptibility. Nature Genetics 43 (8), 761–767.

Everett, E.T., 2011. Fluoride's effects on the formation of teeth and bones, and the influence of genetics. Journal of Dental Research 90 (5), 552–560.
Fabiani, L., Leoni, V., Vitali, M., 1999. Bone-fracture incidence rate in two Italian regions with different fluoride concentration levels in drinking water. Journal of trace elements in medicine and biology 13 (4), 232–237.
Fejerskov, O., Manji, F., Baelum, V., 1990. The nature and mechanisms of dental fluorosis in man. Journal of Dental Research 69 (2), 692–700.
Fejerskov, O., Thylstrup, A., Larsen, M.J., 1977. Clinical and structural features and possible pathogenic mechanisms of dental fluorosis. Scandinavian Journal of Dental Research 85, 510–534.
Fejerskov, O., Richards, A., DenBesten, P.K., 1996. The effect of fluoride on tooth mineralization. In: Fejerskov, O., Ekstrand, J., Burt, B.A. (Eds.), Fluoride in dentistry, 2nd Ed. Munksgaard, Copenhagen, pp. 112–152.
Flensborg, G., Suby, J.A., Martínez, G., 2013. A case of adult osteomyelitis in a Final Late Holocene hunter-gatherer population, eastern Pampa–Patagonian transition, Argentina. International Journal of Paleopathology 3 (2), 128–133.
Flotow, M., 1983. The Archaeological, Osteological, and Paleodemographical Analyses of the Ray Site: A Biocultural Perspective. Unpublished Master's thesis, Department of Anthropology, University of Missouri, Columbia.
Flotow, M., 2006. The Archaeological, Osteological, and Paleodemographical Analyses of the Ray Site: A Biocultural Perspective. Rediscovery 5, 1–119.
Freni, S.C., 1994. Exposure to high fluoride concentrations in drinking water is associated with decreased birth rates. Journal of Toxicology and Environmental Health, Part A Current Issues 42 (1), 109–121.
Gennari, L., Gianfrancesco, F., Di Stefano, M., Rendina, D., Merlotti, D., Esposito, T., Gallone, S., Fusco, P., Rainero, I., Fenoglio, P., Mancini, M., 2010. SQSTM1 gene analysis and gene-environment interaction in Paget's disease of bone. Journal of Bone and Mineral Research 25 (6), 1375–1384.
Goal, M.C.L., 2006. Fluoride in drinking water: a scientific review of EPA's standards. National Research Council of the National Academies Press, Washington, D.C.
Hackett, C.J., 1975. An introduction to diagnostic criteria of syphilis, treponarid and yaws (treponematoses) in dry bones, and some implications. Virchows Archiv A 368 (3), 229–241.
Hamersma, J., Cardon, L.R., Bradbury, L., Brophy, S., Der Horst-Bruinsma, V., Calin, A., Brown, M.A., 2001. Is disease severity in ankylosing spondylitis genetically determined? Arthritis & Rheumatism 44 (6), 1396–1400.
Hocking, L.J., Lucas, G.J., Daroszewska, A., Mangion, J., Olavesen, M., Cundy, T., Nicholson, G.C., Ward, L., Bennett, S.T., Wuyts, W., Van Hul, W., 2002. Domain-specific mutations in sequestosome 1 (SQSTM1) cause familial and sporadic Paget's disease. Human molecular genetics 11 (22), 2735–2739.
Horowitz, H.S., Driscoll, W.S., Meyers, R.J., Heifetz, S.B., Kingman, A., 1984. A new method for assessing the prevalence of dental fluorosis-the tooth surface index of fluorosis. Journal of the American Dental Association 109, 37–41.
Izuora, K., Twombly, J.G., Whitford, G.M., Demertzis, J., Pacifici, R., Whyte, M.P., 2011. Skeletal fluorosis from brewed tea. The Journal of Clinical Endocrinology & Metabolism 96 (8), 2318–2324.
Jadon, D.R., Ramanan, A.V., Sengupta, R., 2013. Juvenile versus adult-onset ankylosing spondylitis—clinical, radiographic, and social outcomes. A systematic review. The Journal of Rheumatology 40 (11), 1797–1805.
Jha, S.K., Mishra, V.K., Sharma, D.K., Damodaran, T., 2011. Fluoride in the environment and its metabolism in humans. In: In: Whitacre, D.M. (Ed.), Reviews of Environmental Contamination and Toxicology Volume 211. Springer, New York, pp. 121–142.
Jolly, S.S., Sing, B.M., Mathur, O.C., 1969. Endemic fluorosis in Punjab (India). The American journal of medicine 47 (4), 553–563.
Jolly, S.S., Prasad, S., Sharma, R., Chander, R., 1973. Endemic fluorosis in Punjab. 1. Skeletal aspect. Fluoride 6 (1), 4–18.
Kilborn, L.G., Outerbridge, T.S., Lei, H.P., 1950. Fluorosis. Canadian Medical Association Journal 62 (2), 135–141.
Kumar, B.V., 2013. Osteomyelitis: an overview. Kerala Journal of Orthopaedics 26 (1), 70–76.
Largent, E.J., 1961. Fluorosis. The Health Aspects of Fluorine Compounds. Fluorosis. The Health Aspects of Fluorine Compounds.
Lew, D.P., Waldvogel, F.A., 2004. Osteomyelitis. The Lancet 364 (9431), 369–379.
Littleton, J., 1999. Paleopathology of skeletal fluorosis. American Journal of Physical Anthropology 109 (4), 465–483.
Lories, R.J., Luyten, F.P., 2013. Overview of joint and cartilage biology. In: Thakker, R.V., Whyte, M.P., Eisman, J., Igarashi, T. (Eds.), Genetics of Bone Biology and Skeletal Disease. Academic Press, New York, NY, pp. 35–51.
Lukacs, J.R., 1984. Dental fluorosis in early Neolithic Pakistan. American Journal of Physical Anthropology 63 (2), 188.
Mader, R., Buskila, D., Verlaan, J.J., Atzeni, F., Olivieri, I., Pappone, N., Di Girolamo, C., Sarzi-Puttini, P., 2013. Developing new classification criteria for diffuse idiopathic skeletal hyperostosis: back to square one. Rheumatology 52 (2), 326–330.
Mays, S., 2008. Metabolic bone disease. In: Pinhasi, R., Mays, S. (Eds.), Advances in Human Paleopathology. Wiley and Sons, Chichester, UK, pp. 215–251.
Meirelles, E.S., Borelli, A., Camargo, O.P., 1999. Influence of disease activity and chronicity on ankylosing spondylitis bone mass loss. Clinical Rheumatology 18 (5), 364–368.
Meindl, R.S., Russell, K.F., Lovejoy, C.O., 1990. Reliability of age at death in the hamann-todd collection: Validity of subselection procedures used in blind tests of the summary age technique. American Journal of Physical Anthropology 83 (3), 349–357.
Mills, B.G., Singer, F.R., Weiner, L.P., Suffin, S.C., Stabile, E., Holst, P., 1984. Evidence for both respiratory syncytial virus and measles virus antigens in the osteoclasts of patients with Paget's disease of bone. Clinical orthopaedics and related research 183, 303–311.
Møller, P.F., Gudjonsson, S.V., 1932. Massive fluorosis of bones and ligaments. Acta Radiologica 13 (3-4), 269–294.
Mosher, G.M., Smith, M.O., Albrecht, J.L., Salaka, V.P., 2013. Treponemal Disease, Tuberculosis and Subsistence-settlement Pattern in the Late Woodland Period West-central Illinois. International Journal of Osteoarchaeology 25 (5), 776–787.
Mousny, M., Omelon, S., Wise, L., Everett, E.T., Dumitriu, M., Holmyard, D.P., Grynpas, M.D., 2008. Fluoride effects on bone formation and mineralization are influenced by genetics. Bone 43 (6), 1067–1074.
Olivieri, I., D'Angelo, S., Palazzi, C., Padula, A., Mader, R., Khan, M.A., 2009. Diffuse idiopathic skeletal hyperostosis: differentiation from ankylosing spondylitis. Current Rheumatology Reports 11 (5), 321–328.
Omueti, J.A.I., Jones, R.L., 1977a. Fluoride adsorption by Illinois soils. Journal of Soil Science 28 (4), 564–572.
Omueti, J.A.I., Jones, R.L., 1977b. Regional distribution of fluorine in Illinois soils. Soil Science Society of America Journal 41 (4), 771–774.
Ortner, D.J., 2003. Identification of Pathological Conditions in Human Skeletal Remains, 2nd ed. Academic Press, Amsterdam.
Petrone, P., Giordano, M., Giustino, S., Guarino, F.M., 2011. Enduring fluoride health hazard for the Vesuvius Area population: the case of AD 79 Herculaneum. PloS one 6 (6), e21085.
Petrone, P., Guarino, F.M., Giustino, S., Gombos, F., 2013. Ancient and recent evidence of endemic fluorosis in the Naples area. Journal of Geochemical Exploration 131, 14–27.
Phenice, T.W., 1969. A newly developed visual method of sexing the os pubis. American Journal of Physical Anthropology 30 (2), 297–301.
Pignolo, R.J., Shore, E.M., Kaplan, F.S., 2011. Fibrodysplasia ossificans progressiva: clinical and genetic aspects. Orphanet Journal of Rare Diseases 6, 80.
Powell, M.L., Cook, D.C. (Eds.), 2005. The myth of syphilis: the natural history of treponematosis in North America. University Press of Florida, Gainesville, Florida.
Ralston, S.H., 2013. Paget's disease of bone. New England Journal of Medicine 368 (7), 644–650.
Ralston, S.H., Langston, A.L., Reid, I.R., 2008. Pathogenesis and management of Paget's disease of bone. The Lancet 372 (9633), 155–163.
Reddy, D.R., 2009. Neurology of endemic skeletal fluorosis. Neurology India 57 (1), 7–12.
Resnick, D., Niwayama, G., 1988. Parathyroid disorder and renal osteodystrophy. In: Resnick, D., Niwayama, G. (Eds.), Diagnosis of Bone and Joint Disorders. WB Saunders, Philadelphia, pp. 3102–3114.
Resnick, D., Niwayama, G., 1983. Entheses and enthesopathy. Anatomical, pathological, and radiological correlation. Radiology 146 (1), 1–9.
Rich, C., Ensinck, J., 1961. Effect of sodium fluoride on calcium metabolism of human beings. Nature 191 (4784), 184–185.
Rich, C., Ensinck, J., Ivanovich, P., 1964. The effects of sodium fluoride on calcium metabolism of subjects with metabolic bone diseases. Journal of Clinical Investigation 43 (4), 545–556.
Roberts, C.A., Manchester, K., 2007. The archaeology of disease. Cornell University Press, Ithica, New York.
Roholm, K., 1937. Fluorine Intoxication. A Clinical Hygienic Study with a Review of the Literature and some Experimental Investigations. Fluorine Intoxication. A Clinical Hygienic Study with a Review of the Literature and some Experimental Investigations.
Roodman, G.D., Windle, J.J., 2005. Paget disease of bone. The Journal of clinical investigation 115 (2), 200–208.
Rubin, M.R., Dempster, D.W., Zhou, H., Shane, E., Nickolas, T., Sliney, J., Silverberg, S.J., Bilezikian, J.P., 2008. Dynamic and structural properties of the skeleton in hypoparathyroidism. Journal of Bone and Mineral Research 23 (12), 2018–2024.
Russell, A.L., 1962. Dental fluorosis in Grand Rapids during the seventeenth year of fluoridation. Journal of the American Dental Association 65, 608–612.
Scheuer, L., Black, S., 2004. The Juvenile Skeleton. Academic Press, London, UK.
Siddiqui, A.H., 1955. Fluorosis in Nalgonda district, Hyderabad-Deccan. British Medical Journal 2 (4953), 1408–1413.
Singh, S., Saha, S., Singh, S., Shukla, N., Reddy, V.K., 2018. Oral health-related quality of life among 12–15-year children suffering from dental fluorosis residing at endemic fluoride belt of Uttar Pradesh, India. Journal of Indian Association of Public Health Dentistry 16 (1), 54.
Singh, A., Dass, R., Hayreh, S.S., Jolly, S.S., 1962. Skeletal changes in endemic fluorosis. The Journal of Bone and Joint Surgery. British volume 44 (4), 806–815.
Smith, B.D., Yarnell, R.A., 2009. Initial formation of an indigenous crop complex in eastern North America at 3800 BP. Proceedings of the National Academy of Sciences 106 (16), 6561–6566.
Smith, B.D., 2006. Eastern North America as an independent center of plant domestication. Proceedings of the National Academy of Sciences 103 (33), 12223–12228.
Smith, B.D., 1992. Hopewellian farmers of eastern North America. Rivers of Change: Essays on Early Agriculture in Eastern North America. Smithsonian Institution Press, Washington, DC, pp. 201–248.
Sobacchi, C., Schulz, A., Coxon, F.P., Villa, A., Helfrich, M.H., 2013. Osteopetrosis: genetics, treatment and new insights into osteoclast function. Nature Reviews Endocrinology 9 (9), 522–536.
Stark, Z., Savarirayan, R., 2009. Osteopetrosis. Orphanet Journal of Rare Diseases 4, 5.
Tamer, M.N., Köroğlu, B.K., Arslan, Ç., Akdoğan, M., Köroğlu, M., Çam, H., Yildiz, M., 2007. Osteosclerosis due to endemic fluorosis. Science of the total environment 373 (1), 43–48.
Teotia, S.P.S., Teotia, M., 1988. Endemic skeletal fluorosis: clinical and radiological variants. Fluoride 21 (1), 39–44.
Teotia, M., Teotia, S.P.S., Kunwar, K.B., 1971. Endemic skeletal fluorosis. Archives of Disease in Childhood 46 (249), 686–691.
Thylstrup, A., Fejerskov, O., 1979. A scanning electron microscopic and

Evidence of Skeletal Fluorosis at the Ray Site, Illinois, USA: a pathological assessment, discussion of environmental factors.

SUPPLEMENTARY INFORMATION

Authors: Elizabeth A. Nelson[1,2], Christine L. Halling[3], Jane E. Buikstra[4]

Affiliations:

[1]Department of Archaeogenetics, Max Planck Institute for the Science of Human History, Germany

[2]Institute for Archaeological Sciences, University of Tübingen, Germany

[3]Department of Justice, State of Louisiana, USA

[4]Center for Bioarchaeological Research, Arizona State University, Arizona, USA

SI Figure 1. Ray 3-71, male, left tibia displaying sclerotic periosteal deposition on the anterior and medial aspect of the diaphysis and the medial malleolus. Abnormal bone formation highlighted with red outlines.

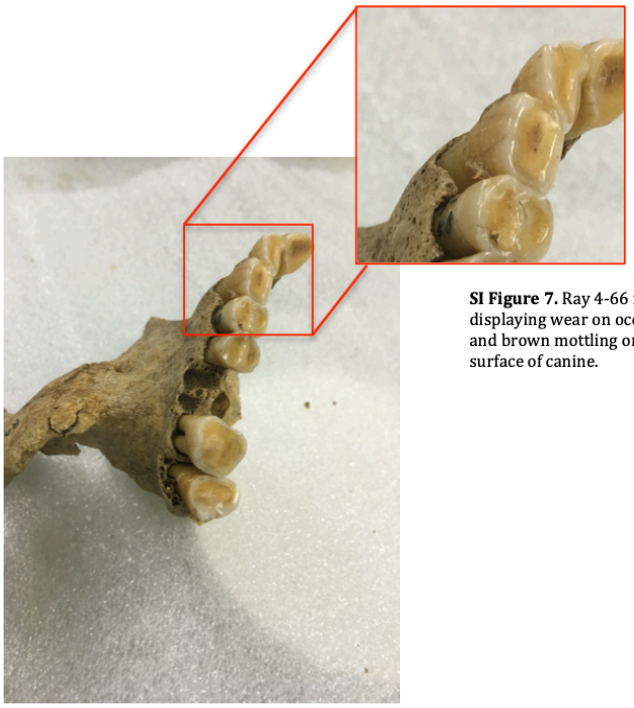
SI Figure 2. Ray 3-18, thoracic vertebra displaying ossification of the anterior longitudinal ligament.

SI Figure 3. Ray 3-18, right radius displaying abnormal surface of posterior diaphysis with raised sharp ossification of tendinous and ligamentous attachments at the the posterior distal surface (indicated by red arrow).

SI Figure 4. Ray 2-69 mandible displaying significant dental wear.

SI Figure 5. Ray 2-77 maxilla displaying resorbed alveolus resulting in antemortem loss of teeth.

SI Figure 6. Ray 3-71, left clavicle displaying abnormal bone formation on the diaphysis of the clavicle.



SI Figure 7. Ray 4-66 right maxilla displaying wear on occlusal surface and brown mottling on lingual surface of canine.

SI Figure 8. Ray 4-66 right third maxillary molar displaying hard soil deposition obscuring view for dental assessment.

SI Figure 9. Ray 3-38 mandible displaying severe wear of teeth.

SI Figure 10. Ray 3-38 left third mandibular molar displaying brown mottling.

Paper II

E. A. Nelson, J. E. Buikstra, A. Herbig, T.A. Tung, K.I. Bos (2020).

Advances in Paleopathology of Tuberculosis in Pre-contact Andean South America

International Journal of Paleopathology, 29, 128-140.

International Journal of Paleopathology 29 (2020) 128–140

Contents lists available at ScienceDirect

International Journal of Paleopathology

journal homepage: www.elsevier.com/locate/ijpp

Advances in the molecular detection of tuberculosis in pre-contact Andean South America

Elizabeth A. Nelson[a,b], Jane E. Buikstra[c], Alexander Herbig[a], Tiffiny A. Tung[d,**], Kirsten I. Bos[a,*]

[a] Department of Archaeogenetics, Max Planck Institute for the Science of Human History, Kahlaische Str 10, 07745 Jena, Germany
[b] Eberhard Karls Universität Tübingen, Geschwister-Scholl-Platz, 72074 Tübingen, Germany
[c] Center for Bioarchaeological Research, Arizona State University, 1151 S. Forest Ave., Tempe, AZ, 85281, USA
[d] Department of Anthropology, Vanderbilt University, VU Station B #356050, Nashville, TN 37235, USA

ARTICLE INFO

Keywords:
Andean South America
Ancient DNA
Huari
Molecular paleopathology
Tuberculosis

ABSTRACT

Andean paleopathological research has significantly enhanced knowledge about the geographical distribution and evolution of tuberculosis (TB) in pre-Columbian South America. In this paper, we review the history and progress of research on ancient tuberculosis (TB) in the Andean region, focusing on the strengths and limitations of current approaches for the molecular detection of ancient pathogens, with special attention to TB. As a case study, we describe a molecular screening approach for the detection of ancient *Mycobacterium tuberculosis* in individuals from Late Intermediate Period (1000–1400 CE) contexts at the site of Huari, Peru. We evaluate 34 commingled human vertebrae and combine morphological assessments of pathology with high throughput sequencing and a non-selective approach to ancient pathogen DNA screening. Our method enabled the simultaneous detection of ancient *M. tuberculosis* DNA and an evaluation of the environmental microbial composition of each sample. Our results show that despite the dominance of environmental DNA, molecular signatures of *M. tuberculosis* were identified in eight vertebrae, six of which had no observable skeletal pathology classically associated tuberculosis infection. This screening approach will assist in the identification of candidate samples for downstream genomic analyses. The method permits higher resolution disease identification in cases where pathology may be absent, or where the archaeological context may necessitate a broad differential diagnosis based on morphology alone.

1. Introduction

Archaeological collections of Andean South America have been an important influence on methodological advances in the field of paleopathology. The environmental conditions of certain regions naturally promote the preservation of human remains, yielding an abundance of biological material including bone, teeth, hair, soft tissue and coprolites, all of which are rich resources for the study of past disease (Bos et al., 2014; Harkins et al., 2015; Salo et al., 1994; Verano, 1997). Pioneers of the field in the 19th and 20th centuries, such as Hrdlička and Moodie, took special interest in the anatomy, health, and disease of ancient Peruvians (Hrdlička, 1914; Moodie, 1923; Tello, 1909; Weiss et al., 1984), though their focus centered heavily on individual case studies. This early attention set the stage for the many historical and ethnographic descriptions of disease in the region (Hoyle et al., 1945; de Ayala, 1980; Tello and Williams, 1930; Urteaga-Ballon, 1991), which have yielded a wealth of information concerning the health of its past peoples.

In 1997, John Verano published *Advances in the Paleopathology of Andean South America*. In this work, he discussed the foundations of paleopathology in the Andean region and highlighted the advances research on Andean populations stimulated in this increasingly multidisciplinary field of investigation. In the 20 years following its publication, Andean collections have continued to play a central role in anthropological theory (Lozada and Tantaleán, 2019) and bioarchaeological methods (Blom et al., 2005; Klaus, 2014) as related to paleopathology, while promoting interdisciplinary collaboration through the incorporation of new technologies to better understand population health and disease in the past (Aufderheide et al., 2004; Bos et al., 2014; Wilbur and Buikstra, 2006).

Notable progress in paleopathology has come from the considerable contributions of genetic methods as applied to infectious disease

* Corresponding author.
** Corresponding author at: Department of Anthropology, Vanderbilt University, 212 Garland Hall, Nashville, TN 37235, USA.
E-mail addresses: nelson@shh.mpg.de (E.A. Nelson), buikstra@asu.edu (J.E. Buikstra), herbig@shh.mpg.de (A. Herbig), t.tung@vanderbilt.edu (T.A. Tung), bos@shh.mpg.de (K.I. Bos).

https://doi.org/10.1016/j.ijpp.2019.12.006
Received 28 February 2019; Received in revised form 6 December 2019; Accepted 13 December 2019
Available online 20 January 2020

detection and characterization in archaeological specimens, which now extend as far back as the Neolithic (Bos et al., 2014; Mühlemann et al., 2018; Salo et al., 1994; Schuenemann et al., 2018; Spyrou et al., 2018; Vågene et al., 2018). The recent introduction of high throughput sequencing greatly assists our quest to address medically and historically relevant questions. Genome-level analyses permit a glimpse at complex evolutionary history, thus yielding a better understanding of disease presence, geographic range, and ecology. The addition of genetic analyses also provide a means to move beyond morphologically detectable diseases to those that leave only molecular remnants in archaeological tissues such as infections of *Salmonella enterica* and hepatitis B virus (Mühlemann et al., 2018; Vågene et al., 2018). In the case of Andean populations, genome-level investigations of tuberculosis (TB) have altered our perceptions of ancient disease ecology, where the potential of zoonotic transmissions from marine mammals are now entering into discussions on disease origin in the region (Bos et al., 2014).

Here we explore how molecular analyses of archaeological material from the Andes are contributing to new and much needed standards in the detection of *Mycobacterium tuberculosis* DNA in ancient remains. We begin by describing the disease itself in a social and biological framework and then examine its known distribution in the Andes of South America by drawing upon previously published paleopathological studies. Finally, we explore the benefits of incorporating molecular data into morphology-based research on skeletal tuberculosis, along with a consideration of archaeological context to establish an interpretive framework. This is demonstrated via the presentation of TB screening data for individuals from Late Intermediate Period (LIP) contexts at the site of Huari, the former capital city of the Wari Empire, located in the Central Andes of Peru.

2. Tuberculosis past and present

Tuberculosis is caused by infection of any member of the *Mycobacterium tuberculosis* complex (MTBC). Members of this complex show extensive host ranges that include humans and some mammals (Brosch et al., 2002). Transmission occurs most frequently via the inhalation of infectious droplets from an individual with active disease; less common is infection via the consumption of contaminated animal products. The bacterium primarily infects the lungs, but it can also disseminate to extra-pulmonary sites through haematogenous spread of the bacilli following the initial infection or through the lymphatic system by way of infected macrophages or by liberated TB bacilli (Mousa, 2007). Socio-cultural factors that contribute to the spread of TB include socio-political disruption, migration, population crowding, poverty, animal interaction, compromised immune systems and malnutrition (Lönnroth et al., 2009; Wilbur and Buikstra, 2006). Today, TB is considered a re-emerging disease and is a global threat, largely due to the development of antibiotic resistant strains.

2.1. Tuberculosis skeletal morphology

Descriptions of skeletal lesions consistent with TB are plentiful in the archaeological record, with earliest evidence from contexts on or near the Mediterranean dating to 7250 BCE (Hershkovitz et al., 2015; El-Najjar et al., 1996). The diagnostic pattern of lytic lesions characteristic of TB allow for detection of the disease in archaeological contexts. These skeletal changes can, however, be similar to those caused by other parasitic, fungal, and bacterial infections (Buikstra, 1976; Roberts and Buikstra, 2019). Primary bone destruction involves intracellular infection of multinuclear osteoclasts that lead to pathological osteoclast activation (Hoshino et al., 2014). As the disease progresses, osteoblasts may form new woven bone around the destructive lesions; bone destruction with little to no new bone formation characterizes the active phase of TB (Mousa, 2007; Resnick and Niwayama, 1995). Characteristic changes to the vertebrae are primarily destructive, resorptive lesions with a loss of the extracellular matrix, most frequently circular in shape with rounded edges, penetrating the cortical bone and exposing trabeculae (Roberts and Buikstra, 2019). Areas of haematopoietic marrow with rich blood supply are the most frequent sites to show physical changes. In adults, this includes the vertebrae, sternum, ribs and the epiphyses of long bones; for subadults, the red marrow is distributed over much more of the skeleton, and thus the infection typically involves bones of the hands and feet (Buikstra, 1976; Roberts and Buikstra, 2019; Subasi et al., 2004). In both adults and subadults, thoracic and lumbar regions of the spine are most frequently affected. Subadults are more likely to exhibit non-specific changes such as periosteal reactions on long bones (Cieślik, 2017; Roberts and Buikstra, 2019). It is common for two or more adjacent vertebrae to be affected, but noncontiguous vertebral involvement is also possible (Polley and Dunn, 2009). Lesions of spinal tuberculosis are typically located on the superior, inferior, and/or anterior aspects of vertebral bodies. They may be small in size (< 3 mm) or may progressively increase in size to ultimately involve the entire vertebral body, which may lead to vertebral collapse, resulting in kyphosis (Roberts and Buikstra, 2019).

Importantly, many individuals suffering from TB do not present skeletal changes classically associated with active infection. Though this may vary between populations, skeletal tuberculosis occurs in approximately 3 % of individuals affected with pulmonary TB and 30 % of individuals with extrapulmonary TB (Kastert and Uehlinger, 1964; Roberts and Buikstra, 2019). Likewise, skeletal changes are not associated with early phases of the infection but rather occur in chronic states. Some cases could, therefore, go undetected in standard skeletal surveys of archaeological material (Roberts and Buikstra, 2003).

3. Tuberculosis paleopathology of Andean South America

Tuberculosis research in paleopathology has focused on cases that present skeletal lesions, with inclusion of non-skeletal pathology such as lung nodules (Kim et al., 2016; Lunardini et al., 2012; Salo et al., 1994). A number of reports explore the distribution of TB in bioarchaeological collections from pre-Columbian North and South America, and the majority of these cases cluster in the U.S. Southwest, the U.S. Midwest, and the Andean region of South America (Buikstra, 1981, 1999; Roberts and Buikstra, 2003). Due to mitigating factors such as inadequate skeletal preservation, incomplete archaeological recovery and looting, it is not possible to assess the true geographic range of TB in the pre-contact Americas. However, the desert environment in parts of the Andes is highly conducive to the preservation of mummified and skeletal tissues, which has increased the visibility of TB infections in the region (Arriaza et al., 1995; Arrieta et al., 2014). Excellent preservation also results from the mortuary practices employed by pre-Hispanic Andean peoples, via both intentional and unintentional mummification from wrapping the dead in textiles and placing them in cool, dry, stone-lined tombs (Tello, 1909; Vreeland, 1998). Given the sizeable corpus of mummies from the Andes (Arriaza, 1995; Guillén, 2004), many have been radiographed to examine soft tissue lesions and mineralized structures such as lung nodules and skeletal changes that may or may not be apparent through morphological observation (Friedrich et al., 2010).

Skeletal observations of tuberculosis-like lesions from ancient South American contexts were recorded as early as 1940 (García-Frías, 1940; Requena, 1945). Following this, Marvin Allison and Enrique Gerszten investigated evidence for TB in ancient Peru through analyses of mummies and skeletons, with contributions from others such as Jose García Frias (Guillén, 2012). In 1973, Allison published the first diagnosis of pre-Columbian TB in the Andes based on skeletal and soft tissue observations (Allison et al., 1973). This study of a mummified sub-adult combined skeletal observations of tuberculosis-like lesions and the microscopic identification of acid-fast bacilli from mummified organs (lung, pleura, liver, and right kidney) morphologically consistent with *M. tuberculosis* (Allison et al., 1973). Additionally, the individual

Fig. 1. Map of pre-contact TB cases in the Andes. This map illustrates locations of reported ancient tuberculosis in Andean South America (adapted from Roberts and Buikstra, 2003). The Andean cultural region is shaded with black lines (adapted from Verano, 1997). Red dots indicate sites where TB was identified in 1 individual; blue triangles indicate sites where TB was identified in 2 or more individuals. The yellow star indicates the location of Huari, the archaeological site from which the individuals included in this study were recovered. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

presented skeletal features consistent with TB, including lytic foci of the vertebrae. Having been recovered from a Nasca cemetery in southern Peru dating to 200–800 CE, analysis of this individual yielded the strongest evidence at the time for the existence of TB prior to European contact in the Americas. Furthermore, this pioneering study provided the first diagnosis of pre-Columbian TB based on soft tissue and skeletal evidence (Allison et al., 1973). Since then, many more cases of pre-contact TB have been identified through macroscopic and microscopic approaches (Allison et al., 1981; Buikstra and Williams, 1991; Arriaza et al., 1995; Burgess, 1992; Lombardi and García Cáceres, 2000; Salo et al., 1994; Klaus et al., 2010).

Fig. 1 illustrates South American archaeological sites where published cases of TB have been identified based on morphology. Limiting our discussion to those archaeological sites where more than one individual shows skeletal changes consistent with TB, the earliest New World examples are from Andean South America. While there exists a report of possible TB from Chile dating to 290 CE (Allison et al., 1981), the earliest uncontested cases of New World TB date to 700 CE from Chile (Allison et al., 1981) and Peru (Allison et al., 1981; Buikstra and Williams, 1991; Burgess, 1992; García-Frías, 1940; Lombardi and García Cáceres, 2000; Owen, 1993), followed by later reports from Colombia (Rivas, 1988; Correal and Flórez, 1992; Martinez de Arateco Hoyo, 1999; Rodriguez, 1988; Romero Arateco, 1998). Evidence of TB continues to gradually increase in the Andean Middle Horizon

(600–1000 CE), where TB has been identified in multiple sites including Montegrande (~890 CE) and Los Médanos, Nasca, Peru, with some presence in San Geronimo of the Osmore Valley of southern coastal Peru (Allison et al., 1981; Lombardi and García Cáceres, 2000; Burgess, 1992) and the north coast in the Lambayeque Valley, dating to 900–1100 CE (Klaus et al., 2010). There appears to be an increase in identifiable TB cases in later periods with the majority of examples coming from the LIP (1000–1400 CE). Estuquiña, in the Moquegua valley of southern Peru dating to ~1350 CE, provided the largest skeletal series of ancient TB (37 individuals), and made possible the first paleoepidemiological assessment of this disease in the Americas (Buikstra and Williams, 1981). Additional reports from the LIP include those recovered from sites in the Azapa Valley, Chile (Arriaza et al., 1995), and sites such as Huayari in the Palpa of south-central Peru (Allison et al., 1981) (~1250 CE), La Caleta de San José in the Lambayeque department (Klaus et al., 2010) (1375–1450 CE), and Algodonal from the Osmore River Valley (Owen, 1993; Burgess, 1992; Roberts and Buikstra, 2003). Following this peak, there is a significant decrease in reports of skeletal tuberculosis from the Late Horizon, the time of Inka imperial rule (1450–1532 CE).

Despite the abundance of tuberculosis-like lesions reported from a number of pre-Columbian sites, the existence of TB in the Americas prior to European contact remained contentious. Early on, more conservative perspectives on the topic stated there were too few examples in the literature, differential diagnoses were of poor quality, and/or the described pathology was non-specific for TB infection (Hrdlička, 1909; Moodie, 1923; Morse, 1961; Stead et al., 1995). Likewise, molecular characterization of modern MTBC strains across the globe showed that the European lineage (L4) currently dominates in the Americas (Gagneux et al., 2006), originally leading scholars to interpret this as evidence that TB was introduced through European contact.

A comprehensive understanding of the skeletal changes associated with TB and the geographic and ecological histories of this disease in Andean South America is, as of yet, limited. Multiple factors, both cultural and environmental, may reduce the recovery and preservation of skeletons and mummies; these factors compromise paleopathological and epidemiological studies of TB and other diseases that affect the skeleton (Roberts and Buikstra, 2003). Furthermore, the likely circumstance of overlooking multiple cases due to the absence of skeletal changes, even in the presence of TB infection, challenges population-wide studies of tuberculosis from archaeological contexts. It is in response to this challenge that specialized methods involving molecular analyses have been adapted and applied to studies of ancient diseases.

4. Molecular methods and pre-Columbian TB

With the first reports on recovery of preserved DNA in ancient tissues (e.g. Pääbo, 1989), the application of molecular methods to paleopathology has led to significant advances in the detection of ancient pathogens. TB has featured prominently in these studies, where the first genetic analyses of ancient disease involved the detection of TB in the Andean cultural region (Salo et al., 1994) and beyond (Spigelman and Lemma, 1993).

4.1. Sample selection in molecular studies

Molecular paleopathologists draw upon multiple sources of data to identify archaeological material that may be well suited for ancient pathogen DNA analysis. Morphological descriptions of skeletal lesions are key for discernable diseases, but historical records and/or mortuary contexts that suggest altered burial programs in response to outbreaks may also be informative about when and where past epidemics occurred (DeWitte, 2016; Margerison and Knüsel, 2002; Signoli et al., 2002). Pathophysiology of the suspected organism, identified through morphological changes or other methods (i.e., historical documents, burial programs, etc.), should be considered when selecting skeletal elements for analysis. For example, the ideal source material for detection of *Yersina pestis*, the causative agent of plague, in archaeological samples is thought to be dental remains due to the diffuse sepsis that can occur in the terminal phase of *Y. pestis* infection: the densely vascularized dental pulp chamber may harbor residues of a pathogen present in the blood at the time of death (Drancourt et al., 1998). In contrast, vertebrae are the preferred elements to sample for *M. tuberculosis* detection due to their presentation of pathognomonic lesions, which is related to the organism's venous spread, affinity for oxygen rich environments, and colonization of bone marrow (Buikstra, 1976). Likewise, lung nodules, though less frequently detected in the archaeological record, are strong candidates for harboring TB DNA (Sabin et al., 2019; Salo et al., 1994).

4.2. M. tuberculosis detection via the Polymerase Chain Reaction (PCR)

At its outset, PCR-based detection techniques were applied to ancient DNA and quickly became a popular method for TB detection in archaeological material (Salo et al., 1994; Spigelman and Lemma, 1993). In 1994, Salo et al. sampled two soft tissue sites (a lung nodule and a lymph node) of a naturally mummified female from Chiribaya Alta (~1000 – 1300 CE). By employing aDNA methods of extraction and molecular detection via PCR, the authors identified genetic sequences purportedly unique to the MTBC. This was the first genetic evidence for TB infection in the Americas prior to European contact. Since then, additional investigations of Andean paleopathology have applied PCR detection to identify *M. tuberculosis* in several other individuals from the region (Arriaza et al., 1995; Konomi et al., 2002; Klaus et al., 2010). These methods, however, did not permit identification of the ancient MTBC strain or its relationship to modern forms. The question of TB presence in the pre-Columbian Americas had been answered, but the mystery of its origin remained.

Thirty years on, PCR continues to be the most common method for detection of ancient *M. tuberculosis* due to its affordability and ease of use (Harkins and Stone, 2015; Stone and Ozga, 2019). This method usually targets genetic regions believed to be multi-copy and specific to the MTBC such as the mobile elements IS6110 and IS1081 (Wilbur et al., 2009). The utility of this detection approach in a clinical setting as a means of rapid diagnosis has been confidently demonstrated (Eisenach et al., 1991). However, archaeological material presents a rather different situation, where the majority of the extracted DNA stems not from the host, but rather from the environment. Host DNA is often limited to less than 1 % of total DNA content (Skoglund et al., 2012) and the dominant environmental signal is composed of a variety of bacterial species, many of which are non-pathogenic members of the mycobacterial taxonomic group. The existence of these insertion sequences in environmental species of mycobacteria remains underexplored and thus raises the possibility of false positive PCR-based detections (Müller et al., 2015). In 2015, Harkins and colleagues presented an improved assay to target MTBC members that offers increased accuracy for PCR-based detection. This assay includes primers that target the mobile elements IS6110, IS1081, as well as two primers that target the single copy gene *rpoB* (rpoB1 and rpoB2) (Harkins et al., 2015). While methods of TB detection via PCR are improving, the approach remains limited in that these data do not offer the resolution needed to facilitate larger discussions related to the pathogen's evolution or its relationship to other strains. These themes are better explored via approaches that provide genome-wide data.

4.3. The high-throughput sequencing revolution

The introduction of Next Generation Sequencing (NGS) revolutionized DNA research, and is regarded as the current gold standard for DNA analyses of archaeological material. High throughput sequencing technologies make possible the simultaneous sequencing of millions of DNA molecules. As such, they yield an abundance of genetic

information from a single archaeological sample (Green et al., 2010; Rasmussen et al., 2010; Prüfer et al., 2014). The most straightforward method, which indiscriminately sequences a portion of the DNA present in an extract, is commonly referred to as "shotgun sequencing". In archaeological samples, this provides a glimpse of the molecular complexity of the material under study. Usually, an archaeological sample contains DNA from a multitude of sources including human, fungal, plant, environmental microbes, and potentially pathogens. Shotgun sequencing allows for a broad scale survey of the DNA contents of a given sample, be it soft tissue, bone, or tooth.

In most examples, pathogen DNA constitutes a very small amount of the overall composition of the extracted sample, thus making it difficult to detect in a typical shotgun sequenced dataset. Methods that facilitate the evaluation of community profiles of large genetic datasets have been developed, and some of these have shown promise as pathogen screening tools for archaeological samples. The MEGAN Alignment Tool (MALT), for example, is a sequence alignment program that compares DNA fragments to a database of genomic sequences and provides a taxonomic assignment of all DNA sequences that can be identified (Herbig et al., 2016; Vågene et al., 2018). Their taxon assignments, and similarities to the reference sequences upon which these assignments are based, can be easily viewed in MEGAN, the Metagenome Analyzer tool (Huson et al., 2016). In addition to identifying DNA that is endogenous to the archaeological material, MALT permits an evaluation of the environmental background profile, which, in the context of MTBC, can be very informative in the detection of false positives (Warinner et al., 2017). In general, MALT can detect pathogens in a sample with limited ascertainment bias in comparison to other molecular detection techniques that screen for a target pathogen thought to be present *a priori* (Schuenemann et al., 2011; Devault et al., 2014; Bos et al., 2015) or that rely on parallel capture of several pathogenic species based on derived genomic characters (Bos et al., 2015). Therefore, identification of pathogenic organisms not previously suspected to be present in a biological sample is made possible (Vågene et al., 2018), as well as the potential detection of co-morbidities.

Ancient DNA is highly fragmented with a typical length of fewer than 100 base pairs. It is also typically "damaged" where cytosine (C) nucleotides can deaminate to form uracil, which results in the incorporation of thymine (T) by polymerases in subsequent amplifications. These C to T transitions accumulate on the terminal ends of ancient molecules and their frequency increases with time. The pattern and distribution of C to T changes in molecules from an archeological sample can thus be used to authenticate the recovered DNA as ancient (Briggs et al., 2009; Krause et al., 2010). This method of authentication is limited with PCR data because analyses are restricted to the center of the fragment. Next generation sequencing approaches, however, offer the added benefit of sequencing the full fragment, thus making such authentication possible.

4.4. The power of reconstructing ancient pathogen genomes

Once pathogen DNA has been detected and is confirmed to be ancient, whole-genome capture can be performed to enrich for DNA of a specified target (Spyrou et al., 2019). This is done through hybridization of the target DNA to a collection of synthetically produced probes that are complementary to the DNA of the desired pathogen. Subsequently, hybridized DNA is separated from the background DNA pool and liberated from the probes, leaving an enriched DNA fraction (Bos et al., 2019). This reduces the background and increases the proportion of target pathogen DNA in the sequenced portion. When sufficient target DNA is captured, whole-genome reconstruction is possible. Whole-genome analysis enables an investigation of evolutionary changes and relationships between lineages through the evaluation of small changes in DNA sequences known as single nucleotide polymorphisms (SNPs). The evolutionary history of a pathogenic bacterium can be further investigated through molecular dating analyses of ancient sequences, which can be used as calibration points to determine more reliable mutation rates and extend our understanding of bacterial evolution beyond those determined from modern epidemics alone.

By employing high throughput sequencing and capture methods, Bos et al. (2014) were able to recover and reconstruct three pre-Columbian MTBC genomes from the southern coast of Peru dating to 1028–1280 CE associated with the Chiribaya culture. The genomic reconstruction of these pre-contact MTBC strains supported earlier evidence that revealed the presence of TB in the New World prior to European contact (Salo et al., 1994), though revealed two novel details. First, comparison against a modern global TB dataset permitted a Bayesian molecular dating analysis, which suggested a much younger age for the most recent common ancestor (MRCA) of MTBC than previously thought, with an estimated date of less than 6000 years. A previous estimation for the age of the MTBC MRCA was based on similarity observed between phylogenetic structure of human mitochondrial lineages and human-adapted lineages of MTBC: This led to the assumption that the MTBC left Africa with humans during Pleistocene migrations, with a best MRCA estimate of 70,000 years ago (Comas et al., 2013). The younger MRCA date, however, has gained further support through dating analyses that have incorporated ancient European TB genomes as calibration points (Bos et al., 2014; Kay et al., 2015; Sabin et al., 2019). These results are incompatible with models that suggest a terrestrial introduction of TB to the Americas prior to European contact (Hershkovitz et al., 2015; Rothschild and Martin, 2003). Second, the position of the ancient Peruvian genomes in the global TB phylogeny revealed the strains to be most closely related to a member of the MTBC currently associated with seals and sea lions, *M. pinnipedii* (Bos et al., 2014). This finding gave rise to the hypothesis that the introduction of MTBC in the Americas occurred through zoonotic transfer of a pathogen from pinnipeds to humans. This is consistent with the inferred level of interaction between humans and seals, and the reliance on seal meat as a food source (DeFrance, 2009; Reitz and Sandweiss, 2001).

The hypothesis of transmission from pinnipeds to humans, however, warrants further exploration. The evaluation of additional ancient TB genomes from pre-contact populations in various regions of the Americas will provide the opportunity to evaluate the ecology of TB amongst communities not known to exploit pinnipeds. A first step in this pursuit is the development of robust techniques for screening skeletal material from the vast South American archaeological record.

5. A demonstration of metagenomic analysis from shotgun data: case study from the Terminal Wari and Late Intermediate period, Huari, Peru

5.1. Biocultural context

Through application of the advanced methods described above, we analyzed human vertebrae from Terminal Wari (ca. 1000–1100 CE) and the second half of the LIP (1275–1400 CE) contexts at the site of Huari, the former capital city of the Wari Empire (Fig. 2). The Wari Empire (600–1000/1100 CE), with its core in the Ayacucho Basin of central, highland Peru, was the first expansive empire in the ancient Andes (Isbell and Schreiber, 1978; Schreiber, 1992; Tung, 2012). The Wari Empire influenced Andean cultural traditions such as ceramic styles and iconography (Cook and Glowacki, 2003), architectural forms (Isbell and McEwan, 1991), and also likely altered language patterns (Heggarty and Beresford-Jones, 2012; Mannheim, 1981). Wari extended their influence through trade networks and by establishing new agricultural complexes and administrative centers across the Peruvian Andes, from Cajamarca in the north to the Moquegua valley in the south, and to many Pacific coastal valleys (Schreiber, 1992; Tung, 2012; Williams, 2001). Although the reasons for Wari imperial collapse are still not fully understood, skeletal evidence for violence points to socio-political turmoil within the imperial core (Tung, 2014b) combined with climate

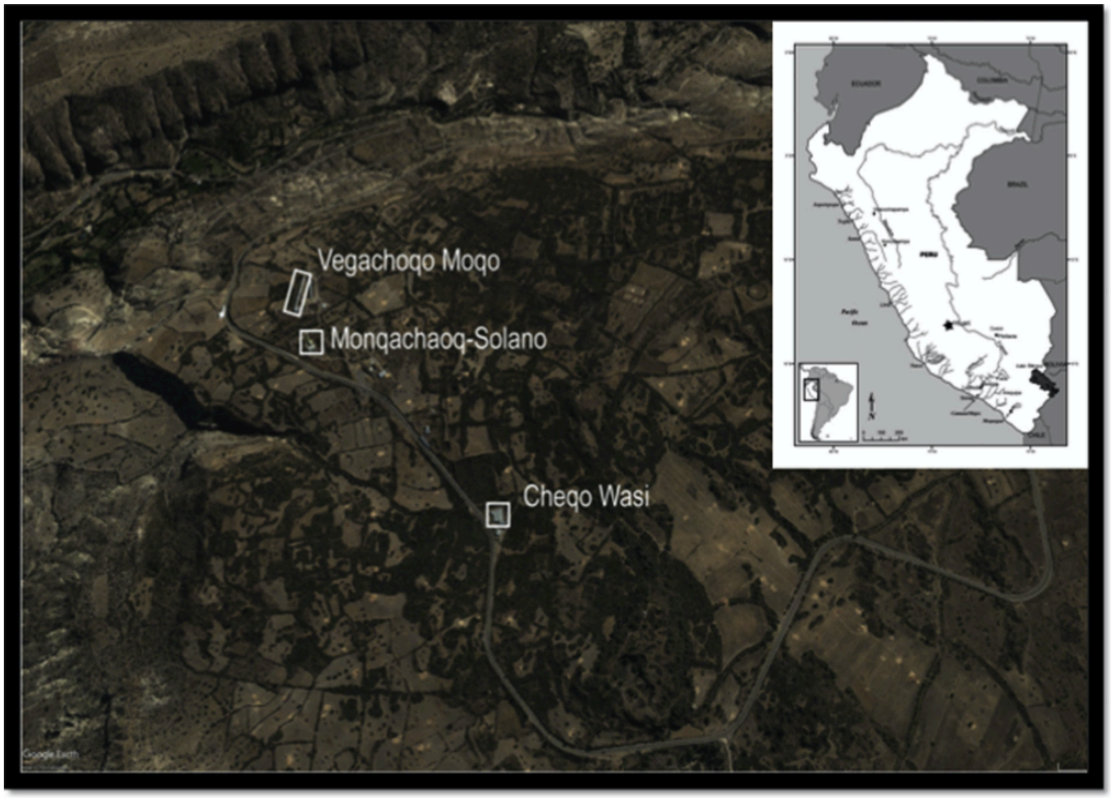


Fig. 2. Map of Huari. Location of Huari with the identification of the three sectors included in this study: Vegachayoq Moqo, Monqachayoq, and Cheqo Wasi. The map inset shows Peru with the location of Huari indicated by the black star. Image by Tiffiny Tung.

stress in the form of a severe and extended drought (Bird et al., 2011; Thompson et al., 2013). These and other factors contributed to the decline of the Wari Empire ca. 1000–1100 CE.

The political decline of Wari ushered in the LIP, a time characterized by a shift to more dispersed settlements (many of which were constructed as defensive sites), the cessation of monumental building projects, and increased violence, particularly in the central and southern Peruvian Andes (Arkush and Tung, 2013; Tung, 2008). Stable isotope analyses demonstrate that diets significantly changed in the 200 years after Wari imperial decline; particularly noteworthy is the significant decline in the childhood consumption of the socially valued food, maize (Tung et al., 2016). Coincident with these increases in violence and dietary changes is an increase in skeletal lesions consistent with infectious disease.

5.2. Materials and methods

To better understand population health during and after the period of imperial decline, we conducted a gross morphological evaluation of pathology and a metagenomic screen for *M. tuberculosis* DNA in vertebral samples from the Terminal Wari context of Cheqo Wasi, (ca. 1000–1150 CE; n = 4), and the post-Wari sectors of Monqachayoq and Vegachayoq Moqo (ca. 1275–1400 CE; n = 30). We collected bone powder from each vertebra for DNA extraction and sequencing. Sequencing data of each sample was then analyzed for ancient human DNA content and ancient *M. tuberculosis* DNA content. Morphological assessments and molecular methods and analyses are described in detail in our Supplemental Materials (Supplementary Appendix A).

5.3. Results

All vertebrae displayed excellent preservation with little to no taphonomic damage. Two (HUA002, HUA004) of the four Cheqo Wasi elements displayed lytic foci with advanced destruction that penetrated deep into the trabecular matrix (Fig. 3a and 3b). Vertebra HUA002 presented one cavity of lytic destruction on the anterior aspect that involved greater than half of the anterior surface and penetrated into the vertebral body (Fig. 3a). Likewise, HUA004 presented large, well-developed lytic lesions; however, these occured on the lateral and superior aspects of the vertebral body (Fig. 3b). Both vertebrae displayed smooth surfaces within the foci (Fig. 3a). These lesions are consistent with bacterial infection, specifically tuberculosis and/or brucellosis, but could also be attributed to paracoccidoidimycosis or echninococcosis (Roberts and Buikstra, 2019).

Vertebrae from Vegachayoq Moqo and Monqachayoq included elements that showed no external macroscopically observable pathological changes, some that exhibited hypervascularity on the vertebral bodies, some that presented non-diagnostic skeletal changes, and others that exhibited lesions suggestive of bacterial infection. Vertebrae with pathological skeletal changes displayed a spectrum of presentation from very slight reactive bone on the anterior surface (Fig. 4a) and irregular pitting to hypervascularization on the anterior aspect of the vertebral body. However, two vertebrae from Monqachayoq displayed TB-like lesions in the form of circular, lytic foci involving < 1/3 of the vertebral body (Fig. 4b). Although atypical changes such as superficial irregular pitting have previously been associated with early stage TB (Baker et al., 1988; Pálfi, 2012; Maczel, 2003; Mays et al., 2002; Zink et al., 2007), our molecular screening approach is broad and would permit identification of either TB or brucellosis, which must be considered due to the commingled nature of this skeletal sample (Mays et al., 2001).

For the 34 vertebrae considered here, human DNA content ranged from 0.034 to 11.9 %, with most samples containing 3 % or less (Table 1). Whole microbial content analysis in MALT was performed

Fig. 3. Tuberculosis-like lesions. Two thoracic vertebrae from Cheqo Wasi displaying characteristic tuberculosis-like lesions with large lytic foci invading the trabeculae on the anterior (a) and lateral (b) surfaces. The vertebrae shown are samples HUA002 (a) and HUA004 (b).

using a database of all bacterial genomes available on NCBI RefSeq accessed December 2016 with a 95 % sequence identity threshold. This resulted in the identification of 41 to 1163 MTBC reads per sample (Table 1). Fig. 5 shows a taxonomic tree of assigned reads as well as a DNA alignment of MTBC-assigned reads against the modern TB reference genome called by MALT (NC_009565.1). In addition to TB DNA, our MALT results reveal an abundance of various mycobacterial species. These include members of the MTBC, as well as the *Mycobacterium avium* complex, the *Mycobacterium fortuitum* complex, and environmental mycobacteria such as *Mycobacterium chubuense* and *Mycobacterium intercellulare*. All samples show signatures of mycobacterial background likely originating from the depositional environment, storage conditions, and other contexts.

Subsequent mapping to the TB ancestor reference yielded between 1403 and 3432 putative MTBC DNA reads per sample library, constituting less than 0.1 % of the total. These yields are consistent with expectations of ancient endogenous content from shotgun datasets, as are the short average fragment lengths of approximately 43–57 base-pairs (Table 1). The number of MTBC-assigned molecules in the MALT analysis is consistently lower than the corresponding number of reads that map to the reference. This trend results from a genetic similarity of environmental mycobacterial species to the reference. MALT, therefore, offers higher specificity in read assignment than a mapping approach.

Analyses of damage patterns provided a means of authentication for both human and MTBC DNA. C to T changes at the terminal 5′ ends ranged from approximately 3 %–7 % for the mapped MTBC fragments (Table 1). Human DNA damage was notably higher, from 5 % to 23 %. The lower damage signal in the pathogen DNA could result from protection conferred by the thick, lipid-rich hydrophobic cell wall that could reduce hydrolytic DNA damage (Lindahl, 1993). Such a phenomenon has been proposed for the superior DNA preservation observed in other ancient mycobacteria (Schuenemann et al., 2011). However, it could equally result from modern DNA fragments that are similar enough to the MTBC reference that they map to it, and thus reduce the overall signal of ancient damage. In general, samples with higher amounts of human DNA damage also presented higher TB DNA damage. However, this was not consistent across all samples: HUA025, for example, had an estimated 10.9 % human DNA damage and only 3.5 % for MTBC DNA. Likewise, there was no regular agreement between endogenous human DNA recovery and the amount of MTBC DNA detected. This is best demonstrated by HUA041, which yielded the highest amount of human endogenous DNA at 11.9 %, though with only 0.05 % MTBC DNA. Similar percentages of MTBC DNA were reported for other samples, and these were consistently lower than the amounts reported for human DNA.

The two vertebrae that showed TB-like lesions, HUA002 and HUA004, each yielded 0.041 % MTBC DNA, with 1883 reads and 2547 reads mapping to the TB reference genome, respectively. Although these vertebrae display well-developed lesions consistent with tuberculosis, they did not yield more MTBC DNA than the other specimens that showed no diagnostic skeletal changes. Mapping to the human reference genome yielded 0.034 % endogenous human DNA for HUA002, which is on par with the amounts reported for MTBC DNA. However, HUA004 yielded 0.84 % endogenous human DNA, which exceeds the amounts in HUA002 by over an order of magnitude. These results indicate that increased human DNA preservation does not directly translate into increased MTBC preservation.

Estimation of the minimum number of individuals is challenging

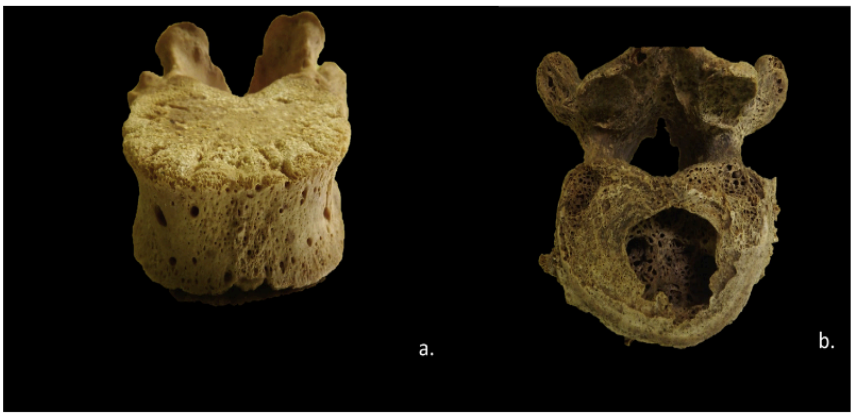


Fig. 4. Morphological variation displayed by the vertebrae recovered from Huari. Two thoracic vertebrae from the Monqachayoq-Solano sector of Huari present the variation shown across the 34 vertebrae selected for this study. The image on the left, (a) HUA041, displays slight reactive bone on the anterior aspect of the thoracic vertebral body. The image on the right, (b) HUA035, presents a large lytic focus on the inferior aspect of the thoracic vertebral body that is well developed and extends to the internal vertebral body, showing a smooth internal surface of the exposed trabeculae.

Table 1
Mapping statistics of sequencing data to the TB ancestor reference with archaeological context and MALT assigned reads to the MTBC.

Sample Name	Site/Sector	Archaeological Date	Raw Reads	Endogenous Human DNA %	Human DNA damage	Mapped Reads to TB	Endogenous TB DNA %	TB DNA damage	Average fragment length	% GC content	MALT Assigned Reads to MTB complex
HUA002	Cheqo Wasi	ca. 1000–1150 CE	10,125,640	0.034	20.1 %	1,883	0.041	6.2 %	43.60	62.32	104
HUA004	Cheqo Wasi	ca. 1000–1150 CE	13,565,300	0.843	23.5 %	2,547	0.041	7.4 %	45.22	61.39	174
HUA006	Vegachayoq Moqo	1275 – 1375 CE	12,484,864	3.325	9.1 %	3,257	0.051	3.7 %	55.98	61.37	620
HUA016	Vegachayoq Moqo	1275 – 1375 CE	12,318,062	0.339	8.2 %	2,242	0.045	7.9 %	53.33	58.62	287
HUA024	Vegachayoq Moqo	1275 – 1375 CE	12,603,366	0.312	10.8 %	1403	0.023	4.9 %	55.23	58.94	41
HUA025	Monqachayoq-Solano	1275 – 1375 CE	11,738,298	2.162	10.9 %	2,329	0.042	3.5 %	53.15	60.54	71
HUA037	Monqachayoq- Solano	1275 – 1375 CE	10,410,440	1.682	7.5 %	3432	0.072	4.6 %	57.98	61.87	1163
HUA041	Monqachayoq -Solano	1275 – 1375 CE	11,097,868	11.893	5.5 %	2,658	0.052	3.1 %	53.93	61.81	566

Sequencing was performed using an Illumina HiSeq 4000 with a paired end sequencing kit to a depth of approximately 5 million reads per sample. Raw reads indicates the number of genetic reads before processing. Human and TB endogenous DNA percentages shown are values given after quality filtering. The number of mapped reads to the TB ancestor is showing mapped reads after quality filtering. The percent damage DNA presented is the percent for the 1 st base of the 5′ end. The average fragment length of these samples is within expectation of ancient DNA and the GC content is near what is expected of MTBC (~65 %).

from commingled remains and was done for the set of vertebrae considered here based on age at death. From this assessment, the eight vertebrae identified as having detectable levels of MTBC DNA represent a minimum of five individuals: one from Cheqo Wasi (based on 2 vertebrae), two from Monqachayoq (based on three vertebrae) and two from Vegachayoq Moqo (based on three vertebrae).

6. Discussion

Paleopathological analysis often requires complete skeletons for confident diagnoses. The commingled context of the skeletal remains from the terminal and post-Wari populations necessitated the formulation of a broad differential diagnosis that included bacterial, parasitic, and fungal infections. Our study demonstrates that in such cases, molecular screening provides a robust alternative for identification of ancient infections.

The skeletal and molecular results presented here reveal the presence of MTBC in these Terminal Wari and post-Wari communities from a time period that encompasses imperial decline and the two centuries that followed. Of the 34 vertebrae analyzed, eight yielded detectable MTBC DNA through the application of our pathogen screening method. Together, our results provide yet another example of molecular survival in skeletal material in the Andes (Fehren-Schmitz et al., 2017; Llamas et al., 2016; Posth et al., 2018). Our results more than double the number of ancient Andean individuals for which MTBC infection has been demonstrated via next generation sequencing. These cases expand our understanding of the geographic and temporal presence of TB and further demonstrate its temporal correlation with increases in sublethal trauma relative to the height of Wari rule (Tung, 2014b). Together, these reveal a decline in population level health, which may have played a role in Wari collapse. The multiple lines of bioarchaeological data suggest that in the 200 years following Wari political decline, the post-Wari communities in the former heartland were experiencing significant increases in violence (Tung, 2008; Tung et al., 2016), profound changes in mortuary practices, new forms of violent body desecration (Tung, 2014a), and significant shifts in diets (Tung et al., 2016). This conforms to the general characterization of the LIP as a time of political turmoil plagued by internecine warfare and poor health (Arkush and Tung, 2013; Kurin et al., 2016; Covey, 2008). The presence of TB within the post-Wari communities in this study is consistent with the current understanding of factors influencing TB infection such as malnutrition (Macallan, 1999), stress (Prince et al., 2007), and political conflict (Barr and Menzies, 1994; Lönnroth et al., 2009). An established correlation exists between rises in TB incidence and political instability (Ali et al., 2012; Canetti et al., 2014; WHO, 2014), and products of political instability such as stress and malnutrition, specifically protein deficiency, are known to directly influence immune response to TB infections (Hood, 2013; Zachariah et al., 2002).

Importantly, our analyses reveal the presence of MTBC DNA in individual vertebrae that were absent of macroscopically visible pathological lesions. This further affirms the utility of genetic screening as applied to the study of diseases in the archaeological record that occur without skeletal involvement (Vågene et al., 2018). In fact, of the eight vertebrae molecularly identified as putatively positive for TB, only two, HUA002 and HUA004, display pathognomonic osteological changes. While HUA041 does show slight reactive bone on the anterior aspect of the vertebral body (Figs. 4a and 6), others display hypervascularization that is either pronounced (HUA037), or is slight enough to be considered within normal variation (HUA006, HUA016, HUA024, HUA025). Of note, hypervascularization was also observed in several vertebrae in our skeletal series for which no TB DNA was detected. Further to this, we report a lack of detectable MTBC DNA in vertebrae that did show pathology that is characteristic of TB infection (e.g. HUA035, Fig. 4b). While an alternative diagnosis could be considered for these individual cases, biases in DNA preservation will also have an influence on the recovery of TB DNA.

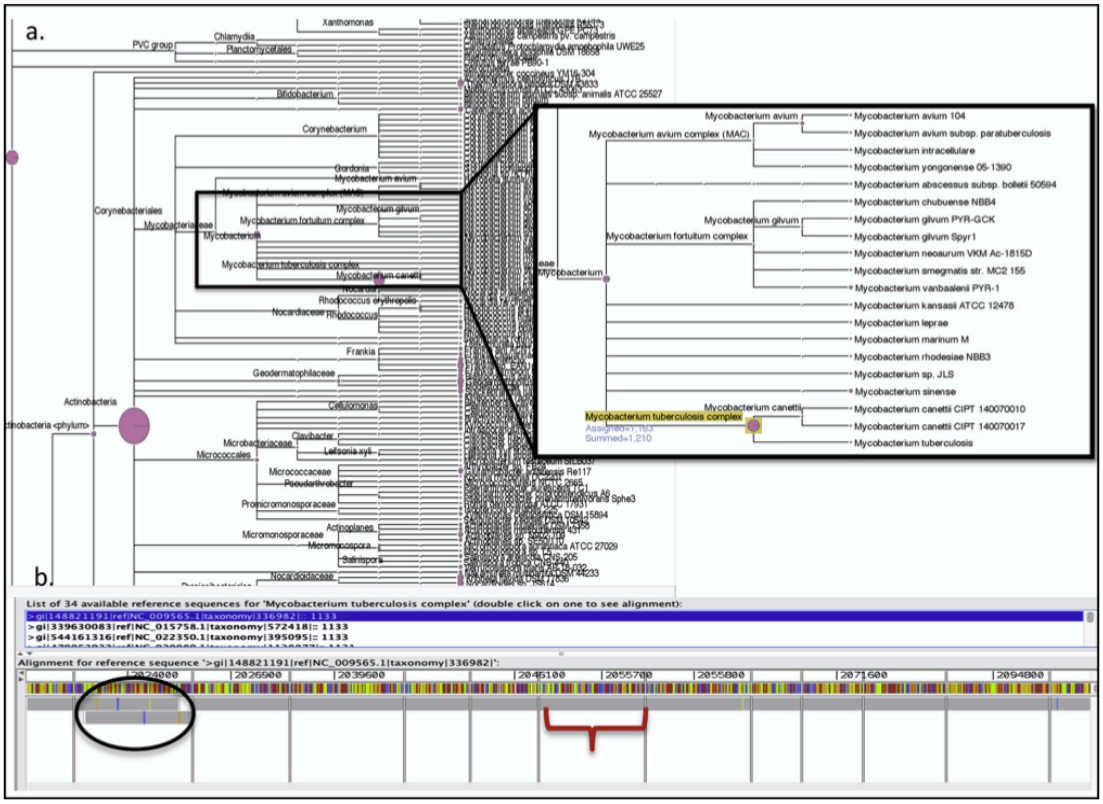


Fig. 5. Metagenomic screening for MTBC DNA using MALT. a) MALT output of sample HUA037, visualized in MEGAN. The taxonomic tree provides various organisms to which genetic reads of our sample are assigned. The inset box provides a zoomed in view of the results, which show 1,163 reads assigning to the MTBC. b) Using the alignment viewer we were able to assess the quality of the alignment to a reference genome. The first three of the 34 reference genomes to which sequences from our sample align are listed in the top box with their accession numbers, and the selected reference is highlighted in blue. The nucleotide sequence of the reference is displayed below the nucleotide coordniate positions, where colors denote different bases. Genetic reads from our sample that align to the reference are presented below it. Light grey indicates all genetic positions that match the reference, and the bracket shows an example of a DNA read that is a full match to the reference. Colored sections within the light grey indicate mismatches in the nucleotide sequence of our sample compared to the selected reference. A region of the sequence with several mismatches is indicated by the oval. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Fig. 6. The eight vertebrae in which MTBC DNA was detected by using MALT for pathogen screening of non-enriched DNA sequence data. Only two of the vertebrae (HUA002 and HUA004) showed characteristic skeletal lesions suggestive of tuberculosis; all other vertebrae display only slight non-specific changes of either reactive bone (HUA041) or increased vascularity in comparison to the rest of the population of these three sectors; however, this may also be considered within normal variation.

Our metagenomic approach allowed us to investigate the community of microbes that colonized the skeletal material, with a special interest in identifying ancient MTBC DNA. This analysis revealed the presence of environmental taxa that are genetically related to the MTBC, but are not considered pathogenic to humans. This is expected of tissues interred in soil, where mycobacterial content is high. Despite this dominant signal, MTBC DNA could still be confidently detected in eight samples. Although the presence of mycobacterial background can pose challenges in the identification of genuine MTBC DNA, investigation of genetic similarity to a reference genome can assist in authentication and in prioritizing samples to be investigated at the genomic level. Greater chance of genomic recovery is likely correlated with strong signals of molecular preservation based on mapping to a reference, and fewer genetically related bacterial species in the background (Warinner et al., 2017; Key et al., 2017). Unexpectedly, human DNA content did not seem to be a predictor of MTBC preservation. Likewise, MTBC DNA recovery was not positively correlated with lesion progression based on our morphological assessment, which suggests a bias in molecular preservation.

This case study demonstrates the application of recommended protocols for TB screening of shotgun data from archaeological tissues. Using these methods we were able to detect traces of MTBC DNA; however, these data do not currently offer the resolution needed to establish their relationship to previously published strains, and thus we cannot determine if the TB varieties circulating in the Huari populations are related to the *M. pinnipedii*-like strain previously identified from the pre-contact sites of coastal Peru (Bos et al., 2014). By employing targeted capture of TB DNA in future, genomic reconstruction may eventually better characterize the strain(s) we have detected from Huari. Additionally, such analyses would provide the opportunity to evaluate evolutionary changes occurring within the strains and to detect any existing signals of adaptation to either host or environment. We consider all eight samples as promising candidates for subsequent genomic capture.

Genetic analyses of ancient pathogens can benefit greatly from the extensive growing databases that catalogue genome-level diversity in medically relevant microbes. The genetic databases available in MTBC research are no exception, and are heavily dominated by the most globally represented TB lineages L2 and L4, which achieved their current distribution in the relatively recent past (Liu et al., 2018; Stucki et al., 2016). Further exploration of diversity in the animal-adapted lineages of the MTBC would be helpful for detections of zoonotic diseases in archaeological material using detection methods such as MALT. Furthermore, ancient infectious diseases that may be extinct or are not currently genomically characterized will not be detected in this pathogen screening strategy. Likewise, the absence of detectable pathogen DNA does not necessarily indicate absence of infection: it could be that DNA was not recovered, did not preserve, or remains undetected via the techniques employed here.

7. Conclusions

In 1997, John Verano's seminal work *Advances in the Paleopathology of Andean South America* discussed the foundations of paleopathology and highlighted the advances in this increasingly multidisciplinary field. In the 20 years following this publication, Andean paleopathology has continued to play a central role in the advancement of paleopathology method and theory. TB has frequently been at the forefront of these advancements. While skeletal and soft tissue studies have contributed valuable information on the identification of tuberculosis in the Andean region and beyond, DNA analyses have increased the potential of this detection. High-resolution screening, such as that demonstrated here, is instrumental for identifying, and potentially prioritizing, candidate samples for hybridization capture for genomic evaluation. For the limited number of ancient TB cases investigated thus far at the molecular level, capture has been a necessary step for the reconstruction of ancient TB genomes. By employing our pathogen screening method we have been able to detect candidates for targeted MTBC capture that could ostensibly produce genomic data.

The detection of TB in the Terminal Wari and post-Wari populations from Huari provide additional insight into the geographic distribution of TB in the pre-contact Americas, which may later inform on hypotheses of its origin. This preliminary research also adds to our understanding of the Terminal Wari and post-Wari periods and sets the foundation for further research on the geographic reach and evolution of tuberculosis in the LIP of the Andes. More specifically, it highlights the synergism between disease susceptibility and climatic and dietary shifts (Wilbur et al., 2008). Though the contributions of molecular detection of pathogens to paleopathology have been substantial, we highlight its use as a tool to characterize disease in the past. While molecular paleopathology can assist in identification of infections, its combination with morphology based assessments, historical documents, and contextual evidence can produce rigorous, well developed, confident diagnoses and holistic reconstructions of past population health and disease (Klaus et al., 2010; Wilbur et al., 2009). Integration of data from various angles will facilitate interpretations of past disease experience within a biosocial framework, which remains our ultimate goal.

Acknowledgements

Many thanks to the Ministerio de Cultura in Ayacucho, Peru, and to the Universidad Nacional de San Cristóbal de Huamanga for approving requests to examine the skeletal collections. Thank you to the many participants of the Ayacucho Bioarchaeology Project who have contributed to the research of these post-Wari communities. We thank the National Science Foundation-Archaeology and Biological Anthropology Divisions (grant #1420757) for financial support of the bioarchaeological aspects of the project. Genetic work was funded by the Max Planck Society and European Research Council Starting Grant CoDisEASe (grant number 805268 to KIB). We thank Dr. Anne Grauer and two anonymous reviewers for their insightful comments towards the improvement of this paper. A very special thanks to the organizers and editors of this special volume, Dr. J. Marla Toyne, Dr. Melissa Murphy, and Dr. Haagen Klaus. We would also like to Dr. John Verano for his constant contribution to the field of biological anthropology and for his inspiration leading to this special volume.

Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at doi:https://doi.org/10.1016/j.ijpp.2019.12.006.

References

Ali, A., Nisar, M., Idrees, M., Ahmad, H., Hussain, A., Rafique, S., Sabri, S., ur Rehman, H., Ali, L., Wazir, S., Khan, T., 2012. Prevalence of HBV infection in suspected population of conflict-affected area of war against terrorism in North Waziristan FATA Pakistan. Infect. Genet. Evol. 12 (8), 1865–1869.

Allison, M.J., Gerszten, E., Munizaga, J., Santoro, C., Mendoza, D., 1981. Tuberculosis in pre-Columbian Andean populations (chapter 4). In: Buikstra, J.E. (Ed.), Prehistoric Tuberculosis in the Americas. Northwestern University Archaeological Program, Evanston, IL, pp. 49–61.

Allison, M.J., Mendoza, D., Pezzia, A., 1973. Documentation of a case of tuberculosis in pre-Columbian America. Am. Rev. Respir. Dis. 107 (6), 985–991.

Arkush, E., Tung, T.A., 2013. Patterns of war in the Andes from the Archaic to the late horizon: insights from settlement patterns and cranial trauma. J. Archaeol. Res. 21 (4), 307–369.

Arriaza, B.T., 1995. Chinchorro bioarchaeology: chronology and mummy seriation. Lat. Am. Antiq. 6 (1), 35–55.

Arriaza, B.T., Salo, W., Aufderheide, A.C., Holcomb, T.A., 1995. Pre-Columbian tuberculosis in Northern Chile: molecular and skeletal evidence. Am. J. Phys. Anthropol. 98 (1), 37–45.

Arrieta, M.A., Bordach, M.A., Mendonça, O.J., 2014. Pre-columbian tuberculosis in Northwest Argentina: skeletal evidence from rincón chico 21 cemetery. Int. J. Osteoarchaeol. 24 (1), 1–14.

Aufderheide, A.C., Salo, W., Madden, M., Streitz, J., Buikstra, J., Guhl, F., Arriaza, B.,

Renier, C., Wittmers, L.E., Fornaciari, G., Allison, M., 2004. A 9,000-year record of Chagas' disease. Proc. Natl. Acad. Sci. 101 (7), 2034–2039.
Baker, B.J., Armelagos, G.J., Becker, M.J., Brothwell, D., Drusini, A., Geise, M.C., Kelley, M.A., Moritoto, I., Morris, A.G., Nurse, G.T., Powell, M.L., Rothschild, B.M., Saunders, S.R., 1988. The origin and antiquity of syphilis: paleopathological diagnosis and interpretation. Curr. Anthropol. 29 (5), 703–737.
Barr, R.G., Menzies, R., 1994. The effect of war on tuberculosis: results of a tuberculin survey among displaced persons in El Salvador and a review of the literature. Tuber. Lung Dis. 75 (4), 251–259.
Bird, B.W., Abbott, M.B., Vuille, M., Rodbell, D.T., Stansell, N.D., Rosenmeier, M.F., 2011. A 2,300-year-long annually resolved record of the South American summer monsoon from the Peruvian Andes. Proc. Natl. Acad. Sci. 108 (21), 8583–8588.
Blom, D.E., Buikstra, J.E., Keng, L., Tomczak, P.D., Shoreman, E., Stevens-Tuttle, D., 2005. Anemia and childhood mortality: Latitudinal patterning along the coast of pre-Columbian Peru. Am. J. Phys. Anthropol. 127 (2), 152–169.
Bos, K.I., Harkins, K.M., Herbig, A., Coscolla, M., Weber, N., Comas, I., Forrest, S.A., Bryant, J.M., Harris, S.R., Schuenemann, V.J., Campbell, T.J., Majander, K., Wilbur, A.K., Guichon, R.A., Wolfe Steadman, D.L., Cook, D.C., Niemann, S., Behr, M.A., Zumarraga, M., Bastida, R., Huson, D., Nieselt, K., Young, D., Parkhill, J., Buikstra, J.E., Gagneux, S., Stone, A.C., Krause, J., 2014. Pre-Columbian mycobacterial genomes reveal seals as a source of New World human tuberculosis. Nature 514 (7523), 494–497.
Bos, K.I., Jäger, G., Schuenemann, V.J., Vågene, Å., Spyrou, M.A., Herbig, A., Nieselt, K., Krause, J., 2015. Parallel detection of ancient pathogens via array-based DNA capture. Philos. Trans. Biol. Sci. 370 (1660), 20130375.
Bos, K.I., Kühnert, D., Herbig, A., Esquivel-Gomez, L.R., Valtueña, A.A., Barquera, R., Giffin, K., Kumar Lankapalli, A., Nelson, E.A., Sabin, S., Spyrou, M.A., 2019. Paleomicrobiology: diagnosis and evolution of ancient pathogens. Annu. Rev. Microbiol. 73, 639–666.
Briggs, A.W., Stenzel, U., Meyer, M., Krause, J., Kircher, M., Pääbo, S., 2009. Removal of deaminated cytosines and detection of in vivo methylation in ancient DNA. Nucleic Acids Res. 38 (6), e87.
Brosch, R., Gordon, S.V., Marmiesse, M., Brodin, P., Buchrieser, C., Eiglmeier, K., Garnier, T., Gutierrez, C., Hewinson, G., Kremer, K., Parsons, L.M., 2002. A new evolutionary scenario for the *Mycobacterium tuberculosis* complex. Proc. Natl. Acad. Sci. 99 (6), 3684–3689.
Buikstra, J.E., 1976. The Caribou Eskimo: general and specific disease. Am. J. Phys. Anthropol. 45 (3), 351–367.
Buikstra, J.E. (Ed.), 1981. Prehistoric Tuberculosis in the Americas (No. 5). Northwestern University Archeological Program.
Buikstra, J.E., 1999. Paleoepidemiology of Tuberculosis in the Americas. Tuberculosis: Past and Present. Golden Book Publisher Ltd, Szeged, Hungary, pp. 479–494.
Buikstra, J.E., Williams, S., 1991. Tuberculosis in the americas: current perspectives. In: Ortner, D., Aufderheide, A. (Eds.), Human Paleopathology, Current Syntheses and Future Options, pp. 161–172.
Burgess, S.D., 1992. Chiribayan Skeletal Pathology on the South Coast of Peru: Patterns of Production and Consumption. Ph.D. dissertation. University of Chicago, Chicago, Illinois.
Canetti, D., Russ, E., Luborsky, J., Gerhart, J.I., Hobfoll, S.E., 2014. Inflamed by the flames? The impact of terrorism and war on immunity. J. Trauma. Stress 27 (3), 345–352.
Cieślik, A.I., 2017. Evidence of tuberculosis among children in medieval (13th–15th century) Wrocław: a case study of hip joint tuberculosis in a juvenile skeleton excavated from the crypt of the St. Elizabeth church. Anthropol. Rev. 80 (2), 219–231.
Cook, A.G., Glowacki, M., 2003. Pots, Politics, and Power. In the Archaeology and Politics of Food and Feasting in Early States and Empires. Springer, Boston, MA, pp. 173–202.
Comas, I., Coscolla, M., Luo, T., Borrell, S., Holt, K.E., Kato-Maeda, M., Parkhill, J., Malla, B., Berg, S., Thwaites, G., Yeboah-Manu, D., 2013. Out-of-Africa migration and Neolithic coexpansion of Mycobacterium tuberculosis with modern humans. Nat. Genet. 45 (10), 1176.
Correal, G., Flórez, I., 1992. Estudio de las momias guanes de la Mesa de los Santos, Santander, Colombia. Rev Acad Col Cienc Exac Fis y Nat 18, 283–290.
Covey, R.A., 2008. Multiregional perspectives on the archaeology of the Andes during the Late Intermediate Period (c. AD 1000–1400). J. Archaeol. Res. 16 (3), 287–338.
De Ayala, F.G.P., 1980. Nueva Corónica Y Buen Gobierno (Vol. 2). Fundacion Biblioteca Ayacucho.
DeFrance, S.D., 2009. Zooarchaeology in complex societies: political economy, status, and ideology. J. Archaeol. Res. 17, 105–168.
Devault, A.M., McLoughlin, K., Jaing, C., Gardner, S., Porter, T.M., Enk, J.M., Thissen, J., Allen, J., Borucki, M., DeWitte, S.N., Dhody, A.N., 2014. Ancient pathogen DNA in archaeological samples detected with a Microbial Detection Array. Sci. Rep. 4, 4245.
DeWitte, S.N., 2016. The anthropology of plague: insights from bioarcheological analyses of epidemic cemeteries. Mediev. Globe 1 (1), 6.
Drancourt, M., Aboudharam, G., Signoli, M., Dutour, O., Raoult, D., 1998. Detection of 400-year-old Yersinia pestis DNA in human dental pulp: an approach to the diagnosis of ancient septicemia. Proc. Natl. Acad. Sci. 95 (21), 12637–12640.
Eisenach, K.D., Sifford, M.D., Cave, M.D., Bates, J.H., Crawford, J.T., 1991. Detection of Mycobacterium tuberculosis in sputum samples using a polymerase chain reaction. Am. Rev. Respir. Dis. 144, 1160–1163.
El-Najjar, M., Al-Shiyab, A., Al-Sarie, I., 1996. Cases of tuberculosis at'Ain Ghazal, Jordan. Palorient 123–128.
Fehren-Schmitz, L., Harkins, K.M., Llamas, B., 2017. A paleogenetic perspective on the early population history of the high altitude Andes. Quat. Int. 461, 25–33.
Friedrich, K.M., Nemec, S., Czerny, C., Fischer, H., Plischke, S., Gahleitner, A., Viola, T.B., Imhof, H., Seidler, H., Guillen, S., 2010. The story of 12 Chachapoyan mummies through multidetector computed tomography. Eur. J. Radiol. 76 (2), 143–150.
Gagneux, S., DeRiemer, K., Van, T., Kato-Maeda, M., De Jong, B.C., Narayanan, S., Nicol, M., Niemann, S., Kremer, K., Gutierrez, M.C., Hilty, M., 2006. Variable host–pathogen compatibility in Mycobacterium tuberculosis. Proc. Natl. Acad. Sci. 103 (8), 2869–2873.
García-Frías, J.E., 1940. La tuberculosis en los antiguos peruanos. Actualidad Médica Peruana 5 (10), 274–291.
Green, R.E., Krause, J., Briggs, A.W., Maricic, T., Stenzel, U., Kircher, M., Patterson, N., Li, H., Zhai, W., Fritz, M.H.Y., Hansen, N.F., 2010. A draft sequence of the Neandertal genome. Science 328 (5979), 710–722.
Guillén, S.E., 2004. Artificial mummies from the Andes. Coll. Antropol. 28 (2), 141–157.
Guillén, S.E., 2012. A history of paleopathology in Peru and Northern Chile: from head hunting to head counting. In: Buikstra, J.E., Roberts, C. (Eds.), The Global History of Paleopathology: Pioneers and Prospects. Oxford University Press, New York, pp. 312–328.
Harkins, K.M., Buikstra, J.E., Campbell, T., Bos, K.I., Johnson, E.D., Krause, J., Stone, A.C., 2015. Screening ancient tuberculosis with qPCR: challenges and opportunities. Philos. Trans. Biol. Sci. 370 (1660), 20130622.
Harkins, K.M., Stone, A.C., 2015. Ancient pathogen genomics: insights into timing and adaptation. J. Hum. Evol. 79, 137–149.
Heggarty, P., Beresford-Jones, D., 2012. Archaeology and Language in the Andes. PhilPapers, British Academy.
Herbig, A., Maixner, F., Bos, K.I., Zink, A., Krause, J., Huson, D.H., 2016. MALT: fast alignment and analysis of metagenomic DNA sequence data applied to the Tyrolean Iceman. BioRxiv, 050559.
Hershkovitz, I., Donoghue, H.D., Minnikin, D.E., May, H., Lee, O.Y.C., Feldman, M., Galili, E., Spigelman, M., Rothschild, B.M., Bar-Gal, G.K., 2015. Tuberculosis origin: the Neolithic scenario. Tuberculosis 95, S122–S126.
Hood, M.L.H., 2013. A narrative review of recent progress in understanding the relationship between tuberculosis and protein energy malnutrition. Eur. J. Clin. Nutr. 67 (11), 1122.
Hoshino, A., Hanada, S., Yamada, H., Mii, S., Takahashi, M., Mitarai, S., Yamamoto, K., Manome, Y., 2014. Mycobacterium tuberculosis escapes from the phagosomes of infected human osteoclasts reprograms osteoclast development via dysregulation of cytokines and chemokines. Pathog. Dis. 70 (1), 28–39.
Hoyle, R.L., Uhle, M., Kroeber, A.L., 1945. Los mochicas (pre-chimu, de Uhle y early chimu, de Kroeber). Sociedad Geográfica Americana.
Hrdlička, A., 1909. Tuberculosis Among Certain Indian Tribes of the United States (No. 42). U.S. Government Printing Office, Washington D.C.
Hrdlička, A., 1914. Anthropological Work in Peru, in 1913: With Notes on the Pathology of the Ancient Peruvians, With Twenty-six Plates (Vol. 61, No. 18). Smithsonian Institution.
Huson, D.H., Beier, S., Flade, I., Górska, A., El-Hadidi, M., Mitra, S., Ruscheweyh, H.J., Tappu, R., 2016. MEGAN community edition-interactive exploration and analysis of large-scale microbiome sequencing data. PLoS Comput. Biol. 12 (6), e1004957.
Isbell, W.H., Schreiber, K.J., 1978. Was Huari a state? Am. Antiq. 43 (3), 372–389.
Kastert, J., Uehlinger, E., 1964. Skelettuberkulose: Mit einen Beitragtiber Allgemeine Pathologie und Pathologische Anatomie der Skelettuberkulose. In: In: Hein, J., Kleinschmidt, H., Uehlinger, E. (Eds.), Handbuch der Tuberkulose 4. Thieme, Stuttgart, pp. 443–532.
Kay, G.L., Sergeant, M.J., Zhou, Z., Chan, J.Z.M., Millard, A., Quick, J., Szikossy, I., Pap, I., Spigelman, M., Loman, N.J., Achtman, M., 2015. Eighteenth-century genomes show that mixed infections were common at time of peak tuberculosis in Europe. Nat. Commun. 6, 6717.
Key, F.M., Posth, C., Krause, J., Herbig, A., Bos, K.I., 2017. Mining metagenomic data sets for ancient DNA: recommended protocols for authentication. Trends Genet. 33 (8), 508–520.
Kim, Y.S., Lee, I.S., Oh, C.S., Kim, M.J., Cha, S.C., Shin, D.H., 2016. Calcified pulmonary nodules identified in a 350-year-old-joseon mummy: the first report on ancient pulmonary tuberculosis from archaeologically obtained pre-modern Korean samples. J. Korean Med. Sci. 31 (1), 147–151.
Klaus, H.D., Wilbur, A.K., Temple, D.H., Buikstra, J.E., Stone, A.C., Fernandez, M., Wester, C., Tam, M.E., 2010. Tuberculosis on the north coast of Peru: skeletal and molecular paleopathology of late pre-Hispanic and postcontact mycobacterial disease. J. Archaeol. Sci. 37 (10), 2587–2597.
Klaus, H.D., 2014. Frontiers in the bioarchaeology of stress and disease: cross-disciplinary perspectives from pathophysiology, human biology, and epidemiology. Am. J. Phys. Anthropol. 155 (2), 294–308.
Konomi, N., Lebwohl, E., Mowbray, K., Tattersall, I., Zhang, D., 2002. Detection of mycobacterial DNA in Andean mummies. J. Clin. Microbiol. 40 (12), 4738–4740.
Krause, J., Fu, Q., Good, J.M., Viola, B., Shunkov, M.V., Derevianko, A.P., Pääbo, S., 2010. The complete mitochondrial DNA genome of an unknown hominin from southern Siberia. Nature 464 (7290), 894.
Kurin, D.S., Lofaro, E.M., Gómez Choque, D.E., Krigbaum, J., 2016. A bioarchaeological and biogeochemical study of warfare and mobility in Andahuaylas, Peru (ca. AD 1160–1260). Int. J. Osteoarchaeol. 26 (1), 93–103.
Lindahl, T., 1993. Instability and decay of the primary structure of DNA. Nature 362 (6422), 709–715.
Liu, Q., Ma, A., Wei, L., Pang, Y., Wu, B., et al., 2018. China's tuberculosis epidemic stems from historical expansion of four strains of Mycobacterium tuberculosis. Nature Ecol. Evol. 1 (12), 1982–1992. https://doi.org/10.1038/s41559-018-0680-6.
Llamas, B., Fehren-Schmitz, L., Valverde, G., Soubrier, J., Mallick, S., Rohland, N., Nordenfelt, S., Valdiosera, C., Richards, S.M., Rohrlach, A., Romero, M.I.B., 2016. Ancient mitochondrial DNA provides high-resolution time scale of the peopling of the Americas. Sci. Adv. 2 (4), e1501385.
Lombardi, G.P., García Cáceres, U., 2000. Multisystemic tuberculosis in a pre-columbian Peruvian mummy: four diagnostic levels, and a paleoepidemiological hypothesis.

ChungarÃ¡ 32 (1), 55–60.
Lönnroth, K., Jaramillo, E., Williams, B.G., Dye, C., Raviglione, M., 2009. Drivers of tuberculosis epidemics: the role of risk factors and social determinants. Soc. Sci. Med. 68 (12), 2240–2246.
Lozada, M.C., Tantaleán, H. (Eds.), 2019. Andean Ontologies: New Archaeological Perspectives. University Press of Florida.
Lunardini, A., Costantini, L., Biasini, L.C., Caramella, D., Fornaciari, G., 2012. Evidence of congenital syphilis and tuberculosis in a XIX century mummy (Perugia, Italy). J. Biol. Res. 85 (1), 241.
Macallan, D.C., 1999. Malnutrition in tuberculosis. Diagn. Microbiol. Infect. Dis. 34 (2), 153–157.
Maczel, M., 2003. On the Traces of Tuberculosis. Diagnostic Criteria of Tuberculous Affection of the Human Skeleton and Their Application in Hungarian and French Anthropological Series. University of Szeged.
Margerison, B.J., Knüsel, C.J., 2002. Paleodemographic comparison of a catastrophic and an attritional death assemblage. Am. J. Phys. Anthropol. 119 (2), 134–143.
Martinez de Arateco Hoyo, R., 1999. Paleoepidemiologia Y Salud Publica De La Tuberculosis En Colombia: Trabajo De Ingreso Como Miembro Activo a La Sociedad Colombiana.
Mays, S., Fysh, E., Taylor, G.M., 2002. Investigation of the link between visceral surface rib lesions and tuberculosis in a medieval skeletal series from England using ancient DNA. Am. J. Phys. Anthropol. 119 (1), 27–36.
Mays, S., Taylor, G.M., Legge, A.J., Young, D.B., Turner-Walker, G., 2001. Paleopathological and biomolecular study of tuberculosis in a medieval skeletal collection from England. Am. J. Phys. Anthropol. 114 (4), 298–311.
Moodie, R.L., 1923. Paleopathology: an Introduction to the Study of Ancient Evidences of Disease. University of Illinois Press, Urbana.
Morse, D., 1961. Prehistoric tuberculosis in America. Am. Rev. Respir. Dis. 83 (4), 489–504.
Mousa, H.A., 2007. 'Bones and joints tuberculosis. Bahrain Med Bulletin 29, 1–9.
Mühlemann, B., Jones, T.C., de Barros Damgaard, P., Allentoft, M.E., Shevnina, I., Logvin, A., Usmoanova, E., Panyushkina, I.P., Boldgiv, B., Bazartseren, T., Tashbaeva, K., Merz, V., Lau, N., Smrčka, V., Voyakin, D., Kitov, E., Epimakhov, A., Pokutta, D., Vicze, M., Price, D., Moiseyev, V., Hansen, A.J., Orlando, L., Rasmussen, L., Sikora, M., Vinner, L., Osterhaus, A.D.M.E., Smith, D.J., Glebe, D., Fouchier, R.A.M., Drosten, Sjörgen, K., Kristiansen, K., Willerslev, E., 2018. Ancient hepatitis B viruses from the Bronze Age to the medieval period. Nature 557 (7705), 418.
Müller, R., Roberts, C.A., Brown, T.A., 2015. Complications in the study of ancient tuberculosis: non-specificity of IS6110 PCRs. Star Sci. Technol. Archaeol. Res. 1 (1), 1–8.
Owen, B.D., 1993. A Model of Multiethnicity: State Collapse, Competition, and Social Complexity From Tiwanaku to Chiribaya in the Osmore Valley, Peru. Doctoral dissertation. University of California, Los Angeles.
Pääbo, S., 1989. Ancient DNA: extraction, characterization, molecular cloning, and enzymatic amplification. Proceedings of the National Academy of Sciences 86 (6), 1939–1943.
Pálfi, G., 2012. Juvenile cases of skeletal tuberculosis from the Terry Anatomical collection (Smithsonian Institution, Washington, DC, USA). Acta Biologica Szegediensis 56 (1), 1–12.
Polley, P., Dunn, R., 2009. Noncontiguous spinal tuberculosis: incidence and management. Eur. Spine J. 18 (8), 1096–1101.
Posth, C., Nakatsuka, N., Lazaridis, I., Skoglund, P., Mallick, S., Lamnidis, T.C., et al., 2018. Reconstructing the deep population history of central and south America. Cell 175 (5), 1185–1197. https://doi.org/10.1016/j.cell.2018.10.027. PMID: 30415837.
Prince, M., Patel, V., Saxena, S., Maj, M., Maselko, J., Phillips, M.R., Rahman, A., 2007. No health without mental health. Lancet 370 (9590), 859–877.
Prüfer, K., Racimo, F., Patterson, N., Jay, F., Sankararaman, S., Sawyer, S., Heinze, A., Renaud, G., Sudmant, P.H., De Filippo, C., Li, H., 2014. The complete genome sequence of a Neanderthal from the Altai Mountains. Nature 505 (7481), 43.
Rasmussen, M., Li, Y., Lindgreen, S., Pedersen, J.S., Albrechtsen, A., Moltke, I., Metspalu, M., Metspalu, E., Kivisild, T., Gupta, R., Bertalan, M., 2010. Ancient human genome sequence of an extinct Palaeo-Eskimo. Nature 463 (7282), 757.
Reitz, E.J., Sandweiss, D.H., 2001. Environmental change at Ostra Base Camp, a Peruvian pre-ceramic site. J. Archaeol. Sci. 28 (10), 1085–1100.
Requena, A., 1945. Evidencia de tuberculosis en la América pre-Columbia. Acta Venezolana 1–20.
Resnick, D., Niwayama, G., 1995. Osteomyelitis, septic arthritis, and soft tissue infection: organisms. In: Resnick, D. (Ed.), Diagnosis of Bone and Joint Disorders, 3rd ed. Saunders, Philadelphia, pp. 2448–2558.
Rivas, A.M.B., 1988. Las patologías óseas en la población de Marín. Boletín de Arqueología de la Fian 3, 3–24.
Roberts, C.A., Buikstra, J.E., 2003. The Bioarchaeology of Tuberculosis: a Global Perspective on a Re-emerging Disease. University Press of Florida, Gainesville, FL.
Roberts, C.A., Buikstra, J.E., 2019. Bacterial infections. In: Buikstra, J.E. (Ed.), Ortner's Identification of Pathological Conditions in Human Skeletal Remains. Academic Press, pp. 321–439.
Rodriguez, J.V., 1988. Acera de la supusesta debilidad mental y fúscia de los músicas posible causa de su conquista y posterior extinction. Arqueologia 1, 42–46.
Romero Arateco, W.M., 1998. Mal de Pott en momia de la colección del museo arqueológico Marqué de San Jorge. Maguare 99–115.
Rothschild, B.M., Martin, D., 2003. Frequency of pathology in a large natural sample from Natural Trap Cave with special remarks on erosive disease in the Pleistocene. Reumatismo 55, 58–65.
Sabin, S., Herbig, A., Vågene, Å.J., Ahlström, T., Bozovic, G., Arcini, C., Kühnert, D., Bos, K.I., 2019. A seventeenth-century *Mycobacterium tuberculosis* genome supports a Neolithic emergence of the *Mycobacterium tuberculosis* complex. BioRxiv, 588277.
Salo, W.L., Aufderheide, A.C., Buikstra, J., Holcomb, T.A., 1994. Identification of Mycobacterium tuberculosis DNA in a pre-Columbian Peruvian mummy. Proc. Natl. Acad. Sci. 91 (6), 2091–2094.
Schreiber, K.J., 1992. Wari Imperialism in Middle Horizon Peru. Museum of Anthropology University of Michigan, Ann Arbor.
Schuenemann, V.J., Bos, K., DeWitte, S., Schmedes, S., Jamieson, J., Mittnik, A., Forrest, S., Coombes, B.K., Wood, J.W., Earn, D.J., White, W., 2011. Targeted enrichment of ancient pathogens yielding the pPCP1 plasmid of *Yersinia pestis* from victims of the Black Death. Proc. Natl. Acad. Sci. 108 (38), E746–E752.
Schuenemann, V.J., Lankapalli, A.K., Barquera, R., Nelson, E.A., Hernández, D.I., Alonzo, V.A., Bos, K.I., Morfín, L.M., Herbig, A., Krause, J., 2018. Historic Treponema pallidum genomes from Colonial Mexico retrieved from archaeological remains. PLoS Negl. Trop. Dis. 12 (6), e0006447.
Signoli, M., Séguy, I., Biraben, J.N., Dutour, O., Belle, P., 2002. Paleodemography and historical demography in the context of an epidemic. Population 57 (6), 829–854.
Skoglund, P., Malmström, H., Raghavan, M., Storå, J., Hall, P., Willerslev, E., Gilbert, M.T., Götherström, A., Jakobsson, M., 2012. Origins and genetic legacy of Neolithic farmers and hunter-gatherers in Europe. Science 336 (6080), 466–469.
Spigelman, M., Lemma, E., 1993. The use of the polymerase chain reaction (PCR) to detect Mycobacterium tuberculosis in ancient skeletons. Int. J. Osteoarchaeol. 3 (2), 137–143.
Spyrou, M.A., Tukhbatova, R.I., Wang, C.C., Valtueña, A.A., Lankapalli, A.K., Kondrashin, V.V., Tsybin, V.A., Khokhlov, A., Kühnert, D., Herbig, A., Bos, K.I., 2018. Analysis of 3800-year-old Yersinia pestis genomes suggests Bronze Age origin for bubonic plague. Nat. Commun. 9 (1), 2234.
Spyrou, M.A., Bos, K.I., Herbig, A., Krause, J., 2019. Ancient pathogen genomics as an emerging tool for infectious disease research. Nat. Rev. Genet. 20, 323–340.
Stead, W.W., Eisenach, K.D., Cave, M.D., Beggs, M.L., Templeton, G.L., Thoen, C.O., Bates, J.H., 1995. When did Mycobacterium tuberculosis infection first occur in the New World? An important question with public health implications. Am. J. Respir. Crit. Care Med. 151 (4), 1267–1268.
Stone, A.C., Ozga, A.T., 2019. Ancient DNA in the study of ancient disease. In: Buikstra, J.E. (Ed.), Ortner's Identification of Pathological Conditions in Human Skeletal Remains. Academic Press, pp. 183–210.
Stucki, D., Brites, D., Jeljeli, L., Coscolla, M., Liu, Q., Trauner, A., Fenner, L., Rutaihwa, L., Borrell, S., Luo, T., Gao, Q., 2016. Mycobacterium tuberculosis lineage 4 comprises globally distributed and geographically restricted sublineages. Nat. Genet. 48 (12), 1535–1543.
Subasi, M., Bukte, Y., Kapukaya, A., Gurkan, F., 2004. Tuberculosis of the metacarpals and phalanges of the hand. Ann. Plast. Surg. 53 (5), 469–472.
Tello, J.C., 1909. La Antigüedad De La Sífilis En El Perú. Sanmartí, Lima.
Tello, J.C., Williams, H.U., 1930. Ancient syphilitic skull from Paracas in Peru. Ann. Med. Hist. 2, 515–529.
Thompson, L.G., Mosley-Thompson, E., Davis, M.E., Zagorodnov, V.S., Howat, I.M., Mikhalenko, V.N., Lin, P.N., 2013. Annually resolved ice core records of tropical climate variability over the past ~1800 years. Science 340 (6135), 945–950.
Tung, T.A., 2008. Violence after imperial collapse: a study of cranial trauma among Late Intermediate period burials from the former Huari capital, Ayacucho, Peru. Nawpa Pacha 29 (1), 101–117.
Tung, T.A., 2012. Violence, Ritual, and the Wari Empire: a Social Bioarchaeology of Imperialism in the Ancient Andes. University Press of Florida, Gainesville, Florida.
Tung, T.A., 2014a. Agency,'til death do us part? Inquiring about the agency of dead bodies from the ancient Andes. Cambridge Archaeol. J. 24 (3), 437–452.
Tung, T.A., 2014b. Making warriors, making war: violence and militarism in the wari empire. In: Scherer, A.K., Verano, J. (Eds.), Embattled Bodies, Embattled Places: War in Pre-Columbian America. Dumbarton Oaks, Washington, DC, pp. 227–256.
Tung, T.A., Miller, M., De Santis, L., Sharp, E.A., Kelly, J., VanDerwarker, A., 2016. Patterns of violence and diet among children during a time of imperial decline and climate change in the ancient Peruvian Andes. In: Wilson, G. (Ed.), Food and Warfare. Springer, pp. 193–228.
Urteaga-Ballon, O., 1991. Medical Ceramic Representation of Nasal Leishmaniasis and Surgical Amputation in Ancient Peruvian Civilization. Human Paleopathology: Current Synthesis and Future Options. Smithsonian Institution Press, Washington, DC, pp. 95–101.
Vågene, Å., Herbig, A., Campana, M.G., García, N.M.R., Warinner, C., Sabin, S., Spyrou, M.A., Valtueña, A.A., Huson, D., Tuross, N., Bos, K.I., 2018. *Salmonella enterica* genomes from victims of a major sixteenth-century epidemic in Mexico. Nat. Ecol. Evol. 2 (3), 520–528.
Verano, J.W., 1997. Advances in the paleopathology of Andean South America. J. World Prehist. 11 (2), 237–268.
Vreeland, J.M., 1998. Mummies of Peru. in: Cockburn a, Cockburn E, Reyman TA, Editors. Mummies, Disease and Ancient Cultures. Cambridge University Press, Cambridge, UK, pp. 154–189.
Warinner, C., Herbig, A., Mann, A., Fellows Yates, J.A., Weiß, C.L., Burbano, H.A., Orlando, L., Krause, J., 2017. A robust framework for microbial archaeology. Annu. Rev. Genomics Hum. Genet. 18, 321–356.
Weiss, K.M., Ferrell, R.F., Hanis, C.L., 1984. A new world syndrome of metabolic diseases with a genetic and evolutionary basis. Am. J. Phys. Anthropol. 27 (S5), 153–178.
Wilbur, A.K., Buikstra, J.E., 2006. Patterns of tuberculosis in the Americas: how can modern biomedicine inform the ancient past? Memã³rias Do Inst. Oswaldo Cruz 101, 59–66.
Wilbur, A.K., Farnbach, A.W., Knudson, K.J., Buikstra, J.E., 2008. Diet, tuberculosis, and the paleopathological record. Curr. Anthropol. 49 (6), 963–991.
Wilbur, A.K., Bouwman, A.S., Stone, A.C., Roberts, C.A., Pfister, L.A., Buikstra, J.E., Brown, T.A., 2009. Deficiencies and challenges in the study of ancient tuberculosis DNA. J. Archaeol. Sci. 36 (9), 1990–1997.

Williams, P.R., 2001. Cerro Baúl: a wari center on the Tiwanaku frontier. Lat. Am. Antiq. 12 (1), 67–83.
World Health Organization, 2014. Injuries and Violence: The Facts. 2014. WHO, Geneva.
Zachariah, R., Spielmann, M.P., Harries, A.D., Salaniponi, F.M.L., 2002. Moderate to severe malnutrition in patients with tuberculosis is a risk factor associated with early death. Trans. R. Soc. Trop. Med. Hyg. 96 (3), 291–294.
Zink, A.R., Molnár, E., Motamedi, N., Pálfy, G., Marcsik, A., Nerlich, A.G., 2007. Molecular history of tuberculosis from ancient mummies and skeletons. Int. J. Osteoarchaeol. 17 (4), 380–391.

Advances in Paleopathology of Tuberculosis in Pre-contact Andean South America.

SUPPLEMENTARY INFORMATION

Authors: Elizabeth A. Nelson[1,2], Jane E. Buikstra[4] Alexander Herbig, Tiffiny A. Tung, Kirsten I. Bos

Affiliations:

[1]Department of Archaeogenetics, Max Planck Institute for the Science of Human History, Germany

[2]Institute for Archaeological Sciences, University of Tübingen, Germany

[3]Center for Bioarchaeological Research, Arizona State University, Arizona, USA

[4]Department of Anthropology, Vanderbilt University, Nashville, TN, USA

Supplementary Appendix A

Nelson et al

Advances in the Molecular Detection of Tuberculosis in Pre-contact Andean South America

Thirty-four vertebrae were collected from the site of Huari, Peru for DNA extraction and sequencing. Vertebrae that originated from the Cheqo Wasi (n=4) and Vegachayoq Moqo (n=14) sectors were excavated by Enrique Bragayrac (Bragayrac, 1991). Vertebrae from the Monqachayoq sector (n=16) were excavated by Francisco Solano (Solano, 1981). All three sites contained mass burials with commingled remains (Tung 2014). Vertebrae were macroscopically analyzed for preservation level, estimation of age (Scheuer and Black, 2004), and presence of pathological changes (Buikstra and Ubelaker, 1994).

Each vertebra was sampled for DNA extraction in the dedicated ancient DNA laboratory at Max Planck Institute for the Science of Human History (Jena, Germany). Sampling was performed using a dental drill to yield approximately 50mg of pulverized bone, with a sampling location preference for lesions on vertebral bodies. Extraction of ancient DNA was performed following established protocols (Dabney et al., 2013). Double indexed libraries were subsequently manufactured from 10µl of extract following established methods (Meyer et al., 2012). Libraries were sequenced on the Illumina platform (Illumina HiSeq 4000) to a depth of ca. 5 million reads per sample.

To assess host DNA preservation, shotgun data were first mapped to the hg19 human reference genome (GRch37.p13) with the Burrows Wheeler Aligner (BWA) (Li and Durbin, 2010) as implemented in EAGER (Peltzer et al., 2016). We applied BWA mapping parameters of low stringency, which allow for short read lengths (-l 16), limited allowance of nucleotide mismatches to the reference (-n 0.01), and a map quality filtering (-q) of 37. These parameters are suitable for ancient DNA datasets where taphonomic damage in DNA fragments has not been enzymatically repaired. The proportion of detectable DNA damage in the Huari samples was evaluated with mapDamage2 (Jónsson et al., 2013).

Datasets were further screened for pathogens with MALT (Vågene et al, 2018) using a database constructed from all bacterial genomes available on NCBI RefSeq as of December 2016. Read alignments against curated references were inspected for authentication purposes. Though some genetic difference from the reference is expected in ancient DNA, particularly at the ends of DNA fragments where chemical damage can accumulate, true MTBC reads should show very few (one or two) genetic differences from the reference sequence (Key et al., 2017; Warinner et al., 2017).

Reads were mapped to a computationally reconstructed genomic sequence made to represent a hypothetical ancestor of the MTBC (Comas et al., 2010), again using BWA (Li and Durbin, 2010) as implemented in EAGER (Peltzer et al., 2016). This mapping was performed using the parameters described above for the human DNA analysis. Authentication of ancient DNA via damage pattern analysis followed, with methods described above (Figure S1). Both human and MTBC mapping results for the samples putatively positive for *M. tuberculosis* DNA are reported in Table 1.

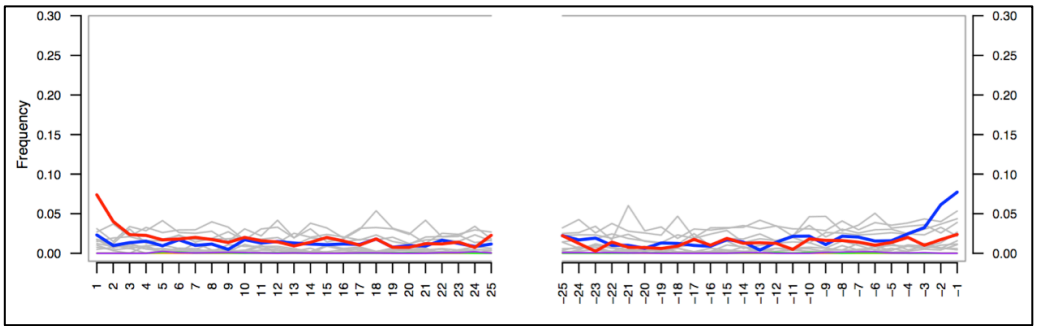


Figure S1. Ancient DNA damage. Damage profile produced by mapDamage2.0 (Jónsson et al., 2013) showing the DNA deamination pattern of MTBC DNA extracted from vertebra HUA004. Visual inspection reveals damage of the DNA fragments in a pattern typical of ancient DNA: the red, curved lined on the left of the image indicates C to T substitutions occurring at the 5 prime end at a frequency of approximately 7%; likewise, the blue, curved line on the right indicates G to A substitutions occurring at the 3 prime end of the same value.

References cited

Bragayrac, D.E. (1991). Archaeological excavations in the Vegachayoq Moqo sector of Huari. In W.H. Isbell & G.F. McEwan (Eds.), *Huari administrative structure: Prehistoric monumental architecture and state government*. Washington, D.C.: Dumbarton Oaks. pp. 71-80.

Buikstra, J.E. & Ubelaker, D.H., 1994. Standards for Data Collection from Human Skeletal Remains. Arkansas Archeological Survey Research Series No. 44, Arkansas, USA.

Comas, I., Chakravartti, J., Small, P.M., Galagan, J., Niemann, S., Kremer, K., Ernst, J.D. & Gagneux, S. (2010). Human T cell epitopes of Mycobacterium tuberculosis are evolutionarily hyperconserved. *Nature genetics*, *42*(6), 498.

Dabney, J., Knapp, M., Glocke, I., Gansauge, M.T., Weihmann, A., Nickel, B., Valdiosera, C., García, N., Pääbo, S., Arsuaga, J.L., & Meyer, M. (2013). Complete mitochondrial genome sequence of a Middle Pleistocene cave bear reconstructed from ultrashort DNA fragments. *Proceedings of the National Academy of Sciences*, *110*(39), 15758-15763.

Jónsson, H., Ginolhac, A., Schubert, M., Johnson, P.L., & Orlando, L. (2013). mapDamage2. 0: fast approximate Bayesian estimates of ancient DNA damage parameters. *Bioinformatics*, *29*(13), 1682-1684.

Key, F. M., Posth, C., Krause, J., Herbig, A., & Bos, K. I. (2017). Mining metagenomic data sets for ancient DNA: recommended protocols for authentication. *Trends in Genetics*, *33*(8), 508-520.

Kircher, M., Sawyer, S., and Meyer, M. (2012). Double indexing overcomes inaccuracies in multiplex sequencing on the Illumina platform. *Nucleic Acids Research, 40*(1), e3.

Li, H., & Durbin, R. (2010). Fast and accurate long-read alignment with Burrows–Wheeler transform. *Bioinformatics*, *26*(5), 589-595.

Peltzer, A., Jäger, G., Herbig, A., Seitz, A., Kniep, C., Krause, J., & Nieselt, K. (2016). EAGER: efficient ancient genome reconstruction. *Genome biology*, *17*(1), 60.

Scheuer, L., Black, S., 2004. The Juvenile Skeleton. Academic Press, London, UK.

Solano Ramos, F., & Anaya, V. G. (1981). Estudio arqueológico en el sector de Monqachayoq-Wari. *Unpublished bachelor's thesis, Universidad Nacional de San Cristóbal de Huamanga, Ayacucho*.

Tung, T. A. (2014). Agency,'til death do us part? Inquiring about the agency of dead bodies from the ancient Andes. *Cambridge Archaeological Journal*, *24*(3), 437-452.

Vågene, Å.J., Herbig, A., Campana, M.G., García, N.M.R., Warinner, C., Sabin, S., Spyrou, M.A., Valtueña, A.A., Huson, D., Tuross, N., & Bos, K.I., (2018). Salmonella enterica genomes from victims of a major sixteenth-century epidemic in Mexico. *Nature ecology & evolution*, *2*(3), 520-528.

Warinner, C., Herbig, A., Mann, A., Fellows Yates, J.A., Weiß, C.L., Burbano, H.A., Orlando, L., & Krause, J., (2017). A robust framework for microbial archaeology. *Annual review of genomics and human genetics*, *18*, 321-356.

Paper III

E. A. Nelson, A.K Lankapalli, M. A. Spyrou, Susanna J. Sabin, Å. Vågene, A. Childebeyeva, B. Rohrlach, J.A. Fellows Yates, M. Cabrera, J. Ochotama, T.A. Tung, A. Herbig, K. I. Bos (2020).

Tuberculosis in the wake of the Wari empire

Manuscript

Tuberculosis in the Wake of the Wari Empire

Authors: Elizabeth A. Nelson[1,2], Aditya Kumar Lankapalli[1], Maria Spyrou[1], Susanna Sabin[1,] Åshild Vågene[3], Ainash Childebeyeva[1], Ben Rohrlach[1], James A. Fellows Yates[1], Martha Cabrera[5], Jose Ochotama[5], Tiffiny A. Tung[4], Alexander Herbig[1], Kirsten I. Bos[1]

Affiliations:

[1]Department of Archaeogenetics, Max Planck Institute for the Science of Human History, Jena, Germany

[2]Institute for Archaeological Sciences, University of Tübingen, Germany

[3]Section for Evolutionary Genomics, The GLOBE Institute, Faculty of Health and Medical Sciences, University of Copenhagen, Denmark

[4]Department of Anthropology, Vanderbilt University, Tennessee, USA

[5] Universidad Nacional San Cristóbal de Huamanga, Ayacucho, Peru

Abstract

Archaeological evidence suggests a significant increase in tuberculosis (TB) infection during the Late Intermediate Period (~1000-1400 CE, LIP) in the Andean region of South America. Osteological identification and subsequent molecular analyses of LIP human remains from three sites in the Osmore River Valley (ORV) revealed the LIP MTBC strain as *M. pinnipedii*, a variety circulating in modern seals and sea lions. The rise in TB during the LIP coincides with significant cultural and climatic transitions including the wane of the Wari Empire, one of the most influential cultures of the LIP. Populations residing in the former administrative center of Wari, Huari, during the decline and in the wake of the empire display osseous lesions consistent with TB. To better understand the geographic spread and evolution of pre-colonial Andean MTBC strains we perform molecular analysis of 92 vertebral bodies and 11 ribs from three sectors of Huari, located in the Peruvian Andes. Here, we present seven ancient *Mycobacterium tuberculosis* complex (MTBC) genomes recovered from Huari, permitting the evaluation of MTBC strain diversity on a local scale. Results of phylogenetic analysis indicate the MTBC strains at Huari as being closely related to the

previously identified ancient MTBC strains, *M. pinnipedii*. Genomic analyses reveal unexpected diversity with extensive microevolution of different contemporaneous clades existing within a narrow geographic range. The molecular confirmation of TB at Huari in the wake of imperial decline supports bioarchaeological data that show an increase in skeletal tuberculosis during the LIP and suggests this extinct strain was widely circulating amongst pre-contact Andean populations.

Significance Statement

Although the rise of TB in the Late Intermediate Period (LIP, 1000-1400 CE) is evidenced in the archaeological record, only three TB genomes, have been characterized revealing a variant adapted for seals and sea lions, *M. pinnipedii,* circulating in coastal sites. Thus, the geographic expanse of *M. pinnipedii* and MTBC strains in the LIP Andes remained unknown. Here, we present seven LIP TB genomes from Huari, the former Wari imperial core. Our results allow for the evaluation of MTBC strain diversity on a local scale and provide insights into the evolution of MTBC in a population experiencing socio-political and climatic stressors. Our work expands the known geographic distribution of pre-colonial *M. pinnipedii* strains to the highland Andean region suggesting this strain was widely circulating.

Introduction

Tuberculosis (TB) is a socially disruptive, deadly disease that has been a central focus of research investigation into the human past. Today, TB is one of the largest global disease burdens with recent estimates of 10 million new infections and 1.5 million deaths in 2018 alone (World Health Organization, 2019). In humans TB is caused by a group of closely related mycobacterial species, the *Mycobacterium tuberculosis* complex (MTBC), that is comprised of seven main globally distributed lineages (L1-L7) and one newly identified lineage (L8) whose regional distribution suggestions a long history of co-adaption with humans (Gagneux et al., 2006; Ngabonziza et al., 2020). Additionally, a clade of MTBC strains has been isolated from a diverse group of mammalian species, of which several have caused disease in humans as well (Orgeur and Brosch, 2018; Gonzalo-Asensio et al., 2014). Members of the MTBC possess features such as latency (Blaser and Kirchner, 2007) and broad transmissibility which are influenced more by demography and contact rates than by host species (Brites et

al., 2018). These features have made the MTBC one of the most globally ubiquitous and successful groups of pathogenic bacteria.

The tendency of advanced MTBC infections to cause diagnostic skeletal and mummified soft tissue lesions has facilitated identification of TB in archaeological contexts (Allison et al., 1981; Roberts and Buikstra, 2003), which led to molecular detection (Salo et al., 1994) and subsequent genomic characterization of ancient strains (Bos et al., 2014; Chan et al, 2013; Kay et al, 2105; Sabin et al, 2020). However, clinical studies report that diagnostic skeletal changes occur in only a small percentage of infected individuals (CDC, 2018; Cowie and Sharpe, 1997; Lešić et al., 2010; Tuli, 2016). This implies the disease was more prevalent in the past than can be estimated based solely on morphological analysis.

Osteological observations have long been reported from Andean South American contexts as early as 200 CE, with the most robust evidence being from the Chilean Atacama Desert as early as 700 CE (Allison et al., 1981). The presence of skeletal TB increases slightly in the Middle Horizon (~600 to 1000 CE) (Roberts and Buikstra, 2003), which coincides with the expansive phase of the Tiwanaku and Wari empires, as well as Lambayeque on the North Coast of Peru (Jennings, 2006; Schreiber, 1991; Vogel, 2018). Although there are accounts of purported skeletal TB in various regions of Peru dating to the Middle Horizon (MH, ~600 – 1000 CE), the majority of paleopathological observations of skeletal TB from the Andean region date to the Late Intermediate Period (LIP, ~1000 – 1400 CE) (Allison et al., 1981; Arriaza et al., 1995; Rivas Boada, 1988; Bos et al., 2014; Buikstra and Williams, 1991; Burgess 1992, 2000; Klaus et al., 2010; Owen, 1993; Roberts and Buikstra, 2003; Salo et al., 1994; Toyne et al., 2020). This includes the three cases of TB confirmed via genomic evidence from the Chiribaya culture of the Osmore River Valley (ORV), where molecular investigation identified *Mycobacterium pinnipedii*, an MTBC strain adapted for seals and sea lions, circulating in LIP coastal populations (Bos et al., 2014). The identification of *M. pinnipedii* in human populations suggests a zoonotic origin for pre-colonial Andean TB. Archaeological evidence of maritime practices and utilization of seals as resources in ancient Andean populations supports the possibility of a zoonotic transmission of *M. pinnipedii* (Alcalde, 2005; Quilter, 1990). However, the genetic

diversity of the MTBC strains circulating in the Andean highland region during the LIP TB rise remains largely unexplored.

The LIP is ushered in by the decline of the prolific Wari Empire that, during the MH, exercised its influence from the highland capital of Huari, 3,000 meters above sea level and approximately 500km from the ORV coastal sites. During the LIP an intense climatic shift to severe aridity coincident with cultural transitions led by the decline of the Tiwanaku state and Wari empire evidenced by a movement to more dispersed settlements (Fehren-Schmitz et al., 2014; Seltzer and Hastorf, 1990; Thompson et al., 2013). For those who remained at Huari during the terminal-Wari and post-Wari eras of the LIP, the intense drought led to a shift in diet and the disintegration of the Wari empire with socio-political upheaval signaled by an increase in lethal and sublethal violence (Schittek et al., 2014; Tung 2008, 2014; Tung et al., 2016). Such environmental, dietary and socio-political factors have been implicated in modern epidemiological studies as drivers of TB epidemics (Lönnroth et al., 2009). Likewise, LIP Huari populations display skeletal lesions suggestive of bacterial infection, namely tuberculosis (paper II of this thesis). In order to investigate the presence of MTBC strains and strain diversity in Huari populations experiencing the decline and fall of the empire in the midst of climatic, socio-political and dietary shifts we applied molecular methods and analyses.

Here we present the molecular detection and phylogenetic characterization of ancient MTBC strains circulating in terminal and post-Wari populations recovered from the former capital of the Wari Empire, Huari. Using a broad pathogen screening method to analyze these commingled remains, we were able to identify MTBC reads in 14 vertebrae. We demonstrate the use of filtering methods to authenticate the genetic variation observed and subsequently reconstruct the largest of number of published ancient TB genomes from one site to date allowing for the evaluation of MTBC strain diversity on a local scale. Our work provides a deeper understanding of TB across the LIP Andes and insight to the evolution and ecology of MTBC strains in this archaeological community experiencing climatic variability, dietary shifts, and socio-political instability, from the heart of the former Wari Empire, in the Peruvian Andes.

Ethics Statement

In this study we apply pathogenomic analyses to answer archaeological questions of biomedical relevance. In effort to understand the health and disease of post-Wari populations we incorporate archaeological studies, paleopathological analysis and genomic methods to analyze human remains. Our work was performed with active communication and involvement with Peruvian archaeologists and community of archaeologists and students at the Universidad Nacional de San Cristóbal de Huamanga (UNSCH). Our research was conducted with the permission of UNSCH and research approval and authorization were given by the Peruvian Ministry of Culture through which the appropriate permits were obtained. Our research demonstrates the application of molecular methods to a well-studied archaeological context to inform local Peruvians, and by extension, a global audience, on the history of TB, a disease that continues to burden the world today that we may shed more light on our understanding the evolution, ecology, and success of this pathogen.

Results

Skeletal Analysis. We examined 2,226 vertebrae from LIP commingled contexts from three sectors of the site of Huari: Cheqo Wasi (CW, 1044 – 1155 CE), Vegachayoq Moqo (VM, 1305 CE – 1405 CE) and Monqachayoq-Solano (MQS, 1321 CE – 1425 CE) (Figure 1). Dates for each sector were obtained through AMS dating (Supplementary Table 1). Our results show 132 vertebrae presented pathological lesions and abnormal skeletal changes. Furthermore, we find that juvenile thoracic vertebrae were more frequently affected with pathological skeletal changes (n=83, 12.2% of thoracic vertebrae examined) compared to adult thoracic vertebrae (n= 20, 1.3% of thoracic vertebrae). Conversely, adult lumbar vertebrae were more frequently affected than juvenile vertebrae. Of the 132 vertebrae, only 5 showed signs confidently associated with tuberculosis while the remaining 127 displayed a spectrum of skeletal presentation including slight, non-specific skeletal changes such as slight reactive bone on the vertebral body, resorptive lesions on the superior-anterior aspect of the vertebral body, and hyper-vascularity of the vertebral body which may or may not be considered within normal variation.

As TB is known to cause skeletal changes in only a small percentage of people experiencing TB infections (Jaffe, 1972) the low frequency of skeletal changes suggestive of TB is expected and represents a small percentage of the infected population (~1-20%). Therefore, due to the low frequency of skeletal changes in an infected population, we selected skeletal elements which display severe pathological manifestation as well as skeletal elements showing only slight changes and those which may be considered within normal anatomical variation. The resulting sample set included a total of 103 skeletal elements, 92 vertebrae and 11 ribs, across multiple age groups from three sectors of Huari (CW, MQS, and VM) (Supplementary Table 2).

Broad Pathogen Screening and Detection of MTBC DNA. Here we employed high throughput sequencing and broad pathogen screening methods to 69 skeletal samples that showed some evidence of pathology, 58 of which were vertebrae and 11 were ribs. All skeletal elements were sampled in a laboratory dedicated to aDNA using a dental drill to collect approximately 50mg of bone powder. Ancient DNA extractions were prepared using protocols customized for ancient DNA (see methods). Subsequent amplified libraries were sequenced to a depth of 5 million reads on Illumina platforms (NextSeq 500 or HiSeq 4000). Pathogen screening was performed using the HOPS pipeline which incorporates the taxonomic identification of preprocessed sequencing reads using the Megan Alignment Tool (MALT) (version 0.3.8) (Vågene et al., 2018) and authentication of ancient DNA (Hübler et al., 2019). MALT taxonomic assignments of microbial content were performed against the NCBI Nucleotide (nt) database (November 2017) using a minimum sequence identity of 95% and results were visualized in MEGAN (Huson et al., 2016). Six vertebrae harbored reads that assigned to the MTBC ranging in number from 13 to 8,308 (Table 1). Further visual inspection of the alignments supported this identification. These results were combined with screening data from an additional 34 Huari samples presented elsewhere (Nelson et al., 2020), which yielded a total of 14 vertebrae that demonstrated potential molecular survival of MTBC DNA. This subset consisted of at least two vertebrae from each of the three sectors of Huari sampled for this project (Table 1) (Supplementary Figure 1).

To further explore MTBC DNA preservation, subsequent reference-based mapping of pre-processed reads to a reconstructed MTBC ancestral reference genome (Comas et al., 2010) was performed using BWA as part of the EAGER pipeline (Peltzer et al.,

2016) (Table 1). This process revealed between 1,131 (HUA051) to 16,993 (HUA093) unique mapping reads per library, with percent MTBC DNA comprising 0.02% - 0.4% of total DNA content. These results are on par with DNA yields that permitted genome reconstruction in material from similar geographical locations and time periods (Bos et al., 2014). Sequenced reads were also mapped to the human reference genome (Hg19) which yielded between 252 (HUA037) – 259,993 (HUA041) unique reads after quality filtering with 0.006% - 7.06% endogenous human DNA content (Supplementary Table 4). Authenticity of ancient DNA was assessed by observing cytosine deamination patterns on the 5' and 3' ends of DNA molecules. Observations for both human and TB mapping showed damage patterns typical of ancient DNA (Table 1 and Supplementary Tables 3 and 4) (Supplementary Figure 2) (Jónsson et al., 2013).

Hybridization capture and MTBC genomic recovery. All libraries that putatively demonstrated MTBC DNA preservation, along with a positive control of known MTBC DNA content, were prepared for targeted in-solution hybridization enrichment of MTBC DNA (see methods). Uracil-DNA-Glycosylase (UDG) was used excise the uracil residues (Briggs et al., 2010) DNA "damage" typical of aDNA. The removal of this damage allows for more confident SNP calling in downstream evolutionary analyses. In-solution hybridization of MTBC DNA was performed using single-stranded DNA probe set based on the reconstructed MTBC ancestor (Comas et al., 2010). We performed two rounds of in-solution capture and sequenced the capture products to a depth of 10,000 reads (see methods).

After in-solution capture and sequencing all data were mapped to the MTBC ancestor reference (Comas et al., 2010) using BWA and the EAGER pipeline (Peltzer et al., 2016). Results revealed MTBC endogenous DNA values ranging from 1.13% – 19.02% of the total sequencing data and mean coverage of the reference genome ranging from 0.09 – 6.26. Our positive control showed expected values of 37.67% endogenous TB DNA and a mean coverage of 21.2, thus indicating good experimental conditions (Supplementary Table 5).

Non-pathogenic mycobacteria are ubiquitous in the environment, with an abundance of species identified in water and soil (Falkinham, 2002; Torvinen et al., 2010). The

porous nature of bone allows for soil mycobacteria to infiltrate the sample and by extension contaminate the extraction. This similar genetic composition of environmental mycobacteria from soils can complicate pathogen studies of archaeological material through becoming unintentionally included in targeted capture and genome reconstruction (Warriner et al., 2017). To investigate the influence of non-target environmental mycobacteria on genomic MTBC coverage, we performed MALT analysis to calculate estimations of target MTBC reads in all capture products. Thus, permitting the evaluation of nontarget environmental mycobacteria inadvertently enriched during the TB capture and subsequently incorporated in the reconstructed genomes. The samples varied in the percent of reads able to be taxonomically assignment ranging from 0.85% to 12.8% (Supplementary Table 6). However, calculations estimating the percentage of MTBC specific reads within each sequenced library revealed several datasets to contain a high amount of non-target environmental mycobacterial DNA (Supplementary Table 6). In this analysis, all samples that showed an upper boundary estimation of at least 85% of MTBC reads and lower boundary estimation of at least 50% of MTBC or greater would be included in downstream analysis (Supplementary Table 6). We identified seven samples which met or exceeded the threshold of upper and lower boundary MTBC estimates. However, estimates indicate six samples contained high levels of non-target mycobacterial background of the capture product; although there was target MTBC DNA, the confident recovery of this DNA would likely be challenged by the abundance of non-target mycobacteria. Based on the results of this analysis, individuals HUA004, HUA006, HUA016, HUA037, HUA041, HUA051, and HUA057 only were carried forward for downstream analyses (Figure 2).

The selected vertebrae (HUA004, HUA006, HUA016, HUA037, HUA051, HUA057) were resampled taking approximately 50 mg of bone powder that was extracted and prepared into UDG treated libraries for in-solution capture of MTBC DNA and subsequent genome reconstruction. Recovery of high-quality data is challenged by the previously mentioned incorporation of environmental background and while human aDNA studies have worked to identify skeletal elements yielding efficient human DNA, sampling strategies have not yet been optimized for ancient MTBC DNA. Therefore, in this second sampling we tested the influence of sampling location on MTBC DNA

recovery via comparison of MTBC DNA yield from several anatomical locations of the vertebrae, normalized against approximately 50mg of starting material. The seven vertebrae were systematically sampled an additional four to ten times from the following anatomical aspects of the vertebral body: anterior aspect, lateral aspects (left and right), superior, inferior, the trabecular bone of the vertebral body, and lesion(s) (if present) and trabecular bone posterior to any present lesion(s). The number of sampling locations per vertebrae varied with the presence and number of morphologically detectable lesions consistent with TB. Extracted DNA from each sampled site was prepared as individual UDG treated libraries with unique barcodes such that data DNA yields could be compared. Together with the initial seven libraries generated previously, a total of 53 unique libraries from various sampling sites formed the basis of this analysis. Libraries were all captured for MTBC DNA as described above, and comparisons were based on the number of reads that mapped to the reconstructed ancestor. Although there is some variation in recovery of MTBC DNA across the different sampling locations (Supplementary Table 7), our statistical analysis of the resulting libraries showed no significant *Spearman's correlation* existed between sampling site and TB endogenous DNA recovery. Therefore, endogenous yield of TB DNA seems to be independent of sampling location.

Ancient genome reconstruction. All sequencing data of the in-solution hybridization capture was pooled by individual vertebrae and mapped to the MTBC ancestor reference (MTB_anc) revealing endogenous TB DNA percentages ranging from 2.034 to 9.402 (Table 2). In addition, the mean coverage of each reconstructed genome ranged from 5.308-fold to 33.529-fold coverage, with 76.89% of the MTBC ancestor reference covered at least 2-fold in all reconstructed genomes (Table 2, extended: Supplementary Table 8). As all vertebrae were recovered from commingled depositions, estimation of the minimum number of individuals represented is based on age at death assessments and archaeological context (sector of origin). From this we estimate our dataset to represent at least 5 individuals.

Authenticating Genetic Variation in Ancient MTBC Genomes. The samples that were selected for full genome capture and reconstruction showed relatively low to moderate environmental mycobacterial background with upper estimations of MTBC DNA composition of the sample ranging between 77% - 98% (Supplementary Table

6). However, due to the highly conserved regions of mycobacteria this background of closely related mycobacterial species can map to the MTBC ancestor reference and lead to false positive SNPs producing longer branches in the reconstructed genome (Bos et al., 2019). False positive SNPs can complicate evolutionary and phylogenetic analyses and lead to wrong topologies and the formation of artificial branches in phylogenetic trees (Keller et al., 2019). To filter out variant calls from co-enriched non-target mycobacterial species we evaluated the unique and shared variant positions in our dataset via an evaluation tool that compares coverage at a given position under conditions of 1) stricter and 2) more permissive mapping (https://github.com/andreasKroepelin/SNP_evaluation) (Keller et al., 2019). In this process, coverage in the 50bp window around each SNP is considered under the two mapping parameters. SNPs that exceed a predefined threshold of proportional coverage increase under relaxed mapping parameters are removed (see methods). Furthermore, we excluded each position in which we observe the existence of additional heterozygous positions in the same read within the 50 bp window. Additionally, we filtered out any position which lacked full genomic coverage in the 50bp window as deletions might result in mapping errors (Keller et al., 2019). This analysis was performed on all ancient Peruvian genomes, which included all UDG-treated Huari libraries and the previously published ORV TB genomes (Bos et al., 2014). Our evaluation resulted in the removal of 1566 out of 6302 shared positions and 402 out of 514 unique positions (Supplementary Table 9).

Phylogenetic analysis. A set of 30 genomes composed of 11 modern *M. microti* genomes, 9 modern *M. pinnipedii* genomes and 10 ancient MTBC genomes dating to the LIP, three previously published ancient coastal TB genomes (Bos et al., 2014) and our seven reconstructed MTBC genomes from Huari, were included in our phylogenetic analysis (Supplementary Table 10). The *Mycobacterium microti* clade was used as the outgroup.

Our Maximum Likelihood (ML) tree constructed from the SNP alignment include 2,026 variant positions resulted in the phylogenetic placement of the ancient MTBC genomes from Huari as closely related to modern *M. pinnipedii* strains and the previously published ancient TB genomes from the ORV (Bos et al., 2014). This reveals

that the pinniped strain of MTBC circulated not only in ancient coastal Peruvian populations but also contemporaneously circulated in ancient highland populations from the Andean Basin, over 500 km away. The topology shows variation in the strains present at Huari with genomes from all three sectors of Huari clustering separately and radiating out from a polytomy. The MTBC strain from the Cheqo Wasi (HUA004) sector is unique and has no shared derived position. Furthermore, Vegachayoq Moqo (HUA006, HUA016, HUA051, HUA057) and Monqachayoq-Solano (HUA037, HUA041) are contemporaneous and geographically located closer to each other than the earlier Terminal Wari site of Cheqo Wasi (Figure 1). Yet, our phylogeny shows the MTBC strains recovered from the Vegachayoq Moqo sector (HUA006, HUA016, HUA051, HUA057) being more closely related to the ORV MTBC strains (54U, 58U, and 64U) (Bos et al., 2014) with bootstrap support of 90% (Figure 3).

To evaluate the effect of the SNP filtering process on the topology we constructed an alignment generated from the filtered dataset in addition to an alignment that was generated from a dataset which had not been filtered. The results of this analysis show all genomes now present reduced branch lengths (Supplementary Figure 4). This affect is most noticeable in HUA004 in which the shared branch between HUA004 (CW) and HUA037 (MQ) and HUA041 (MQ) was reduced to now show a polytomy. Examination of the snpTable generated by MultiVCFAnalyzer (Bos et al., 2014) confirmed there are no shared informative positions between these genomes (Supplementary Table 11). Further evaluation of the snpTable identified 299 variant positions that occur in at least one of the ancient genomes from Huari. A total of 42 variants are shared among the *M. pinnipedii* genomes and the Huari and ORV ancient TB genomes.

Functional analysis of variant positions. Functional analysis of variant positions was carried out on the same dataset of 30 genomes which was also used for phylogenetic analysis using SnpEff (v. 3.1) (Cingolani et al., 2012). For this analysis we focus on variants that would lead to the formation of new amino acids (NON_SYNONYMOUS_CODING) or variants which may disrupt the function of a gene through the formation of a STOP codon (STOP_GAINED) or variants causing a loss of STOP codons (STOP_LOST) and the loss of start codons (START_LOST) thereby forming pseudogenes. Of the 299 variant positions, 116 are unique to the

ancient genomes and do not occur in the *M. microti* or *M. pinnipedii* genomes and within those, 56 are nonsynonymous coding variants. Only one position occurring in a coding region was unique to ancient MTBC strains and is confidently shared between all ancient MTBC genomes (Huari and ORV). This SNP (S601R), located on position (1050212), creates a non-synonymous amino acid change in a gene (Rv0939). Detailed results of SnpEff analysis can be found in (Supplementary Table 11).

The CW genome (HUA004) has 13 unique variant positions five of which are categorized as non-synonymous and include variants in genes involved with the formation of conserved hypothetical proteins (Rv2390c), cell wall processes (mmpL8), intermediary metabolism and respiration (rbsK), and information pathway (sigJ). Functional changes of the variant at position 2683723 also causes an upstream effect (about 11 bases) on gene rpfD (Rv2389c) which is involved in cell wall processes, specifically resuscitation and growth of dormant, non-growing cells and the disruption of this gene has been demonstrated to provide a growth advantage for in vitro growth of H37Rv strains (De Jesus et al., 2017). Furthermore, SnpEff analysis of HUA004 also revealed a variant causing a STOP codon (STOP_GAINED), in a gene (Rv2303c) associated with virulence and detoxification, the product of which is a probable antibiotic resistance protein (Rustad et al., 2014). This suggests the gene is disrupted and non-functional. Additional SNP effects occurring within HUA004 include an upstream effect (72 bases) in a gene known to be involved in lipid degradation, fadD5 (RV0166).

The MQS genomes (HUA037 and HUA041) possess 56 SNPs which are unique to their group and 1 additional SNP (position: 1164571) that is shared between the MQS HUA037 genome and the ORV ancient MTBC genomes. It should be noted that this position is not well covered in the other ancient TB genomes (HUA004, HUA051, HUA057), including HUA041 so it may be possible this variant is not uniquely shared between HUA037 and the ORV genomes. In fact, there exist no variant positions unique to either of the MQS genomes (HUA037 or HUA041) in which the other does not present an ambiguous call and consequently we cannot confidently identify SNPs truly unique to either genome. However, between these two genomes there exist 28 SNPs categorized as NON-SYNONYMOUS_CODING, five of these involved in the

probable formation of hypothetical proteins, six involved in cell wall and cellular processes including the formation of integral proteins (Rv2180c), membrane proteins (Rv0012), secreted proteins (RV3605c), and ATP-binding and transport proteins involved in active transport of glutamine (Rv0073) and other unidentified substrates (Rv1447). Additionally, these genomes possess several non-synonymous variants in genes involved in intermediary metabolism and respiration, lipid metabolism, and regulatory proteins, with variants in genes *blal* and Rv2011c both of which are involved in transcription repression. MQS genes also possess SNPs in genes associated with virulence, detoxification, and adaptation such as Rv2190c of which the function is not fully known in addition to *vapC15* and *mazF7* which are part of the well characterized toxin-antitoxin (TA) systems of MTBC (Solano-Gutierrez et al., 2019; Tandon et al., 2020).

Analysis of variant positions in the VM genomes (HUA006, HUA016, HUA051, HUA057) revealed a total of 43 positions that occurred in at least one of the VM genomes. As the highest coverage genome, all variants were present in HUA016, however, these show as ambiguous sites in the other 3 VM genomes. These SNPs included a total of 22 nonsynonymous coding variants, eight of which were coding for hypothetical conserved proteins, one (Rv0674) of which has been found to be an essential gene (Sassetti and Rubin, 2003). Four variant positions are involved in cell wall and cellular processes, including *corA* which is thought to be involved in the transport of magnesium and cobalt ions across the membrane, and *ctpC* which functions to produce metal cation-transporting ATPase possibly catalyzing the transport of metal cations. The remaining two variant positions occur in genes coding for membrane proteins essential for in vitro growth (Rv0102 and Rv2609c) (DeJesus et al., 2017). Five nonsynonymous coding variants occur in genes involved in intermediary metabolism and respiration including *gnd1* and Rv1882c in which disruption of the gene is known to cause a growth advantage for H37Rv (DeJesus et al., 2017). Four nonsynonymous coding SNPs were identified in genes involved with lipid metabolism (mas, tesB2, tgs4, ppt) with only ppt having been demonstrated to be essential for in vitro growth (DeJesus et al., 2017). Analysis reveals HUA016 possesses a variant at the position 2201019 affecting the gene *vapC14* (Rv1953), which is involved in virulence, detoxification and adaptation and is part of the toxin-antitoxin operon

demonstrated to have a growth advantage for in vitro growth in association with gene disruption (DeJesus et al., 2017).

The VM genomes and ORV genomes share variant positions including within a highly conserved intergenic region with downstream effects (33 bases) in the gene *mrp* (Rv1229c) involved with intermediary metabolism and respiration through likely producing ATP binding proteins. This gene has been demonstrated to be essential for in vitro growth of H37Rv (DeJesus et al., 2017). HUA016 and 64U share a SNP at the position 1345945 effecting the gene Rv1202, involved in biosynthesis of succinyl-diaminopimelate desuccinylase (*dapE*) and is an essential gene for in vitro growth (DeJesus et al., 2017; Sassetti and Rubin, 2003; Griffin et al., 2011).

Analyzing selection pressure on S601R. To investigate selective pressure and possible signals of adaptation in the ancient Peruvian TB strains, we analyzed the non-synonymous SNP (S601R) that was identified to only be present in ancient isolates. This codon is conserved in the animal-adapted *M. tuberculosis* strains (bovis_ravenel, M_orygis, capraeRW044, Pinnipedii7739, capraeD028, capraeRW079, Pinnipedii_G0149, Pinnipedii_G01222, Pinnipedii7011, MicrotiERR027294, chimpanzee_bacillus, Pinnipedii_G01491, Pinnipedii_G01492), as well as the other human-specific strains of MTBC. The function of the gene Rv0939 is largely unknown. According to the Mycobrowser database, Rv0939 is a non-essential gene, however, it is possibly involved in cellular metabolism, degradation, and respiration.

We performed a codon-by-codon analysis of the gene Rv0939 in ancient MBTC and *M. microti* and *M. pinnipedii* animal strains. We used the Fast Unconstrained Bayesian Approximation (FUBAR) method from the datamonkey server (Pond and Frost, 2005), and identified codon S601R as being potentially under selection with posterior probability of positive selection at the site (p= 0.873) and the posterior probability of negative selection at the site (p=0.080). We also performed a site-wise analysis of positive selection using codeml as implemented in the PAML package for the gene Rv0939 (Yang, 2007). We compared models M8a (no site can have a dN/dS > 1) and M8 (sites are allowed to have dN/dS >1) using a likelihood ratio test, which showed that the two models were not significantly different for the gene Rv0939 (ΔLRT=0.36,

p-value= 0.55 based on X^2 df =1). The site S601R was identified as potentially under selection via the Bayes Empirical Bayes (BEB) analysis in codeml (Supplementary Table 12).Although S601R was shown as being potentially significant by FUBAR (0.873) and the codeml Bayes Empirical Bayes (BEB) analyses, we cannot rule out the possibility of genetic drift acting on this codon.

Molecular dating. Molecular dating analysis was performed on 2,026 SNPs curated from a genome-wide alignment of 30 genomes. The alignment consists of ancient genomes recovered from human samples and modern *M. pinnipedii* and *M. microti* strains isolated from animals. An ML tree was reconstructed using RAxML (Stamatakis, 2014) for the SNP alignment with a GTRGAMMA model of substitution for 1000 bootstrap replicates. TempEst analysis assessed the presence of a temporal signal in the historic genomes given the ML tree and radiocarbon dates. The root to tip analysis of sample dates and genetic distances of the rooted ML tree (after selecting best fitting root) revealed a strong linear relationship with R^2 = 0.96 correlation coefficient of 0.97. This hinted at the presence of a temporal signal between the ancient and modern genomes. In order to estimate the emergence of the subclade including the ancient MTBC, *M. microti* and *M. pinnipedii*, we performed a Bayesian Skyline analysis using BEAST2 version 2.5.2 (Bouckaert et al., 2019). The mean values of radiocarbon ages (Supplementary Table 1) were used to calibrate the tips of ancient genomes while 2010 was considered as the tip age for the remaining 20 modern genomes (Sabin et al., 2020). The distribution of rate heterogeneity among the sites was estimated with a Gamma site model for 8 categories, wherein the shape (1.0) and proportion of invariant (0.1) parameters were estimated. The rate of nucleotide changes, and nucleotide frequencies were estimated with a GTR model of substitution (rate of change from C to T was set to default 1.0). The rate of conversion along the branches of the tree is modelled as an uncorrelated relaxed clock model, with rates sampled from a log-normal probability distribution for a set of discrete rates (-1) equal to the number of branches in the tree. An initial clock rate of 4.6E-8 was set for the analysis as prior (Bos et al., 2014). A Maximum Likelihood phylogeny was provided as an initial tree. A Coalescent Bayesian Skyline model was considered as tree prior and all the associated parameters were fixed as default. Additionally, two priors were assigned for all *M. microti* strains fixed as a monophyletic clade and other genomes as another

monophyletic clade. An MCMC chain was run for 300,000,000 iterations with a burn-in of 1,000,000 and logged the tree states at every 1000 steps. Tracer v 1.7.1 (Rambaut et al., 2018) was used to evaluate the convergence of all the parameters and reconstruct the demographic history. The run converged and a Maximum Clade credibility tree was determined and the node ages of both the *M. microti* and ancient genomes were estimated.

Using the Bayesian skyline model, we estimated a mutation rate of 1.359E-4 substitutions per site per year (1.1305E-4 – 1.6094E-4 95% HPD). The most recent common ancestor (MRCA) age estimate for the historic and *M. pinnipedii* was 973 yBP (890 – 1036 yBP 95% HPD). The age of the MRCA *of M. microti* was 1394 yBP (1313 – 1512 yBP 95% HPD) (Supplementary Table13).

Our Skyline analysis further revealed evidence of a fluctuation in population demographics of the ancient MTBC and *M. microti* strains (Supplementary Figure 6). This fluctuation is estimated to have occurred between 1300 – 1400 CE years, possibly suggestive of a bottleneck event, then followed by a rapid population expansion between 1400 – 1500 CE that stabilizes throughout the modern era.

Discussion

Detecting and recovering MTBC DNA. We present the largest set of reconstructed pre-colonial MTBC genomes from a single site to date and subsequently more than doubled the current existing dataset of ancient South American TB genomes. There was no identification of MTBC in rib samples, however, recovery of MTBC DNA was successful from vertebrae which did not show pathological changes that are broadly considered to be diagnostic and therefore may have gone overlooked. This was made possible through exerted effort in skeletal sampling, DNA extraction, genome sequencing, capture and possibly most importantly, efforts in mitigating against environmental mycobacterial contaminants for authentication of the observed genetic variation. Among the many challenges in ancient pathogen DNA research is the existence of environmental microbes and complications in reconstructing ancient pathogen genomes (Warriner et al., 2017). Archaeological contexts primarily yield

samples which have been surrounded by environmental matrices of soil, water, plants, and more. The breadth of microbial organisms existing in the environment infiltrate porous skeletal samples and reside on sample surfaces resulting in the unavoidable inclusion of exogenous and non-target microbial DNA in extracted samples. This challenge is particularly emphasized in studies of ancient mycobacterial diseases, i.e. tuberculosis and Hansen's disease. The strong conservation of genomic regions within mycobacteria results in shared genetic information across the genus in both pathogenic and environmental organisms. Here we circumvent the challenge of reconstructing genomes which include non-specific mycobacterial reads by utilizing MALT (Vågene et al., 2018) as a tool to evaluate the metagenomic background and specifically evaluate the mycobacterial composition of samples with special attention to non-target mycobacterial reads within extracts (Spyrou et al., 2019). By evaluating the proportion of non-target mycobacterial reads to MTBC reads we were able to infer which samples would most likely yield sufficient target MTBC reads in hybridization capture and allow for a "cleaner" reconstructed MTBC genome. Furthermore, for those samples selected for capture, we filtered the reconstructed genomes removing questionable positions most likely deriving from environmental mycobacteria and allowed for a more conservative and thus more confident reconstruction of ancient MTBC genomes (Keller et al., 2019).

MTBC pinniped strains at Huari. Unlike the Chiribaya archaeological contexts of the low land ORV ancient TB genomes (Bos et al., 2014), Huari is located the approximately 500km away in the central highland region near present-day Ayacucho, Peru. At the Andean site of Huari, there is evidence of interaction and domestication of camelids and cuy (guinea pig), yet no evidence of pinniped exploitation or exchange. The Huari faunal assemblages found in contexts contemporary with those from which we identified TB possibly indicate additional mammalian hosts and/or strains. Likewise, the expanse of the Wari and later evidence of post-Wari population dispersal would facilitate movement of the pathogen via infected people, and by extension animals, across a large geographical range resulting in the geographic spread of TB. In fact, during the Middle Horizon there appears to be little genetic exchange between lowland and highland populations of the Andes. Genetic and archaeological evidence suggests that it is not until the LIP that there is visible movement with increased

population densities in lowland sites (Reindel and Wagner, 2009) and population homogenization as seen by a decrease in genetic distances between lowland and highland populations (Fehren-Schmitz et al., 2014). This increase in population movement between the highlands would most certainly facilitate the movement of pathogen and spread of tuberculosis.

The shared variant positions, similar genomic architecture, and phylogenetic placement of the Huari MTBC genomes as closely related to the contemporaneous Chiribaya ORV strains reveals a broad geographical presence of *M. pinnipedii* in human hosts during the LIP. Although all TB genomes from Huari contexts are closely related to the ORV TB strains, the phylogenetic analysis reveals somewhat unexpected genetic diversity within the Huari genomes. While the genomes do cluster together reflecting the sector from which they were recovered, TB genomes from the two sectors most closely located, being less than one kilometer apart in distance, and dated to be contemporaneous sectors, Vegachayoq Moqo (1305 CE – 1405 CE) and Monqachayoq-Solano (1321 CE – 1425 CE), are unexpectedly distantly related from one another. Additionally, Vegachayoq Moqo is more closely related to the ORV Chiribaya TB genomes (Bos et al., 2014) than any of the other Huari strains. The diversity observed within this geographic and temporally narrow range may suggest tuberculosis outbreaks were widespread in the LIP Andean cultural region thus supporting archaeological observations which show and increase in skeletal signs of tuberculosis in the LIP, indicating a greater number of infected people than what can be observed through osteological analysis. Indeed, our recovery of MTBC genomes from vertebrae which show little to no pathological change further supports this premise.

Molecular dating and currently characterized ancient Andean MTBC strains. Previously, tuberculosis in the pre-colonial Americas was understood by the presence of skeletal changes, yet now with the reconstruction of ancient MTBC genomes we move closer to a better understanding on the evolution and ecology of the disease prior to European colonization (Bos et al., 2014). The identification of these MTBC strains as closely related to the seal adapted species, *M. pinnipedii*, suggests the appearance of TB in these communities may have been the result of a zoonotic event and provides a possible explanation as to the initial introduction of TB to the pre-colonial Americas.

However, results of our Bayesian molecular dating results suggest the currently characterized ancient MTBC strains of the Americas represent an only recently emerged subclade with an MRCA age estimate for the ancient and modern *M. pinnipedii* at approximately 973 yBP (890 – 1036 yBP 95% HPD). Furthermore, age of the MRCA of *M. microti* which is basal to all currently characterized *M. pinnipedii* genomes was estimated to be 1394 yBP (1313 – 1512 yBP 95% HPD). The age of the earliest radiocarbon date for our Huari samples is 1044 CE (HUA004, Cheqo Wasi) which suggests our genomes are some of the earliest representatives of this subclade. However, skeletal evidence of TB has been recovered from contexts dating much earlier than estimated MRCA dates of this subclade (Buikstra and Roberts, 2003). Some of the earliest evidence dates to 200 CE from Tarapaca, Chile and the most robust evidence being from the Chilean Atacama Desert as early as 700 CE (Allison et al., 1981). Although more molecular research with a broader time transect is needed, these data may suggest that earlier skeletal manifestations of TB developed from an ancestral species contributing similar paleopathological manifestations which was ultimately replaced by ancient *M. pinnipedii* strains. Likewise, with European colonization, we see complete replacement of these MTBC strains that once broadly infected human Andean populations with L4 strains known to be strongly associated with European populations (Gagneux et al., 2006).

The Late Intermediate Period rise in tuberculosis. The Late Intermediate Period (1000-1400 CE) of the South American Andes is a time of climatic and cultural transitions with a severe drought leading to a shift in diet and lifeways, to the rise in skeletal tuberculosis and the fall of this first expansive empire of the Andean cultural region. Cessation of the Wari Empire came after approximately 500 years of power across a large portion of the vast Andean region with evidence of Wari influence reaching to the northern Peruvian Andean highlands across to the northern and southern Peruvian coastlines with some limited evidence of influence in the lower valleys (Tung, 2012). During the Middle Horizon (~600-1100 AD), the Wari exercised authority out from the established administrative center of Huari, located approximately 2900masl (meters above sea level) in the Ayacucho Basin of the central Andes. The Wari empire achieved a large sphere of influence incorporating many cultures over various Andean landscapes evidenced by ceramics with Wari style artistic depictions (Jennings, 2006;

Williams et al., 2019), agricultural developments (Valencia Zegarra, 2005), language diffusion (Heggarty and Beresford-Jones, 2012), and Wari imperial monumental architecture (Isbell and McEwan, 1991).

For reasons not yet fully understood, the decline of the Wari Empire began around 1000 to 1100 CE in the terminal Wari era of the LIP (Tung, 2012). While it remains unclear what lead to the decline of the Empire it is certainly clear that Wari influence spread far and wide across the Andes. Similarly, the rise of skeletal tuberculosis during the LIP is still attempting to be understood. Previous reports summarizing identification of skeletal TB from archaeological populations across the Americas show less than five cases during the Middle Horizon, yet in there are over 68 published identifications of skeletal TB from the LIP (Roberts and Buikstra, 2003). However, it must be considered that the observance of an increase in cases of skeletal tuberculosis during the LIP may in fact be an artifact of archaeologically invisible populations which may not have been excavated or analyzed as of yet. Yet, if we consider the previously discussed factors associated with the emergence and spread of TB (i.e. malnutrition, stress, social and political instability, and population displacement) we find many of them are consistent with the events of the LIP (Tung, 2012; Tung et al., 2016; Fehren-Schmitz et al., 2014; Thompson, 2011).

Furthermore, the impressive reach the Wari had across the Andean landscape evidenced by cultural and artistic influences as well as constructions of sites and of road networks would support the movement of people and pathogens. Likewise, the terminal-Wari and post-Wari TB genomes illustrate that this MTBC strain, *M. pinnipedii,* had a geographic presence away from the coast and into the Andes within cultures that have no association with seal exploitation but are known to have culture of expansion and large networks of cultural influence. It is likely with the movement away from monumental and administrative centers to more dispersed settlements that disease was able to migrate with people and successfully spread across the Andean regions which would be consistent with archaeological evidence from Peru, Chile, and Colombia (Rivas Boada, 1988; Correal and Florez, 1992; Romero Arateco, 1998; Allison et al., 1981). While the existence of this strain in the Andean highlands may be the result of human-to-human transmission, more data and research are needed to confidently conclude

adaptation for human-to-human transmission and the existence of additional mammalian vectors remains a possibility. Tuberculosis has been identified in a number of animals and MTBC animal adapted strains do not appear to hold a strict sense of host-species specificity (Brites et al., 2018) and thus, further complicate investigations of MTBC origin and transmission in the pre-colonial Americas. Although, much remains to be understood and further research is needed, the detection of diverse MTBC strains at Huari likely reflects biological and cultural interplay with socio-cultural and environmental pressures influencing MTBC spread and evolution.

Conclusion

Our application of aDNA methods to address bioarchaeological questions concerning the health and disease of terminal and post-Wari communities and the rise of TB during the LIP not only support archaeological observations but, also, add to our understanding of population health and disease in the wake of the Wari Empire and the provide insight into the ecology and evolution of MTBC in the LIP Andes. This is the first study in which ancient MTBC diversity has been molecularly evaluated on a local scale. As such our work offers, for the first time, a view into the disease-scape, focused on TB, of a community experiencing the decline and wake of an expansive empire through the application of molecular methods. We present in this paper the largest number of ancient MTBC genomes from one site, more than doubling the number of pre-colonial MTBC genomes published to date. We demonstrate methods to overcome the challenge of environmental mycobacterial background and increase confidence that genetic variation observed in reconstructed ancient genomes derives from MTBC rather than environmental mycobacteria. Our contribution further adds to the current understanding of the geographic range human TB infections caused by *M. pinnipedii* and to the diversity in strains existing during the LIP.

Our results show evidence of microevolution of different contemporaneous clades recovered from populations that resided in a single site. Furthermore, our results show that in the wake of sociopolitical turbulence and climatic variability, tuberculosis was present in communities residing in what was once the administrative core of the Wari Empire. The social and environmental features of the LIP are known to promote this

very infectious, gradually debilitating and wasting disease. With the incorporation of previously published datasets we provide evidence for the existence of MTBC strains closely related to *M. pinnipedii* across contemporaneous yet variable environmental and geographic contexts throughout the Andes. The existence of ancient *M. pinnipedii* strain diversity recovered from archaeological communities experiencing climatic variability, disruption in socio-political systems, dietary changes, and shifts in settlement patterns suggests this observation may be the result of environmental and cultural transitions promoting the spread of TB. However, with additional ancient South American MTBC genomes from archaeologically and historically contextualized sources we will be able to shed further light on defining how cultural practices, environmental shifts, dietary change, and animal reservoirs play a role in the success of this devastating infectious disease.

Materials and Methods

Sectors of the site and samples. The skeletal samples come from populations of three separate sectors within the site of Huari: Cheqo Wasi (CW), Monqachayoq-Solano (MQS), and Vegachayoq Moqo (VM). These sectors are located within an approximately 1 km stretch of Huari (Supplementary image – map) and provide a time transect dating from the Terminal Wari era (CW, 1044 – 1155 CE) to the post-Wari era (1321 – 1425 CE for MQS and 1305-1405 CE for VM) of the LIP. Dating was performed using ASM and all samples were calibrated with OxCal, results are presented in detail in Supplementary Table 1. All individuals at these sectors were commingled, therefore, skeletal analyses were limited to a single element. The samples included a total of 92 vertebrae and 11 ribs with estimated age-at-death ranges from infant to adult. The selected skeletal elements included those which presented some pathological bony changes, including those consistent with TB and those with non-specific pathological changes, to those that presented no external macroscopically observable pathological change. This variation is described in previously published analyses (Nelson et al., 2020). Detailed site descriptions with a table of all screened samples and radiocarbon dating results can be found in the SI appendix.

Drilling of skeletal samples and extraction. All laboratory work was performed in ancient DNA dedicated facilities in Jena, Germany at the Max Planck Institute for the Science of Human History. A total of 92 vertebrae and 11 ribs were drilled using a dental drill to collect approximately 50mg of bone powder for DNA extraction. Drilling locations of the ribs were focused on the presence of bony exostoses and periosteal reactive bone that appeared consistent with descriptions of pathological changes to ribs associated with TB (Baker et al., 1999). These exostoses primarily appeared on the pleural surface of the ribs. The sample aspects of vertebrae are outlined in Supplementary Table 7.

Extractions were performed using a protocol customized for ancient DNA recovery (Dabney et al., 2013). The bone powder was collected in a 2ml tube and was suspended with 1ml of extraction buffer (0.45M EDTA, pH 8.0, and 0.25 mg/ml proteinase K) and incubated for approximately 18 hours at 37° Celsius. Upon being removed from incubation, the tubes were centrifuged, and the supernatant was collected from each sample. The remaining bone pellet was saved and stored at -20 ° Celsius. The supernatant was added to 10 ml of GuHCl-based binding buffer in High Pure Extender Assembly Roche silica membrane spin columns. The columns were spun for 9 min at 1500 rpm setting. Once the DNA was bound to the silica membrane, the DNA was then purified using the High Viral Nucleic Acid Kit (Roche). The DNA that was now bound to the silica membrane of the column and purified was eluted in 50ul of TET (10mM Tris-HCl, 1mM EDTA pH 8.0, 0.05% Tween20) with a saturation time of 2 minutes before centrifugation at 14,000 rpm for 1 min. The elution step was repeated to collect a total of 100ul elute. A sample of a cave bear bone of known DNA concentration served as the positive control for all extraction batches. Likewise, extraction blanks were included to with every batch to act as negative controls.

Library preparation. All extracts were first prepared as libraries without UDG treatment in order to authenticate the antiquity of the skeletal samples by observation of DNA damage consistent with ancient DNA. Ten µl of extracted DNA from each sampled were prepared as individual double-stranded DNA libraries following a modified protocol (Meyer and Kircher, 2010). This protocol begins with the repair of DNA fragment overhands using a blunt -end repair method whereby 10 µl of each

extract is treated with 10 µl of H_2O, 5 µl of NEB Buffer 2 (New England Biolabs), .2 µl of dNTP mix, 2 µl BSA, 5 µl ATP (10mM), 2 µl T4 PNK (polynucleotide kinase), and 0.4 µl T4 polymerase to equal 50 µl of total reaction for each sample. Each reaction was incubated at 15°C for 15 minutes, then at 25°C for 15 minutes in the thermocycler. The reactions were then subjected to a purification step in which they were then combined in a 1.5 ml LoBind tube with 600 µl of PB buffer (Qiagen). Each reaction was then loaded onto a MinElute column and incubated at room temperature for 1-2 minutes. The columns were then centrifuged for 30 seconds at 13000 rpm and upon completion the flow through supernatant was discarded. To remove salts and purify the DNA bound to the MinElute column we then added 650 PE wash buffer (Qiagen) to each column and immediately centrifuged these samples for 30 seconds at 13000 rpm. The flow through was again discarded and the columns were then "dry spun" in the centrifuge for 1 min at 13000 rpm, the columns were then rotated 180 ° and spun again to ensure the removal of all flow through. The column was then inserted into a new 1.5 ml LoBind tube and eluted in 20 µl of elution buffer containing .05% Tween. Adapter ligation was then performed by adding the elute to 20 µl of Quick Ligase Buffer, 1 µl of adapter mix, and 1 µl of Quick ligase to create an adapter ligation reaction that was incubated at 22 ° C for 20 minutes. The reaction was then purified as outlined above with a final elute of 22 µl of EB and Tween. Adapter fill in was then performed by adding 4 µl of Thermopol Buffer, 2.1 µl of dNTPs, 2 µl of Bst polymerase, and 13.8 of H_2O to each purified eluate of the ligation assay. This was then incubated at 37 ° C for 30 minutes and then 80 ° C for 10 minutes in the thermocycler. The process resulted in 40 µl of DNA library. An extraction blank and library blank were included for each batch to evaluate the cleanliness of the process and reagents used. The positive control sample from extraction was also carried forward. Libraries were then quantified with a qPCR reaction with 1 µl of DNA from each prepared ancient DNA library, 10 µl of Dynamo, 1 µl of each primer (IS7 and IS8), 7 µl of H_2O, and 1 µl of 12 DNA standards from 10^3 to 10^8 quantities along with two qPCR blanks. The remaining amount of prepared library was stored at -20 ° C for storage in the ancient DNA "clean lab". After quantification the prepared libraries were then barcoded individually giving each sample library a distinct 8 base pair identifier index for later reference, evaluation, and analysis of the DNA unique to that sample. Indexing was performed across two days, separating the library into two equal halves to ensure the successful indexing and

recovery of DNA. Indexing was performed using Pfu Turbo Cx Hotstart DNA Polymerase (Agilent) and modified to be most efficient by splitting the library to a number of reactions based on the concentration of DNA (a maximum of 2 x 10^{10} molecules of DNA) in the library to ensure the appropriate amount of index primers to be added so that all DNA fragments would be recovered. Libraries were split a minimum for 4 times and adjusted for higher concentrations for maximal efficiency of the indexing reaction. Based on a split of four the library would be divided across four reactions with each reaction receiving 10 µl of Pfu Turbo Buffer, 1.5 µl of BSA, 1 µl of dNTPs, 1 µl of Pfu Turbo Polymerase, 9 µl of H_2O and 2 µl of each primer (P5 and P7). The reaction was amplified with an initial denaturation at 95 ° C for 2 minutes and 10 cycles of 95 ° C for 30 seconds with a following 58 ° C for 30 seconds and 72 ° C for 1 minute ending with an elongation at 72 ° C for 10 minutes and holding the reactions at 10 ° C. All indexed reactions were then purified using a MinElute Kit using one column for a maximum of four reactions. All indexed libraries were quantified using qPCR with IS5 and IS6 primers to calculate the efficiency of the indexing reaction. The total result of all indexing resulted in 50 µl of index library. The prepared libraries were then amplified to 10nM and sequenced as paired-end 75 base pair reads using a HiSeq 4000 or sequenced as single-end 75 base pair using a Next-Seq 500.

MALT analysis and detection of ancient MTBC DNA. These libraries were sequenced without enrichment to allow for the screening of the samples microbial content in order to detect infectious agents, as well as, produce a metagenomic profile of the samples and thus evaluate the environmental background of non-pathogenic mycobacteria. Pathogen DNA screening was performed using the HOPS pipeline (Hübler et al., 2019) which incorporates the Megan Alignment Tool (MALT) (version 0.3.8) (Vågene et al., 2018) with evaluations by quality measures customized for ancient DNA and provides visual outputs for the detection of pathogens. Samples were screened using a minimum of 95 percent identity to be matched with an organism against the NCBI full nucleotide database (nt) (Nov. 2017) which includes bacteria, viruses, and eukaryotes. These methods are further outlined in our previously published paper discussing the shotgun pathogen screening of these samples (Nelson et al. 2020).

UDG treated library preparation. All extracts of samples in which MTBC DNA was

detected were prepared as double-stranded DNA libraries using a volume of 50 ul of DNA. The UDG treatment and library protocol differ in the first stage of library preparation where we first treated the libraries with uracil-DNA-glycosylase (UDG) to remove postmortem damage (Briggs et al., 2010). Using a total of 50 µl of DNA we then added to the reaction a master mix of 15 µl of NEB Buffer 2, 15 µl of ATP, .75 µl of BSA, 1.8 µl of dNTPs, 6 µl of T4 PNK, 9 µl of USER enzyme and 22.45 µl of H_2O. The reaction was incubated at 37 ° C for 3 hours. The reaction then received 6 µl of T4 polymerase to each library and was incubated for an additional 20 minutes at 25 ° C and 10 minutes at 12 ° C. The reaction was purified using the MinElute kit as previously described and the remainder of the protocol follows the traditional library preparation from adapter ligation forward. The UDG-treated libraries were quantified on a quantitative PCR (qPCR) using IS7/IS8 primer combinations. The quantification of DNA was used to calculate the necessary quantity of barcodes needed for each sample. Subsequently, libraries were divided into multiple PCR reactions for double indexing (Meyer and Kircher, 2010; Kircher et al., 2012) based on their initial quantification and measurement, in order to ensure maximal amplification efficiency. Indexing reactions were split across two days to safeguard against any error that may be introduced in the process. Index combinations consisting of unique 8 base pair identifier for each library was used to ensure identification of DNA from each sample. Index combinations were ligated to DNA library molecules and run in a 10-cycle amplification reaction described above. All reactions were purified using the MinElute DNA purification kit as previously described, and eluted in TET (10mM Tris-HCl, 1mM EDTA pH 8.0, 0.05% Tween20). Indexed libraries were then quantified on a qPCR using IS5/IS6 primer combinations and the indexing efficiency was calculated to ensure maximum efficiency. The indexing procedure was the repeated on the remaining UDG library and efficiency was calculated and the two indexed libraries were combined. After indexing all libraries, the concentration of each library was measured and then calculations for amplifications were performed to split each library across 3 reactions so the combined products would meet the required concentration of 200-300 ng/ul for in-solution TB capture. Amplification was performed on the indexed libraries using Herculase II Fusion DNA Polymerase (Agilent). The amplified indexed libraries were subsequently purified using the MinElute DNA purification kit (Qiagen), and eluted in TET (10mM Tris-HCl, 1 mM EDTA pH 8.0, 0.05% Tween20).

In-solution Whole genome MTBC capture. In-solution MTBC capture was performed using single stranded DNA probes computationally designed to reflect the genome of the MTBC ancestor (MTB_anc) (Comas et al., 2010) by reverting phylogenetically informative derived positions to the ancestral alleles in the H37Rv reference genome (NC_000962.1). The reconstructed ancestor is equidistant to all MTBC lineages with the exception of rate variation and thereby provides the breadth of sequence diversity of the MTBC. Capture probes are 60 base pairs in length with 52 bases as the complimentary MTBC sequences and 8 bases as a linker sequence (5' CACTGCGG 3'). Capture probes have a 5 bp overlap and upon removal of duplicate and low complexity probes, provide a set of unique 852,164 probes. Preparation of the capture was carried out according to previously published protocols (Fu et al., 2013; Vågene et al., 2018).

MTBC enrichment was carried out on all samples for two rounds of capture with a positive control (58U) of known ancient TB quantity previously demonstrated (Bos et al., 2014). Each sample was captured separately in a single well of a 96 well plate. Blanks with non-overlapping index combinations were captured together in a single well on the same plate.

Sequencing data of captured products. All captured libraries were sequenced on Illumina platforms as either paired-end 75 bp or 50 bp reads using a HiSeq 4000 or sequenced as single-end 75 bp using a Next-Seq 500. The sequence data were then sorted, assigned, and demultiplexed. Using the EAGER pipeline (v.1.92.55) (Peltzer et al., 2016) reads were prepared for analysis beginning with removal of Illumina adapters using AdapterRevmoal v2 (Schubert et al., 2016). All data from each TB positive sample were concatenated after adapter removal and merging (if paired-end). The pooled data were mapped to the MTBC ancestor reference genome with BWA-aln (v. 0.7.12)(Li and Durbin, 2009) using strict mapping parameters (-l 32, -n 0.1,) for TB captured samples and positive controls and reads with low quality mapping were removed using SAMtools quality filtering (-q 37). Relaxed mapping parameters and lower threshold for SAMtools quality filtering (-q 24) were used for negative controls (-l 16, -n 0.01). The removal of duplicate reads was performed using MarkDuplicates.

A small subset of captured libraries that were not treated the UDG and thus the deamination of cytosine to uracil would remain. Therefore, these libraries were evaluated using mapDamage2.0 (Jónsson et al., 2013) in order to visualize DNA damage patterns (Supplementary Figure 2). All reconstructed ancient MTBC genomes which showed at least 5-fold mean coverage were carried forward for further analyses.

Removing reads from contaminant reads, evaluating genomes for heterozygous calls and authentication of informative positions. All MTBC genomes reconstructed from UDG treated libraries with a minimum of 5 -fold coverage were included in an evaluation of variant positions to validate authenticity of the genetic variation observed. We estimated the number of heterozygous variant calls within our reconstructed genomes by considering the "haploid" nature of *Mycobacterium tuberculosis* complex members, in which we assume "heterozygous" SNPs are the result of closely related environmental bacteria mapping to the TB reference or evidence of a mixed infection. We performed SNP calling with the UnifiedGenotyper in GATK (DePristo et al., 2011). We then constructed a table of all variant positions, including heterozygous positions, across our dataset using MultiVCFAnalyzer v0.85 (Bos et al., 2014). Heterozygous variant positions were called at a minimum of 5-fold coverage. Using R v3.4.1 we produced histograms of allele frequencies for all SNPs with allele frequencies between 10-90% read support for all reconstructed genomes. (Supplement Figure 3).

Using the SnpEvaluation tool (https://github.com/andreasKroepelin/SNP_evaluation) we filtered out any positions originating from contaminant environmental mycobacteria and to authenticate the informative variant positions as deriving from MTBC members so they may be confidently included in phylogenetic analysis. For this evaluation we followed the use of this tool as outlined by Keller et al., (2019) with some customization for MTBC based analyses whereby we used a relaxed threshold ratio of 1.1. For this analysis we mapped sequencing reads using BWA (Li and Durbin, 2009) twice: all Huari TB positive samples to the MTBC ancestor reference were mapped first using strict mapping parameters (-n 0.1, -l 32) that allow for few mismatched positions and secondly mapped using relaxed mapping parameters (-n 0.01, -l 32) which allow for more mismatched positions. Using the snpTable generated by MutliVCFAnalyzer we compared VCF files of the two mappings within the SNPEvaluation tool, evaluating each position within a 50-bp window, observing for coverage within the window,

additional heterozygous position, and the coverage of the position in both strict and relaxed mapping parameters. Positions would be considered of low confidence and filtered out if within the 50-bp window there were regions that were not covered, there were heterozygous positions, or the ratio of mean coverage between strict and relaxed mapping was higher than 1.1 (Supplementary Table 9). The evaluation criteria were consistent for both shared and unique positions within each genome.

Following the removal of contaminant of questionable SNPs, we again produced histograms of allele frequencies illustrating the number of heterozygous sites to compare the effect of SNP filtering using the same parameters on the genomes (Supplementary Figure 3).

Phylogenetic Analysis. All reconstructed MTBC genomes recovered from Huari were included in phylogenetic analysis after the filtering of questionable SNPs and authentication of informative positions through the use of SNPEvaluation. The Huari MTBC genomes were analyzed together with using a dataset including the three previously published ancient coastal TB genomes (Bos et al., 2014) and 20 modern MTBC genomes: 11 *M. microti* genomes and 9 *M. pinnipedii* genomes. We performed calling of variant SNPs within our reconstructed genomes using the tool UnifiedGenotyper (v. 3.5) from the Genome Analysis Toolkit (GATK) (DePristo et al., 2011). Using MultiVCFAnalyzer (https://github.com/alexherbig/MultiVCFAnalyzer) we produced tables for downstream SNP analysis including a table of all variant positions across all genomes and a SNP alignment for phylogenetic tree construction. Our parameters selected for a minimum of five-fold coverage of all variant positions and a minimum genotype quality of 30. In addition to these quality measures, our SNP analysis and construction of the SNP alignment also included the removal of regions such as repetitive elements, tRNAs, mRNAs, and rRNAs. The Maximum Likelihood (ML) tree was constructed using RAxML v.8 (Stamatakis et al., 2014) using the GTRGAMMA substitution model. The ML tree was generated using the SNP alignment with authenticated informative positions generated by MultiVCFAnalyzer (Bos et al., 2014) with criteria of a minimum 5-fold coverage of each genome and incorporating all of the data (using the full SNP alignment = 2,026 sites) with inclusion of missing and ambiguous (N) sites.

Phylogenetic trees were constructed on both filtered and unfiltered datasets using consistent parameters to compare the effect of SNP filtering on phylogenetic analysis and tree topology (Supplementary Figure 4).

Functional Analysis of SNPs and Adaptive Signals. All variant positions in Huari MTBC genomes were analyzed for functional changes using output generated by MultiVCFAnalyzer (snpTableForsnpEff.tsv). This analysis included the dataset used for phylogenetic analysis (n=30) (Supplementary Table 11). Variant positions were called at a minimum of 5-fold coverage with at least 90% read support. All positions that were previously authenticated in the SNP filtering process described above, were included in the SNP table for analysis. Functional analysis was performed using SNPEff (v 3.1) (Cingolani et al., 2012) whereby the annotation for variant calls within all genomes were generated by including an annotated custom reference database of protein-coding and non-protein coding genes of the MTBC ancestor (MTB_anc) reference. Special attention was given to NON_SYNONYMOUS_CODING, STOP_GAIN and START_LOST. Investigation of upstream or downstream variants was limited to 100 bp (-ud 100). Both positions unique to each genome and positions shared across Huari genomes were investigated.

Bayesian molecular dating of MTBC. Molecular dating analysis was performed on the snpAlignment.fasta output from MultiVCF analyzer generated for 30 genomes (2026) sites that consist of ancient genomes recovered from human samples and modern *M. pinnipedii* and *M. microti* isolated from animals. Dates used in the molecular dating analysis were calculated using the mean values of AMS radiocarbon dates calibrated (cal) using OxCal. In the case of HUA051, no AMS dates were provided therefore dates were assumed (Supplementary information). We performed Coalescent Bayesian Skyline analysis using Beauti setup version 2.5.2. We used the Gamma site model with a default substitution rate (1.0) and gamma category count of 8 with a shape estimate of 1.0, proportion invariant of 0.1 (estimate) and substitution model – GTR (all 1.0 and estimated, rate CT is constant 1.0). Estimations were created using a relaxed clock log with -1 as the number of discrete rates (normalized) and a clock rate of 4.6^{-8} based on Bos et al., 2014. Three independent chains were run

with lengths of 300,000,000. Once all three chains converged, log chains were not combined and the log files were analyzed using Tracer and the Bayesian Skyline analysis was performed using BEAST2 version 2.5.2 (Bouckaert et al., 2019).

ACKNOWLEDGEMENTS. We thank the Ministerio de Cultura in Ayacucho, Peru and to the Universidad Nacional de San Cristóbal de Huamanga for approving requests to examine collections. We are grateful to Dr. Johannes Krause, Andres Köplin, Aida Andrades Valteuña, Cody Parker, Ron Huebler, Dr. Felix M. Key and all members of the Department of Archaeogenetics, Max Planck Institute for the Science of Human History for support and valuable feedback in this project. We thank the Max Planck Institute for the Science of Human History laboratory staff for their assistance. Thank you to Marcel Keller of the University of Estonia for support and advice in SNP analysis. Thank you to the many participants of the Ayacucho Bioarchaeology Project who have contributed to the research of these post-Wari communities. We thank the National Science Foundation-Archaeology and Biological Anthropology Divisions (grant number 1420757 to TAT) for financial support of the bioarchaeological aspects of the project. Genetic work was funded by the Max Planck Society and European Research Council Starting Grant CoDisease (grant number 805268 to KIB).

Figures 1 - 3

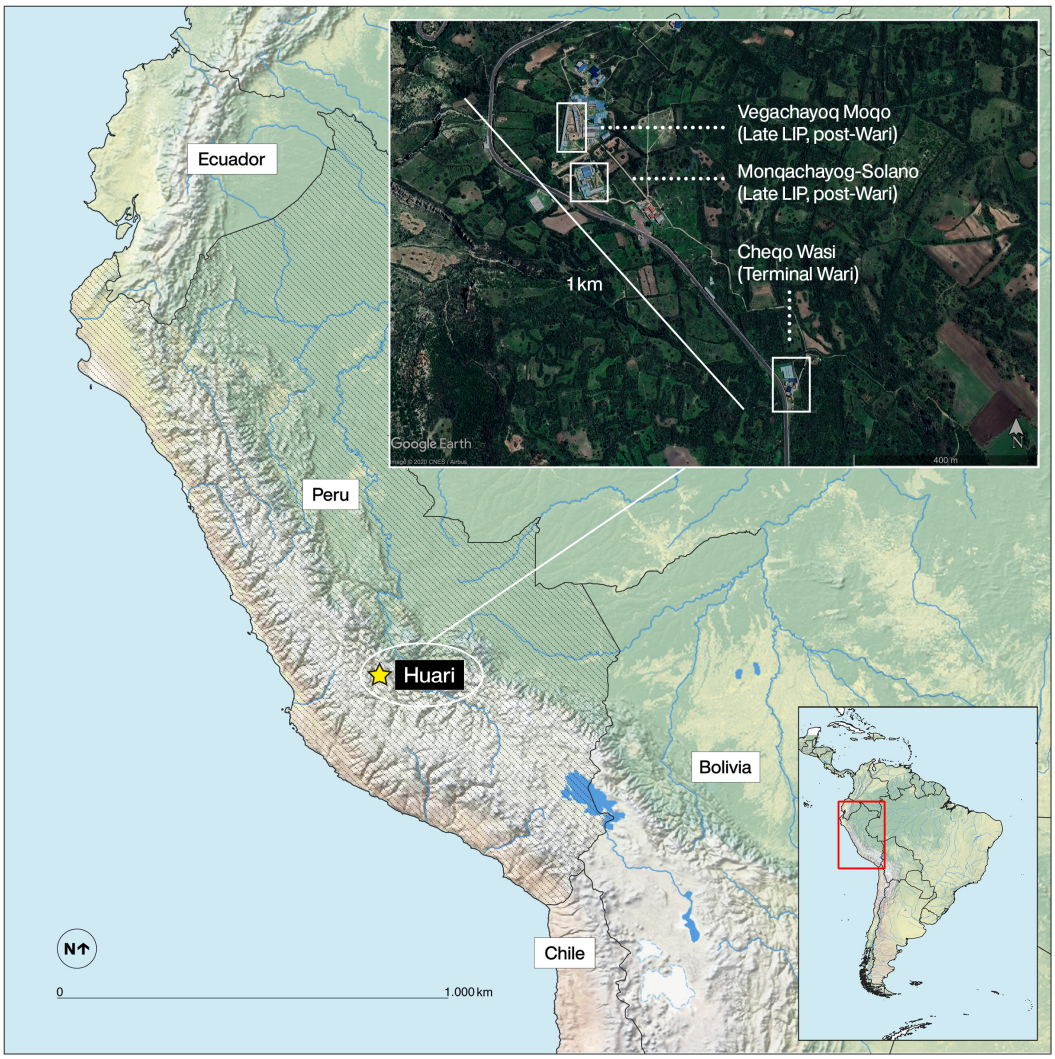


Figure 1. Map of the location of the site of Huari, the capital of the Wari Empire. The inset shows the location of the three LIP sectors of Huari from which skeletal samples were included in this research: 1) Cheqo Wasi 1044 – 1155 CE 2) Vegachayoq Moqo 1305 CE – 1405 CE and 3) Monqachayoq 1321 CE – 1425 CE. These sectors cover a time transect across the LIP and are within approximately 1 km of each other.

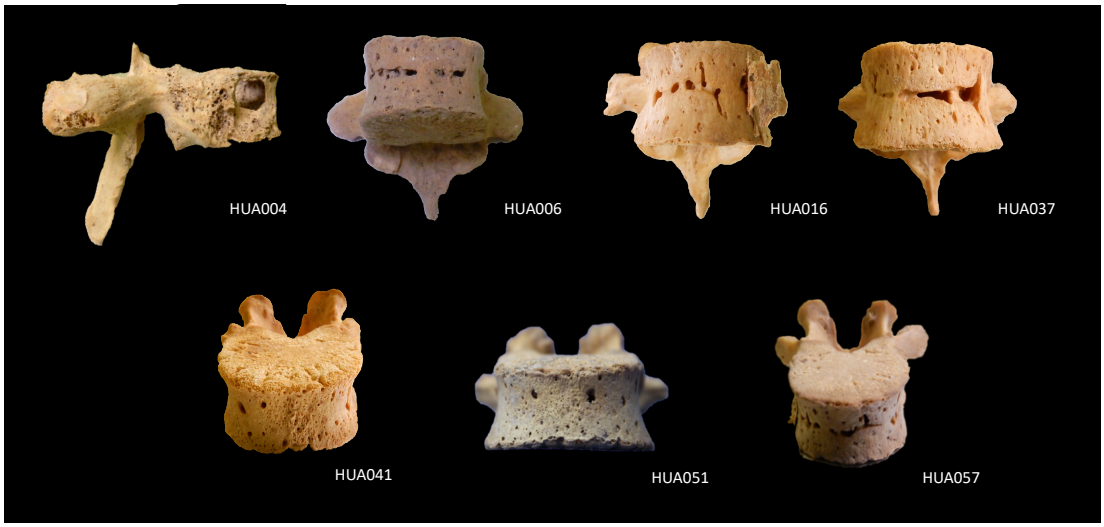


Figure 2. Skeletal elements of the reconstructed genomes. The seven vertebrae from Huari positive for MTBC from which we extracted DNA and reconstructed ancient MTBC genomes. Each sector and time period are represented: Cheqo Wasi (HUA004), Vegachayoq Moqo (HUA006, HUA016, HUA052, HUA057) and Monqachayoq-Solano (HUA037, HUA041).

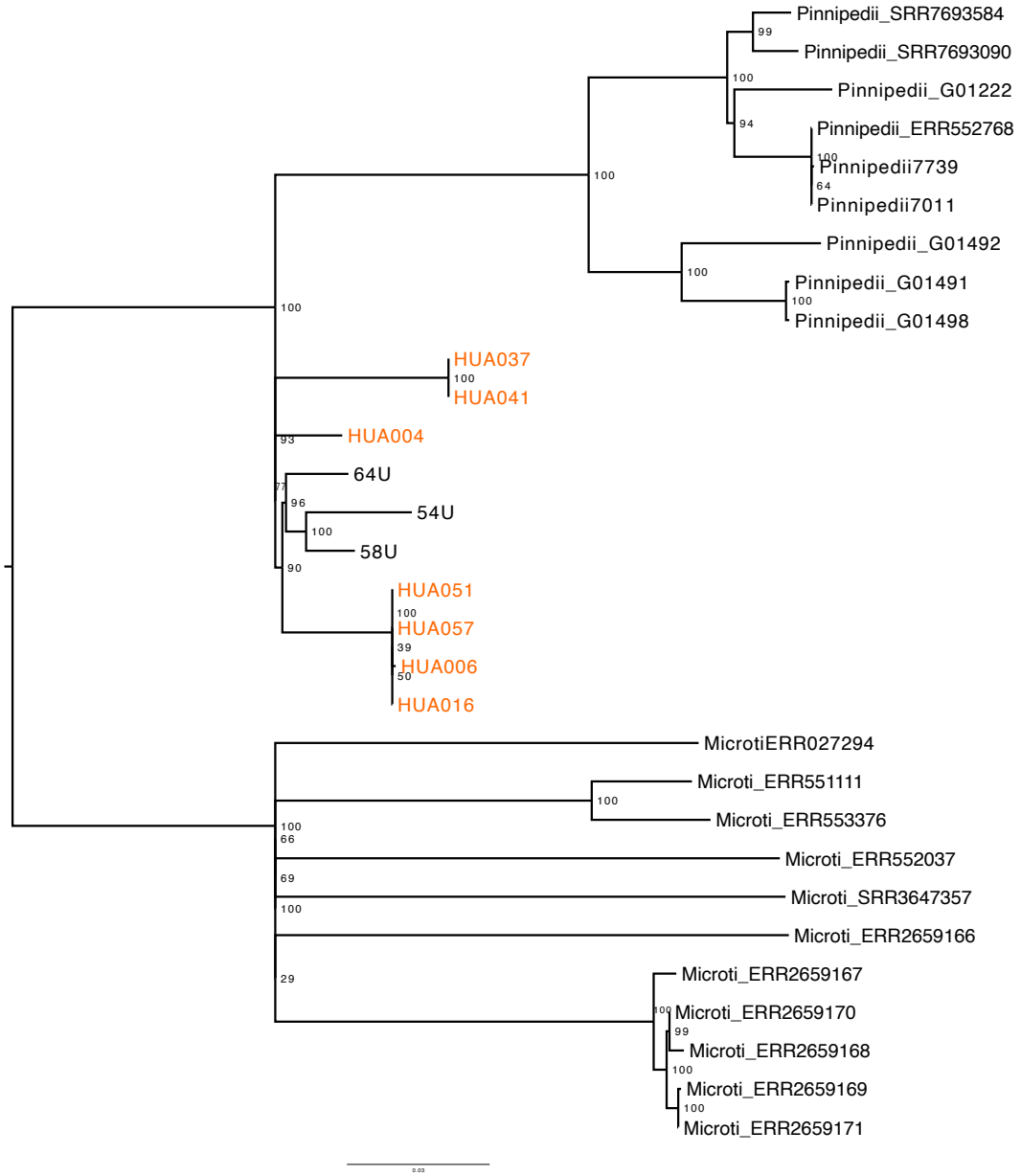


Figure 3. Phylogenetic Tree. Maximum Likelihood tree with full SNP alignment of 2,026 variant positions from a set of 30 genomes including 11 modern *M. microti* genomes, 9 modern *M. pinnipedii* genomes, three ancient LIP TB genomes (64U, 54U, 58U) (Bos. et al., 2014) and the 7 ancient TB genomes from Huari marked in orange. The node labels are showing the bootstrap values (1000 iterations). The *M. microti* subclade was used as the outgroup.

Table 1. Mapping statistics of positive sample non-UDG libraries to the Tb ancestor reference (MTB_anc). For expanded and full mapping statistics of all samples included in study using EAGER see SI Table 3. Samples HUA002-HUA041 were analyzed in Nelson et al., 2020.

Sample Name	Sector of Huari	Raw sequencing reads	Unique mapped reads, quality filtered	Endogenous DNA (%)	Mean fold coverage	DMG 1st Base 3'	DMG 2nd Base 3'	DMG 1st Base 5'	DMG 2nd Base 5'	Average Fragment Length	GC Content (%)	MALT Assigned Reads to MTBC
HUA002	**Cheqo Wasi**	10125640	1883	0.04	0.02	0.09	0.05	0.06	0.04	44	62.32	104
HUA004	**Cheqo Wasi**	13565300	2547	0.04	0.03	0.08	0.06	0.07	0.04	45	61.39	174
HUA006	**Vegachayoq Moqo**	12484864	3257	0.05	0.04	0.03	0.03	0.04	0.03	56	61.37	620
HUA016	**Vegachayoq Moqo**	12318062	2242	0.05	0.03	0.07	0.11	0.08	0.07	53	58.62	287
HUA024	**Vegachayoq Moqo**	12603366	1403	0.02	0.02	0.06	0.09	0.05	0.06	55	58.94	41
HUA025	**Vegachayoq Moqo**	11738298	2329	0.04	0.03	0.03	0.04	0.04	0.03	53	60.54	71
HUA037	**Monqachayoq-Solano**	10410440	3432	0.07	0.05	0.04	0.04	0.05	0.03	58	61.87	1163
HUA041	**Monqachayoq-Solano**	11097868	2658	0.05	0.03	0.03	0.04	0.03	0.04	54	61.81	566
HUA050	**Vegachayoq Moqo**	6204936	1604	0.05	0.02	0.05	0.06	0.04	0.05	56	58.23	24
HUA051	**Vegachayoq Moqo**	13130294	1131	0.02	0.02	0.06	0.06	0.02	0.04	62	60.09	316
HUA055	**Vegachayoq Moqo**	11539042	1480	0.03	0.02	0.06	0.07	0.08	0.08	59	58.92	13
HUA057	**Vegachayoq Moqo**	12227372	2536	0.05	0.04	0.05	0.07	0.05	0.07	64	59.16	318
HUA080	**Monqachayoq-Solano**	5648967	1491	0.04	0.02	0.03	0.03	0.04	0.03	45	61.45	28
HUA093	**Monqachayoq-Solano**	5374126	16993	0.40	0.17	0.05	0.04	0.06	0.04	45	63.38	8308

Table 2. Mapping statistics of reconstructed Huari MTBC genomes mapped to the TB ancestor (MTB_anc). For full mapping statistics from EAGER see SI Table 8.

Sample Name	Site	Raw sequencing reads	Unique mapped reads, quality filtered	Endogenous DNA (%)	Mean fold coverage	Average fragment length	GC content (%)
HUA004	**Huari - Cheqo Wasi**	60085834	381968	2.034	5.3077	61.3	62.02
HUA006	**Huari - Vegachayoq Moqo**	47413794	766883	3.376	10.8606	62.48	62.25
HUA016	**Huari - Vegachayoq Moqo**	59610950	2431074	9.402	33.5289	60.84	63.09
HUA037	**Huari- Monqachayoq-Solano**	66361797	473535	2.206	6.4448	60.04	62.23
HUA041	**Huari- Monqachayoq-Solano**	59637032	1026711	4.135	13.4027	57.59	62.51
HUA051	**Huari - Vegachayoq Moqo**	29438650	541805	2.697	8.4115	68.49	62.36
HUA057	**Huari - Vegachayoq Moqo**	34667789	599142	2.417	10.8851	80.15	62.34

References

Alcalde, C. (2004). Leaders, healers, laborers, and lovers: Reinterpreting women's roles in Moche society. In Ungendering Civilization (pp. 150–169). Routledge.

Allison, M. J., Gerszten, E., Munizaga, J., Santoro, C., & Mendoza, D. (1981). Tuberculosis in pre-Columbian Andean populations. Prehistoric Tuberculosis in the Americas, 49–61.

Arateco, W. M. R. (1998). Mal de Pott en momia de la colección del museo arqueológico Marqués de San Jorge. Maguaré, 0(13), 99–115.

Arriaza, B. T., Salo, W., Aufderheide, A. C., & Holcomb, T. A. (1995). Pre-Columbian tuberculosis in northern Chile: Molecular and skeletal evidence. American Journal of Physical Anthropology, 98(1), 37–45.

Blaser, M. J., & Kirschner, D. (2007). The equilibria that allow bacterial persistence in human hosts. Nature, 449(7164), 843–849. https://doi.org/10.1038/nature06198

Bos, K. I., Harkins, K. M., Herbig, A., Coscolla, M., Weber, N., Comas, I., Forrest, S. A., Bryant, J. M., Harris, S. R., & Schuenemann, V. J. (2014). Pre-Columbian mycobacterial genomes reveal seals as a source of New World human tuberculosis. Nature, 514(7523), 494–497.

Bos, K. I., Kühnert, D., Herbig, A., Esquivel-Gomez, L. R., Andrades Valtueña, A., Barquera, R., Giffin, K., Kumar Lankapalli, A., Nelson, E. A., & Sabin, S. (2019). Paleomicrobiology: Diagnosis and evolution of ancient pathogens. Annual Review of Microbiology, 73, 639–666.

Bouckaert, R., Heled, J., Kühnert, D., Vaughan, T., Wu, C.-H., Xie, D., Suchard, M. A., Rambaut, A., & Drummond, A. J. (2014). BEAST 2: A software platform for Bayesian evolutionary analysis. PLoS Comput Biol, 10(4), e1003537.

Briggs, A. W., Stenzel, U., Meyer, M., Krause, J., Kircher, M., & Pääbo, S. (2010). Removal of deaminated cytosines and detection of in vivo methylation in ancient DNA. Nucleic Acids Research, 38(6), e87–e87.

Brites, D., Loiseau, C., Menardo, F., Borrell, S., Boniotti, M. B., Warren, R., Dippenaar, A., Parsons, S. D. C., Beisel, C., Behr, M. A., Fyfe, J. A., Coscolla, M., & Gagneux, S. (2018). A New Phylogenetic Framework for the Animal-Adapted Mycobacterium tuberculosis Complex. Frontiers in Microbiology, 9. https://doi.org/10.3389/fmicb.2018.02820

Buikstra, J. E., & Williams, S. (1991). Tuberculosis in the Americas: Current perspectives. Human Paleopathology, Current Syntheses and Future Options, Edited by D. Ortner and A. Aufderheide, Págs, 161–172.

Burgess, S. (1992). Health and El Algodonal: A preliminary report. Annual Meeting of the Society for American Archaeology, Pittsburgh, Pennsylvania.

Burgess, S. D. (1992). Chiribayan Skeletal Pathology on the South Coast of Peru: Patterns of Production and Consumption. [Ph.D. dissertation]. University of Chicago.

Chan, J. Z.-M., Sergeant, M. J., Lee, O. Y.-C., Minnikin, D. E., Besra, G. S., Pap, I., Spigelman, M., Donoghue, H. D., & Pallen, M. J. (2013). Metagenomic analysis of tuberculosis in a mummy. New England Journal of Medicine, 369(3), 289–290.

Cingolani, P., Platts, A., Wang, L. L., Coon, M., Nguyen, T., Wang, L., Land, S. J., Lu, X., & Ruden, D. M. (2012). A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff. Fly, 6(2), 80–92. https://doi.org/10.4161/fly.19695

Comas, I., Chakravartti, J., Small, P. M., Galagan, J., Niemann, S., Kremer, K., Ernst, J. D., & Gagneux, S. (2010). Human T cell epitopes of Mycobacterium tuberculosis are evolutionarily hyperconserved. Nature Genetics, 42(6), 498–503. https://doi.org/10.1038/ng.590

Correal, G., & Flórez, I. (1992). Estudio de los guanes de la Mesa de Los Santos Santander (Colombia). In Actas del primer Congreso Internacional de Estudios Sobre Momias (pp. 307–313). Museo Arqueológico y etnográfico de Tenerife.

Cowie, R. L., & Sharpe, J. W. (1997). Extra-pulmonary tuberculosis: A high frequency in the absence of HIV infection. The International Journal of Tuberculosis and Lung Disease, 1(2), 159–162.

DeJesus, M. A., Gerrick, E. R., Xu, W., Park, S. W., Long, J. E., Boutte, C. C., Rubin, E. J., Schnappinger, D., Ehrt, S., Fortune, S. M., Sassetti, C. M., & Ioerger, T. R. (2017). Comprehensive Essentiality Analysis of the Mycobacterium tuberculosis Genome via Saturating Transposon Mutagenesis. MBio, 8(1). https://doi.org/10.1128/mBio.02133-16

Falkinham, J. O. (2002). Nontuberculous mycobacteria in the environment. Clinics in Chest Medicine, 23(3), 529–551. https://doi.org/10.1016/s0272-5231(02)00014-x

Fehren-Schmitz, L., Haak, W., Mächtle, B., Masch, F., Llamas, B., Tomasto Cagigao, E., Sossna, V., Schittek, K., Isla Cuadrado, J., Eitel, B., & Reindel, M. (2014). Climate change underlies global demographic, genetic, and cultural transitions in pre-Columbian southern Peru. Proceedings of the National Academy of Sciences, 111(26), 9443. https://doi.org/10.1073/pnas.1403466111

Gagneux, S., DeRiemer, K., Van, T., Kato-Maeda, M., Jong, B. C. de, Narayanan, S., Nicol, M., Niemann, S., Kremer, K., Gutierrez, M. C., Hilty, M., Hopewell, P.

C., & Small, P. M. (2006). Variable host–pathogen compatibility in Mycobacterium tuberculosis. Proceedings of the National Academy of Sciences, 103(8), 2869–2873. https://doi.org/10.1073/pnas.0511240103

Gonzalo-Asensio, J., Malaga, W., Pawlik, A., Astarie-Dequeker, C., Passemar, C., Moreau, F., Laval, F., Daffé, M., Martin, C., Brosch, R., & Guilhot, C. (2014). Evolutionary history of tuberculosis shaped by conserved mutations in the PhoPR virulence regulator. Proceedings of the National Academy of Sciences, 111(31), 11491. https://doi.org/10.1073/pnas.1406693111

Griffin, J. E., Gawronski, J. D., DeJesus, M. A., Ioerger, T. R., Akerley, B. J., & Sassetti, C. M. (2011). High-Resolution Phenotypic Profiling Defines Genes Essential for Mycobacterial Growth and Cholesterol Catabolism. PLOS Pathogens, 7(9), e1002251. https://doi.org/10.1371/journal.ppat.1002251

Heggarty, P., & Beresford-Jones, D. (2012). Conclusion: A Cross-Disciplinary Prehistory for the Andes? Surveying the State of the Art. In Archaeology and Language in the Andes (p. 407).

Hübler, R., Key, F. M., Warinner, C., Bos, K. I., Krause, J., & Herbig, A. (2019). HOPS: Automated detection and authentication of pathogen DNA in archaeological remains. Genome Biology, 20(1), 1–13.

Huson, D. H., Beier, S., Flade, I., Górska, A., El-Hadidi, M., Mitra, S., Ruscheweyh, H.-J., & Tappu, R. (2016). MEGAN Community Edition—Interactive Exploration and Analysis of Large-Scale Microbiome Sequencing Data. PLoS Computational Biology, 12(6), e1004957–e1004957. PubMed. https://doi.org/10.1371/journal.pcbi.1004957

Isbell, W. H., McEwan, G. F., & Oaks, D. (1991). Huari Administrative Structure: Prehistoric Monumental Architecture and State Government. Dumbarton Oaks.

Jaffe, H. L. (1972). Skeletal lesions caused by certain other infectious agents. Metabolic, Degenerative and Inffammatory Diseases of Bone and Joints. Lea & Febiger, Philadelphia, Pa, 1015–1031.

Jennings, J. (2006). Understanding Middle Horizon Peru: Hermeneutic spirals, interpretative traditions, and Wari administrative centers. Latin American Antiquity, 17(3), 265–285.

Jónsson, H., Ginolhac, A., Schubert, M., Johnson, P. L. F., & Orlando, L. (2013). mapDamage2.0: Fast approximate Bayesian estimates of ancient DNA damage parameters. Bioinformatics, 29(13), 1682–1684. https://doi.org/10.1093/bioinformatics/btt193

Kay, G. L., Sergeant, M. J., Zhou, Z., Chan, J. Z.-M., Millard, A., Quick, J., Szikossy, I., Pap, I., Spigelman, M., Loman, N. J., Achtman, M., Donoghue, H. D., & Pallen, M. J. (2015). Eighteenth-century genomes show that mixed infections were common at time of peak tuberculosis in Europe. Nature Communications, 6(1), 6717. https://doi.org/10.1038/ncomms7717

Keller, M., Spyrou, M. A., Scheib, C. L., Neumann, G. U., Kröpelin, A., Haas-Gebhard, B., Päffgen, B., Haberstroh, J., i Lacomba, A. R., & Raynaud, C. (2019). Ancient Yersinia pestis genomes from across Western Europe reveal early diversification during the First Pandemic (541–750). Proceedings of the National Academy of Sciences, 116(25), 12363–12372.

Klaus, H. D., Wilbur, A. K., Temple, D. H., Buikstra, J. E., Stone, A. C., Fernandez, M., Wester, C., & Tam, M. E. (2010). Tuberculosis on the north coast of Peru: Skeletal and molecular paleopathology of late pre-Hispanic and postcontact mycobacterial disease. Journal of Archaeological Science, 37(10), 2587–2597.

Lešić, A. R., Pešut, D. P., Marković-Denić, L., Maksimović, J., Čobeljić, G., Milošević, I., Atkinson, H. D. E., & Bumbaširević, M. (2010). The challenge of osteo-articular tuberculosis in the twenty-first century: A 15-year population-based study. The International Journal of Tuberculosis and Lung Disease, 14(9), 1181–1186.

Lönnroth, K., Jaramillo, E., Williams, B. G., Dye, C., & Raviglione, M. (2009). Drivers of tuberculosis epidemics: The role of risk factors and social determinants. Social Science & Medicine, 68(12), 2240–2246.

Ngabonziza, J. C. S., Loiseau, C., Marceau, M., Jouet, A., Menardo, F., Tzfadia, O., Antoine, R., Niyigena, E. B., Mulders, W., & Fissette, K. (2020). A sister lineage of the Mycobacterium tuberculosis complex discovered in the African Great Lakes region. Nature Communications, 11(1), 1–11.

Orgeur, M., & Brosch, R. (2018). Evolution of virulence in the Mycobacterium tuberculosis complex. Current Opinion in Microbiology, 41, 68–75. https://doi.org/10.1016/j.mib.2017.11.021

Owen, B. (1993). A Model of Multiethnicity: State Collapse, Competition, and Social Complexity from Tiwanaku to Chiribaya in the Osmore Valley [Ph.D. dissertation]. University of Los Angeles.

Peltzer, A., Jäger, G., Herbig, A., Seitz, A., Kniep, C., Krause, J., & Nieselt, K. (2016). EAGER: Efficient ancient genome reconstruction. Genome Biology, 17(1), 60. https://doi.org/10.1186/s13059-016-0918-z

Pond, S. L. K., & Frost, S. D. (2005). Datamonkey: Rapid detection of selective pressure on individual sites of codon alignments. Bioinformatics, 21(10), 2531–2533.

Quilter, J. (1990). The Moche revolt of the objects. Latin American Antiquity, 42–65.

Rambaut, A., Drummond, A. J., Xie, D., Baele, G., & Suchard, M. A. (2018). Posterior summarization in Bayesian phylogenetics using Tracer 1.7. Systematic Biology, 67(5), 901.

Reindel, M., & Wagner, G. A. (2009). New technologies for archaeology: Multidisciplinary investigations in Palpa and Nasca, Peru. Springer Science & Business Media.

Reported TB in the US 2018 | Data & Statistics | TB | CDC. (2020, August 4). https://www.cdc.gov/tb/statistics/reports/2018/default.htm

Rivas Boada, A. M. (1988). Las patologías óseas en la población de Marín. Boletín de Arqueología de La Fian, 3(1), 3–24.

Roberts, C. A., & Buikstra, J. E. (2003a). The bioarchaeology of tuberculosis: A global perspective on a re-emerging disease. University Press of Florida.

Roberts, C. A., & Buikstra, J. E. (2003b). The bioarchaeology of tuberculosis: A global perspective on a re-emerging disease. University Press of Florida.

Rustad, T. R., Minch, K. J., Ma, S., Winkler, J. K., Hobbs, S., Hickey, M., Brabant, W., Turkarslan, S., Price, N. D., Baliga, N. S., & Sherman, D. R. (2014). Mapping and manipulating the Mycobacterium tuberculosis transcriptome using a transcription factor overexpression-derived regulatory network. Genome Biology, 15(11), 502. https://doi.org/10.1186/s13059-014-0502-3

Sabin, S., Herbig, A., Vågene, Å. J., Ahlström, T., Bozovic, G., Arcini, C., Kühnert, D., & Bos, K. I. (2020). A seventeenth-century Mycobacterium tuberculosis genome supports a Neolithic emergence of the Mycobacterium tuberculosis complex. Genome Biology, 21(1), 1–24.

Salo, W. L., Aufderheide, A. C., Buikstra, J., & Holcomb, T. A. (1994). Identification of Mycobacterium tuberculosis DNA in a pre-Columbian Peruvian mummy. Proceedings of the National Academy of Sciences, 91(6), 2091–2094.

Sassetti, C. M., & Rubin, E. J. (2003). Genetic requirements for mycobacterial survival during infection. Proceedings of the National Academy of Sciences, 100(22), 12989–12994. https://doi.org/10.1073/pnas.2134250100

Schittek, K., Mächtle, B., Schäbitz, F., Forbringer, M., Wennrich, V., Reindel, M., & Eitel, B. (2014). Holocene environmental changes in the highlands of the southern Peruvian Andes (14 S) and their impact on pre-Columbian cultures. Clim Past Discuss, 10, 1707–1746.

Schreiber, K. J. (1991). The association between roads and polities: Evidence for Wari roads in Peru. Ancient Road Networks and Settlement Hierarchies in the New World, 42–53.

Seltzer, G. O., & Hastorf, C. A. (1990). Climatic change and its effect on prehispanic agriculture in the central Peruvian Andes. Journal of Field Archaeology, 17(4), 397–414.

Solano-Gutierrez, J. S., Pino, C., & Robledo, J. (2019). Toxin–antitoxin systems shows variability among Mycobacterium tuberculosis lineages. FEMS Microbiology Letters, 366(1), fny276.

Spyrou, M. A., Bos, K. I., Herbig, A., & Krause, J. (2019). Ancient pathogen genomics as an emerging tool for infectious disease research. Nature Reviews Genetics, 20(6), 323–340. https://doi.org/10.1038/s41576-019-0119-1

Stamatakis, A. (2014). RAxML version 8: A tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics, 30(9), 1312–1313. https://doi.org/10.1093/bioinformatics/btu033

Tandon, H., Melarkode Vattekatte, A., Srinivasan, N., & Sandhya, S. (2020). Molecular and structural basis of cross-reactivity in M. tuberculosis toxin–antitoxin systems. Toxins, 12(8), 481.

Thompson, L. G., Mosley-Thompson, E., Davis, M. E., Zagorodnov, V. S., Howat, I. M., Mikhalenko, V. N., & Lin, P.-N. (2013). Annually Resolved Ice Core Records of Tropical Climate Variability over the Past ~1800 Years. Science, 340(6135), 945. https://doi.org/10.1126/science.1234210

Torvinen, E., Torkko, P., & Rintala, A. N., Helena. (2010). Real-time PCR detection of environmental mycobacteria in house dust. Journal of Microbiological Methods, 82(1), 78–84. https://doi.org/10.1016/j.mimet.2010.04.007

Toyne, J. M., Esplin, N., & Buikstra, J. E. (2020). Examining variation in skeletal tuberculosis in a late pre-contact population from the eastern mountains of Peru. International Journal of Paleopathology, 30, 22–34.

Tuli, S. M. (2016). Tuberculosis of the skeletal system. JP Medical Ltd.

Tung, T. A. (2008). Violence after imperial collapse: A study of cranial trauma among Late Intermediate period burials from the former Huari capital, Ayacucho, Peru. Nawpa Pacha, 29(1), 101–117.

Tung, T. A. (2012). Violence, ritual, and the Wari Empire: A social bioarchaeology of imperialism in the ancient Andes. University Press of Florida.

Tung, T. A., Miller, M., DeSantis, L., Sharp, E. A., & Kelly, J. (2016). Patterns of Violence and Diet Among Children During a Time of Imperial Decline and Climate Change in the Ancient Peruvian Andes. In A. M. VanDerwarker & G. D. Wilson (Eds.), The Archaeology of Food and Warfare: Food Insecurity in Prehistory (pp. 193–228). Springer International Publishing. https://doi.org/10.1007/978-3-319-18506-4_10

Vågene, Å. J., Campana, M. G., García, N. M. R., Warinner, C., Spyrou, M. A., Valtueña, A. A., Huson, D., Tuross, N., Herbig, A., & Bos, K. I. (2017). Salmonella enterica genomes recovered from victims of a major 16th century epidemic in Mexico. BioRxiv, 106740.

Valencia Zegarra, A. (2005). Wari hydraulic works in the Lucre Basin. In Pikillacta: The Wari Empire in Cuzco (pp. 85–97). University of Iowa Press.

Vogel, M. A. (2018). New Research on the Late Prehistoric Coastal Polities of Northern Peru. Journal of Archaeological Research, 26(2), 165–195. https://doi.org/10.1007/s10814-017-9108-0

Warinner, C., Herbig, A., Mann, A., Fellows Yates, J. A., Weiß, C. L., Burbano, H. A., Orlando, L., & Krause, J. (2017). A Robust Framework for Microbial Archaeology. Annual Review of Genomics and Human Genetics, 18(1), 321–356. https://doi.org/10.1146/annurev-genom-091416-035526

Williams, P., Nash, D., Cook, A., Isbell, W., & Speakman, R. (2019). Wari Ceramic Production in the Heartland and Provinces. Ceramics of the Indigenous Cultures of South America: Studies of Production and Exchange through Compositional Analysis, 125–133.

Yang, Z. (2007). PAML 4: Phylogenetic analysis by maximum likelihood. Molecular Biology and Evolution, 24(8), 1586–1591.

Tuberculosis in the wake of the Wari empire

SUPPLEMENTARY INFORMATION

Authors: Elizabeth A. Nelson[1,2], Aditya Kumar Lankapalli[1], Maria Spyrou[1], Susanna Sabin[1], Åshild Vågene[3], Ainash Childebeyeva[1], Ben Rohrlach[1], James A. Fellows Yates[1], Martha Cabrera[5], Jose Ochotama[5], Tiffiny A. Tung[4], Alexander Herbig[1], Kirsten I. Bos[1]

Affiliations:

[1]Department of Archaeogenetics, Max Planck Institute for the Science of Human History, Jena, Germany

[2]Institute for Archaeological Sciences, University of Tübingen, Germany

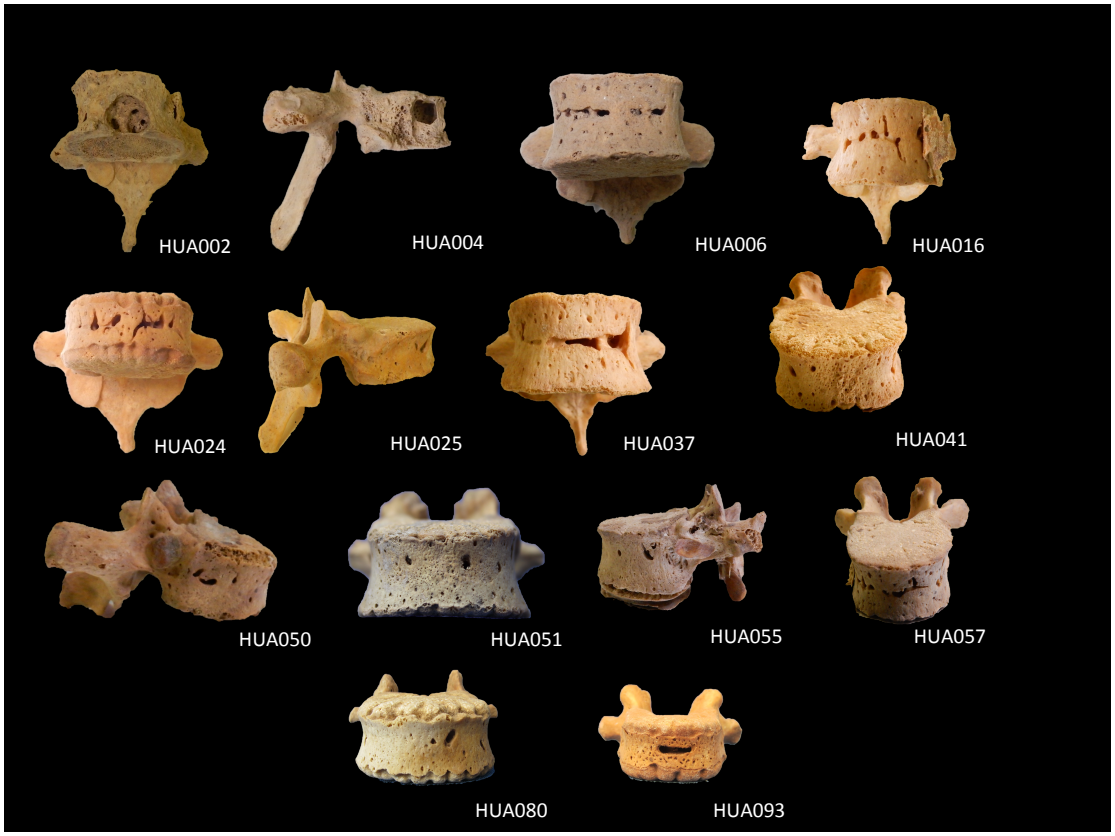
[3]Section for Evolutionary Genomics, The GLOBE Institute, Faculty of Health and Medical Sciences, University of Copenhagen, Denmark

[4]Department of Anthropology, Vanderbilt University, Tennessee, USA

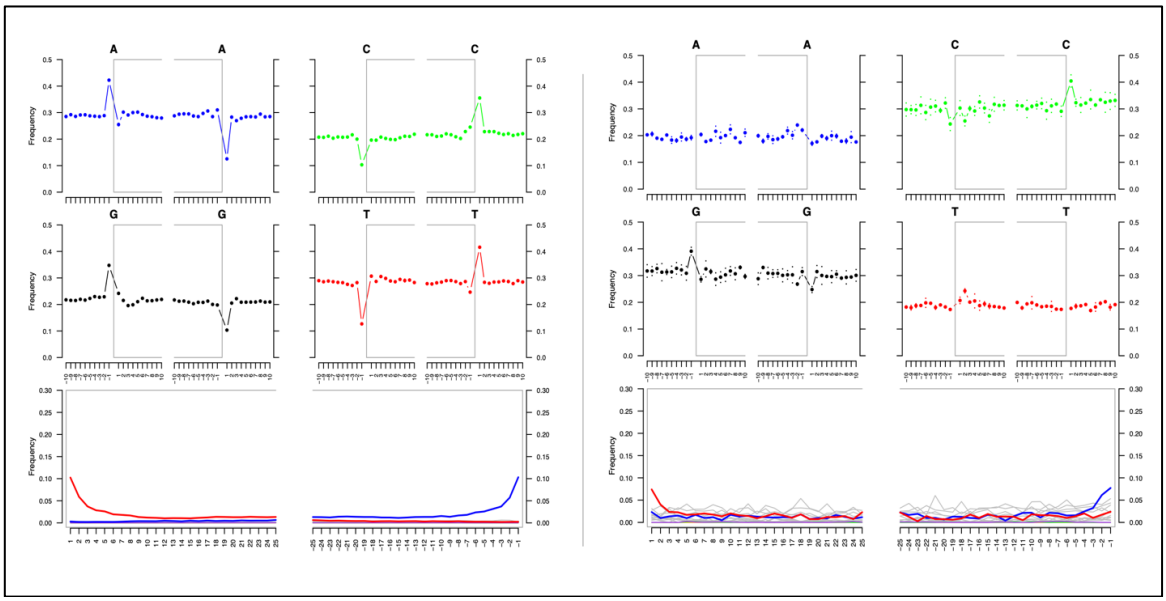
[5] Universidad Nacional San Cristóbal de Huamanga, Ayacucho, Peru

CONTENTS:

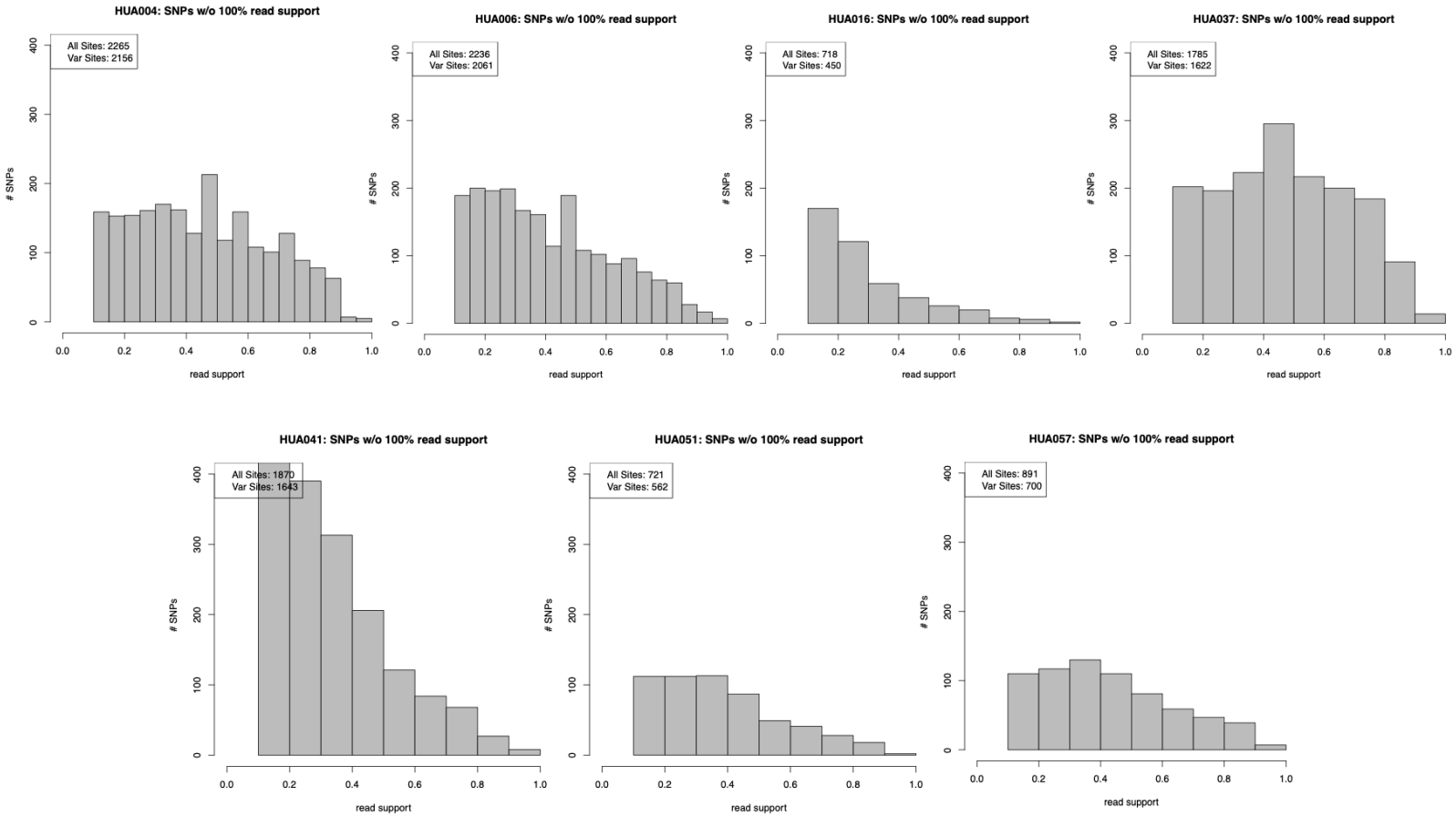
Refer to CD for Supplementary Tables 1- 13



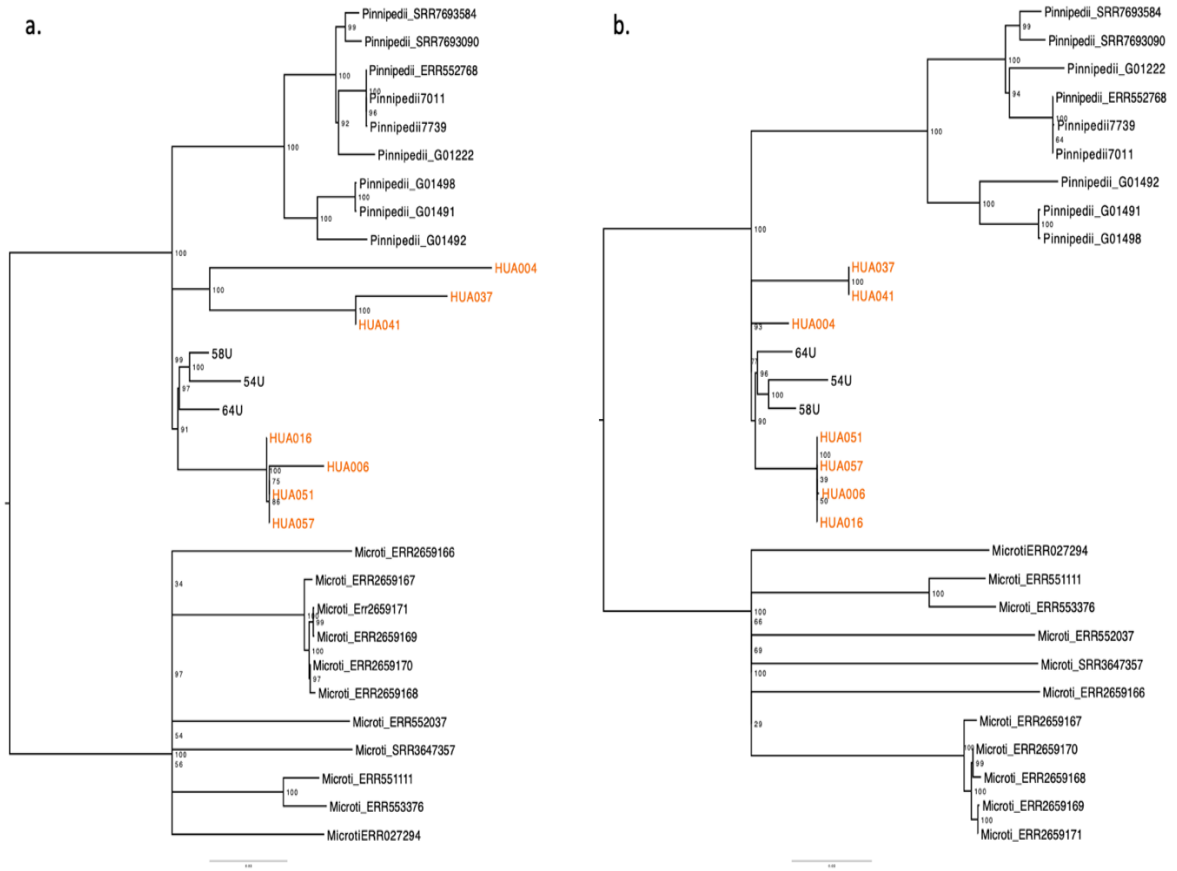
Supplementary Figure 1. The 14 samples from Huari in which MTBC DNA was detected using MALT. These include both sub-adult and adult vertebrae. Only HUA002 and HUA004 display typical skeletal lesions suggestive of tuberculosis. The remaining 12 vertebrae show variation in skeletal presentation with some hypervascularity (HUA016, HUA024, HUA037) that may be considered outside of normal variation, in addition to slight reactive bone (HUA041, HUA051), and some vertebrae that may be considered within normal human variation (HUA025, HUA050, HUA055, HUA057, HUA080, HUA093).



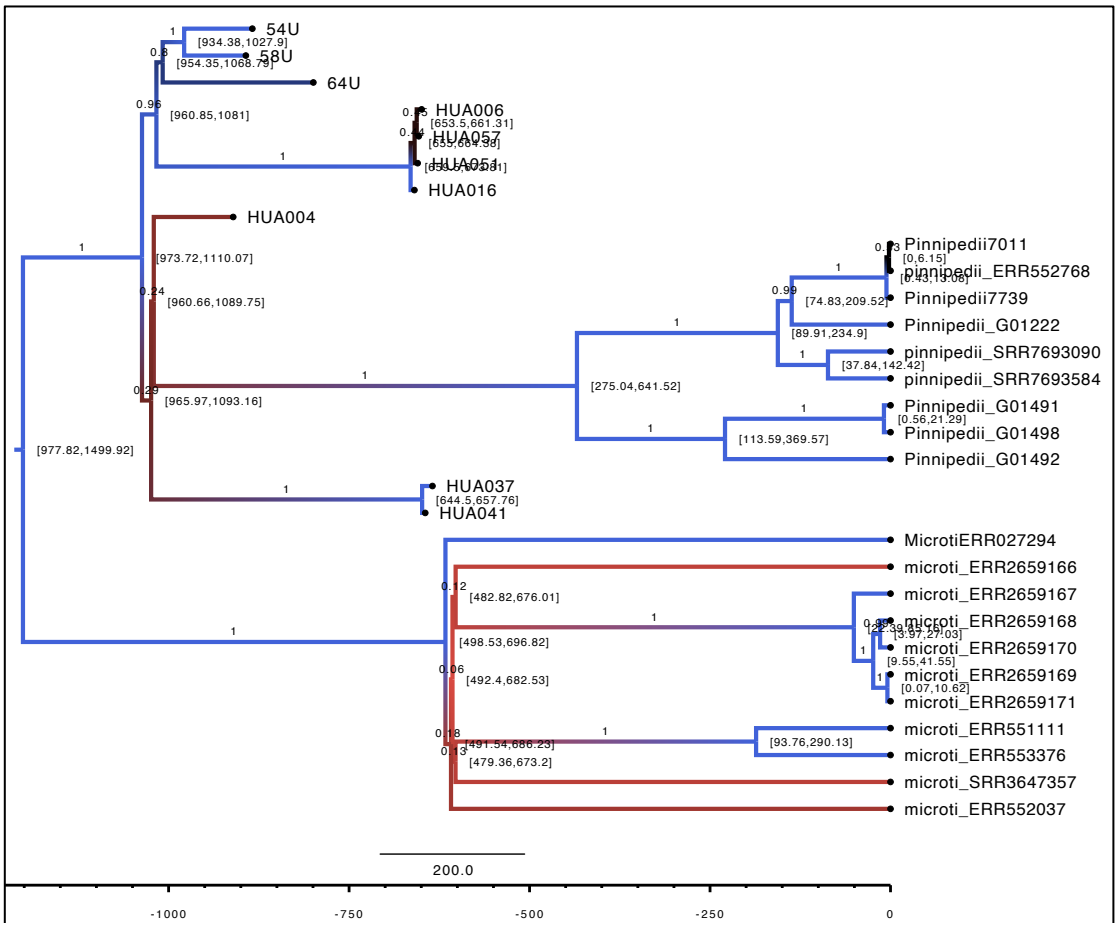
Supplementary Figure 2. DNA damage plots of HUA004. Damage plot generated from HUA004 mapping directly to the human reference (Hg19) mapping (left) and to the MTBC ancestor (MTB_anc) (right) reference genome. Characteristic patterns of DNA damage are displayed for both analyses. Damage profiles were produced using mapDamage2.0 (Jónsson et al., 2013).



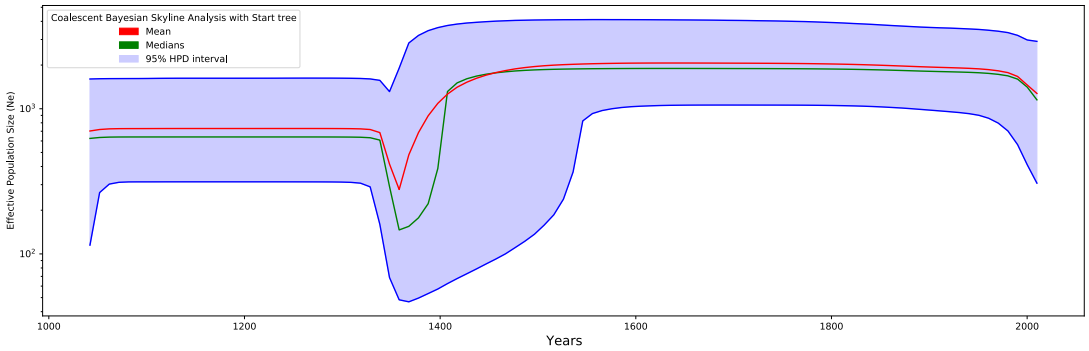
Supplementary Figure 3. Histograms of SNP allele frequency distributions for the Huari TB genomes. The x-axis represents the percent of read support, or frequency of reads, covering a SNP position in which the SNP was detected. The y-axis represents the number of SNPs with the respect to the read support. : All Sites: represents the SNPs which deviate from the reference. : Var sites: are those SNPs that are considered multiallelic according to the threshold (below 90% homozygous). A reduction of SNPs occurs in the filtering process with a noticeable reduction in the number of multi-allelic positions.



Supplementary Figure 4. Phylogenetic trees before and after SNP evaluation and filtering. Our dataset prepared as maximum likelihood phylogenetic trees with both the a) unfiltered dataset and the b) filtered dataset after SNP evaluation and removal of non-confident positions. Here we show a reduction in branch length and correction of branches formed from false positive informative SNPs.



Supplementary Figure 5. The Maximum Clade Credibility tree for trees skyline with starting tree. The burn-in for the MCC tree is 10%. The 95% HPD date ranges are specified at the internal nodes. The branches are colored based on posterior probability (red [0.0] –blue [1.0], black [0.5]). The time scale is specified at the bottom, with the recent age 2010 (here as 0 and goes backwards in time). The dates are proportional to the age estimated from the present. The tree is not rooted.



Supplementary Figure 6. Bayesian skyline plot illustrating the demographic history of ancient MTBC and *M. microti* strains. The time is specified in years before the present on x-axis, and the change in effective population size (Ne) is specified on y-axis in log-scale. The mean (red) and median (green) line represents the estimates of effective population size through time. The blue area represented the 95% highest posterior density interval of the effective infections at time. This skyline plot indicates the demographic change starting from 2010. The population demographics appears to fluctuation between 1300 – 1400 CE, probably a bottleneck event, then followed by rapid population expansion between 1400 – 1500 CE that stabilizes throughout the modern era.

Supplementary Text

Archaeological Context

Our research focused on terminal and post-Wari populations within the site of Huari, the former administrative center of the Wari empire located in the 2900 meters above sea level in the Andean highlands of Peru within the Department de Ayacucho. Huari is a large site spanning 3-5 km in size when considering the central architecture but can be expanded to nearly 15 square kilometers when including habitations surrounding Huari. (Isbell et al., 1991; Tung, 2008). A total of 103 skeletal elements from Huari were included in this study were collected from the Universidad de San Cristobal de Huamanga archaeological collections in Ayacucho, Peru (Supplementary Table 2). We focused our study on terminal and post-imperial contexts from three sectors of the site: Cheqo Wasi (CW, Terminal Wari ca. 1000 – 1150 CE) and the contemporary Vegachayoq Moqo (VM) and Monqachayoq-Solano (MQ, post-Wari ca. 1200-1450). The two sectors of VM and MQ are centrally located within the site only 150 meters apart in distance and are some of the earliest constructions in Huari (Isbell, 1997) yet had continued use into the LIP. However, the samples we collected from both correspond to the post-Wari era of the last half of the Late Intermediate Period.

Cheqo Wasi. Cheqo Wasi was excavated by Benavides in the late 1880's. This sector was an area for high status individuals with ceremonial spaces and craft production. Four vertebrae were obtained from CW for this study, most likely reflecting elite individuals from the terminal Wari period experience imperial decline.

Vegachayoq Moqo. Vegachayoq Moqo was first excavated by Enrique Bragayrac in the 1980s (Bragayrac, 1991) and is one of the largest Huari excavations (González and Bragayrac, 1986;, Gonzales, 1997; Pérez Calderón, 1999). This sector has been classified as a ceremonial center with gateway access and bordered by stone worked walls containing a D-shaped structure and a deep courtyard or mesa for displaying ceremonial objects (Isbell, 1997). In the later years of the Wari, a long rectangular stone wall with burial niches was constructed on the eastern border of the courtyard where it still stands today. However, archaeological analyses report the burial program of the analyzed samples as varying from the traditional post-Wari mortuary program wherein the skeletons were commingled and deposited, or "dumped" (Kuzminsky et al., 2016) as secondary

burial bundles along the wall of the D-shaped structure located in Vegachayoq Moqo (González Carre et al., 1996; Tung ,2014).

Monqochayoq-Solano. Excavations at Monqachayoq were led by Francisco Solano in 1977-1978 (Solano and Guerrero 1981). This central site includes spacious underground gallery tombs, four of which have been excavated (Solano and Guerrero, 1981; Tung, 2008). Solano reports that two of the galleries which contained human remains were badly looted and disturbed (Solano and Guerrero 1981: 186; Tung, 2008). Therefore, the remains were commingled and provenance of original interment within the galleries is not known. However, multiple lines of evidence, including ceramic styles (Solano and Guerrero, 1981) and AMS dating (Tung, 2008), have revealed the continued use of the underground mortuary galleries into the post-Wari era of the LIP.

Detailed information on the archaeological contexts has been previously published (Tung, 2012; Bragayrac, 1991; Solano, 1981) and discussed in regard to sample selection for this project (Nelson et al., 2020).

Skeletal Analysis

All skeletal samples (vertebrae and ribs) displayed excellent preservation with little to no taphonomic damage. Each skeletal element was macroscopically analyzed using a 10x handheld loop and documented using skeletal recording standards as outlined by Buikstra and Ubelaker, 1994. Skeletal analysis included documentation of preservation estimation, cultural modification (cinnabar), and the presence of pathological changes including trauma. Estimations for age-at-death were performed by observing neural arch development and fusion, neural arch and centrum fusion, and epiphyseal appearance and fusion (Scheuer and Black, 2004; Baker et al., 2005). Because skeletal observations were limited to a single element, age estimations are limited and broad after reaching the age category of adult (A). Difference in appearance and fusion for the skeletal elements were considered when analyzing thoracic versus lumbar vertebrae (Scheuer and Black, 2004).

Pathological analysis was performed using standard pathology recording codes (Buikstra and Ubelaker, 1994) accompanied by detailed descriptions of lesion location, appearance, if the

process was active at time of death or healing, and extent of skeletal involvement (Buikstra and Ubelaker, 1994). In this paleopathological assessment we chose to include not only pathological changes consistent with advanced-stages of skeletal tuberculosis, such as penetrating lytic lesions, but also include minor osseous lesions which may be reactive bone associated with early stages of tuberculosis skeletal involvement. Likewise, we also chose to select skeletal elements displaying other skeletal changes that are not diagnostic but have more recently been attributed to tuberculosis as early stage osseous changes. These "early stage osseous changes" include: hypervascularization (abnormal severe pitting) of the vertebral body (Pálfi, 2002), hypertrophic osteoarthropathy (HOA) and periosteal development on the visceral aspect of the ribs and associated exostoses (Baker, 1999; Kelley and Micozzi, 1984; Roberts et al., 1998). Because we know skeletal manifestations of tuberculosis occur in small percentages of the affected population and likely reflect a much larger infected population, we also include a set of vertebrae from the same contexts which display vascularization of the body which may be considered within normal variation.

References

Baker, B.J., 1999. Early manifestations of tuberculosis in the skeleton. Tuberculosis: Past and Present.–Golden Book and Tuberculosis Foundation, Budapest-Szeged 301–311.

Baker, B.J., Dupras, T.L., Tocheri, M.W., 2005. The Osteology of Infants and Children. Texas A&M University Press.

Bragayrac, D.E., 1991. Archaeological excavations in the Vegachayoq Moqo sector of Huari. Huari administrative structure: Prehistoric monumental architecture and state government 71–80.

Buikstra, J.E., Ubelaker, D.H., 1994. Standards for data collection from human skeletal remains. 1994. Arkansas: Arkansas Archaeological Survey Research Series.

Buikstra, J.E., Williams, S., 1991. Tuberculosis in the Americas: Current Perspectives, in: Human Paleopathology: Current Syntheses and Future Options Edited by D. Ortner and A. Aufderheide.

Correal, G., Flórez, I., 1992. Estudio de los guanes de la Mesa de Los Santos Santander (Colombia)., in: Actas Del Primer Congreso Internacional de Estudios Sobre Momias. Museo Arqueológico y etnográfico de Tenerife, Santa Cruz de Tenerife, pp. 307–313.

Gonzáles Carré, E.G., Urrutia, J., Lévano, J., 1997. Ayacucho: San Juan de la Frontera de Huamanga. Banco de Crédito del Perú.

González Carré, E., Bragayrac Dávila, E., Vivanco Pomacanchari, C., Tiesler Blos, V., López Quispe, M., 1986. El Templo Mayor en la ciudad de Wari: estudios arqueológicos en Vegachayoq Moqo. Universidad Nacional de San Cristóbal de Huamanga, Ayacucho.

Isbell, W.H., 2004. Palaces and politics in the Andean Middle Horizon. Palaces of the Ancient New World 191.

Kelley, M.A., Micozzi, M.S., 1984. Rib lesions in chronic pulmonary tuberculosis. American Journal of Physical Anthropology 65, 381–386.
Kuzminsky, S.C., Tung, T.A., Hubbe, M., Villaseñor-Marchal, A., 2016. The application of 3D geometric morphometrics and laser surface scanning to investigate the standardization of cranial vault modification in the Andes. Journal of Archaeological Science: Reports 10, 507–513. https://doi.org/10.1016/j.jasrep.2016.11.007
Nelson, E.A., Buikstra, J.E., Herbig, A., Tung, T.A., Bos, K.I., 2020. Advances in the molecular detection of tuberculosis in pre-contact Andean South America. International Journal of Paleopathology.
Pálfi, G., 2012. Juvenile cases of skeletal tuberculosis from the Terry Anatomical collection (Smithsonian Institution, Washington, DC, USA). Acta Biologica Szegediensis 56, 1–12.
Pérez Calderón, I., 1999. Huari: misteriosa ciudad de piedra. Universidad Nacional de San Cristóbal de Huamanga, Facultad de Sociales Ciencias, Ayacucho.
Roberts, C.A., Boylston, A., Buckley, L., Chamberlain, A.C., Murphy, E.M., 1998. Rib lesions and tuberculosis: the palaeopathological evidence. Tubercle and Lung Disease 79, 55–60.
Scheuer, L., Black, S., 2004. The Juvenile Skeleton. Elsevier, London.
Solano Ramos, F., Anaya, V.G.,
1981. Estudio arqueológico en el sector de Monqachayoq-Wari. Unpublished bachelor's thesis, Universidad Nacional de San Cristóbal de Huamanga, Ayacucho.
Tung, T.A., 2012. Violence, ritual, and the Wari Empire: A social bioarchaeology of imperialism in the ancient Andes. University Press of Florida.
Tung, T.A., 2008. Violence after imperial collapse: A study of cranial trauma among Late Intermediate period burials from the former Huari capital, Ayacucho, Peru. Nawpa Pacha 29, 101–117.
Tung, T.A., Miller, M., DeSantis, L., Sharp, E.A., Kelly, J., 2016. Patterns of Violence and Diet Among Children During a Time of Imperial Decline and Climate Change in the Ancient Peruvian Andes, in: VanDerwarker, A.M., Wilson, G.D. (Eds.), The Archaeology of Food and Warfare: Food Insecurity in Prehistory. Springer International Publishing, Cham, pp. 193–228. https://doi.org/10.1007/978-3-319-18506-4_10

Paper IV

E. A. Nelson, E. Guevara, J. M. Toyne, A. Herbig, J. Krause, K. I. Bos (2020).

Tuberculosis in a Pre-colonial Chachapoya Individual from the northeastern Peruvian Andes

Manuscript

Tuberculosis in a Pre-colonial Chachapoya Individual from the northeastern Peruvian Andes

Authors: Elizabeth A. Nelson[1,2], Evelyn Guevara[1,3], J. Marla Toyne[4], Alexander Herbig[1], Johannes Krause[1], Kirsten I. Bos[1]

Affiliations:
[1]Department of Archaeogenetics, Max Planck Institute for the Science of Human History, Germany
[2]Institute for Archaeological Sciences, University of Tübingen, Germany
[3]Department of Forensic Medicine, University of Helsinki, Finland
[4]Department of Anthropology, University of Central Florida, USA

ABSTRACT

The application of molecular methods to paleopathology has revealed a strain of tuberculosis (TB) closely related to a variety currently adapted to seals and sea lions that caused human infection in the western Andes of pre-colonial South America. Our understanding of ancient TB distribution in terms of geography and genetic diversity is, however, limited since genome-level evaluations have thus far been restricted to only a small number of individuals of the western Andean region. Here we present a reconstructed ancient TB genome from the eastern Andean slopes recovered from the Chachapoya funerary site Diablo Huasi, located in the subtropical mountain forests of Amazonas, Peru. Because this skeleton displayed pathological changes that were non-specific in lesion distribution and morphology, we employed a broad pathogen screening method with limited ascertainment bias. TB was the only pathogen confidently detected and its preservation permitted the subsequent reconstruction of a greater than 15-fold TB genome. Our phylogenetic analysis reveals that the Diablo Huasi TB strain is closely related to those from the neighboring ancient coastal and highland populations. Our results demonstrate the utility of molecular methods for paleopathological analyses and expand the known geographic range of ancient TB strains thus providing insight into their local ecology and evolution.

INTRODUCTION

Molecular paleopathology is a synergistic field of study bringing together osteological and anatomical observations with molecular methods. Morphologically based paleopathological studies can be limited by factors which obfuscate the detection or diagnosis of infectious disease.

Archaeological and environmental conditions can negatively affect bone preservation or skeletal completeness prohibiting the development of a differential diagnosis. Furthermore, many pathogenic infections do not manifest in skeletal tissues or may result in unspecific osseous responses which are shared across multiple conditions (Buikstra, 2010). Likewise, infectious skeletal response may vary between hosts and/or populations. Application of molecular methods to paleopathology have served to assist in the identification of diseases which may have otherwise gone unidentified. The utility of molecular methods has been well demonstrated in challenging paleopathological cases which included remains which did not present any morphologically detectable pathology (Spyrou et al., 2016; Vågene et al., 2018) or displayed nonspecific skeletal changes (Schuenemann et al., 2018). Even more impressive is the application of NGS methods applied to molecular paleopathology which permits the identification of co-infections of ancient pathogens through the use of broad pathogen screening methods (Giffin et al., 2020). With the advent of Next Generation Sequencing (NGS) methods, ancient DNA paleopathology has begun to generate an abundance of informative data contributing to evolutionary analyses and providing insight to ancient human pathogen evolution and ecology. In American paleopathology, characterizing the pre-colonial disease-scape has taken spotlight to better understand host-pathogen coevolution and migration in the Americas prior to European colonization.

The study of ancient tuberculosis (TB) in pre-colonial American contexts serves as an example to the utility and benefit of interdisciplinary studies employing both skeletal and molecular methods. Tuberculosis, caused by members of the *Mycobacterium tuberculosis* complex (MTBC), is one of the deadliest and most socially disruptive infectious diseases and as such, it has been at the center of medical and anthropological research. The MTBC is composed of eight globally distributed lineages with lineage 4 (L4) predominantly associated with European populations, currently dominating across Europe and the Americas (Gagneux et al., 2006; Ngabonziza et al., 2020). At first glance, it would seem these modern data suggests TB was first introduced to the Americas through European colonization. Even before the characterization of modern strain distribution, many argued that TB was not present prior to European arrival due to the limited skeletal evidence, lack of consensus on skeletal evidence and supposed low population densities of Indigenous Americans (Buikstra, 1999; Hrdlicka, 1909; Morse, 1961; Cockburn, 1963; Roberts and Buikstra, 2003). However, there exists an abundance of skeletal evidence of tuberculosis across the pre-colonial Americas (Allison et al., 1981, 1973; Buikstra and Williams, 1991; Garcia Frias, 1940; Klaus et al., 2010; Lombardi and García Cáceres, 2000). Application of PCR methods to pre-colonial skeletal and soft tissue evidence of TB resulted in the detection of genetic sequences reported to be specific to TB (Salo et al., 1994). However, tuberculosis continued to be contested by critics who argued the disease to have a European origin. The susceptibility of Indigenous Americans to TB during European colonization was used to support this claim (Stead 1997, 2000). Conversely, a body of osteological and molecular research continued to grow further demonstrating the existence of pre-colonial TB in South America (Arriza et al., 1995; Konomi et al., 2002; Klaus et al., 2010; Guichón et al., 2015). In 2014, by combining morphologically based

paleopathology and molecular methods to samples from pre-colonial South American contexts researchers not only identified the presence of MTBC DNA but also provided species level identification of the causative MTBC member (Bos et al.).

The first genomic characterization of ancient TB from the Americas (Bos et al., 2014) came from three vertebrae with TB lesions recovered from three LIP Chiribaya sites in the Osmore River Valley, not far from the coast of Peru (Owen, 1993; Burgess, 1992). Phylogenetic analysis revealed the member of MTBC that infected these archaeological populations was closely related to a strain associated with modern seals and sea lions, *M. pinnipedii*. The identification of *M. pinnipedii* led to the hypothesis that TB had been introduced to the Americas through a zoonotic event, perhaps from marine mammals. Since that time, additional studies have applied those methods to contexts to generate additional genome-level data on TB across the pre-colonial Americas (paper III and V) to evaluate strain diversity in the Americas prior to European arrival and to gain a better understanding of MTBC ecology and evolution. Thus far *M. pinnipedii* has been genomically characterized from coastal and highland sites from the Andean cultural region. However, this strain has yet to be identified in the eastern region of the Andes.

Recently, skeletal tuberculosis was identified in 13 individuals (6.3% in the total population) dating to pre-colonial contexts from the Chachapoyas site of Kuelap (Toyne et al., 2020). According to studies of modern populations, the development of skeletal lesions in association with tuberculosis occurs in a small percentage of affected populations, with conservative estimates at approximately 5-7% (Steinbock, 1976). This suggests that TB was circulating in high numbers in this population. Toyne and colleagues describe the Kuelap population-based study of advanced mutli-focal tuberculosis (TB) with distributions of skeletal lesions that were not consistent across the group (Toyne et al., 2020). Likewise, there appears to be variation in skeletal manifestation of TB compared to what has been observed in previously described contemporaneous populations (Buikstra and Williams, 1991; Toyne et al., 2020). Therefore, they conclude TB may manifest differently in the Chachapoya populations as compared to Peruvian populations further west and south, in which skeletal TB has been well described (Buikstra and Williams, 1991; Klaus et al., 2010).

This paper investigates one such case in which the individual was recovered from a Chachapoya cliff side funerary site and presents diffuse lytic skeletal lesions with an unusual distribution and atypical morphology leading to a broad differential diagnosis (Toyne et al., in press) (Supplementary Figure 1). To accommodate a broad differential diagnosis and limit ascertainment bias, we sampled multiple anatomical locations and employed a non-targeted approach with a broad pathogen screening method in effort to detect infectious agents which may be involved in the observed skeletal response. Here we present the first pre-colonial MTBC genome recovered from the Chachapoyas region in the northeastern Andean high-altitude subtropical rainforest.

INDIVIDUAL UNDER STUDY AND CONTEXT

Bioarchaeological Context

On the eastern slopes of the Andes in the North of Peru lies the Chachapoyas region covering approximately 150,000 km^2 of jungled terrain (Church and Guengerich, 2017; Toyne et al., 2020). The archaeological culture of the Chachapoya is still being understood due to the limited amount of systematic intensive archaeological investigations (Toyne et al., in press). However, iconographic programs, architectural style, mortuary practices, and genetic data suggest the communities of the Chachapoya region show some level of shared identity across what was likely a confederation of autonomous chiefdoms (Nystrom, 2009; Schjellerup, 1997; Toyne and Narváez Vargas, 2014; Toyne et al., 2020). The culture thrived from the late Middle Horizon across the Late Intermediate Period (CE 800-1470) and although the geographic location is relatively isolated there is evidence of interaction between Amazon and western Andean groups.

Diablo Huasi is a Chachapoya funerary complex that is made up of more than 40 tombs etched out of cliffside limestone escarpments. The individual included in our study comes from *Estructura Funeraria* (EF01), which appeared to be disturbed prior to archaeological investigation despite the challenges of accessing this cliff side construction due to its vertical orientation (Toyne et al., in prep). The individual included in this study was found with an additional 11-13 mummified individuals. Radiocarbon dating estimates the individual's interment to the LIP (CE 1296-1396 cal.). Skeletal analysis and morphological pathology assessments were performed prior to this study (Toyne et al., in prep). Results of the skeletal analysis estimate the individual was likely a male between 30-39 years of age at death. Observations of pathological changes describe involvement of the spinal column with multi-focal lytic lesions on the anterior aspect of vertebral bodies from the 7^{th} thoracic (T7) vertebra to the 12^{th} thoracic (T12) vertebra. The individual also displays kyphosis at T11 and T12 due to lytic destruction of the vertebral bodies (Toyne et al., in prep). Posterior aspects of the vertebral body and neural arch do not show any pathological involvement. However, lytic lesions were present on the necks of the left and right 11^{th} ribs. Additionally, radiographs show evidence of cranial involvement with extensive unhealed macroporosity in the endocranial surface. Pathological skeletal changes appear to be active at time of death, but some skeletal signs exist of healing (Toyne et al., in prep). Based on observations the authors conclude tuberculosis, brucellosis and mycotic infections as the likely candidates in their differential diagnosis with mycosis being most consistent with the observed skeletal changes.

Skeletal elements included in study

Three samples from the individual interred at Diablo Huasi, DIA002, were included in this study. Three separate anatomical elements were sampled including the right mandibular first molar (DIA002.A), the petrous portion of the left temporal bone (DIA002.B), a thoracic vertebra (DIA002.C) (Supplementary Figure 2). Petrous portion of the temporal bone and teeth are known to be rich sources of endogenous host DNA (Pinhasi et al., 2015), however the petrous portion is

not yet known to harbor ancient pathogens likely due to the limited vascularization of the skeletal feature (Anson et al., 1966; Margaryan et al., 2018). In contrast, teeth are also excellent sources of pathogen DNA from septicemic cases where pathogens become blood borne and travel to teeth via the inferior and superior alveolar blood vessels (Rombouts et al., 2017; Spyrou et al., 2019). However, the pathophysiology of some microbes, including MTBC members, make it unlikely that the microbe will be identified in teeth as they favor high-oxygen, marrow rich environments such as vertebrae. Therefore, a vertebra displaying pathological lesions was also selected for sampling. Three anatomical elements (tooth, petrous portion, vertebra) from the individual, DIA002, were sampled for DNA extraction, converted into libraries and shotgun sequenced.

RESULTS

Sequencing

The selection of multiple anatomical elements (petrous portion of the temporal bone, tooth and vertebra) decreases the bias introduced by sampling location for this study. Data from DIA002.B (petrous portion) was intended for human population genetic analysis and therefore the library was prepared with partial UDG treatment which would allow for partial removal of damage which had accumulated over time, typical of ancient DNA (Rohland et a., 2015). The partial removal of DNA damage maintains an amount of damage which does not influence the accuracy for typical human analyses. Therefore, these libraries are economically beneficial particularly when large sample sets are included as they can be used for downstream analyses. However, the tooth (DIA002.A) and the vertebrae (DIA002.C) were prepared as libraries which were not treated with UDG for the observation of damage patterns in pathogen DNA which may be detected permitting the authentication of the pathogen's antiquity (Jonsson et al., 2013). The resultant libraries were double indexed, giving each a unique barcode for downstream reference and sequenced using an Illumina HiSeq 4000. Libraries prepared from the tooth (DIA002.A) and the petrous portion (DIA002.B) were sequenced using a paired-end 50 base pair kit to a depth of 5 million reads. The library from the vertebra (DIA002.C) was sequenced to a depth of 10 million reads using a single-end 75 base pair kit. Sequencing data was then prepared using the EAGER pipeline (Peltzer et al., 2016) and mapped to the human refence genome (hg19) with BWA. Results show a range of 419,174 to 1,995,340 sequencing reads mapping to the human genome after quality filtering and removal of duplicates, with the petrous portion yielding the greatest amount of human DNA (Table 1). Authentication of ancient DNA was performed through the observation of damage patterns using mapDamage2 (Jonsson et al., 2013) which shows a damage percent of approximately 6 – 11% on DNA fragments from the three samples, typical of aDNA.

Pathogen screening

Microbial composition of the samples was evaluated using MALT (Vågene et al., 2018) within the HOPS pipeline. Analysis was performed in reference to a database constructed to include complete sequences and assemblies of bacteria, viruses, some eukaryotic parasites, and a human reference

(hg19) (full-bac-full-vir-etal-nov_2017). The HOPS (Hübler et al., 2020) pipeline evaluated each sample for the presence of ancient pathogens. Ancient MTBC reads were detected in the vertebral sample (DIA002.C) with a total of 1145 summarized reads on the MTBC node. However, no evidence of MTBC was found in samples taken from the tooth or petrous portion of the same individual. The tooth of the individual shows evidence for *Salmonella enterica* subspecies *enterica* infection, however, the HOPS analysis indicates this identification lacks the appropriate edit distance distribution to be considered confident. Surprisingly, HOPS analysis of the sample taken from the petrous portion yields evidence of a parasitic infection by *Schistosoma mansoni* with 151 reads assigning to the organism at 100% identity.

Evaluation of detected pathogens

Schistosoma mansoni is a water-borne parasite also known as a blood fluke primarily infecting the lower intestinal system (Barsoum et al., 2013). However, *S. mansoni* ova may gain access to other regions of the body including the central nervous system and result in cerebral schistosomiasis. In order to investigate the detection of *Schistosoma mansoni* from the petrous portion sample, sequencing data was mapped to a complete genome reference of the organism (NC_0315102.1) using EAGER (Pelzer et al., 2016). Mapping metrics showed 0.03% endogenous *S. mansoni* DNA with a mean genomic coverage of 0.0007. Further inspection of the reads in IGV revealed a few sporadic sequencing reads mapping to the reference genome in regions not specific enough to confidently confirm the presence of this organism. *Schistosoma mansoni* was not detected in the tooth or vertebral sample.

Detection of MTBC reads in sample DIA002.C (vertebra) was further confirmed by mapping of the shotgun data to the TB ancestor (MTB_anc) reference genome (Comas et al., 2010) using EAGER (Peltzer et al., 2016). Results show a total of 3209 reads mapping to the TB reference with an estimated 0.0419 mean coverage of the reference genome (Table 2). Mapping values also showed good distribution of coverage across the genome with percentage values typical for positive shotgun detection of MTBC (1-fold = 2.21, 2-fold = 0.25, and 3-fold = 0.14) (Table 2). However, the mapping to the TB ancestor showed unexpectedly low damage with values of 0.0139 for the 3' end of the sequencing reads and 0.0147 for the 5' end. Because so many bacteria that are closely related to the MTBC exist in the environment they subsequently may be in the extraction and library of the vertebra (DIA002.C). The inclusion of genetically similar environmental microbes can compromise evaluation of damage of the MTBC as they may also map to the reference in some capacity. To investigate this further, this library was included in an in-solution hybridization capture of MTBC reads for the evaluation of damage. This library was partnered with a UDG treated library from the same extract to permit phylogenetic analysis.

Ancient Chachapoya MTBC genome

Detection of MTBC reads in the vertebral sample (DIA002.C) was followed by the preparation of a double indexed UDG treated libraries and subsequent in-solution hybridization capture for

MTBC reads from both the non-UDG treated library (DIA002.C0101) and UDG treated library (DIA002.C0102). The capture product was sequenced to a depth of 10 million reads and sequencing data was processed and mapped to the TB ancestor using EAGER (Peltzer et al., 2016). Two rounds of in-solution capture of the UDG treated library (DIA002.C0102) resulted in the recovery of MTBC DNA with endogenous DNA values after quality filtering of 5.06% and a mean MTBC genome coverage of 5.7 fold (Table 3). However, TB mapping values also show that only 85% of the reference genome had a coverage greater than 1-fold. However, the low ratio of mapped reads prior to duplicate removal to the number of mapped reads after duplicate removal (1.58), known as the cluster factor, indicate adequate genetic diversity remains within the library. This would thereby permit deeper sequencing to improve coverage across the genome. Deeper sequencing to a depth of 20 million reads and concatenation of all sequenced capture data for the UDG treated library (DIA002.C0102) enabled the reconstruction of an MTBC genome with 15.67-fold coverage (Table 4). The amount of the reference genome that is now covered greater than 1-fold is 94.35%, permitting phylogenetic analysis.

Concerning the library for the vertebra (DIA002.C0101), evaluation of DNA damage after enrichment revealed a slight increase with a percentage of approximately 2% (Table 3). Damage profiles produced by mapDamage2 (Jónsson et al., 2013) show background signals which may still be interfering with accurate detection of post-mortem damage patterns in the DNA fragments. This was investigated further using MALT to evaluate the composition of the library, specifically non-MTBC environmental mycobacterial contaminants which may be included in the sample (Table 5). Results of this analysis show MTBC reads comprise between 94% – 97% of all mycobacterial sequencing data from DIA002.C0101.

Authentication of Genetic Variation in the Reconstructed Genome

The challenge of environmental microbes continues with the inclusion of genetically similar environmental mycobacteria in the in-solution hybridization product. In order to include genetic reads belonging to MTBC members with high confidence and thereby authenticate genetic diversity several steps were taken. An evaluation of variant positions was performed using the SNPEvaluation tool (https://github.com/andreasKroepelin/SNP_evalutation) (methods). In this analysis each position must have a minimum coverage of five-fold to be included in downstream analyses. Evaluation of each position took place within a 50bp window, in which coverage of each SNP is considered under two mapping parameters: relaxed mapping parameters and stringent mapping parameters. Positions in which we observe the existence of additional heterozygous positions in the same sequencing read were excluded. Additionally, because deletions might result in mapping errors, we filtered out any position in which the 50 base pair window lacked full coverage. This evaluation resulted in the removal of a total of 75 positions in DIA002: 40 (out of 116) positions that were unique to the sample and 35 (out of 156) positions which were shared between this genome and other MTBC genomes.

Phylogenetic Analysis

Our phylogenetic analysis was performed using a set of 279 genomes (Supplementary Table 1) including representatives of all lineages and 9 modern *M. pinnipedii* genomes and 11 ancient MTBC genomes dating to the LIP, including three previously published ancient coastal TB genomes (Bos et al., 2014), seven TB genomes from the central Andean highland site of Huari (Nelson et al., in prep: paper III), and the newly reconstructed TB genome (DIA002.C) from the Chachapoyas site on the eastern slopes of the Andes. *Mycobacterium canetti* was used as the outgroup.

Our maximum parsimony tree was constructed with an alignment prepared with minimum 5-fold calling on all positions and the removal of SNPs that have ≥ 5% of ambiguous calls (N) in the alignment (95% partial deletion) resulting in 45,498 out of 53,816 variant positions (Figure 1). This phylogenetic analysis reveals DIA002.C as closely related to the previously described ancient strains and as closely related to the modern *M. pinnipedii* strains. This reveals that the pinniped-adapted strain had an expansive presence across the pre-colonial Andean cultural region and circulated not only in ancient coastal and ancient highland populations of Peru but also existed in the eastern slopes of the Andes in the Amazonas region.

DISCUSSION

Diagnosing the past: skeletal markers and molecular data

The recovery of MTBC DNA in individual DIA002 provides molecular evidence for TB in an individual who displays an unusual distribution of skeletal lesions. This atypical distribution and appearance led to the development of a differential diagnosis that included tuberculosis as well as other infectious diseases with emphasis on mycotic infection (Toyne et al., in press). However, authors note that the unusual presentation of pathology could be the product of a comorbidity. The application of ancient DNA methods allowed for the confident species level identification of TB in this individual, however, there was no confident identification of another pathogen suggesting comorbidity. This may be due to a variety of factors which influence molecular paleopathological analyses including differences in pathogen DNA preservation. The previously studied (Toyne et al., 2020) Chachapoya population of Kuelap which is located not far from Diablo Huasi, also presents skeletal lesions consistent with TB and suggests TB was endemic in this area during the Late Intermediate Period. This supports the body of paleopathological research demonstrating an increase in TB infections across the Andes during the LIP (Buikstra and Roberts, 2003; Buikstra and Williams, 1981; Burgess, 1992; Klaus et al., 2010; Lombardi and García Cárceres, 2000;). Yet, the reported atypical variation in TB skeletal evidence and lesion distribution in the population from Kuelap (Toyne et al., 2020) suggests other Chachapoya populations may also present variable distribution in TB skeletal pathology. However, the Kuelap cases have not been confirmed through molecular analyses.

MTBC from the Chachapoyas Region

Mycobacterium pinnipedii strains have been recovered from LIP sites including the Chiribaya archaeological contexts of the Osmore River Valley, located not far from the coast and from Huari, located in the Ayacucho Basin of the central Andes (Bos et al., 2014; paper III of this thesis). However, the identification of *M. pinnipedii* at Diablo Huasi indicates a much broader range of this strain than previously understood. This funerary site, Diablo Huasi, of the LIP Chachapoya provides identification of *M. pinnipedii* in the eastern piedmont of the Andes in the high-altitude subtropical cloud forest. Although these three different Peruvian site locations (Osmore River Valley, Huari, and Diablo Huasi) from which ancient TB has been recovered are contemporaneous in time period (LIP) they represent very different ecological contexts. The distance of the Chachapoya from the coast, which is even further inland than Huari, suggests alternative vectors of transmission, whether human or animal remains unknown. However, like the Wari, and many other contemporary Andean cultures and polities, the Chachapoya were not restricted or isolated by their terrain (Quilter, 2014; Toyne et al., 2020; Tung, 2012). The Chachapoya established road networks and utilized the river systems to travel quickly and efficiently permitting cultural interaction and exchange. In fact, the Chachapoya show evidence of Amazonian and Wari influence in architectural style, iconographic associations, and in socio-political structure as chiefdoms. Furthermore, archaeological evidence of Wari expansion into this area during the Middle Horizon (600 – 1000 CE) with the establishment of centers to capitalize on resources also supports themes of cultural interaction (Church and Von Hagen, 2008). Likewise, the Inka ultimately capitalized on the geographic location of the Chachapoya cultural region due to the accessibility of cultural networks from the Peruvian Andes to the Amazon Basin (Schjellerup, 2015; Toyne et al., 2020). Indeed, many features of the Chachapoya culture and archaeological evidence of the region suggest a high degree of interaction. Such mobility and interaction would surely facilitate the movement of infectious disease, by infected person or animal, yet the resolution of understanding on mechanisms and directions of this is, as of yet, unclear.

CONCLUSION

Here we have presented the identification of strains related to those found in other archaeological material of pre-contact Peru, which show greatest homology to extant *M. pinnipedii* strains from the eastern slopes of the Andes, thus providing a greater understanding of the broad geographic and ecological presence of this MTBC strain across the LIP Andes. Recovery of ancient strains closely related to *M. pinnipedii* from coastal LIP sites in the lower Osmore River Valley provided a plausible scenario of a zoonotic event as archaeological coastal populations were known to exploit seals and sealions. Yet, as we continue to molecularly explore ancient Andean TB and characterize ancient MTBC strains from the Andean region of Peru, we find *M. pinnipedii* strains existing across multiple ecological zones, likely reflecting the ability of ancient Andean people to

conquer a multitude of ecosystems and environments with continual interaction. Not only does this case of unusual skeletal presentation demonstrate the utility of molecular methods for complementing paleopathological analyses but also reveals closely related MTBC strains spread out across a large geographic expanse with presence in the coast, the highlands, and the eastern slopes of the subtropical forest. The identification of these ancient strains from multiple contemporaneous LIP sites, located within significant distance from one another and in diverse ecological zones suggests the ancient *M. pinnipedii* strain may be the causative agent in the LIP rise in TB. Only with further research and the recovery and characterization of additional pre-colonial Andean TB strains can we truly understand the evolution, ecology, and success of this strain during the LIP.

METHODS

Sampling methods for different elements

All elements were sampled in the Max Planck Institute for the Science of Human History ancient DNA laboratory, Jena, Germany. The right mandibular molar (DIA002.A) was sectioned at the cemento-enamel junction to access the dental pulp chamber by using a hand saw and vice to secure and stabilize the tooth. Once the dental pulp chamber was exposed, approximately 50mg of dental powder was collected from the chamber by drilling the internal surface with a dental drill. The fragmented petrous portion of the temporal bone was wrapped in foil with the location for sampling exposed. Approximately 50 mg of bone was collected by drilling with a dental drill into the bone approximately 2 mm deep. The bone of the vertebra was thin and fragile with some fragments already existing. In order to minimize exposure to heat and additional damage to the bone, small fragments of the bone were pulled off using tweezers and weighed until the sample reached approximately 50 mg. The fragments were then pulverized into powder using a mortar and pestle.

Extraction of Ancient DNA

All samples (DIA002.A, DIA002.B, DIA002.C) were extracted for DNA using a protocol optimized for the retrieval of short fragment typical of ancient DNA (Dabney et al., 2013). A total of 1 ml lysis buffer was added to all samples to liberate the DNA from the mineral matrix of bone. Samples were then left to incubate over night at 37° Celsius. Each sample was then carried through a purification process whereby extracted DNA would be bound to a silica membrane of spin columns using 10 ml of GuHCL-based binding buffer and unwanted extract products would be washed away using the Viral Nucleic Acid Kit (Roche). The DNA would be eluted from the silica membrane using 100ul TET (10mM Tris-HCL, 1mM EDTA pH 8.0, 0.05% Tween20). Controls for the experiment, both negative and positive, were carried along with the samples.

Library Preparation of Extracted Ancient DNA

Library preparation methods used for these DNA extracts varied depending on anatomical sampling location. All samples were converted into double stranded, double indexed libraries and the DNA extractions from the vertebra and tooth were prepared without UDG treatment (Meyer and Kircher, 2010). However, as the extract taken from the petrous portion fragment was intended for human population genetics purposes and was partially treated with Uracil-Glycosolase to excise uracils which may be products of post-mortem damage to DNA (Rohland et a., 2015). For each sample, the process resulted in 40 µl of DNA library. An extraction blank and library blank with a positive control from extraction were included for each batch (2) to evaluate the process and reagents used. Libraries were then quantified with a qPCR reaction with 1 µl of DNA from each prepared ancient DNA library, 10 µl of Dynamo, 1 µl of each primer (IS7 and IS8), 7 µl of H_2O, and 1 µl of 12 DNA standards from 10^3 to 10^8 quantities along with two qPCR blanks. After quantification the prepared libraries were then barcoded giving each sample library a distinct 8 base pair identifier index for later reference, evaluation, and analysis of the DNA unique to that sample. Indexing was performed using *Pfu Turbo Cx Hotstart DNA Polymerase* (Agilent) in which each library was split to a number of reactions based on the concentration of DNA (a maximum of 2 x 10^{10} molecules of DNA) to ensure all DNA fragments would be efficiently indexed and recovered. All indexed reactions were then purified using a MinElute Kit using one column for a maximum of four reactions. All indexed libraries were quantified using qPCR with IS5 and IS6 primers to calculate the efficiency of the indexing reaction and amount of quantifiable DNA. This resulted in a total of 50 µl of index library. The three prepared libraries were then amplified to 10nM and sequenced as paired-end 50 base pair reads (DIA002.A, DIA002.B) or as single-end 75 base pair reads (DIA002.C) using a HiSeq 4000.

Preparing and mapping sequencing reads

All sequencing data was processed and prepared using the EAGER pipeline (Peltzer et al., 2016). This pipeline incorporates the removal of adapters and merging of reads (if necessary), mapping to a selected reference, removal of duplicate reads, SNP calling, and allows for the evaluation of damage patterns through the generation of damage plots (Peltzer et al., 2016; Schubert et al., 2016; Li and Durbin, 2009; McKenna et al., 2010; Jónsson et al., 2013). All data from the three samples were concatenated after adapter removal and merging (if paired-end) (Schubert et al., 2016). The data were mapped to the human (Hg19) reference genome with BWA-aln (v. 0.7.12) (Li and Durbin, 2009) using the following mapping parameters (-l 16, -n 0.01) for samples and negative controls. Reads with low quality mapping were removed using SAMtools quality filtering (-q 24). The removal of duplicate reads was performed using MarkDuplicates. In order to visualize DNA damage patterns libraries were evaluated using mapDamage2.0 (Jónsson et al., 2013).

MALT analysis and detection of ancient MTBC DNA

The libraries were sequenced with sequencing depth to allow for the screening of the samples microbial content in order to detect infectious agents, as well as, produce a metagenomic profile

of the samples and thus evaluate the environmental background of non-pathogenic mycobacteria. Pathogen DNA screening was performed using the HOPS pipeline (Hübler et al., 2019) which incorporates the MEGAN Alignment Tool (MALT) (version 0.3.8) (Vågene et al., 2018) with evaluations by quality measures customized for ancient DNA and provides visual outputs for the detection of pathogens. Samples were screened using a minimum of 95 percent identity to be matched with an organism against a custom database of full genomes for bacterial, viral, and some eukaryotic organisms (Nov. 2017).

Upon detection of pathogens (MTBC and *Schistosoma mansoni*) the appropriate samples were subsequently mapped using EAGER (Peltzer et al., 2016) to the reference genomes: the reconstructed TB ancestor (MTB_anc) for MTBC read detection and NW_017386861.1 for *Schistosoma mansoni*. Mapping parameters for initial pathogen screening were used (-l 16, -n 0.01, -q 24).

UDG treated library preparation

The library for DIA002.C was prepared as double-stranded DNA libraries using a volume of 50 ul of DNA. The UDG treatment and library protocol differs in the first stage of library preparation where we first treated the libraries with uracil-DNA-glycosylase (UDG) to remove postmortem damage (Briggs et al., 2010). Using a total of 50 µl of DNA we then added to the reaction a master mix of 15 µl of NEB Buffer 2, 15 µl of ATP, .75 µl of BSA, 1.8 µl of dNTPs, 6 µl of T4 PNK, 9 µl of USER enzyme and 22.45 µl of H_2O. The reaction was incubated at 37 degrees C for 3 hours. The reaction then received 6 µl of T4 polymerase to each library and was incubated for an additional 20 minutes at 25° C and 10 minutes at 12° C. The reaction was then purified using the MinElute kit (Qiagen) and the remainder of the protocol follows the traditional library preparation from adapter ligation forward. The UDG-treated libraries were quantified on a quantitative PCR (qPCR) using IS7/IS8 primer combinations. This quantification was used to calculate the appropriate quantity of index needed for the UDG treated library of DIA002.C. Subsequently, the library was divided into multiple PCR reactions for double indexing (Meyer and Kircher, 2010; Kircher et al., 2012) based on their initial quantification and measurement, in order to ensure maximum efficiency of the indexing reaction. Indexing reactions were split across two days to safeguard against any error that may be introduced in the process. Index combinations consisting of unique 8 base pair identifier for each library was used to ensure identification of DNA from each sample. Index combinations were ligated to DNA library molecules using *Pfu Turbo Cx Hotstart DNA Polymerase* (Agilent) and run in a 10-cycle amplification reaction described above. The reaction was purified using the MinElute DNA purification kit (Qiagen) as previously described, and eluted in TET (10mM Tris-HCl, 1mM EDTA pH 8.0, 0.05% Tween20). The indexed UDG treated library was then quantified on a qPCR using IS5/IS6 primer combinations and concentration was calculated to ensure maximum efficiency. After the concentration of the indexed UDG treated library was measured, calculations for amplifications for capture and sequencing were performed to identify the number of amplification reactions needed to meet the

required concentration of 200-300 ng/ul for in-solution TB capture. Amplification was performed using *Herculase II Fusion DNA Polymerase* (Agilent). The amplified library was subsequently purified using the MinElute DNA purification kit (Qiagen), and eluted in TET (10mM Tris-HCl, 1 mM EDTA pH 8.0, 0.05% Tween20).

In-solution Whole genome MTBC capture

In-solution MTBC capture was performed on the indexed UDG library and the non-UDG treated library of the DIA002.C sample using single stranded DNA probes computationally designed to reflect the genome of the MTBC ancestor (MTB_anc) (Comas et al., 2010) by reverting phylogenetically informative derived positions to the ancestral alleles in the H37Rv reference genome (NC_000962.1). Capture probes are 60 base pairs in length with 52 bases as the biological sequence and 8 bases as a linker sequence (5' CACTGCGG 3'). Capture probes have a 5 bp overlap and upon removal of duplicate and low complexity probes, provide a unique set of 852,164 probes. Preparation of the capture was carried out according to previously published protocols (Fu et al., 2013; Vågene et al., 2018).

MTBC enrichment was carried out on each sample for two rounds of capture with a positive control (58U) of known ancient TB quantity previously demonstrated (Bos et al., 2014). Blanks with non-overlapping index combinations were captured together in a single well on the same plate.

Sequencing data of captured products

Both captured libraries were sequenced on Illumina platforms as paired-end 75 base pair reads using a HiSeq 4000. The sequence data were then sorted, assigned, and demultiplexed. Using the EAGER pipeline (v.1.92.55) (Peltzer et al., 2016) reads were prepared for analysis beginning with removal of Illumina adapters using AdapterRemoval v2 (Schubert et al., 2016). The data were mapped to the MTBC ancestor reference genome as described above but using UDG treated appropriate strict mapping parameters (-l 32, -n 0.1) for TB captured samples and positive controls and reads with low quality mapping were removed using SAMtools quality filtering (-q 37). Relaxed mapping parameters and lower threshold for SAMtools quality filtering (-q 24) were used for negative controls (-l 16, -n 0.01). The removal of duplicate reads was performed using MarkDuplicates. The DIA002.C01 library captured was not treated with UDG and thus the deamination of cytosine to uracil would remain. Therefore, the DNA damage patterns for this library was evaluated using mapDamage2.0 (Jónsson et al., 2013).

Evaluating genomes for heterozygous calls and authentication of informative positions

The DIA002.C genome reconstructed from the UDG treated library was evaluated to validate authenticity of the genetic variation observed. We estimated the number of heterozygous variant calls within the reconstructed genome by considering the "haploid" nature of MTBC members, in which we assume "heterozygous" SNPs are the result of closely related environmental bacteria

mapping to the TB reference or evidence of a mixed infection. We performed SNP calling with the UnifiedGenotyper in GATK. We then constructed a table of all variant positions, including heterozygous positions, across our dataset using MultiVCFAnalyzer v0.85 (Bos et al., 2014) with all positions requiring a minimum of 5-fold coverage. Heterozygous variant positions were called at a minimum of 5-fold coverage. Using R v3.4.1 we produced histograms of allele frequencies for all SNPs with allele frequencies between 10-90% read support for all reconstructed genomes.

We utilized the SnpEvaluation tool (https://github.com/andreasKroepelin/SNP_evaluation) to filter out any positions originating from contaminant environmental mycobacteria and authenticate the informative variant positions as deriving from MTBC members so they may be confidently included in phylogenetic analysis. For this evaluation we followed the use of this tool as outlined by Keller et al., (2019) with some customization for MTBC based analyses. We mapped the UDG treated DIA002.C captured product to the MTBC ancestor reference first using strict mapping parameters (-n 0.1, -l 32) that allow for few mismatched positions and relaxed mapping parameters (-n 0.01, -l 32) which maintain length but allow for more mismatched positions. Using the snpTable generated by MutliVCFAnalyzer we compared VCF files of the two mappings within the SNPEvaluation tool, visualizing each position within a 50-bp window, observing for coverage within the window, additional heterozygous position, and the coverage of the position in both strict and relaxed mapping parameters. Positions would be considered of low confidence and filtered out if within the 50-bp window there were regions that were not covered, there were heterozygous positions, and the ratio of mean coverage between strict and relaxed mapping was higher than 1.1. The evaluation criteria were consistent for both shared and unique positions within the genome.

Phylogenetic Analysis

Our phylogenetic analysis was performed using a set of 278 genomes with *Mycobacterium canetti* selected as the outgroup (Supplementary Table 1). Variant positions were called using the tool UnifiedGenotyper (v. 3.5) from the Genome Analysis Toolkit (GATK) (DePristo et al., 2011). The SNP alignment and tables for downstream SNP analysis were produced using MultiVCFAnalyzer (Bos et al., 2014). Parameters selected for a minimum of 5-fold coverage of all variant positions and a minimum genotype quality of 30. The construction of the alignment also included the removal of regions including repetitive elements, tRNAs, mRNAs, and rRNAs. The maximum-parsimony tree was constructed using MEGA7 (Kumar et. al, 2016) using the sub-pruning tree model and selecting for partial deletion of 95% and 500 bootstraps.

Table 1. Mapping statistics of sequence data from DIA002 libraries mapped to the human reference (Hg19).

Sample Name	Sample type	# of Raw sequenced reads	Unique mapped reads after QF	Endogenous DNA QF (%)	Mean coverage	DMG 1st Base 3'	DMG 2nd Base 3'	DMG 1st Base 5'	DMG 2nd Base 5'	median fragment length	GC content in %
DIA002.A	**Tooth**	9621994	419174	18.53	0.0081	0.0942	0.0641	0.0918	0.0606	59	42.83
DIA002.B	**Petrous portion**	11461154	1995340	54.301	0.0356	0.1088	0.0138	0.1101	0.0123	53	47.01
DIA002.C	**Thoracic vertebra**	8616648	835024	11.074	0.0172	0.0625	0.0332	0.0772	0.0517	74	44.12
EXB048.A	Extraction blank	9795904	55	0.022	0.0005	0.1667	0.1429	0	0.0625	39	60.45
LIB036.A	Library blank	3144412	2644	1.984	0.0001	0.0166	0.0107	0.005	0	72	50.96

*QF = indicates values are given after quality filtering of data

Table 2. Mapping statistics of shotgun sequencing data of DIA002.C mapped to the TB ancestor reference (MTB_anc).

Sample Name	Sample type	# of Raw sequenced reads	Unique mapped reads after QF	Endogenous DNA QF (%)	Mean coverage	Coverage >= 1X %	Coverage >= 2X %	Coverage >= 3X %	DMG 1st Base 3'	DMG 2nd Base 3'	DMG 1st Base 5'	DMG 2nd Base 5'	median fragment length	GC content in %
DIA002.C	**Thoracic Vertebra**	8616648	3209	0.044	0.0419	2.21	0.25	0.14	0.0139	0.013	0.0147	0.018	58	62.47
EXB048.A	Extraction blank	9795904	55	NF	0.0005	0.04	0.01	0	0.1667	0.1429	0	0.0625	39	60.45
LIB036.A	Library blank	3144412	54	0.038	0.0005	0.05	0	0	0.08	0.1333	0.0476	0.1053	41	63.27

*QF = indicates values are given after quality filtering of data

Table 3. Mapping statistics of TB captured UDG and non-UDG libraries of sample DIA002.C mapped to the TB ancestor reference (MTB_anc).

Sample Name	library type	# of Raw sequenced reads	Unique mapped Reads after QF	Endogenous DNA QF (%)	Cluster Factor	Mean coverage	std. dev. Coverage	Coverage >= 1X in %	Coverage >= 2X in %	Coverage >= 3X in %	DMG 1st Base 3'	DMG 2nd Base 3'	DMG 1st Base 5'	DMG 2nd Base 5'	median fragment length	GC content in %
DIA002.C0102	UDG	22326746	321942	5.06	1.58	5.76	43.41	85.05	72.13	59.71	0.01	0.01	0.01	0.01	75	62.08
DIA002.C0101	nonUDG	11823484	125336	4.50	1.78	2.14	22.10	67.51	40.12	21.29	0.02	0.02	0.02	0.02	70	61.99
EAD001.A0101	pos. control	23249176	2847831	50.46	1.73	30.84	26.32	95.97	94.72	93.3	0.01	0.00	0.01	0.00	46	62.83
EXB048.A1602	Extraction blan	6117598	441	2.37	182.06	0.01	0.15	0.39	0.06	0.03	0.01	0.01	0.01	0.02	55	61.96
LIB039.A1401	Extraction blan	762006	66	1.96	124.49	0.00	0.03	0.08	0	0	0.03	0	0	0	48.5	62.42

*QF = indicates values are given after quality filtering of data

Table 4. Mapping statistics of pooled TB captured UDG libraries of sample DIA002.C mapped to the TB ancestor reference (MTB_anc).

Sample Name	# of Raw sequenced reads	Unique mapped Reads after QF	Endogenous DNA QF (%)	Cluster Factor	Mean coverage	Coverage >= 1X in %	Coverage >= 2X in %	Coverage >= 3X in %	Coverage >= 4X in %	Coverage >= 5X in %	median fragment length	GC content in %
DIA002.C pooled data	59085978	955232	3.994	2.461	15.6716	94.35	92.43	89.94	87.38	84.57	76	62.98

*QF = indicates values are given after quality filtering of data

Table 5. Evaluation of non-MTBC *Mycobacteria* in non-UDG library for DIA002.C using MALT performed at 95 percent identity.

Sample Library	Total Assigned Reads in MALT 95ID	Total Summarized Reads Mycobacteria	Total Assigned Reads Mycobacteria	Total non-MTBC Myocbacterial Reads	Total Summarized Reads MTBC	Total Assigned Reads MTBC	Upper Boundary Calculation Estimation of MTBC reads (%)	Lower Boundary Calculation Estimation of MTBC reads (%)	Middle Calculation Estimation of MTBC reads (%)
DIA002.C0101	1229232	159036	5180	9709	149327	2372	0.97	0.94	0.95

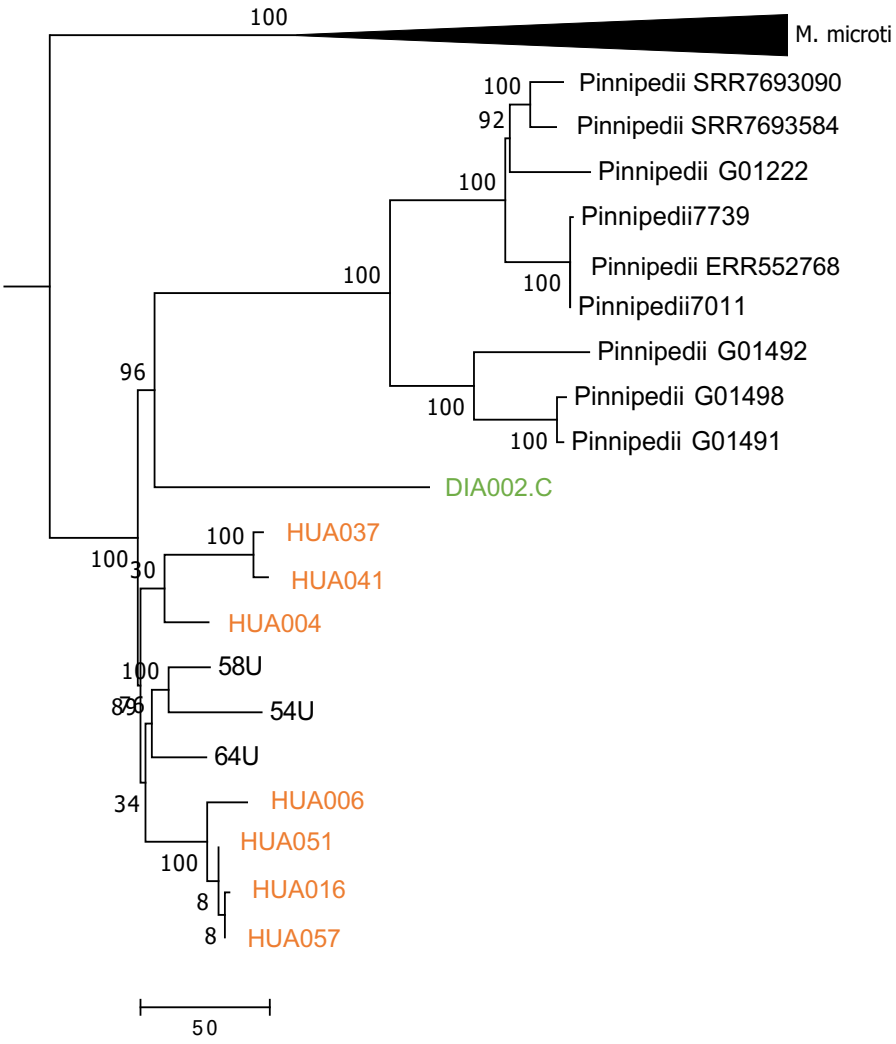


Figure 1. Phylogenetic tree. Maximum Parsimony phylogeny constructed with a partial SNP alignment (95% partial deletion = 45,498 variant positions out of 53,816) of 278 genomes. The portion of the tree shown includes modern *M. microti* genomes *M. pinnipedii* genomes, ancient genomes from Peruvian LIP contexts including Diablo Huasi (DIA002.C) indicated in green, Huari (HUA004, HUA006, HUA016, HUA037, HUA041, HUA051, HUA057) marked in orange and the Osmore River Valley (58U, 54U, 64U). Numbers on nodes are showing bootstrap values (500 iterations). This tree was constructed using MEGA7 (Kumar et al., 2016).

References

Allison, M.J., Gerszten, E., Munizaga, J., Santoro, C., Mendoza, D., 1981. Tuberculosis in pre-Columbian Andean populations. Prehistoric tuberculosis in the Americas 49–61.

Allison, M.J., Mendoza, D., Pezzia, A., 1973. Documentation of a case of tuberculosis in pre-Columbian America. American Review of Respiratory Disease 107, 985–991.

Anson, B.J., Winch, T.R., Warpeha, R.L., Donaldson, J.A., 2016. LXXVI The Blood Supply of the Otic Capsule of the Human Ear: With Special Reference to that of the Cochlea. Annals of Otology, Rhinology & Laryngology. https://doi.org/10.1177/000348946607500402

Arateco, W.M.R., 1998. Mal de Pott en momia de la colección del museo arqueológico Marqués de San Jorge. Maguaré 0, 99–115.

Arriaza, B.T., Salo, W., Aufderheide, A.C., Holcomb, T.A., 1995. Pre-Columbian tuberculosis in northern Chile: Molecular and skeletal evidence. American journal of physical anthropology 98, 37–45.

Bos, K.I., Harkins, K.M., Herbig, A., Coscolla, M., Weber, N., Comas, I., Forrest, S.A., Bryant, J.M., Harris, S.R., Schuenemann, V.J., 2014. Pre-Columbian mycobacterial genomes reveal seals as a source of New World human tuberculosis. Nature 514, 494–497.

Brites, D., Loiseau, C., Menardo, F., Borrell, S., Boniotti, M.B., Warren, R., Dippenaar, A., Parsons, S.D.C., Beisel, C., Behr, M.A., Fyfe, J.A., Coscolla, M., Gagneux, S., 2018. A New Phylogenetic Framework for the Animal-Adapted Mycobacterium tuberculosis Complex. Front. Microbiol. 9. https://doi.org/10.3389/fmicb.2018.02820

Buikstra, J.E., 2010. Paleopathology: a contemporary perspective. A companion to biological anthropology 395.

Buikstra, J.E., Williams, S., 1991. Tuberculosis in the Americas: Current Perspectives, in: Human Paleopathology: Current Syntheses and Future Options Edited by D. Ortner and A. Aufderheide.

Burgess, S., 1992. Health and El Algodonal: a preliminary report. Presented at the Annual Meeting of the Society for American Archaeology, Pittsburgh, Pennsylvania.

Burgess, S.D., 1992. Chiribayan Skeletal Pathology on the South Coast of Peru: Patterns of Production and Consumption. (Ph.D. dissertation). University of Chicago, Chicago, Illinois.

Chan, J.Z.-M., Sergeant, M.J., Lee, O.Y.-C., Minnikin, D.E., Besra, G.S., Pap, I., Spigelman, M., Donoghue, H.D., Pallen, M.J., 2013. Metagenomic analysis of tuberculosis in a mummy. New England Journal of Medicine 369, 289–290.

Church, W., Guengerich, A., 2017. La (re) construcción de Chachapoyas a través de la historia e histografía. Boletín de Arqueología PUCP 5–38.

Church, W.B., Von Hagen, A., 2008. Chachapoyas: Cultural development at an Andean cloud forest crossroads, in: The Handbook of South American Archaeology. Springer, pp. 903–926.

Cockburn, A., 1963. The Evolution and Eradication of Infectious Diseases. The Evolution and Eradication of Infectious Diseases.

Comas, I., Chakravartti, J., Small, P.M., Galagan, J., Niemann, S., Kremer, K., Ernst, J.D., Gagneux, S., 2010. Human T cell epitopes of Mycobacterium tuberculosis are evolutionarily hyperconserved. Nature Genetics 42, 498–503. https://doi.org/10.1038/ng.590

Correal, G., Flórez, I., 1992. Estudio de los guanes de la Mesa de Los Santos Santander (Colombia)., in: Actas Del Primer Congreso Internacional de Estudios Sobre Momias. Museo Arqueológico y etnográfico de Tenerife, Santa Cruz de Tenerife, pp. 307–313.

Dabney, J., Knapp, M., Glocke, I., Gansauge, M.-T., Weihmann, A., Nickel, B., Valdiosera, C., García, N., Pääbo, S., Arsuaga, J.-L., Meyer, M., 2013. Complete mitochondrial genome sequence of a Middle Pleistocene cave bear reconstructed from ultrashort DNA fragments. PNAS 110, 15758–15763. https://doi.org/10.1073/pnas.1314445110

Fu, Q., Meyer, M., Gao, X., Stenzel, U., Burbano, H.A., Kelso, J., Pääbo, S., 2013. DNA analysis of an early modern human from Tianyuan Cave, China. Proceedings of the National Academy of Sciences 110, 2223–2227.

Gagneux, S., DeRiemer, K., Van, T., Kato-Maeda, M., Jong, B.C. de, Narayanan, S., Nicol, M., Niemann, S., Kremer, K., Gutierrez, M.C., Hilty, M., Hopewell, P.C., Small, P.M., 2006. Variable host–pathogen compatibility in Mycobacterium tuberculosis. PNAS 103, 2869–2873. https://doi.org/10.1073/pnas.0511240103

García-Frías, J., 1940a. La tuberculosis en los antiguos peruanos. Actualidad Médica Peruana 5, 274–291.

García-Frías, J., 1940b. La tuberculosis en los antiguos peruanos. Actualidad Médica Peruana 5, 274–291.

Giffin, K., Lankapalli, A.K., Sabin, S., Spyrou, M.A., Posth, C., Kozakaitė, J., Friedrich, R., Miliauskienė, Ž., Jankauskas, R., Herbig, A., 2020. A treponemal genome from an historic plague victim supports a recent emergence of yaws and its presence in 15 th century Europe. Scientific Reports 10, 1–13.

Hrdlička, A., Fewkes, J.W., 1909. Tuberculosis among certain Indian tribes of the United States. US Government Printing Office.

Hübler, R., Key, F.M., Warinner, C., Bos, K.I., Krause, J., Herbig, A., 2019. HOPS: automated detection and authentication of pathogen DNA in archaeological remains. Genome Biology 20, 280. https://doi.org/10.1186/s13059-019-1903-0

Huson, D.H., Beier, S., Flade, I., Górska, A., El-Hadidi, M., Mitra, S., Ruscheweyh, H.-J., Tappu, R., 2016. MEGAN Community Edition - Interactive Exploration and Analysis of Large-Scale Microbiome Sequencing Data. PLoS Comput Biol 12, e1004957–e1004957. https://doi.org/10.1371/journal.pcbi.1004957

Jennings, J., 2006. Understanding Middle Horizon Peru: hermeneutic spirals, interpretative traditions, and Wari administrative centers. Latin American Antiquity 17, 265–285.

Jónsson, H., Ginolhac, A., Schubert, M., Johnson, P.L.F., Orlando, L., 2013. mapDamage2.0: fast approximate Bayesian estimates of ancient DNA damage parameters. Bioinformatics 29, 1682–1684. https://doi.org/10.1093/bioinformatics/btt193

Keller, M., Spyrou, M.A., Scheib, C.L., Neumann, G.U., Kröpelin, A., Haas-Gebhard, B., Päffgen, B., Haberstroh, J., i Lacomba, A.R., Raynaud, C., 2019. Ancient Yersinia pestis genomes from across Western Europe reveal early diversification during the First Pandemic (541–750). Proceedings of the National Academy of Sciences 116, 12363–12372.

Kircher, M., Sawyer, S., Meyer, M., 2012. Double indexing overcomes inaccuracies in multiplex sequencing on the Illumina platform. Nucleic acids research 40, e3–e3.

Klaus, H.D., Wilbur, A.K., Temple, D.H., Buikstra, J.E., Stone, A.C., Fernandez, M., Wester, C., Tam, M.E., 2010. Tuberculosis on the north coast of Peru: skeletal and molecular paleopathology of late pre-colonial and postcontact mycobacterial disease. Journal of Archaeological Science 37, 2587–2597.

Li, H., Durbin, R., 2009. Fast and accurate short read alignment with Burrows–Wheeler transform. Bioinformatics 25, 1754–1760. https://doi.org/10.1093/bioinformatics/btp324

Lombardi, G.P., García Cáceres, U., 2000. Multisystemic tuberculosis in a pre-columbian peruvian mummy: four diagnostic levels, and a paleoepidemiological hypothesis. Chungará (Arica) 32, 55–60.
Margaryan, A., Hansen, H.B., Rasmussen, S., Sikora, M., Moiseyev, V., Khoklov, A., Epimakhov, A., Yepiskoposyan, L., Kriiska, A., Varul, L., 2018. Ancient pathogen DNA in human teeth and petrous bones. Ecology and evolution 8, 3534–3542.
McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A., Garimella, K., Altshuler, D., Gabriel, S., Daly, M., 2010. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. Genome research 20, 1297–1303.
Meyer, M., Kircher, M., 2010. Illumina Sequencing Library Preparation for Highly Multiplexed Target Capture and Sequencing. Cold Spring Harb Protoc 2010, pdb.prot5448. https://doi.org/10.1101/pdb.prot5448
Morse, D., 1961. Prehistoric tuberculosis in America. American Review of Respiratory Disease 83, 489–504.
Nelson, E.A., Buikstra, J.E., Herbig, A., Tung, T.A., Bos, K.I., 2020. Advances in the molecular detection of tuberculosis in pre-contact Andean South America. International Journal of Paleopathology.
Ngabonziza, J.C.S., Loiseau, C., Marceau, M., Jouet, A., Menardo, F., Tzfadia, O., Antoine, R., Niyigena, E.B., Mulders, W., Fissette, K., 2020. A sister lineage of the Mycobacterium tuberculosis complex discovered in the African Great Lakes region. Nature Communications 11, 1–11.
Nystrom, K.C., 2009. The reconstruction of identity: A case study from Chachapoya, Peru. Bioarchaeology and Identity in the Americas 82–102.
Owen, B., 1993. A Model of Multiethnicity: State Collapse, Competition, and Social Complexity from Tiwanaku to Chiribaya in the Osmore Valley (Ph.D. dissertation). University of Los Angeles, Los Angeles, California.
Peltzer, A., Jäger, G., Herbig, A., Seitz, A., Kniep, C., Krause, J., Nieselt, K., 2016. EAGER: efficient ancient genome reconstruction. Genome Biology 17, 60. https://doi.org/10.1186/s13059-016-0918-z
Pinhasi, R., Fernandes, D., Sirak, K., Novak, M., Connell, S., Alpaslan-Roodenberg, S., Gerritsen, F., Moiseyev, V., Gromov, A., Raczky, P., 2015. Optimal ancient DNA yields from the inner ear part of the human petrous bone. PloS one 10, e0129102.
Quilter, J., 2013. The ancient central Andes. Routledge.
Rivas Boada, A.M., 1988. Las patologías óseas en la población de Marín. Boletín de Arqueología de la Fian 3, 3–24.
Roberts, C.A., Buikstra, J.E., 2019. Bacterial infections, in: Ortner's Identification of Pathological Conditions in Human Skeletal Remains. Elsevier, pp. 321–439.
Roberts, C.A., Buikstra, J.E., 2003a. The bioarchaeology of tuberculosis: a global perspective on a re-emerging disease. University Press of Florida.
Roberts, C.A., Buikstra, J.E., 2003b. The bioarchaeology of tuberculosis: a global perspective on a re-emerging disease. University Press of Florida.
Rohland, N., Harney, E., Mallick, S., Nordenfelt, S., Reich, D., 2015. Partial uracil–DNA–glycosylase treatment for screening of ancient DNA. Philosophical Transactions of the Royal Society B: Biological Sciences 370, 20130624.

Rombouts, C., Giraud, T., Jeanneau, C., About, I., 2017. Pulp vascularization during tooth development, regeneration, and therapy. Journal of dental research 96, 137–144.
Salo, W.L., Aufderheide, A.C., Buikstra, J., Holcomb, T.A., 1994. Identification of Mycobacterium tuberculosis DNA in a pre-Columbian Peruvian mummy. Proceedings of the National Academy of Sciences 91, 2091–2094.
Schjellerup, I.R., 1997. Incas and Spaniards in the Conquest of the Chachapoyas. Archaeological and Ethnohistorical Research in the North-eastern Andes of Peru.
Schreiber, K.J., 1991. The association between roads and polities: evidence for Wari roads in Peru. Ancient road networks and settlement hierarchies in the New World 42–53.
Schubert, M., Lindgreen, S., Orlando, L., 2016a. AdapterRemoval v2: rapid adapter trimming, identification, and read merging. BMC research notes 9, 1–7.
Schubert, M., Lindgreen, S., Orlando, L., 2016b. AdapterRemoval v2: rapid adapter trimming, identification, and read merging. BMC Research Notes 9, 88. https://doi.org/10.1186/s13104-016-1900-2
Schuenemann, V.J., Kumar Lankapalli, A., Barquera, R., Nelson, E.A., Iraíz Hernández, D., Acuña Alonzo, V., Bos, K.I., Márquez Morfín, L., Herbig, A., Krause, J., 2018. Historic Treponema pallidum genomes from Colonial Mexico retrieved from archaeological remains. PLoS neglected tropical diseases 12, e0006447.
Spyrou, M.A., Bos, K.I., Herbig, A., Krause, J., 2019. Ancient pathogen genomics as an emerging tool for infectious disease research. Nature Reviews Genetics 20, 323–340. https://doi.org/10.1038/s41576-019-0119-1
Spyrou, M.A., Tukhbatova, R.I., Feldman, M., Drath, J., Kacki, S., de Heredia, J.B., Arnold, S., Sitdikov, A.G., Castex, D., Wahl, J., 2016. Historical Y. pestis genomes reveal the European Black Death as the source of ancient and modern plague pandemics. Cell Host & Microbe 19, 874–881.
Stamatakis, A., 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics 30, 1312–1313. https://doi.org/10.1093/bioinformatics/btu033
Steinbock, R.T., 1976. Paleopathological diagnosis and interpretation: bone diseases in ancient human populations. Charles C Thomas Pub Limited.
Toyne, J.M., Esplin, N., Buikstra, J.E., 2020. Examining variation in skeletal tuberculosis in a late pre-contact population from the eastern mountains of Peru. International Journal of Paleopathology 30, 22–34.
Toyne, J.M., Narváez, A., 2014. The fall of Kuelap: Bioarchaeological analysis of death and destruction on the eastern slopes of the Andes. Embattled bodies, embattled places. Washington DC: Dumbarton Oaks Research Library and Collection. p 341–364.
Toyne, J.M., Schow, C., Esplin, N., in press. Beyond tuberculosis: Identifying other pathological spinal conditions in ancient Chachapoya human remains.
Tuli, S.M., 2016. Tuberculosis of the skeletal system. JP Medical Ltd.
Tung, T.A., 2012. Violence, ritual, and the Wari Empire: A social bioarchaeology of imperialism in the ancient Andes. University Press of Florida.
Tung, T.A., Miller, M., DeSantis, L., Sharp, E.A., Kelly, J., 2016. Patterns of Violence and Diet Among Children During a Time of Imperial Decline and Climate Change in the Ancient Peruvian Andes, in: VanDerwarker, A.M., Wilson, G.D. (Eds.), The Archaeology of Food and Warfare: Food Insecurity in Prehistory. Springer International Publishing, Cham, pp. 193–228. https://doi.org/10.1007/978-3-319-18506-4_10

Vågene, Å.J., Herbig, A., Campana, M.G., García, N.M.R., Warinner, C., Sabin, S., Spyrou, M.A., Valtueña, A.A., Huson, D., Tuross, N., 2018. Salmonella enterica genomes from victims of a major sixteenth-century epidemic in Mexico. Nature ecology & evolution 2, 520–528.

Warinner, C., Herbig, A., Mann, A., Fellows Yates, J.A., Weiß, C.L., Burbano, H.A., Orlando, L., Krause, J., 2017. A Robust Framework for Microbial Archaeology. Annual Review of Genomics and Human Genetics 18, 321–356. https://doi.org/10.1146/annurev-genom-091416-035526

World Health Organization, 2019. WHO Global Tuberculosis Report. WHO, Geneva, Switzerland.

Tuberculosis in a Pre-colonial Chachapoya Individual from the northeastern Peruvian Andes

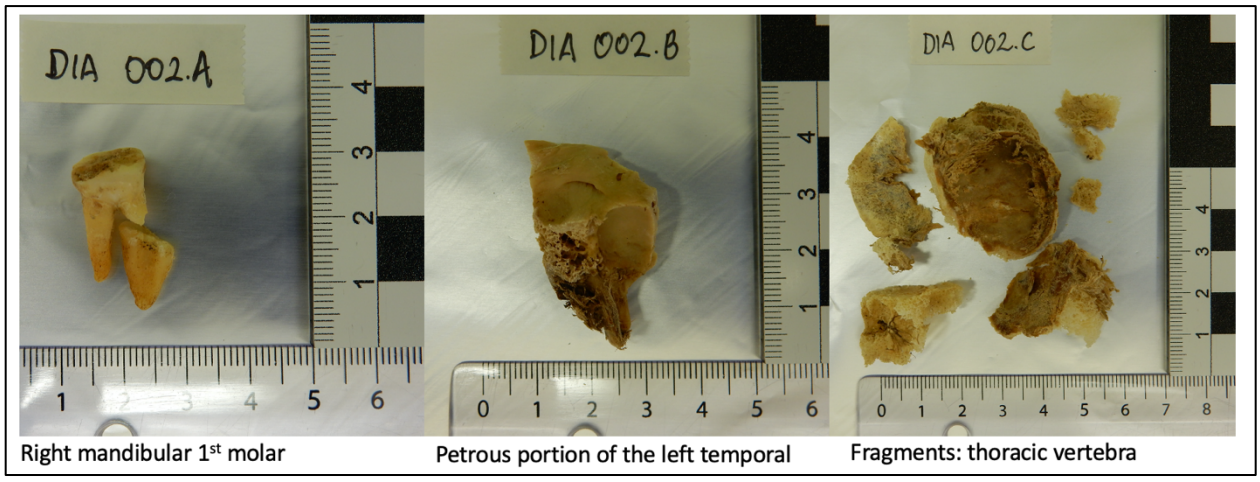
SUPPLEMENTARY INFORMATION

Authors: Elizabeth A. Nelson[1,2], Evelyn Guevara[1,3], J. Marla Toyne[4], Alexander Herbig[1], Johannes Krause[1], Kirsten I. Bos[1]

Affiliations:
[1]Department of Archaeogenetics, Max Planck Institute for the Science of Human History, Germany
[2]Institute for Archaeological Sciences, University of Tübingen, Germany
[3]Department of Forensic Medicine, University of Helsinki, Finland
[4]Department of Anthropology, University of Central Florida, USA

CONTENTS:

Supplementary Figure 1. Images of the individual (DIA002) recovered from Diablo Huasi. The view on the left shows severe kyphosis of the spine due to the pathological destruction of the vertebral body. The image on the right shows consecutive multi-focal lesions of the thoracic vertebrae.



Supplementary Figure 2. Images of the samples from the individual in this study (DIA002).

Supplementary Table 1. List of all genomes included in phylogenetic analysis. All genomes with accession numbers are listed. Names and radiocarbon dates are listed when available.

Name	Study	Sample Accession	Lineage (all *M. microti* and *M. pinnipedii* listed at species level)	Date (ancient only)
DIA002.C	this	*NA*	*M. pinnipedii*	1296 - 1396 CE
HUA004	paper III of thesis	*NA*	*M. pinnipedii*	1044 - 1155 CE
HUA006	paper III of thesis	*NA*	*M. pinnipedii*	1316 - 1405 CE
HUA016	paper III of thesis	*NA*	*M. pinnipedii*	1305 - 1396 CE
HUA037	paper III of thesis	*NA*	*M. pinnipedii*	1326 - 1425 CE
HUA041	paper III of thesis	*NA*	*M. pinnipedii*	1321 - 1410 CE
HUA051	paper III of thesis	*NA*	*M. pinnipedii*	*NA*
HUA057	paper III of thesis	*NA*	*M. pinnipedii*	1314 - 1399 CE
54U	Bos et al. 2014	SRS591994	*M. pinnipedii*	1106 CE
58U	Bos et al. 2014	SRS591992	*M. pinnipedii*	1117 CE
64U	Bos et al. 2014	SRS592075	*M. pinnipedii*	1210.5 CE
Pinnipedii_G01222	Bos et al. 2014	SRS592134	*M. pinnipedii*	
Pinnipedii_G01491	Bos et al. 2014	SRS592135	*M. pinnipedii*	
Pinnipedii_G01492	Bos et al. 2014	SRS592136	*M. pinnipedii*	
Pinnipedii_G01498	Bos et al. 2014	SRS592137	*M. pinnipedii*	
Pinnipedii7011	Bos et al. 2014	SRS592823	*M. pinnipedii*	
Pinnipedii7739	Bos et al. 2014	SRS592824	*M. pinnipedii*	
MP2	Silva-Pereira et al., 2019	SRR7693090	*M. pinnipedii*	
MP1	Silva-Pereira et al., 2019	SRR7693584	*M. pinnipedii*	
pinnipedii SRR552768	Malm et al. 2017	ERR552768	*M. pinnipedii*	
	Malm et al. 2017	ERR551111	*M. microti*	
	Malm et al. 2017	ERR552037	*M. microti*	
	Malm et al. 2017	ERR553376	*M. microti*	
	The Wellcome Sanger Institute	ERR027294	*M. microti*	
	Brites et al., 2018	ERR2659166	*M. microti*	
	Brites et al., 2018	ERR2659167	*M. microti*	
	Brites et al., 2018	ERR2659168	*M. microti*	

	Brites et al., 2018	ERR2659169	*M. microti*	
	Brites et al., 2018	ERR2659170	*M. microti*	
	Brites et al., 2018	ERR2659171	*M. microti*	
	PATRIC	SRR3647357	*M. microti*	
bovis_ravenel	Comas et al. 2013	SRS004666	animal	
capraeD028	Domogalla et al. 2013	SRS386255	animal	
capraeRW044	Domogalla et al. 2013	SRS386245	animal	
capraeRW079	Domogalla et al. 2013	SRS386249	animal	
chimpanzee_bacillus	Coscolla et al. 2013	ERS635608	animal	
L1_1078602	Comas et al. 2013	ERS218206	L1	
L1_1105103	Comas et al. 2013	ERS218207	L1	
L1_1277803	Comas et al. 2013	ERS218208	L1	
L1_157199	Comas et al. 2013	ERS218215	L1	
L1_157508	Comas et al. 2013	ERS218216	L1	
L1_179703	Comas et al. 2013	ERS218220	L1	
L1_468303	Comas et al. 2013	ERS218233	L1	
L1_539606	Comas et al. 2013	ERS218236	L1	
L1_600603	Comas et al. 2013	ERS218242	L1	
L1_745902	Comas et al. 2013	ERS218245	L1	
L1_94701	Comas et al. 2013	ERS218248	L1	
L1_950545	Comas et al. 2013	SRS004759	L1	
L1_BTBS280	Comas et al. 2013	ERS217639	L1	
L1_BTBS493	Comas et al. 2013	ERS217641	L1	
L1_DY28	Comas et al. 2013	ERS218256	L1	
L1_GT281	Comas et al. 2013	ERS218258	L1	
L1_K21	Comas et al. 2013	SRX002001	L1	
L1_K67	Comas et al. 2013	SRX002004	L1	
L1_K93	Comas et al. 2013	SRX002005	L1	
L1_N0014	Comas et al. 2013	ERS218265	L1	
L1_N0043	Comas et al. 2013	ERS218268	L1	
L1_N0062	Comas et al. 2013	ERS217651	L1	
L1_N0065	Comas et al. 2013	ERS217652	L1	
L1_N0067	Comas et al. 2013	ERS217644	L1	
L1_N0079	Comas et al. 2013	ERS217653	L1	
L1_N0127	Comas et al. 2013	ERS218313	L1	
L1_N0132	Comas et al. 2013	ERS218275	L1	
L1_N0134	Comas et al. 2013	ERS218315	L1	
L1_N0141	Comas et al. 2013	ERS218317	L1	

L1_N0182	Comas et al. 2013	ERS217660	L1	
L1_N0196	Comas et al. 2013	ERS218323	L1	
L1_N0203	Comas et al. 2013	ERS217662	L1	
L1_N1004	Comas et al. 2013	ERS217664	L1	
L1_N1006	Comas et al. 2013	ERS217665	L1	
L1_N1009	Comas et al. 2013	ERS217668	L1	
L1_N70	Comas et al. 2013	ERS218286	L1	
L1_N72	Comas et al. 2013	ERS218287	L1	
L1_N73	Comas et al. 2013	ERS218288	L1	
L1_T92	Comas et al. 2013	SRS004759	L1	
L1_V232IO	Comas et al. 2013	ERS218289	L1	
L1_V346IO	Comas et al. 2013	ERS218291	L1	
L1_V372IO	Comas et al. 2013	ERS218292	L1	
L2_1027	Comas et al. 2013	ERS218149	L2	
L2_1336	Comas et al. 2013	ERS218151	L2	
L2_1864	Comas et al. 2013	ERS218153	L2	
L2_72	Comas et al. 2013	ERS218159	L2	
L2_759	Comas et al. 2013	ERS218161	L2	
L2_GQ-1165	Comas et al. 2013	ERS218168	L2	
L2_GQ-1168	Comas et al. 2013	ERS218169	L2	
L2_GQ-1343	Comas et al. 2013	ERS218172	L2	
L2_GQ-1438	Comas et al. 2013	ERS218173	L2	
L2_GQ-1972	Comas et al. 2013	ERS218180	L2	
L2_GQ-50	Comas et al. 2013	ERS218185	L2	
L2_GQ10-3	Comas et al. 2013	ERS218166	L2	
L2_GQ1020	Comas et al. 2013	ERS218165	L2	
L2_GQ1164	Comas et al. 2013	ERS218167	L2	
L2_GQ1331	Comas et al. 2013	ERS218170	L2	
L2_GQ1335	Comas et al. 2013	ERS218171	L2	
L2_GQ1439	Comas et al. 2013	ERS218174	L2	
L2_GQ1580	Comas et al. 2013	ERS218175	L2	
L2_GQ1597	Comas et al. 2013	ERS218176	L2	
L2_GQ1605	Comas et al. 2013	ERS218177	L2	
L2_GQ1645	Comas et al. 2013	ERS218178	L2	
L2_GQ1885	Comas et al. 2013	ERS218179	L2	
L2_GQ1973	Comas et al. 2013	ERS218181	L2	
L2_GQ229	Comas et al. 2013	ERS218182	L2	
L2_GQ254	Comas et al. 2013	ERS218183	L2	

L2_GQ366	Comas et al. 2013	ERS218184	L2	
L2_GQ657	Comas et al. 2013	ERS218186	L2	
L2_GQ73	Comas et al. 2013	ERS218187	L2	
L2_GQ74-3	Comas et al. 2013	ERS218188	L2	
L2_GQ762	Comas et al. 2013	ERS218189	L2	
L2_GQ856	Comas et al. 2013	ERS218190	L2	
L2_GQ859	Comas et al. 2013	ERS218191	L2	
L2_GQ904	Comas et al. 2013	ERS218192	L2	
L2_GT333	Comas et al. 2013	ERS218259	L2	
L2_GT345	Comas et al. 2013	ERS218260	L2	
L2_GT411	Comas et al. 2013	ERS218261	L2	
L2_GT649	Comas et al. 2013	ERS218262	L2	
L2_M4100A	Comas et al. 2013	SRS004757	L2	
L2_N0002	Comas et al. 2013	ERS218296	L2	
L2_N0003	Comas et al. 2013	ERS218297	L2	
L2_N0005	Comas et al. 2013	ERS218298	L2	
L2_N0008	Comas et al. 2013	ERS217645	L2	
L2_N0010	Comas et al. 2013	ERS218264	L2	
L2_N0017	Comas et al. 2013	ERS217647	L2	
L2_N0020	Comas et al. 2013	ERS218266	L2	
L2_N0034	Comas et al. 2013	ERS218299	L2	
L2_N0039	Comas et al. 2013	ERS218267	L2	
L2_N0041	Comas et al. 2013	ERS218300	L2	
L2_N0044	Comas et al. 2013	ERS218301	L2	
L2_N0050	Comas et al. 2013	ERS217649	L2	
L2_N0051	Comas et al. 2013	ERS218303	L2	
L2_N0053	Comas et al. 2013	ERS218304	L2	
L2_N0094	Comas et al. 2013	ERS218307	L2	
L2_N0128	Comas et al. 2013	ERS217658	L2	
L2_N0130	Comas et al. 2013	ERS218314	L2	
L2_N0150	Comas et al. 2013	ERS218320	L2	
L2_N0151	Comas et al. 2013	ERS218321	L2	
L2_N0158	Comas et al. 2013	ERS218322	L2	
L2_N1001	Comas et al. 2013	ERS217663	L2	
L2_N1037	Comas et al. 2013	ERS217674	L2	
L2_N1051	Comas et al. 2013	ERS217676	L2	
L2_rus14	Casali et al. 2012	ERS003236	L2	
L2_rus16	Casali et al. 2012	ERS003237	L2	

L2_V119BJ	Comas et al. 2013	ERS218244	L2	
L2_V212BJ	Comas et al. 2013	ERS218241	L2	
L2_V374BJ	Comas et al. 2013	ERS218293	L2	
L3_1016204	Comas et al. 2013	ERS218204	L3	
L3_1180C	Comas et al. 2013	ERS218150	L3	
L3_155808	Comas et al. 2013	ERS218213	L3	
L3_159008	Comas et al. 2013	ERS218217	L3	
L3_162908	Comas et al. 2013	ERS218218	L3	
L3_173305	Comas et al. 2013	ERS218219	L3	
L3_282801	Comas et al. 2013	ERS218229	L3	
L3_296904	Comas et al. 2013	ERS218230	L3	
L3_449706	Comas et al. 2013	ERS218232	L3	
L3_580605	Comas et al. 2013	ERS218239	L3	
L3_597805	Comas et al. 2013	ERS218240	L3	
L3_657005	Comas et al. 2013	ERS218243	L3	
L3_751B	Comas et al. 2013	ERS218160	L3	
L3_793601	Comas et al. 2013	ERS218246	L3	
L3_855B	Comas et al. 2013	ERS218162	L3	
L3_858405	Comas et al. 2013	ERS218247	L3	
L3_910079	Comas et al. 2013	SRS004758	L3	
L3_BTBH273	Comas et al. 2013	ERS217635	L3	
L3_K49	Comas et al. 2013	SRS001884	L3	
L3_MTB_138001	Comas et al. 2013	ERS218209	L3	
L3_N0033	Comas et al. 2013	ERS217648	L3	
L3_N0054	Comas et al. 2013	ERS218269	L3	
L3_N0056	Comas et al. 2013	ERS218270	L3	
L3_N0197	Comas et al. 2013	ERS218324	L3	
L3_N04	Comas et al. 2013	ERS218283	L3	
L3_N1007	Comas et al. 2013	ERS217666	L3	
L3_N1014	Comas et al. 2013	ERS217669	L3	
L3_N1022	Comas et al. 2013	ERS217671	L3	
L3_N1024	Comas et al. 2013	ERS217672	L3	
L3_N1032	Comas et al. 2013	ERS217673	L3	
L3_N1040	Comas et al. 2013	ERS217675	L3	
L3_N1058	Comas et al. 2013	ERS217679	L3	
L3_N24	Comas et al. 2013	ERS218284	L3	
L3_N37	Comas et al. 2013	ERS218285	L3	
L3_rus15	Casali et al. 2012	ERS003250	L3	

L4_1026403	Comas et al. 2013	ERS218205	L4	
L4_141702	Comas et al. 2013	ERS218210	L4	
L4_155008	Comas et al. 2013	ERS218212	L4	
L4_157108	Comas et al. 2013	ERS218214	L4	
L4_217399	Comas et al. 2013	ERS218222	L4	
L4_219799	Comas et al. 2013	ERS218223	L4	
L4_230799	Comas et al. 2013	ERS218224	L4	
L4_231806	Comas et al. 2013	ERS218225	L4	
L4_233602	Comas et al. 2013	ERS218226	L4	
L4_267903	Comas et al. 2013	ERS218228	L4	
L4_478304	Comas et al. 2010	SRS004761	L4	
L4_549204	Comas et al. 2013	ERS218238	L4	
L4_BTBH587	Comas et al. 2013	ERS217636	L4	
L4_BTBS101	Comas et al. 2013	ERS217638	L4	
L4_BTBS458	Comas et al. 2013	ERS217640	L4	
L4_CDC1551E	Comas et al. 2012	SRS005450	L4	
L4_DY131	Comas et al. 2013	ERS218249	L4	
L4_DY167	Comas et al. 2013	ERS218251	L4	
L4_DY195	Comas et al. 2013	ERS218252	L4	
L4_DY22	Comas et al. 2013	ERS218254	L4	
L4_DY8	Comas et al. 2013	ERS218257	L4	
L4_erdman	Comas et al. 2013	SRS003328	L4	
L4_H37Rv	Comas et al. 2013	ERS218263	L4	
L4_K37	Comas et al. 2010	SRS001887	L4	
L4_MT0001	Gardy et al. 2011	SRS074557	L4	
L4_MT0005	Gardy et al. 2011	SRS084142	L4	
L4_N0011	Comas et al. 2013	ERS217646	L4	
L4_N0046	Comas et al. 2013	ERS218302	L4	
L4_N0101	Comas et al. 2013	ERS218271	L4	
L4_N0103	Comas et al. 2013	ERS218309	L4	
L4_N0107	Comas et al. 2013	ERS218310	L4	
L4_N0109	Comas et al. 2013	ERS218311	L4	
L4_N0120	Comas et al. 2013	ERS218272	L4	
L4_N0125	Comas et al. 2013	ERS217657	L4	
L4_N0126	Comas et al. 2013	ERS218273	L4	
L4_N0131	Comas et al. 2013	ERS218274	L4	
L4_N0135	Comas et al. 2013	ERS218276	L4	
L4_N0136	Comas et al. 2013	ERS218316	L4	

L4_N0137	Comas et al. 2013	ERS218277	L4	
L4_N0138	Comas et al. 2013	ERS217659	L4	
L4_N0142	Comas et al. 2013	ERS218278	L4	
L4_N0143	Comas et al. 2013	ERS218279	L4	
L4_N0144	Comas et al. 2013	ERS218280	L4	
L4_N0146	Comas et al. 2013	ERS218318	L4	
L4_N0148	Comas et al. 2013	ERS218319	L4	
L4_N0149	Comas et al. 2013	ERS218281	L4	
L4_N0163	Comas et al. 2013	ERS218282	L4	
L4_N0185	Comas et al. 2013	ERS217661	L4	
L4_N1008	Comas et al. 2013	ERS217667	L4	
L4_N1015	Comas et al. 2013	ERS217670	L4	
L4_N1052	Comas et al. 2013	ERS217677	L4	
L4_N1057	Comas et al. 2013	ERS217678	L4	
L4_V173IO	Comas et al. 2013	ERS218221	L4	
L4_V293EA	Comas et al. 2013	ERS218231	L4	
L4_V318EA	Comas et al. 2013	ERS218290	L4	
L4_V355EA	Comas et al. 2013	ERS218203	L4	
L4_V367IO	Comas et al. 2013	ERS218234	L4	
L4_V440EA	Comas et al. 2013	ERS218294	L4	
L4_V639EA	Comas et al. 2013	ERS218295	L4	
L4_X581	Comas et al. 2012	SRS004841	L4	
L4_X632	Comas et al. 2012	SRS004831	L4	
L4_X721	Comas et al. 2012	SRS005448	L4	
L5_1001003	Comas et al. 2013	ERS218148	L5	
L5_1047301	Comas et al. 2013	ERS218193	L5	
L5_1048001	Comas et al. 2013	ERS218194	L5	
L5_1182103	Comas et al. 2010	SRS004762	L5	
L5_144902	Comas et al. 2013	ERS218211	L5	
L5_256902	Comas et al. 2013	ERS218195	L5	
L5_257702	Comas et al. 2013	ERS218227	L5	
L5_348203	Comas et al. 2013	ERS218196	L5	
L5_349404	Comas et al. 2013	ERS218197	L5	
L5_533304	Comas et al. 2013	ERS218198	L5	
L5_544404	Comas et al. 2013	SRS004763	L5	
L5_553604	Comas et al. 2013	ERS218200	L5	
L5_DY135	Comas et al. 2013	ERS218250	L5	
L5_DY20	Comas et al. 2013	ERS218253	L5	

L5_DY21	Comas et al. 2013	ERS218164	L5	
L5_DY26	Comas et al. 2013	ERS218255	L5	
L6_414104	Comas et al. 2010	SRS004764	L6	
L6_533604	Comas et al. 2013	ERS218199	L6	
L6_538302	Comas et al. 2013	ERS218235	L6	
L6_541504	Comas et al. 2013	ERS218157	L6	
L6_546802	Comas et al. 2013	ERS218237	L6	
L6_823602	Comas et al. 2013	ERS218201	L6	
L6_ABCG	Comas et al. 2013	ERS218202	L6	
L6_GM0981	Comas et al. 2010	SRS004760	L6	
L6_N0060	Comas et al. 2013	ERS217650	L6	
L6_N0089	Comas et al. 2013	ERS217654	L6	
L6_N0090	Comas et al. 2013	ERS217655	L6	
L6_N0091	Comas et al. 2013	ERS218305	L6	
L6_N0092	Comas et al. 2013	ERS218306	L6	
L6_N0098	Comas et al. 2013	ERS217656	L6	
L6_N0099	Comas et al. 2013	ERS218308	L6	
L6_N0115	Comas et al. 2013	ERS218312	L6	
L7_BTBH1012	Comas et al. 2013	ERS217634	L7	
L7_BTBH935	Comas et al. 2013	ERS217637	L7	
L7_BTBS610	Comas et al. 2013	ERS217642	L7	
L7_BTBS746	Comas et al. 2013	ERS217643	L7	
outgroup_canettii	Comas et al. 2010	SRS002004	outgroup	

Paper V

E. A. Nelson, K. E. Blevins, J. E. Buikstra, J. M. Lory, A. Herbig, J. Krause, A. C. Stone, K. I. Bos (2020).

Preliminary investigations of MTBC across the pre-colonial Americas

Manuscript

Preliminary investigations of MTBC across the pre-colonial Americas

Authors: Elizabeth A. Nelson[1,2†], Kelly E. Blevins[3†], Jane E. Buikstra[3], Josefina Mansilla Lory[4], Alexander Herbig[1], Johannes Krause[1], Anne C. Stone[3,5], and Kirsten I. Bos[1]

Affiliations:
[1]Max Planck Institute for the Science of Human History, Jena, Germany
[2]Institute for Archaeological Sciences, University of Tübingen, Tübingen, Germany
[3]School of Human Evolution and Social Change, Arizona State University, Tempe, Arizona, USA
[4]Dirección de Antropología Física, Insituto Nacional de Antropolgía e Histora, Mexico City, Mexico
[5]Institute of Human Origins, Arizona State University, Tempe, Arizona, USA

[†]These authors contributed equally to this work

ABSTRACT

Skeletal evidence of tuberculosis (TB) is abundant across pre-colonial archaeological contexts in the Americas. Although the existence of pre-colonial "New World" TB was historically debated under the assumption the disease was introduced by European colonists, interdisciplinary analyses incorporating microscopic and molecular data continued to support skeletal evidence. The eventual identification of *Mycobacterium tuberculosis complex* (MTBC) DNA in vertebrae from Late Intermediate Period (LIP, 1000 – 1400 CE) contexts of the Osmore River Valley (ORV) in Peru led to the subsequent reconstruction and characterization of three pre-colonial MTBC genomes. Analyses revealed that the strains infecting the ORV populations were closely related to a strain currently adapted for seals and sea lions, *M. pinnipedii*, thus suggesting a zoonotic introduction of TB to South America. Since this discovery, the characterization of additional MTBC genomes in the Andean highlands of Peru and the high-altitude subtropical cloud forest of the eastern Andean slopes of Peru reveal *M. pinnipedii* was circulating in various LIP Peruvian geographic and ecological zones. However, the full geographic presence of M. *pinnipedii* and its involvement in pre-colonial TB across the Americas, remains unknown. To explore MTBC strain diversity in the Americas we molecularly investigated skeletal evidence of TB recovered from various geographic locations in different environments in North and South America. Here, we present preliminary data of our investigations into MTBC strains across the pre-colonial Americas with a sample set from sites across the North and South American continents representing various environments including Canada, California, Mexico, Peru, Chile, Colombia and Argentina.

INTRODUCTION

Although one thought of as a disease of antiquity, tuberculosis (TB) is currently the leading cause of death due to an infectious agent. In 2019, the World Health Organization (WHO) reported that an estimated total 10.5 million people were infected with TB. That year the disease claimed 1.8 million lives, with highest incidence rates reported in Africa, Asia, and South America (WHO, 2019). It is estimated that today one in three people is infected with some form of latent, active, multi-drug resistant or extensively drug resistant tuberculosis. This socially disruptive and devastating disease is caused by members of the *Mycobacterium tuberculosis* complex (MTBC). Features of TB such as latency, respiratory and dietary transmission, and extensive host ranges which including a variety of mammals, have made the MTBC one of the most successful and ubiquitous pathogen groups in human history (Brites et al., 2018; Brosch et al., 2002). The MTBC is comprised of eight lineages that are phylogeographically distributed with lineage 4 (L4) having the greatest global presence today, dominating Europe and the Americas, (Gagneux et al., 2006; Ngabonziza et al., 2020) and likely reflecting the colonial introduction of this strain.

Paleopathological evidence of TB from archaeological contexts is found across the globe suggesting a long history of human infection (Ortner et al., 2012; Robert and Buikstra, 2003; Stone et al., 2009). The abundance of osteological evidence of TB in the pre-colonial North and South Americas (Allison et al., 1973; Allison et al., 1981; Arriaza et al., 1995; Arateco, 1998; Buikstra, 1976; Buikstra, 1981; Buikstra and Williams, 1991; García-Frías, 1940; Guichón et al., 2015; Guillén, 2012; Klaus et al., 2010; Requena, 1945; Roberts and Buikstra, 2003) has provided the opportunity to understand the evolution and ecology of MTBC members circulating in South America prior European arrival. However, TB in the Americas prior to Columbian contact remained controversial for some time despite the skeletal, soft tissue and molecular evidence (Wilbur and Buikstra, 2006). For example, the first genetic detection of ancient TB from pre-colonial South America was published in 1994 (Salo et al., 1994) and although these novel results were supported by soft tissue and skeletal evidence, they were contested by some contemporaries stating it was likely the detection of environmental mycobacteria (Stead et al., 1995). However, in 2014, the application of high throughput sequencing methods to pre-colonial archaeological human remains from the Osmore River Valley (ORV) in Peru allowed for reconstruction of three ancient TB genomes. Subsequent phylogenetic analysis revealed strains circulating in the ORV to be most closely related to a member of the MTBC currently associated with seals and sea lions, *M. pinnipedii* (Bos et al., 2014). This suggests the introduction of TB to the Americas through a zoonotic event of marine mammal to human transmission thus providing some evidence for the source of pre-colonial TB. Since that time human remains displaying tuberculosis like lesions from multiple sites across Peru, contemporaneous with reconstructed genomes from the ORV, have been analyzed both skeletally and molecularly to better understand the geographic expanse of this strain and MTBC strain diversity in the pre-colonial Americas. Multiple studies have identified

ancient strains closely related to *M. pinnipedii* from various locations and ecologies in South American, though they are currently unpublished (paper III and IV of this thesis; Vågene et al., data unpublished).

Although *M. pinnipedii* has been identified from ancient Peruvian contexts, our understanding of MTBC strain diversity across the Americas prior to European arrival remains underexplored. To that end, in collaboration with Dr. Anne C. Stone, Dr. Jane E. Buikstra, and their graduate student Kelly Blevins from Arizona State University, we use molecular approaches to investigate pre-colonial MTBC strains through the applying ancient DNA methods to skeletal evidence of TB from archaeological contexts across the Americas. Here, I present our preliminary data from the application of high through put methods for molecular screening of 28 samples (Table 1) from South American and North American archaeological contexts (Supplementary Figure 1). I discuss the promising results and precautions of this preliminary analysis and paths forward for this project.

RESULTS

Sequenced capture data mapping statistics

The 28 samples included in this project came from archaeological context across the Americas, with representation from North American contexts, including Mesoamerica, and the majority of samples (n=15) from South American contexts. These samples had previously been screened for MTBC DNA using quantitative PCR methods (qPCR) using TaqMan qPCR assays. These assays target the *rpoB* gene, and two insertion elements found in MTBC: IS6110 and IS1081. The assay targeting the *rpoB* gene includes a TaqMan probe that binds to a sequence in the gene that is likely specific to MTBC (Harkins et al. 2015). All putatively positive samples were then prepared as UDG and non-UDG treated libraries for a total of 56 libraries. Mapping of the sequenced capture data to the TB ancestor reference (MTB_anc) (Comas et al., 2010) revealed that out of 56 libraries, six libraries had greater than 1x coverage of the reference genome (Table 2). For extended mapping results showing all samples please see supplementary tables 1 and 3. This included four UDG treated libraries (18881/427.3, 18883/429.1, 18891/492, 18892/495) and two non-UDG treated libraries which were the counterparts to UDG treated libraries 18891/492 and 18892/495. Concerning the UDG treated libraries were used for downstream phylogenetic analysis, coverage of the TB reference (MTB_anc) ranged from 1.24 to 6.81 mean fold. Three of the bone samples, 18881/427.3 (Tenochtitlan, Mexico), 18883/429.1 (Tlatelolco, Mexico) and 18892/495 (Pica 8, Chile), are recorded to have been collected from vertebrae, however, sample 18891/492 (Pica 8, Chile) is recorded to have been collected from a tooth. This would be the first genome reconstruction of ancient TB from a dental sample. Out of all reconstructed genomes, only 18892/495 from Pica 8, Chile, has met the minimum requirement for downstream analyses (5-fold mean coverage).

In further inspection of the mapping statistics, the damage patterns for the libraries which were prepared without UDG treatment show an unusual pattern. While the non-UDG library for the Chilean dental sample 492 (18852) shows approximately 0.04% damage which is a low percentage but within normal variation, the other three samples present extremely low percentages of DNA damage, ranging from approximately 1-2%. This inhibits the authentication of ancient DNA by traditional methods due the absence of an observable damage pattern. This was investigated further by observing damage plots generated by mapDamage2.0 (Jónsson et al., 2013) which confirm unusually low patterns. No positive control was included in extraction and library preparation and therefore could experimental conditions of extraction and library preparation not be examined.

Although these libraries underwent in-solution capture for MTBC enrichment we mapped sequencing data to the human reference genome (Hg19) to evaluate damage patterns associated with reads mapping to the human genome (Supplementary Table 2 and 4). The results of mapping to the human reference genome (Hg19) show all samples have thousands of reads mapping to Hg19 with a range from 3,343 to 892,471 reads mapping after quality filtering. Human endogenous DNA percentages vary in samples between 0.48 and 9.38. Results also show the human DNA damage is unusually low for typical ancient samples, similar to the observed damage in TB mapping with a slight increase. Of the four samples which produced MTBC genomes with at least 1-fold coverage, human DNA damage percentages range from 2% to 5%. Similar to the TB reference mapping, evaluation of damage after mapping to the human reference reveals the damage patterns for sample 492 are within normal variation while the other three samples are much lower than what is normally observed. The absence of these expected patterns of damage of ancient DNA challenges the authentication of the antiquity of the sample (Briggs et al., 2007; Sawyer et al., 2012).

Metagenomic Analysis of Captured Samples

In order to evaluate the microbial composition of the capture products, specifically the mycobacterial reads previously detected, we performed MALT analysis (Vågene et al., 2017) which would provide a metagenomic profile of our samples of interest. This would allow us to observe modern mycobacteria which because of the abundance of conserved genomic regions of mycobacterial species, may be contributing reads to the reconstructed genomes and influencing the damage patterns. We performed MALT analysis on all samples using a 95% sequence identity. Inspection of the node assignments as visualized in MEGAN (Huson et al., 2018) revealed the four genomes, 18881/427.3 (Tenochtitlan, Mexico), 18883/429.1 (Tlatelolco, Mexico), 18891/492 (Pica 8, Chile), and 18892/495 (Pica 8, Chile), yielded sequencing reads assigning to MTBC ranging from 8582 to 1,049,781 (Supplementary Table 5). Further inspection of the alignments supported the identification of MTBC in the samples. However, evaluation of mycobacterial background showed the composition of the library was not over abundant with non-MTBC mycobacteria (Supplementary Table 5) and therefore environmental mycobacteria was ruled out

as a cause to the low damage patterns we observed. Additionally, results of MALT analysis also showed that although the tooth sample from Pica 8, Chile yielded a TB genome (18891/492) it did not show any DNA evidence of oral microbiome members, however, this absence of oral microbes may be due to decontamination using a bleach dilution and the removal of the outer layer of the tooth root using a Dremel diamond wheel before sampling.

A preliminary phylogenetic analysis

Although ancient MTBC genomic analyses call for a minimum of 5-fold coverage for confident SNP calling in phylogenetics and evolutionary analysis (papers III and IV of this thesis), we lowered our minimum coverage requirement to 3-fold mean coverage in order to build a preliminary phylogeny including the 4 newly reconstructed genomes. In order to meet the required 3-fold coverage, the sequencing reads for 18883/429.1 (Tlatelolco, Mexico) which showed a 1.24 mean fold coverage, was concatenated with itself, doubling the sequencing data to creating an artificial increase. Because 18892/495 (Pica 8, Chile) also had a mean fold coverage below the minimum 3-fold threshold (1.90 mean fold coverage), sequencing reads for this sample were also concatenated to provide an artificial increase to meet the threshold. These modifications would permit the construction of a preliminary phylogeny including the 4 newly reconstructed genomes. However, due to the preliminary nature of these data the phylogeny should be viewed as only as a preliminary phylogenetic position of these genomes.

A Maximum Parsimony tree was constructed using MEGA7 (Kumar et al., 2016) to show a phylogeny of the genomes as viewed in Figure 1. The phylogeny shows the genomes deriving from Pica 8, Chile (18891/492 and 18892/495) as closely related to *M. pinnipedii* and the previously characterized ancient TB genomes coastal and lowland sites of Peru, Osmore River Valley (54U, 58U, 64U from Bos et al., 2014) and those from the Andean highland site of Huari (HUA004, HUA006, HUA016, HUA037, HUA041, HUA051, HUA057 from paper III of this thesis). The genome 18891/492 from Pica 8 is placed ancestral the subclade of ancient MTBC strains and modern *M. pinnipedii,* including the other genome from Pica 8 18892/495. This diversity of the genomes from Pica 8 may be a product of the low coverage of 18891/492 which was artificially increased to meet the threshold of 3-fold minimum coverage for the alignment construction using MultiVCFAnalyzer (Bos et al., 2014). Unexpectedly, the reconstructed genomes from central Mexico, 18881/427.3 and 18883/429.1, are positioned between *M. microti* and the subclade made up of ancient TB and *M. pinnipedii*. This represents a strain of ancient MTBC not previously identified with remarkable difference from the previously identified ancient TB strains which are closely positioned with *M. pinnipedii* (Figure 1). The long branch of 18881/427.3 (Tlatelolco, Mexico) is not unexpected as sequencing data for this captured product, like 18891/492 (Pica 8, Chile), was artificially increased through concatenation of the same sequencing reads meet the 3-fold coverage minimum requirement for alignment build using MultiVCFAnalyzer (Bos et al., 2014). Although the artificial increase allowed for preliminary phylogenetic analysis, it does not increase coverage across the reference genome. Therefore, low

coverage of a genome results in an abundance of ambiguous sites with relatively few phylogenetically informative sites. Although the genome from Pica 8, 18891/492, also is of low coverage (1.90 mean fold coverage), the genome 18881/427.3 (Tlatelolco, Mexico) (1.24 mean fold coverage) has less than 60% of the TB reference covered at 1-fold (Supplementary Table 3).

DISCUSSION

Preliminary data: promise and precautions

Our preliminary data support bioarchaeological reports of tuberculosis in Chile and Mesoamerica prior to European contact. The reconstructed ancient MTBC genome from the library 18852/492 (Pica 8, Chile) presents a 1.60-fold mean genomic coverage in this initial reconstruction with damage patterns characteristic of ancient DNA permitting the authentication of its antiquity. Although the non-UDG treated data of the other Chilean sample (18853/495) does not show a typical damage pattern for DNA, both the UDG treated library (18892) and non-UDG treated library (18853) yield adequate coverage for phylogenetic analysis with 6.81 fold mean coverage and 7.48 mean fold coverage respectively. Phylogenetic analysis of the genomes reconstructed using the UDG treated data (18891/492 and 18892/495) reveals the Chilean samples as harboring an ancient MTBC strains closely related to modern *M. pinnipedii* strain and to the other ancient TB strains recovered from the western region of South America (Figure 1). In particular, genome 18892/495 is positioned as closely related to the previously published MTBC strains from the Osmore River Valley (Bos et al., 2014). This finding further supports previous characterizations of reconstructed ancient MTBC genomes, firmly positioned in the subclade of modern *M. pinnipedii* and ancient MTBC, adding to our understanding of the geographic presence of this subclade.

The results of the preliminary phylogenetic analysis reveal the identification of a unique strain of MTBC circulating in Mexican populations (18881/427.3 and 18883/429.1). Although some variation would be expected due to the geographic distance between these data sets, this finding is unexpected due to the previous identification of ancient pre-colonial strains closely related to *M. pinnipedii*. This finding gave rise to the *pinniped hypothesis* stating the initial introduction of TB to the Americas was from a zoonotic event whereby pinnipeds were the source of pre-colonial human TB infection in the Americas. As the strain identified in the Mexican populations (18881/427.3 and 18883/429.1) represents a previously unknown MTBC strain it could provide additional information on the origin and spread of MTBC in the pre-colonial Americas. This strain has not previously been characterized in studies of modern or ancient TB and will require a significant increase in genome coverage to conservatively reconstruct genomes and alignments with increased quality measures and subsequently appropriately analyze and describe for publication.

The preliminary analyses of this data set reveal promise to add to the known diversity of pre-colonial MTBC strains circulating in the Americas. The geographic expanse and cultural breadth from which these samples were recovered provide the opportunity for a greater understanding of MTBC evolution and ecology. Several unique findings in this preliminary analysis will need further investigation. The recovery of an ancient MTBC genome from a tooth sample should be further investigated as the metagenomic analysis showed no oral microbe signature. However, it should be noted this sample was cleaned with a diluted bleach solution for decontamination before sampling. This is would be the first published report of ancient TB deriving from a dental sample. Furthermore, the identification of novel MTBC strains circulating in pre-colonial Mexico will not only require a significant increase in coverage as previously noted, but additional support for the description of this ancient MTBC strain would be provided by the authentication of ancient DNA through observations of damage patterns. Because mycobacteria are ubiquitous in soil and water it is essential to rule out any environmental contaminants and to authenticate the ancient origin of the recovered MTBC strain.

Paths forward

Ancient DNA has undergone serious criticisms of lack of authentication, possible contamination, and irreproducibility of results (Cooper and Poinar, 2001; Gilbert et al., 2005; Willerslev and Cooper, 2005; Muller et al., 2016). New laboratory and computational tools are improving the ability to manage and even remove environmental contamination, permitting the authentication of observed genetic diversity and allowing confident evolutionary analyses. For example, the application the MALT-HOPS pipeline in the molecular detection and characterization of ancient pathogens, including MTBC, not only permits sensitivity and specificity in the identification of pathogens (Hübler et al., 2020; Vågene et al., 2018) but also allow for the simultaneous confident detection of co-infections (Giffin et al., 2020). Additionally, these methodological advances allow the evaluation of the microbial profile of a skeletal/dental sample. Metagenomic approaches offering the ability to inspect the microbial composition of a sample are particularly valuable to paleopathological studies involving archaeological samples and the development of downstream target DNA recovery strategies (Bos et al., 2019; papers II and III of this thesis; Spyrou et al., 2019). Likewise, evaluation of the microbial composition of a sample of extracted DNA enables the identification of pathogens which may have gone undetected, skeletally or through targeted molecular approaches. Additionally, environmental microbes may often be closely related to a pathogen of interest and complicate analyses through false positive detections when using less specific targeted approaches. It should be noted, however, these soil and water residing microbes can also challenge genomic analysis through the unintentional incorporation of environmental DNA to reconstructed genomes (paper III of this thesis, Nelson et al., in prep). To address this, SNPEvaluation was developed to assist in the authentication of the genetic diversity observed in reconstructed ancient genomes and allow for the removal of contaminant reads (https://github.com/andreasKroepelin/SNP_evaluation) (Keller et al., 2019). These methods allow for sensitivity in analyses and for continual methodological adjustments to project design to enable

efficient recovery of DNA with measures of quality control. These steps that are taken in ancient MTBC research and the subsequent reconstruction of ancient MTBC genomes are certainly required for the production of confident results and analyses.

In order to meet the goals of increased genomic coverage, increased quality measurements, and more confident downstream analyses, we have generated additional sequencing data from a new set of capture data (see methods). For this sample set we have included positive controls for all steps in laboratory methods, to monitor the quality of each step and ensure efficiency in each reaction. Additionally, we have added a step between library preparation and indexing in in which we quantify the concentration of DNA extracted from each sample and calculate the number of barcodes necessary to efficiently recover DNA. We constructed UDG libraries to be captured and sequenced deeper to ensure greater genome coverage for phylogenetic analysis. We also constructed non-UDG libraries to evaluate damage profiles for authentication measures. In the generation of these data all steps were carried out in the cleanroom dedicated to ancient DNA at the Max Planck Institute for the Science of Human History, Jena, Germany thereby removing the need to transport libraries. Analysis is underway on this new dataset which is permitting a more in-depth and confident look at the evolution and ecology of MTBC strain diversity.

CONCLUSION

Here, we present preliminary results which show the detection of ancient MTBC strains from Chile and Mexico. The identification of MTBC DNA in a geographically broad sample set which included various anatomical elements provides insight to the recovery of ancient TB DNA from different environments and sample types. Furthermore, our results reveal a broad geographic presence of MTBC strains prior to European contact with the first pre-colonial MTBC genome reconstructed from Mexico. These data provide new information on the presence of pre-colonial MTBC, adding to the known geographic region and cultural landscape of TB in the Americas prior to European colonization. Although the results of our preliminary phylogeny are intriguing, we need further confirmation through the generation of new libraries with greater coverage that will permit more in-depth evolutionary analysis of MTBC across the Americas.

MATERIALS AND METHODS

ASU: Sample preparation, extraction, and screening

One hundred and twenty-seven anatomical elements deriving from individuals with lesions suggestive of tuberculosis were selected for this project. The sample set is composed of a variety of anatomical elements including ribs, vertebrae, a phalanx, a tooth, and pelvis fragments. These elements were processed in cleanroom facilities at Arizona State university (ASU), USA using a Dremel tool to create subsamples to be made into bone or dental powder. These subsamples were first decontaminated by wiping with 10% bleach solution followed by distilled water and UV

radiated for 1 minute. Subsamples of bone and the whole tooth sample were then pulverized using an 8000M Mixer/Mill (SPEX). Each subsample was then individually extracted for DNA following ancient DNA extraction protocols as outlined by Dabney et al. (2013). All DNA extracts were accompanied by extraction blanks and quantified using the Qubit High Sensitivity assay (Life Technologies). No positive control was included in extraction or downstream generation of data.

DNA extracts were screened for the presence of MTBC DNA with quantitative PCR methods (qPCR) using TaqMan qPCR assays. These assays target the *rpoB* gene, and two insertion elements found in MTBC: IS6110 and IS1081. The assay targeting the *rpoB* gene includes a TaqMan probe that binds to a sequence in the gene that is likely specific to MTBC (Harkins et al. 2015). This analysis was performed using an Applied Biosystems 7900HT thermocycler. Of the initial 127 samples, 28 skeletal samples were selected for library preparation based on PCR screening results.

ASU: Library preparation

DNA extracts that amplified for at least one MTBC qPCR assay (n=28) (Table 1) and negative controls (n=11) were prepared into UDG libraries and non-UDG libraries for a total of 56 libraries with 22 negative controls in the ASU cleanroom facilities. It should be noted that the construction of non-UDG and UDG libraries were in some cases made from different DNA extractions of homogenized bone powder depending on the amount of eluate used for qPCR detection which was sometimes repeated to confirm observations in experiments. All libraries and negative controls were then sent to the Max Planck Institute for the Science of Human History, Jena, Germany to be quantified and amplified for in-solution hybridization capture of target MTBC reads and subsequently sequenced. The samples were packed with dry ice to maintain a frozen state to preserve the DNA library and decrease the accumulation of damage and scheduled to arrive in approximately 3-5 days. Unfortunately, the shipment was delayed in transit and upon their arrival the samples were thawed and in a liquid state as the dry ice had sublimated during shipment. The samples were then stored in a freezer at -20 ˚ Celsius in the laboratory for indexed DNA at the Max Planck Institute for the Science of Human History. Samples were then entered into the MPI-SHH database (Pandora) and given external ID codes for in house processing.

Amplifications

The volume of positive samples and negative controls were measured to evaluate for evaporation of the sample which occurred in transit. The concentration (ng/μl^{-1}) of all samples was measured using a 4200 Tape Station Instrument (Agilent). In order to meet the required concentration for capture (approximately 200 ng/μl) for each sample, calculations were performed to identify the appropriate number of cycles of amplification for each individual sample. Amplifications were then performed using the *Pfu Turbo Cx Hotstart DNA Polymerase* (Agilent) and products were purified using a Qiagen MinElute purification kit and eluted in 20 μl. Following purification,

samples were again measured for concentration. The process was repeated as needed until all samples met the required concentration of DNA for MTBC targeted in-solution capture.

MTBC in-solution capture

In-solution hybridization (Vågene et al., 2018) for MTBC DNA was then performed on the amplified DNA libraries and negative controls. The two libraries for each sample (UDG and non-UDG) underwent two rounds of capture using single-stranded probes that were designed using the computationally reconstructed ancestor of MTBC (Comas et al., 2010). The in-solution capture protocol was performed as presented by Fu et al. (2013) but amended to not include Cot-1 DNA in the mastermix. During capture the sample is enriched for MTBC reads through the binding of MTBC to the complimentary probes and all non-MTBC DNA fragments are washed away in purification. A positive control of known MTBC DNA content values was included for capture to ensure efficiency.

Sequencing data of capture products

All captured products of samples were sequenced using a paired-end 75 bp kit to a depth of 10 million reads on an Illumina HiSeq 4000. Using the EAGER pipeline (Peltzer et al., 2016) sequencing data were processed including removal of Illumina adapters using AdapterRevmoal v2 (Schubert et al., 2016), reference mapping using BWA-aln (v. 0.7.12) (Li and Durbin, 2009), quality filtered using SAMtools, removal of duplicates using Mark Duplicates, and the evaluation of DNA damage patterns with mapDamage2.0 (Jónsson et al., 2013). All samples were mapped to both the human (Hg19) reference and TB ancestor reference (MTB_anc) (Comas et al., 2010). Strict mapping parameters for UDG treated libraries (-l 32, -n 0.1, -q 37) were used as is appropriate for libraries in which DNA damage had been removed. However, more lenient mapping parameters (-l 16, -n 0.01, -q 24) were used for the non-UDG treated libraries in order to accommodate for the mismatches due to DNA damage which accumulates over time. Likewise, all negative controls were mapped using lenient parameters as few MTBC reads are expected. Following this, SNP calling was performed with the UnifiedGenotyper in GATK (DePristo et al., 2011) in both UDG treated and non-UDG treated library sequencing data.

Metagenomic Screening for MTBC using MALT

In order to evaluate the capture efficiency and metagenomic composition of the captured libraries, we employed a metagenomic pathogen screening method using the Megan Alignment Tool (MALT) (version 0.3.8) (Vågene et al., 2018). MALT analysis was carried out using the NCBI full nucleotide database (nt) (Nov. 2017) which includes bacteria, viruses, and eukaryotes. The visualization of MALT analysis in MEGAN (Huson et al., 2016) permits the evaluation of alignments of sequencing reads assigned to specific nodes. All samples were screened using a minimum of 95 percent sequence identity. All other parameters were set to default.

Phylogenetic Analysis

All reconstructed MTBC genomes with a minimum of 3-fold coverage were included in phylogenetic analysis. The MTBC genomes presented in this manuscript were analyzed together using a dataset that includes the three previously published ancient coastal TB genomes (Bos et al., 2014) and 7 unpublished ancient TB genomes reconstructed in house at Max Planck Institute for the Science of Human History (paper III of this thesis) (HUA004, HUA006, HUA016, HUA037, HUA041, HUA051, HUA057). We performed calling of variant SNPs within our reconstructed genomes using the tool UnifiedGenotyper (v. 3.5) from the Genome Analysis Toolkit (GATK) (DePristo et al., 2011). Using MultiVCFAnalyzer (Bos et al., 2014) we produced tables for downstream SNP analysis including a table of all variant positions across all genomes and a SNP alignment for phylogenetic tree construction. Our parameters selected for a minimum of 3-fold coverage of all variant positions and a minimum genotype quality of 30. In addition, our SNP analysis and construction of the SNP alignment also included the removal of regions such as homoplastic sites, repetitive elements, tRNAs, mRNAs, and rRNAs. The alignment for the tree was modified using MEGA7 (Kumar et. al, 2016) to 98% partial deletion of data and the subsequent phylogeny was constructed using the Subtree-Pruning-Regrafting algorithm with 200 bootstrap iterations.

MPI-SHH: Laboratory methods for follow up

Bone powder was collected in the ASU cleanroom. All experiments from this point forward were carried out in the DNA laboratories at Max Planck Institute for the Science of Human History. Ancient DNA samples were not taken out of the cleanroom dedicated to ancient DNA until they were given unique barcodes.

MPI-SHH: DNA Extraction

Powdered bone and dental samples were measured to include approximately 50 mg of powder for DNA extraction. Extraction of DNA was performed using a protocol designed for short fragments of DNA, typical of ancient samples (Dabney et al., 2013). The powder for each sample was collected in a 2ml tube and was suspended with 1ml of extraction buffer (0.45M EDTA, pH 8.0, and 0.25 mg/ml proteinase K) and incubated for approximately 18 hours at 37° C. After incubation, the tubes were centrifuged, and the supernatant was collected from each sample. The remaining bone pellet was saved and stored at -20° C. The supernatant was then added to 10 ml of GuHCl-based binding buffer in High Pure Extender Assembly (Roche) silica membrane spin columns. All columns were spun for 9 min at 1500 rpm setting. The DNA was then purified using the High Viral Nucleic Acid Kit (Roche). The DNA bound to the silica membrane of the column was purified was eluted in 50ul of TET (10mM Tris-HCl, 1mM EDTA pH 8.0, 0.05% Tween20) with a saturation time of 2 minutes before centrifugation at 14,000 rpm for 1 min. The elution step was repeated to collect a total of 100 µl elute for each sample. Extraction blanks and one positive control of known DNA concentration was carried along to ensure the experiment was performed under good conditions.

MPI-SHH: Library preparation for new libraries

Subsequently, extractions were prepared into two sets of uniquely indexed libraries: UDG treated libraries (Briggs et al., 2010) and libraries which did not undergo UDG treatment (Meyer and Kircher, 2010). The libraries without UDG treatment would allow for the observation of DNA damage to authenticate the antiquity of the detected TB DNA. The protocol for library build varies between the two methods in both amount of extract DNA and inclusion of enzymes. The protocols for MPI-SHH differ from those used for the previously constructed libraries in that MPI-SHH calculates out the amount of index combinations for each sample to ensure indexing efficiency. Without this calculation, the extracted DNA may be of a greater concentration than the amount of indexes provided and would then be lost in quantification and sequencing. The protocols carried out at MPI-SHH for the new data set are outlined below.

Library build protocol without UDG treatment: Ten µl of extracted DNA from each sample were prepared as individual double-stranded DNA libraries following a modified protocol (Meyer and Kircher, 2010). This begins with repair of DNA fragment overhands using a blunt -end repair method whereby 10 µl of each extract is treated with 10 µl of H_2O, 5 µl of NEB Buffer 2 (New England Biolabs), .2 µl of dNTP mix, 2 µl BSA, 5 µl ATP (10mM), 2 µl T4 PNK (polynucleotide kinase), and 0.4 µl T4 polymerase to equal 50 µl of total reaction for each sample. Each reaction was incubated at 15°C for 15 minutes, then at 25°C for 15 minutes in the thermocycler. The reactions were then subjected to a purification step in which they were then combined in a 1.5 ml LoBind tube with 600 µl of PB buffer (Qiagen). Each reaction was then loaded onto a MinElute column and incubated at room temperature for 1-2 minutes. The columns were then centrifuged for 30 seconds at 13000 rpm and upon completion the flow through supernatant was discarded. We then added 650 PE wash buffer (Qiagen) to each column and immediately centrifuged these samples for 30 seconds at 13000 rpm. The flow through was discarded and the columns were then "dry spun" in the centrifuge for 1 min at 13000 rpm, the columns were then rotated 180 ˚ and spun again to ensure the removal of all flow through. The column was then inserted into a new 1.5 ml LoBind tube and eluted in 20 µl of elution buffer containing .05% Tween. Adapter ligation was then performed by adding the elute to 20 µl of Quick Ligase Buffer, 1 µl of adapter mix, and 1 µl of Quick ligase to create an adapter ligation reaction that was incubated at 22˚ C for 20 minutes. The reaction was then purified as outlined above with a final elute of 22 µl of EB and Tween. Adapter fill in was then performed by adding 4 µl of Thermopol Buffer, 2.1 µl of dNTPs, 2 µl of Bst polymerase, and 13.8 of H_2O to each purified eluate of the ligation assay. This was then incubated at 37 ˚ C for 30 minutes and then 80 ˚ C for 10 minutes in the thermocycler. The process resulted in a remaining 40 µl of DNA library. The libraries which had not yet been indexed were then quantified with a qPCR reaction with 1 µl of DNA from each prepared ancient DNA library, 10 µl of Dynamo, 1 µl of each primer (IS7 and IS8), 7 µl of H_2O, and 1 µl of 12 DNA standards from 10^3 to 10^8 quantities along with two qPCR blanks. The remaining amount of prepared library was stored at -20 ˚ C for storage in the ancient DNA "clean lab".

Library build protocol with UDG treatment: Samples were prepared as double-stranded DNA libraries using a volume of 50 ul of DNA. The UDG treatment and library protocol differ in the first stage of library preparation where we first treated the libraries with uracil-DNA-glycosylase (UDG) to remove postmortem damage. Using a total of 50 µl of DNA we then added to the reaction a master mix of 15 µl of NEB Buffer 2, 15 µl of ATP, .75 µl of BSA, 1.8 µl of dNTPs, 6 µl of T4 PNK, 9 µl of USER enzyme and 22.45 µl of H_2O. The reaction was incubated at 37˚ C for 3 hours. The reaction then received 6 µl of T4 polymerase to each library and was incubated for an additional 20 minutes at 25 ˚ C and 10 minutes at 12 ˚ C. The reaction was then purified using the MinElute kit as previously described and the remainder of the protocol follows the traditional library preparation from adapter ligation forward. The UDG-treated libraries were quantified on a quantitative PCR (qPCR) using IS7/IS8 primer combinations. This quantification was used to calculate the appropriate quantity of index needed for each sample.

Indexing of prepared libraries: Subsequently, all libraries were divided into multiple PCR reactions for double indexing (Meyer and Kircher, 2010; Kircher et al., 2012) based on their initial quantification and measurement, in order to ensure maximal amplification efficiency. The prepared libraries were barcoded individually giving each sample library a unique 8 base pair identifier index for later reference, evaluation, and analysis of the DNA unique to that sample. Indexing was performed in the cleanroom across two days, separating the library into two equal halves to ensure the successful indexing and recovery of DNA. Indexing was performed using *Pfu Turbo Cx Hotstart DNA Polymerase* (Agilent) and with optimized efficiency by splitting the library to a number of reactions based on the concentration of DNA (a maximum of 2 x 10^{10} molecules of DNA) in the library to ensure the appropriate amount of index primers to be added so that all DNA fragments would be recovered. Libraries were split a minimum for 4 times and adjusted for higher concentrations to ensure efficiency of the indexing reaction. Based on a split of four, the library would be divided across four reactions with each reaction receiving 10 µl of Pfu Turbo Buffer, 1.5 µl of BSA, 1 µl of dNTPs, 1 µl of Pfu Turbo Polymerase, 9 µl of H_2O and 2 µl of each primer (P5 and P7). The reaction was amplified with an initial denaturation at 95 ˚ C for 2 minutes and 10 cycles of 95 ˚ C for 30 seconds with a following 58 ˚ C for 30 seconds and 72 ˚ C for 1 minute ending with an elongation at 72 ˚ C for 10 minutes and holding the reactions at 10 ˚ C. All indexed reactions were then purified using a MinElute Kit using one column for a maximum of four reactions. All indexed libraries were quantified using qPCR with IS5 and IS6 primers to calculate the efficiency of the indexing reaction. The total result of all indexing resulted in 50 µl of index library.

Amplification for capture: After indexing all libraries, the concentration of each library was measured and then calculations for amplifications were performed to split each library across 3 reactions so the combined products would meet the required concentration of 200-300 ng/μl for in-solution TB capture. Amplification was performed as previously described on the indexed libraries using *Herculase II Fusion DNA Polymerase* (Agilent). The amplified indexed libraries

were subsequently purified using the MinElute DNA purification kit (Qiagen) and eluted in 25 μl TET (10mM Tris-HCl, 1 mM EDTA pH 8.0, 0.05% Tween20). After all libraries were appropriately prepared, MTBC targeted in- solution capture was performed for two rounds on both library sets. The capture product was then sequenced on a HiSeq 4000 to a depth of 20 million reads using a paired-end 75 base pair kit.

Table 1. List of samples IDs included in this data set with site name and location and anatomical element from which the sample was taken. The samples which were carried through in downstream analyses are highlighted yellow. It should be noted that MPI-SHH ID's are showing the last five digits, as all external libraries are given the prefix XXX001.A… (XXX001.A18873).

MPI-SHH ID: UDG treated libraries	MPI-SHH ID: library (non-UDG)	ASU sample ID	Site name	Site location	Anatomical element sampled
18873	**18834**	**353.3**	Kuelap	Peru	vertebra
18874	**18835**	**354.1**	Kuelap	Peru	rib
18875	**18836**	**355.1**	Kuelap	Peru	mandibular right 2nd molar
18876	**18837**	**358.2**	Kuelap	Peru	pelvis
18877	**18838**	**361.2**	Kuelap	Peru	vertebra
18878	**18839**	**362**	Kuelap	Peru	pelvis
18879	**18840**	**393.2**	Portalegre 1987	Colombia	vertebra
18880	**18841**	**398.2**	UPTC-Basque Alto Laboratorios	Colombia	vertebra
18881	**18842**	**427.3**	STF 2 7	Mexico	vertebra
18882	**18843**	**427.4**	STF 2 7	Mexico	vertebra
18883	**18844**	**429.1**	Tlatelolco 282	Mexico	vertebra
18884	**18845**	**429.2**	Tlatelolco 282	Mexico	rib
18885	**18846**	**429.3**	Tlatelolco 282	Mexico	rib
18886	**18847**	**430.2**	Tlatelolco 23	Mexico	calculus
18887	**18848**	**483**	Kleinberg	Canada	vertebra
18888	**18849**	**488**	Maurice	Canada	vertebra
18889	**18850**	**489**	Maurice	Canada	vertebra
18890	**18851**	**491**	Pica 8	Chile	vertebra
18891	**18852**	**492**	Pica 8	Chile	maxillary right 2nd molar
18892	**18853**	**495**	Pica 8	Chile	vertebra
18893	**18854**	**499**	Puqueldon I	Chile	vertebra
18894	**18855**	**500**	Chonos, Puquitin	Chile	sacrum
18895	**18856**	**502**	Sunol #2305	California, USA	metacarpal
18896	**18857**	**505**	Sunol #2305	California, USA	Thoracic centrum (unfused)
18897	**18858**	**506**	Sunol #2305	California, USA	Thoracic centrum (unfused)
18898	**18859**	**507**	Sunol #2305	California, USA	rib
18899	**18860**	**508.2**	Saujil-Tinogasta	Argentina	proximal pedal phalanx
18900	**18861**	**508.3**	Saujil-Tinogasta	Argentina	rib

Table 2. Mapping statistics for TB captured data from UDG & non-UDG libraries of samples with at least 1-fold coverage when mapped to the TB Ancestor reference. It should be noted that MPI-SHH ID's are showing the last five digits, as all external libraries are given the prefix XXX001.A… (XXX001.A18873).

MPI-SHH ID	External Library ID	Arch Site	Anatomical element	Raw Sequencing Reads	% Merged Reads	Mapped Reads after Quality Filter	Endogenous DNA after Quality Filter	Mean Coverage	Damage 1st Base 3' end	Damage 2nd Base 3' end	Damage 1st Base 5' end	Damage 2nd Base 5' end	median fragment length	GC content in %
UDG treated														
18881	**427.3**	STF , Mexico	Lumbar vertebra	23252946	97.838	104380	1.49	1.24	0.0040	0.0033	0.0034	0.0032	46	63.18
18883	**429.1**	Tlatelolco, Mexico	Lumbar vertebra	24197828	99.179	358196	4.67	3.97	0.0014	0.0010	0.0013	0.0010	44	62.57
18891	**492**	Pica 8, Chile	Upper Right 2nd Molar	28312858	99.275	189048	4.08	1.90	0.0044	0.0014	0.0036	0.0013	42	62.37
18892	**495**	Pica 8, Chile	Lumbar vertebra	24416656	98.844	708301	8.94	6.81	0.0022	0.0013	0.0018	0.0010	40	61.41
Not UDG treated														
18842	**427.3**	STF , Mexico	Lumbar vertebra	25557968	76.762	1752	0.33	0.02	0.0124	0.0060	0.0127	0.0143	56	60.11
18844	**429.1**	Tlatelolco, Mexico	Lumbar vertebra	23459672	83.358	19734	2.29	0.32	0.0168	0.0081	0.0219	0.0102	76	63.88
18852	**492**	Pica 8, Chile	Upper Right 2nd Molar	24471024	98.274	148172	3.25	1.60	0.0361	0.0170	0.0338	0.0161	45	63.26
18853	**495**	Pica 8, Chile	Lumbar vertebra	26783090	98.465	762193	8.88	7.48	0.0193	0.0113	0.0197	0.0106	41	61.49

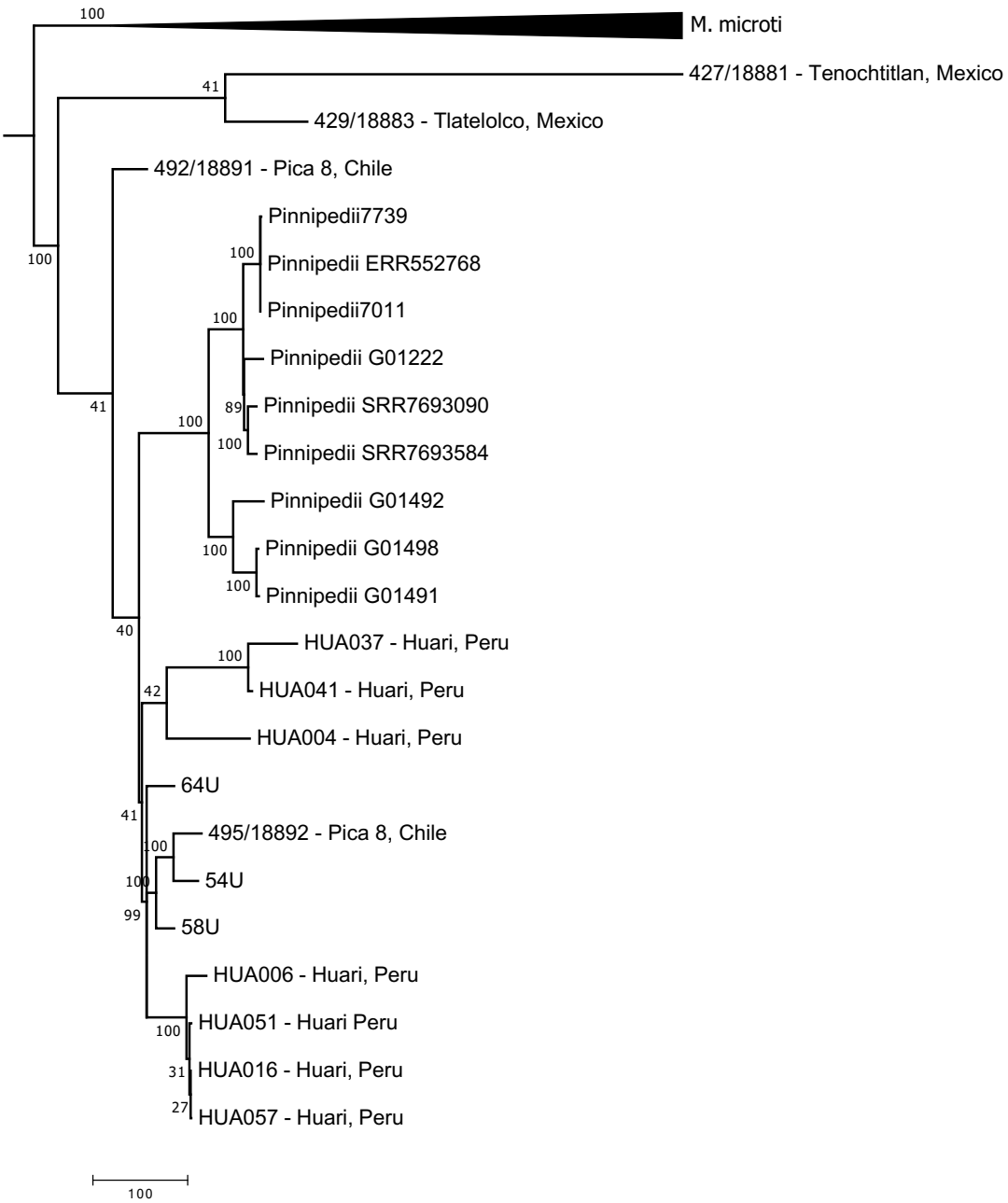


Figure 1. Preliminary phylogeny of pre-colonial American TB genomes. This Maximum Parsimony tree was constructed using 269 genomes with 14 ancient MTBC genomes from the Americas, including the 4 preliminary genomes presented in this work, 7 unpublished genomes from paper III of this thesis and 3 published genomes from Bos et al., 2014 (64U, 54U, 58U). The data were selected for sites which had at least 98% coverage creating an alignment with a total of 37,353 sites out of 60,222 total sites were used. For this phylogeny the minimum fold coverage for calling sites was set to 3-fold to accommodate the low coverage nature of this preliminary data and the tree was constructed using 200 bootstraps. The alignment for this tree was prepared using MultiVCFAnalyzer (Bos et al., 2014) and the tree was constructed using MEGA7 (Kumar et al., 2016) with *M. canetii* as the outgroup.

References

Allison, M. J., Gerszten, E., Munizaga, J., Santoro, C., & Mendoza, D. (1981). Tuberculosis in pre-Columbian Andean populations. *Prehistoric Tuberculosis in the Americas*, 49–61.

Allison, M. J., Mendoza, D., & Pezzia, A. (1973). Documentation of a case of tuberculosis in pre-Columbian America. *American Review of Respiratory Disease*, *107*(6), 985–991.

Arateco, W. M. R. (1998). Mal de Pott en momia de la colección del museo arqueológico Marqués de San Jorge. *Maguaré*, *0*(13), 99–115.

Arriaza, B. T., Salo, W., Aufderheide, A. C., & Holcomb, T. A. (1995). Pre-Columbian tuberculosis in northern Chile: Molecular and skeletal evidence. *American Journal of Physical Anthropology*, *98*(1), 37–45.

Bos, K. I., Harkins, K. M., Herbig, A., Coscolla, M., Weber, N., Comas, I., Forrest, S. A., Bryant, J. M., Harris, S. R., & Schuenemann, V. J. (2014). Pre-Columbian mycobacterial genomes reveal seals as a source of New World human tuberculosis. *Nature*, *514*(7523), 494–497.

Bos, K. I., Kühnert, D., Herbig, A., Esquivel-Gomez, L. R., Andrades Valtueña, A., Barquera, R., Giffin, K., Kumar Lankapalli, A., Nelson, E. A., & Sabin, S. (2019). Paleomicrobiology: Diagnosis and evolution of ancient pathogens. *Annual Review of Microbiology*, *73*, 639–666.

Briggs, A. W., Stenzel, U., Johnson, P. L., Green, R. E., Kelso, J., Prüfer, K., Meyer, M., Krause, J., Ronan, M. T., & Lachmann, M. (2007). Patterns of damage in genomic DNA sequences from a Neandertal. *Proceedings of the National Academy of Sciences*, *104*(37), 14616–14621.

Briggs, A. W., Stenzel, U., Meyer, M., Krause, J., Kircher, M., & Pääbo, S. (2010). Removal of deaminated cytosines and detection of in vivo methylation in ancient DNA. *Nucleic Acids Research*, *38*(6), e87–e87.

Brites, D., Loiseau, C., Menardo, F., Borrell, S., Boniotti, M. B., Warren, R., Dippenaar, A., Parsons, S. D. C., Beisel, C., Behr, M. A., Fyfe, J. A., Coscolla, M., & Gagneux, S. (2018). A New Phylogenetic Framework for the Animal-Adapted Mycobacterium tuberculosis Complex. *Frontiers in Microbiology*, *9*. https://doi.org/10.3389/fmicb.2018.02820

Brosch, R., Gordon, S. V., Marmiesse, M., Brodin, P., Buchrieser, C., Eiglmeier, K., Garnier, T., Gutierrez, C., Hewinson, G., & Kremer, K. (2002). A new evolutionary scenario for the Mycobacterium tuberculosis complex. *Proceedings of the National Academy of Sciences*, *99*(6), 3684–3689.

Buikstra, J. E. (1976). The Caribou Eskimo: General and specific disease. *American Journal of Physical Anthropology*, *45*(3), 351–367.

Buikstra, J. E. (1981). *Prehistoric tuberculosis in the Americas* (Issue 5). Center for Amer Archeology Press.

Buikstra, J. E., & Williams, S. (1991). Tuberculosis in the Americas: Current Perspectives. In *Human Paleopathology: Current Syntheses and Future Options edited by D. Ortner and A. Aufderheide*.

Comas, I., Chakravartti, J., Small, P. M., Galagan, J., Niemann, S., Kremer, K., Ernst, J. D., & Gagneux, S. (2010). Human T cell epitopes of Mycobacterium tuberculosis are evolutionarily hyperconserved. *Nature Genetics*, *42*(6), 498–503. https://doi.org/10.1038/ng.590

Cooper, A., & Poinar, H. N. (2000). Ancient DNA: do it right or not at all. *Science*, *289*(5482), 1139–1139.

Dabney, J., Knapp, M., Glocke, I., Gansauge, M.-T., Weihmann, A., Nickel, B., Valdiosera, C., García, N., Pääbo, S., Arsuaga, J.-L., & Meyer, M. (2013). Complete mitochondrial genome sequence of a Middle Pleistocene cave bear reconstructed from ultrashort DNA fragments. *Proceedings of the National Academy of Sciences*, *110*(39), 15758–15763. https://doi.org/10.1073/pnas.1314445110

DePristo, M. A., Banks, E., Poplin, R., Garimella, K. V., Maguire, J. R., Hartl, C., Philippakis, A. A., Del Angel, G., Rivas, M. A., & Hanna, M. (2011). A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nature Genetics*, *43*(5), 491.

Fu, Q., Meyer, M., Gao, X., Stenzel, U., Burbano, H. A., Kelso, J., & Pääbo, S. (2013). DNA analysis of an early modern human from Tianyuan Cave, China. *Proceedings of the National Academy of Sciences*, *110*(6), 2223–2227.

Gagneux, S., DeRiemer, K., Van, T., Kato-Maeda, M., Jong, B. C. de, Narayanan, S., Nicol, M., Niemann, S., Kremer, K., Gutierrez, M. C., Hilty, M., Hopewell, P. C., & Small, P. M. (2006). Variable host–pathogen compatibility in Mycobacterium tuberculosis. *Proceedings of the National Academy of Sciences*, *103*(8), 2869–2873. https://doi.org/10.1073/pnas.0511240103

García-Frías, J. (1940). La tuberculosis en los antiguos peruanos. *Actualidad Médica Peruana*, *5*(10), 274–291.

Giffin, K., Lankapalli, A. K., Sabin, S., Spyrou, M. A., Posth, C., Kozakaitė, J., Friedrich, R., Miliauskienė, Ž., Jankauskas, R., & Herbig, A. (2020). A treponemal genome from an

historic plague victim supports a recent emergence of yaws and its presence in 15 th century Europe. *Scientific Reports*, *10*(1), 1–13.

Gilbert, M. T. P., Bandelt, H.-J., Hofreiter, M., & Barnes, I. (2005). Assessing ancient DNA studies. *Trends in Ecology & Evolution*, *20*(10), 541–544.

Guichón, R. A., Buikstra, J. E., Stone, A. C., Harkins, K. M., Suby, J. A., Massone, M., Wilbur, A., Constantinescu, F., & Martín, C. R. (2015). Pre-Columbian tuberculosis in Tierra del Fuego? Discussion of the paleopathological and molecular evidence. *International Journal of Paleopathology*, *11*, 92–101.

Guillén, S. E. (2012). A History of paleopathology in Peru and Northern Chile: From head hunting to head counting. *The Global History of Paleopathology: Pioneers and Prospects*, 312–328.

Harkins, K. M., Buikstra, J. E., Campbell, T., Bos, K. I., Johnson, E. D., Krause, J., & Stone, A. C. (2015). Screening ancient tuberculosis with qPCR: challenges and opportunities. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *370*(1660), 20130622.

Hübler, R., Key, F. M., Warinner, C., Bos, K. I., Krause, J., & Herbig, A. (2019). HOPS: Automated detection and authentication of pathogen DNA in archaeological remains. *Genome Biology*, *20*(1), 1–13.

Huson, D. H., Beier, S., Flade, I., Górska, A., El-Hadidi, M., Mitra, S., Ruscheweyh, H.-J., & Tappu, R. (2016). MEGAN Community Edition—Interactive Exploration and Analysis of Large-Scale Microbiome Sequencing Data. *PLoS Computational Biology*, *12*(6), e1004957–e1004957. PubMed. https://doi.org/10.1371/journal.pcbi.1004957

Jónsson, H., Ginolhac, A., Schubert, M., Johnson, P. L. F., & Orlando, L. (2013). mapDamage2.0: Fast approximate Bayesian estimates of ancient DNA damage parameters. *Bioinformatics*, *29*(13), 1682–1684. https://doi.org/10.1093/bioinformatics/btt193

Keller, M., Spyrou, M. A., Scheib, C. L., Neumann, G. U., Kröpelin, A., Haas-Gebhard, B., Päffgen, B., Haberstroh, J., i Lacomba, A. R., & Raynaud, C. (2019). Ancient Yersinia pestis genomes from across Western Europe reveal early diversification during the First Pandemic (541–750). *Proceedings of the National Academy of Sciences*, *116*(25), 12363–12372.

Kircher, M., Sawyer, S., & Meyer, M. (2012). Double indexing overcomes inaccuracies in multiplex sequencing on the Illumina platform. *Nucleic Acids Research*, *40*(1), e3–e3.

Klaus, H. D., Wilbur, A. K., Temple, D. H., Buikstra, J. E., Stone, A. C., Fernandez, M., Wester, C., & Tam, M. E. (2010). Tuberculosis on the north coast of Peru: Skeletal and molecular paleopathology of late pre-Hispanic and postcontact mycobacterial disease. *Journal of Archaeological Science*, *37*(10), 2587–2597.

Kumar, S., Stecher, G., & Tamura, K. (2016). MEGA7: Molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Molecular Biology and Evolution*, *33*(7), 1870–1874.

Li, H., & Durbin, R. (2009). Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics*, *25*(14), 1754–1760. https://doi.org/10.1093/bioinformatics/btp324

Meyer, M., & Kircher, M. (2010). Illumina sequencing library preparation for highly multiplexed target capture and sequencing. *Cold Spring Harbor Protocols*, *2010*(6), pdb. prot5448.

Müller, R., Roberts, C. A., & Brown, T. A. (2015). Complications in the study of ancient tuberculosis: Non-specificity of IS6110 PCRs. *STAR: Science & Technology of Archaeological Research*, *1*(1), 1–8. https://doi.org/10.1179/2054892314Y.0000000002

Ngabonziza, J. C. S., Loiseau, C., Marceau, M., Jouet, A., Menardo, F., Tzfadia, O., Antoine, R., Niyigena, E. B., Mulders, W., & Fissette, K. (2020). A sister lineage of the Mycobacterium tuberculosis complex discovered in the African Great Lakes region. *Nature Communications*, *11*(1), 1–11.

Ortner, D. J., Knüsel, C., & Roberts, C. A. (2012). Special courses in human skeletal palaeopathology. *A Global History of Paleopathology: Pioneers and Prospects*, 684–693.

Peltzer, A., Jäger, G., Herbig, A., Seitz, A., Kniep, C., Krause, J., & Nieselt, K. (2016). EAGER: Efficient ancient genome reconstruction. *Genome Biology*, *17*(1), 60. https://doi.org/10.1186/s13059-016-0918-z

Requena, A. (1946). Evidencia de tuberculosis en la América Precolombina. *Acta Venezolana*, 1–20.

Roberts, C. A., & Buikstra, J. E. (2003). *The bioarchaeology of tuberculosis: A global perspective on a re-emerging disease*. University Press of Florida.

Salo, W. L., Aufderheide, A. C., Buikstra, J., & Holcomb, T. A. (1994). Identification of Mycobacterium tuberculosis DNA in a pre-Columbian Peruvian mummy. *Proceedings of the National Academy of Sciences*, *91*(6), 2091–2094.

Sawyer, S., Krause, J., Guschanski, K., Savolainen, V., & Pääbo, S. (2012). Temporal patterns of nucleotide misincorporations and DNA fragmentation in ancient DNA. *PloS One*, *7*(3), e34131.

Schubert, M., Lindgreen, S., & Orlando, L. (2016). AdapterRemoval v2: Rapid adapter trimming, identification, and read merging. *BMC Research Notes*, *9*(1), 88. https://doi.org/10.1186/s13104-016-1900-2

Spyrou, M. A., Bos, K. I., Herbig, A., & Krause, J. (2019). Ancient pathogen genomics as an emerging tool for infectious disease research. *Nature Reviews Genetics*, *20*(6), 323–340.

Stamatakis, A. (2014). RAxML version 8: A tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics*, *30*(9), 1312–1313.

Stead, W. W., Eisenach, K. D., Cave, M. D., Beggs, M. L., Templeton, G. L., Thoen, C. O., & Bates, J. H. (1995). When did Mycobacterium tuberculosis infection first occur in the New World? An important question with public health implications. *American Journal of Respiratory and Critical Care Medicine*, *151*(4), 1267–1268.

Stone, A. C., Wilbur, A. K., Buikstra, J. E., & Roberts, C. A. (2009). Tuberculosis and leprosy in perspective. *American Journal of Physical Anthropology*, *140*(S49), 66–94.

Vågene, Å. J. (2018). *Genomic insights into pre-and post-contact human pathogens in the New World*. Eberhard-Karls Universität Tübingen.

Vågene, Å. J., Herbig, A., Campana, M. G., García, N. M. R., Warinner, C., Sabin, S., Spyrou, M. A., Valtueña, A. A., Huson, D., & Tuross, N. (2018). Salmonella enterica genomes from victims of a major sixteenth-century epidemic in Mexico. *Nature Ecology & Evolution*, *2*(3), 520–528.

Wilbur, A. K., & Buikstra, J. E. (2006). Patterns of tuberculosis in the Americas: How can modern biomedicine inform the ancient past? *Memórias Do Instituto Oswaldo Cruz*, *101*, 59–66.

Willerslev, E., & Cooper, A. (2005). Ancient dna. *Proceedings of the Royal Society B: Biological Sciences*, *272*(1558), 3–16.

Preliminary investigations of MTBC across the pre-colonial Americas

SUPPLEMENTARY INFORMATION

Authors: Elizabeth A. Nelson[1,2†], Kelly E. Blevins[3†], Jane E. Buikstra[3], Josefina Mansilla Lory[4], Alexander Herbig[1], Johannes Krause[1], Anne C. Stone[3,5], and Kirsten I. Bos[1]

Affiliations:
[1]Max Planck Institute for the Science of Human History, Jena, Germany
[2]Institute for Archaeological Sciences, University of Tübingen, Tübingen, Germany
[3]School of Human Evolution and Social Change, Arizona State University, Tempe, Arizona, USA
[4]Dirección de Antropología Física, Insituto Nacional de Antropolgía e Histora, Mexico City, Mexico
[5]Institute of Human Origins, Arizona State University, Tempe, Arizona, USA

[†]These authors contributed equally to this work

CONTENTS:

Supplementary Table 1. Mapping statistics for the MTBC captured libraries which did not undergo UDG treatment (non-UDG). These data were mapped to the TB ancestor reference. It should be noted that MPI Pandora ID's are showing the last five digits, as all external libraries are given the prefix XXX001.A… (XXX001.A18873).

MPI Pandora ID	ASU ID	Library Type	# of Raw sequencing reads	% Merged reads	Mapped reads after RMDup	Endogenous DNA QF (%)	Cluster factor	Mean coverage	std. dev. coverage	Coverage >= 1X in %	Coverage >= 2X in %	Coverage >= 3X in %	Coverage >= 4X in %	Coverage >= 5X in %	DMG 1st Base 3'	DMG 2nd Base 3'	DMG 1st Base 5'	DMG 2nd Base 5'	average fragment length	median fragment length	GC content in %
18834	353.3	non-UDG	34875534	96.98	98580	1.06	1.79	1.06	27.66	7.75	3.48	2.44	1.93	1.63	0.04	0.04	0.05	0.04	47.63	44	59.85
18835	354.1	non-UDG	12740048	98.41	64983	1.96	1.85	0.72	23.56	4.89	2.50	1.72	1.30	1.06	0.07	0.06	0.07	0.05	48.84	45	58.69
18836	355.1	non-UDG	24233216	99.00	86829	1.92	2.60	0.92	27.28	3.51	1.86	1.46	1.25	1.10	0.05	0.04	0.05	0.04	46.96	43	59.47
18837	358.2	non-UDG	19897988	97.35	89889	1.24	1.31	0.94	14.69	10.21	6.15	4.52	3.65	3.05	0.02	0.03	0.02	0.03	46.28	44	61.04
18838	361.2	non-UDG	22766574	97.54	94742	1.26	1.45	1.01	19.98	8.64	5.09	3.82	3.11	2.63	0.02	0.03	0.02	0.03	47.24	44	60.57
18839	362	non-UDG	22113632	97.13	84591	1.17	1.45	0.89	18.27	8.10	4.61	3.34	2.67	2.24	0.02	0.02	0.02	0.03	46.31	43	60.24
18840	393.2	non-UDG	20949484	97.82	93153	1.84	1.98	1.01	27.83	4.90	2.71	2.03	1.66	1.44	0.05	0.04	0.05	0.04	47.63	44	59.25
18841	398.2	non-UDG	22237094	90.96	49787	1.05	2.02	0.49	12.54	6.12	2.16	1.46	1.14	0.96	0.03	0.03	0.03	0.03	43.30	41	60.43
18842	427.3	non-UDG	25557968	76.76	5207	0.86	15.10	0.06	2.49	1.77	0.54	0.27	0.15	0.12	0.02	0.02	0.02	0.03	53.33	47	59.67
18843	427.4	non-UDG	21593576	85.45	13560	1.21	8.42	0.19	3.65	8.92	2.42	0.86	0.47	0.31	0.02	0.02	0.03	0.02	62.71	55	61.98
18844	429.1	non-UDG	23459672	83.36	27734	2.79	10.34	0.39	10.59	16.63	5.49	1.96	0.68	0.29	0.03	0.02	0.03	0.02	62.72	52	61.94
18845	429.2	non-UDG	21211472	81.04	4364	0.78	15.65	0.05	2.56	0.87	0.30	0.16	0.12	0.09	0.03	0.04	0.03	0.03	50.25	45	58.54
18846	429.3	non-UDG	22617528	81.05	3008	1.02	31.68	0.03	1.95	0.83	0.17	0.10	0.07	0.06	0.02	0.04	0.02	0.04	51.17	45	58.45
18847	430.2	non-UDG	24377072	96.94	63682	1.27	2.32	0.75	28.46	1.74	0.91	0.74	0.64	0.58	0.07	0.07	0.07	0.07	51.84	47	57.88
18848	483	non-UDG	23976124	75.70	18403	0.57	3.09	0.25	10.96	3.81	0.62	0.40	0.32	0.28	0.03	0.05	0.03	0.04	60.57	52	59.75
18849	488	non-UDG	20750050	96.90	56533	1.11	1.83	0.55	12.03	5.68	2.90	2.09	1.66	1.38	0.03	0.04	0.03	0.03	43.14	41	59.74
18850	489	non-UDG	20500602	90.46	46353	1.25	2.27	0.45	9.98	5.31	2.68	1.88	1.46	1.20	0.02	0.04	0.02	0.03	42.55	40	59.93
18851	491	non-UDG	27626674	98.46	39898	1.83	5.83	0.39	14.72	10.83	1.48	0.58	0.44	0.36	0.04	0.04	0.05	0.05	43.65	41	59.34
18852	492	non-UDG	24471024	98.27	180853	5.15	3.38	1.96	22.96	73.51	45.80	24.41	11.46	4.86	0.04	0.02	0.04	0.02	47.87	45	62.37
18853	495	non-UDG	26783090	98.47	892343	9.53	1.30	8.62	11.01	92.15	86.16	79.58	72.78	65.94	0.02	0.01	0.02	0.01	42.61	40	61.22
18854	499	non-UDG	19326322	90.63	27950	0.75	2.41	0.31	9.75	4.08	1.73	1.12	0.83	0.67	0.02	0.03	0.02	0.02	48.23	44	59.96
18855	500	non-UDG	27710236	90.69	123940	2.04	2.07	1.28	25.90	7.43	4.22	3.21	2.68	2.32	0.01	0.02	0.01	0.02	45.49	42	60.88
18856	502	non-UDG	22600968	95.34	48540	0.76	1.62	0.50	12.90	5.30	2.60	1.81	1.40	1.14	0.02	0.03	0.03	0.03	45.34	42	60.18
18857	505	non-UDG	24350442	97.26	67856	0.91	1.51	0.68	14.80	6.25	3.50	2.52	2.01	1.68	0.03	0.03	0.03	0.03	44.30	42	60.29
18858	506	non-UDG	22562888	94.51	97753	1.50	1.56	1.02	19.35	9.18	5.55	4.16	3.36	2.82	0.03	0.04	0.04	0.04	46.14	42	60.51
18859	507	non-UDG	22853492	96.44	50887	0.71	1.52	0.56	13.62	5.96	3.14	2.25	1.76	1.45	0.02	0.03	0.03	0.03	48.30	45	59.94
18860	508.2	non-UDG	27721804	98.13	88409	0.97	1.46	0.93	16.85	7.63	4.74	3.63	2.93	2.47	0.04	0.03	0.04	0.03	46.23	44	60.73
18861	508.3	non-UDG	24493458	98.10	84809	1.10	1.54	0.94	21.31	5.87	3.34	2.46	1.97	1.64	0.02	0.02	0.02	0.02	48.91	45	61.38
Negative Controls																					
18862	EB 1.26.18	non-UDG	1222138	42.82	296	0.17	4.29	0.00	0.21	0.10	0.03	0.02	0.02	0.01	0.04	0.08	0.00	0.01	48.16	44	57.62
18863	EB 11.20.17	non-UDG	1844180	59.18	246	0.30	10.56	0.00	0.20	0.09	0.03	0.02	0.01	0.01	0.04	0.01	0.01	0.05	50.84	46	57.70
18864	EB 2.26.18	non-UDG	1946	84.36	2	0.20	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	49.00	49	67.35
18865	EB 3.17.18	non-UDG	2814288	62.85	1540	0.60	5.37	0.02	0.70	0.38	0.15	0.11	0.08	0.06	0.05	0.05	0.03	0.01	49.00	46	59.31
18866	EB 4.4.18	non-UDG	44558	42.02	57	0.25	1.16	0.00	0.05	0.03	0.01	0.01	0.01	0.01	0.00	0.00	0.00	0.00	48.98	43	58.81
18867	EB 5.7.18	non-UDG	1052782	68.86	398	0.31	4.15	0.00	0.31	0.11	0.04	0.02	0.02	0.02	0.09	0.05	0.03	0.03	49.98	46	57.69
18868	EB 9.11.17	non-UDG	680530	45.89	209	0.21	4.37	0.00	0.21	0.04	0.02	0.01	0.01	0.01	0.04	0.00	0.05	0.04	49.38	45	55.48
18869	EB 9.6.17	non-UDG	5388894	56.35	304	0.25	25.83	0.00	0.27	0.07	0.03	0.02	0.01	0.01	0.04	0.01	0.02	0.02	48.84	45	57.05
18870	LB3.5.18	non-UDG	1761422	60.01	299	0.23	7.84	0.00	0.25	0.07	0.03	0.02	0.01	0.01	0.04	0.03	0.01	0.02	47.20	43	57.02
18871	LB 5.18.18	non-UDG	1073720	50.60	264	0.26	5.22	0.00	0.24	0.07	0.03	0.02	0.01	0.01	0.06	0.02	0.06	0.03	49.66	45	56.42
18872	LB 5.22.18	non-UDG	107062	18.64	84	0.16	1.35	0.00	0.08	0.03	0.01	0.01	0.01	0.00	0.07	0.00	0.08	0.00	45.01	42	56.97

Supplementary Table 2. Mapping statistics for the MTBC captured libraries which did not undergo UDG treatment (non-UDG). These data were mapped to the human reference (Hg19). It should be noted that MPI-SHH ID's are showing the last five digits, as all external libraries are given the prefix XXX001.A… (XXX001.A18873).

MPI Pandora ID	ASU ID	Library Type	# of Raw sequencing reads	% Merged reads	Mapped reads after RMDup	Endogenous DNA QF (%)	Cluster factor	Mean coverage	std. dev. coverage	Coverage >= 1X in %	Coverage >= 2X in %	Coverage >= 3X in %	Coverage >= 4X in %	Coverage >= 5X in %	DMG 1st Base 3'	DMG 2nd Base 3'	DMG 1st Base 5'	DMG 2nd Base 5'	average fragment length	median fragment length	GC content in %
18834	353.3	non-UDG	34875534	96.98	99046	1.04	1.79	1.07	27.95	7.80	3.48	2.44	1.93	1.63	0.04	0.04	0.05	0.04	47.76	44	59.84
18835	354.1	non-UDG	12740048	98.41	65175	1.93	1.85	0.72	23.70	4.90	2.50	1.72	1.31	1.06	0.07	0.05	0.07	0.05	48.92	45	58.69
18836	355.1	non-UDG	24233216	99.00	87011	1.90	2.60	0.93	27.42	3.52	1.86	1.46	1.25	1.10	0.05	0.04	0.05	0.04	47.02	43	59.47
18837	358.2	non-UDG	19897988	97.35	90220	1.21	1.31	0.95	14.87	10.22	6.15	4.53	3.65	3.05	0.02	0.03	0.02	0.03	46.39	44	61.02
18838	361.2	non-UDG	22766574	97.54	95220	1.23	1.45	1.02	20.25	8.65	5.09	3.82	3.12	2.63	0.02	0.03	0.02	0.03	47.39	44	60.55
18839	362	non-UDG	22113632	97.13	85024	1.14	1.46	0.90	18.52	8.11	4.61	3.34	2.68	2.24	0.02	0.02	0.02	0.03	46.45	43	60.23
18840	393.2	non-UDG	20949484	97.82	93546	1.81	1.98	1.01	28.11	4.92	2.71	2.03	1.67	1.44	0.05	0.04	0.05	0.04	47.75	44	59.24
18841	398.2	non-UDG	22237094	90.96	49940	0.95	2.02	0.49	12.66	6.16	2.17	1.46	1.14	0.96	0.03	0.03	0.03	0.03	43.40	42	60.41
18842	427.3	non-UDG	25557968	76.76	6122	0.74	14.41	0.08	2.95	2.04	0.66	0.34	0.21	0.15	0.03	0.02	0.02	0.03	56.67	50	59.52
18843	427.4	non-UDG	21593576	85.45	16751	1.17	7.69	0.25	4.25	11.08	3.21	1.20	0.64	0.43	0.02	0.02	0.02	0.02	65.19	65	61.94
18844	429.1	non-UDG	23459672	83.36	36602	2.97	10.00	0.55	11.42	23.46	9.15	3.63	1.42	0.60	0.02	0.02	0.03	0.02	65.84	68	62.24
18845	429.2	non-UDG	21211472	81.04	4869	0.70	15.56	0.06	2.99	0.98	0.34	0.18	0.13	0.10	0.03	0.04	0.03	0.03	52.89	47	58.58
18846	429.3	non-UDG	22617528	81.05	3343	0.95	32.93	0.04	2.28	0.90	0.20	0.11	0.08	0.06	0.02	0.04	0.02	0.04	53.58	47	58.35
18847	430.2	non-UDG	24377072	96.94	63931	1.24	2.32	0.75	28.66	1.75	0.91	0.74	0.64	0.58	0.07	0.07	0.07	0.07	51.94	47	57.88
18848	483	non-UDG	23976124	75.70	20174	0.58	3.83	0.28	12.11	3.90	0.65	0.42	0.34	0.29	0.03	0.04	0.03	0.04	61.89	55	59.74
18849	488	non-UDG	20750050	96.90	56580	1.07	1.83	0.55	12.06	5.71	2.90	2.10	1.66	1.38	0.03	0.04	0.03	0.04	43.17	41	59.73
18850	489	non-UDG	20500602	90.46	46432	1.13	2.27	0.45	10.04	5.32	2.68	1.88	1.46	1.20	0.02	0.04	0.02	0.03	42.61	40	59.93
18851	491	non-UDG	27626674	98.46	39922	1.80	5.82	0.40	14.76	10.83	1.47	0.58	0.44	0.36	0.04	0.04	0.05	0.05	43.66	41	59.33
18852	492	non-UDG	24471024	98.27	180944	5.07	3.38	1.96	23.01	73.52	45.82	24.43	11.47	4.87	0.04	0.02	0.04	0.02	47.88	45	62.37
18853	495	non-UDG	26783090	98.47	892471	9.38	1.30	8.62	11.03	92.15	86.17	79.59	72.78	65.95	0.02	0.01	0.02	0.01	42.62	40	61.22
18854	499	non-UDG	19326322	90.63	28462	0.69	2.41	0.31	10.11	4.14	1.74	1.12	0.84	0.68	0.02	0.03	0.02	0.03	48.73	44	59.93
18855	500	non-UDG	27710236	90.69	124556	1.86	2.07	1.29	26.28	7.48	4.22	3.22	2.69	2.32	0.01	0.02	0.01	0.02	45.64	42	60.87
18856	502	non-UDG	22600968	95.34	48914	0.73	1.62	0.51	13.16	5.32	2.61	1.81	1.40	1.14	0.02	0.03	0.02	0.03	45.56	42	60.17
18857	505	non-UDG	24350442	97.26	68055	0.89	1.51	0.68	14.95	6.26	3.50	2.53	2.02	1.68	0.03	0.03	0.03	0.03	44.40	42	60.28
18858	506	non-UDG	22562888	94.51	98632	1.44	1.56	1.04	19.82	9.21	5.56	4.17	3.37	2.83	0.03	0.04	0.04	0.04	46.41	43	60.49
18859	507	non-UDG	22853492	96.44	51343	0.69	1.52	0.56	13.87	5.97	3.14	2.26	1.77	1.45	0.02	0.03	0.03	0.03	48.54	45	59.93
18860	508.2	non-UDG	27721804	98.13	88613	0.95	1.46	0.93	16.98	7.62	4.74	3.63	2.93	2.47	0.04	0.03	0.04	0.03	46.29	44	60.73
18861	508.3	non-UDG	24493458	98.10	85235	1.09	1.55	0.95	21.58	5.89	3.34	2.47	1.97	1.64	0.02	0.02	0.02	0.02	49.04	45	61.36

Supplementary Table 3. Mapping statistics for the MTBC captured UDG treated libraries. These data were mapped to the TB ancestor reference (MTBC_anc). It should be noted that MPI-SHH ID's are showing the last five digits, as all external libraries are given the prefix XXX001.A… (XXX001.A18873).

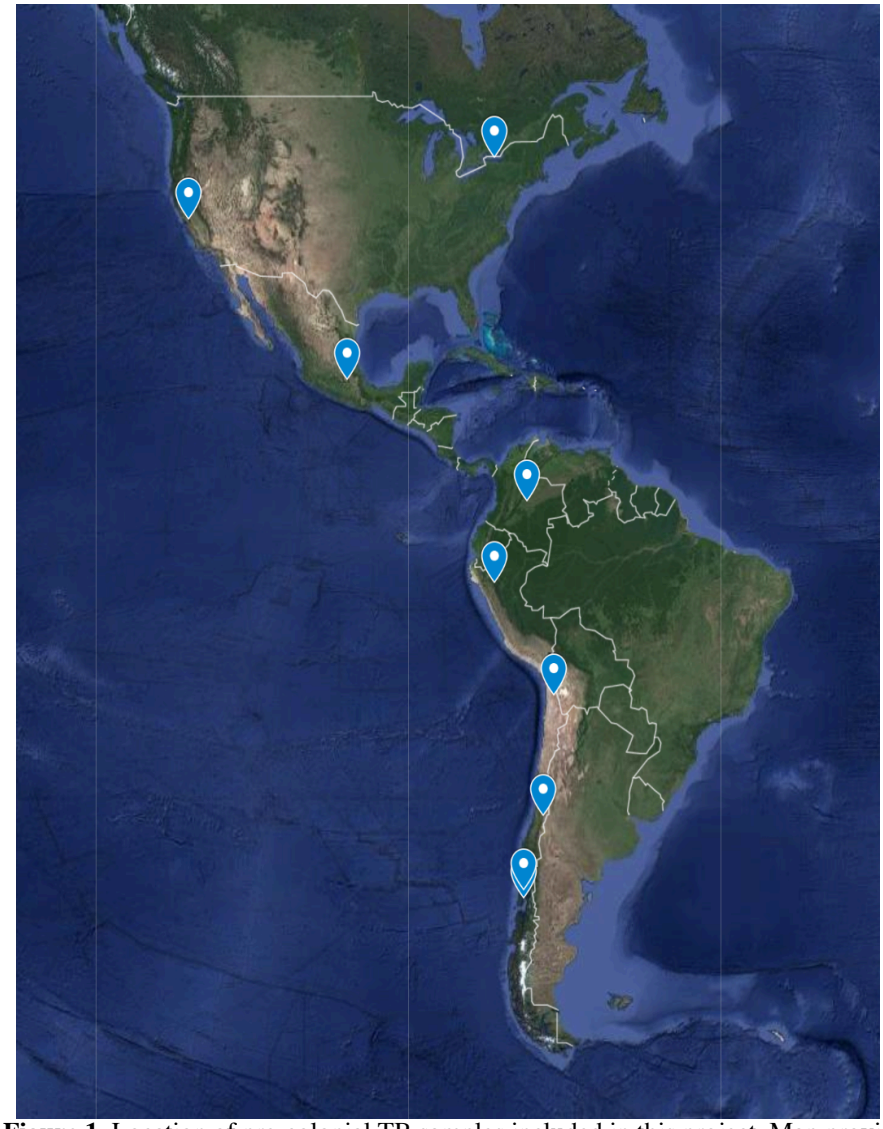
MPI Pandora ID	ASU ID	Library Type	# of Raw sequencing reads	% Merged reads	Mapped reads after RMDup	Endogenous DNA QF (%)	Cluster factor	Mean coverage	std. dev. coverage	Coverage >= 1X in %	Coverage >= 2X in %	Coverage >= 3X in %	Coverage >= 4X in %	Coverage >= 5X in %	DMG 1st Base 3'	DMG 2nd Base 3'	DMG 1st Base 5'	DMG 2nd Base 5'	average fragment length	median fragment length	GC content in %
18873	353.3	UDG	22990494	98.57	14625	0.40	3.00	0.15	2.24	4.21	1.17	0.75	0.58	0.48	0.01	0.01	0.01	0.01	43.91	39	61.07
18874	354.1	UDG	24326394	99.28	24047	1.79	8.24	0.23	2.80	4.16	2.24	1.57	1.20	0.97	0.02	0.01	0.01	0.01	41.73	37	61.44
18875	355.1	UDG	25869886	99.35	13265	1.17	11.01	0.13	2.42	2.59	0.83	0.62	0.52	0.45	0.01	0.02	0.01	0.01	44.10	38	61.25
18876	358.2	UDG	25038084	98.86	27537	0.54	2.36	0.28	3.22	5.00	2.74	1.95	1.51	1.23	0.01	0.01	0.01	0.01	44.77	39	61.37
18877	361.2	UDG	22779122	98.91	20623	0.54	2.88	0.21	2.81	3.62	1.94	1.36	1.07	0.88	0.01	0.01	0.01	0.01	44.01	39	61.09
18878	362	UDG	24264322	97.99	24410	0.66	3.18	0.24	2.95	4.31	2.27	1.58	1.21	0.99	0.01	0.01	0.01	0.01	42.79	38	60.93
18879	393.2	UDG	28337868	98.74	14481	0.78	7.23	0.15	2.60	2.28	1.13	0.80	0.65	0.54	0.01	0.01	0.01	0.01	44.40	38	61.36
18880	398.2	UDG	22608316	98.35	6729	0.18	2.98	0.07	1.52	2.61	0.52	0.33	0.24	0.19	0.01	0.01	0.01	0.01	46.78	42	60.23
18881	427.3	UDG	23252946	97.84	104380	1.49	1.62	1.24	3.64	58.67	27.40	11.28	4.70	2.36	0.00	0.00	0.00	0.00	52.29	46	63.18
18882	427.4	UDG	25244208	98.83	43836	0.92	2.53	0.46	3.74	15.96	3.94	2.26	1.74	1.44	0.01	0.01	0.01	0.01	46.22	40	62.06
18883	429.1	UDG	24197828	99.18	358196	4.67	1.53	3.97	3.89	87.19	74.27	60.58	47.58	36.05	0.00	0.00	0.00	0.00	48.89	44	62.57
18884	429.2	UDG	22855792	99.00	24209	0.94	4.31	0.28	2.45	16.25	2.63	0.84	0.55	0.44	0.00	0.00	0.01	0.00	51.26	45	61.94
18885	429.3	UDG	27954202	98.81	59314	1.17	2.68	0.69	2.92	39.90	12.39	3.83	1.51	0.89	0.00	0.00	0.00	0.00	51.47	45	62.97
18886	430.2	UDG	18940414	98.60	7530	0.78	9.25	0.08	2.22	0.61	0.43	0.35	0.31	0.27	0.02	0.02	0.02	0.02	47.56	40	59.71
18887	483	UDG	24670736	95.68	7920	0.82	12.59	0.09	1.91	2.62	0.59	0.37	0.28	0.24	0.01	0.02	0.01	0.01	48.11	42	60.55
18888	488	UDG	26927666	98.49	15119	0.58	4.70	0.14	2.24	2.54	1.28	0.91	0.71	0.59	0.01	0.01	0.01	0.01	41.38	37	60.70
18889	489	UDG	23878152	98.51	13015	0.46	3.75	0.12	1.97	2.41	1.16	0.80	0.62	0.50	0.01	0.02	0.01	0.01	40.83	37	60.49
18890	491	UDG	22151648	99.15	11331	0.88	8.15	0.11	1.10	7.86	0.66	0.19	0.13	0.11	0.01	0.01	0.01	0.01	41.74	39	60.49
18891	492	UDG	28312858	99.28	189048	4.08	2.96	1.90	2.08	75.01	50.62	30.58	16.77	8.39	0.00	0.00	0.00	0.00	44.30	42	62.37
18892	495	UDG	24416656	98.84	708301	8.94	1.41	6.81	5.34	91.21	84.06	76.25	68.19	60.15	0.00	0.00	0.00	0.00	42.40	40	61.41
18893	499	UDG	29942282	93.18	9604	0.24	3.73	0.11	2.32	2.69	0.90	0.56	0.42	0.33	0.01	0.01	0.01	0.01	50.60	43	60.12
18894	500	UDG	24024250	95.96	19131	0.65	4.03	0.20	2.92	3.28	1.51	1.10	0.89	0.75	0.01	0.01	0.01	0.01	45.05	39	61.09
18895	502	UDG	25975156	95.77	9209	0.21	2.93	0.10	2.16	1.95	0.87	0.58	0.44	0.35	0.01	0.01	0.01	0.01	47.66	40	60.07
18896	505	UDG	22630316	92.94	7206	0.18	2.81	0.09	2.20	1.60	0.71	0.46	0.34	0.27	0.01	0.01	0.01	0.01	53.43	44	59.68
18897	506	UDG	23062290	95.86	27424	0.63	2.56	0.29	3.70	4.19	2.42	1.79	1.43	1.19	0.01	0.01	0.01	0.01	46.54	39	61.06
18898	507	UDG	27495152	99.03	17858	0.32	2.31	0.21	2.92	4.72	2.07	1.35	0.99	0.79	0.09	0.11	0.08	0.10	52.39	47	62.38
18899	508.2	UDG	32586100	98.85	6011	0.20	1.39	0.07	1.32	2.43	0.90	0.52	0.35	0.26	0.08	0.08	0.06	0.09	54.76	48	60.59
18900	508.3	UDG	26138118	97.39	4431	0.23	1.57	0.06	1.41	1.48	0.55	0.34	0.26	0.21	0.10	0.09	0.08	0.08	58.78	60	62.83
Negative Controls																					
18901	EB 11.20.1	UDG	4353688	90.74	1443	0.38	5.65	0.01	0.71	0.32	0.13	0.08	0.06	0.05	0.04	0.03	0.02	0.02	45.05	40	59.30
18902	EB 2.26.18	UDG	6647634	90.12	2195	0.34	5.13	0.02	0.86	0.54	0.18	0.13	0.10	0.09	0.03	0.04	0.02	0.03	45.90	42	60.05
18903	EB 3.17.18	UDG	7319810	92.43	2606	0.52	6.96	0.03	0.94	0.49	0.22	0.16	0.13	0.10	0.03	0.03	0.02	0.02	44.67	40	60.15
18904	EB 4.4.18	UDG	5789336	96.35	2418	0.77	8.64	0.02	0.86	0.43	0.20	0.14	0.11	0.09	0.05	0.03	0.02	0.03	42.88	40	61.06
18905	EB 5.7.18	UDG	6262890	92.20	1896	0.35	5.68	0.02	0.80	0.37	0.16	0.12	0.10	0.08	0.05	0.04	0.02	0.03	44.84	42	59.83
18906	EB 8.7.18	UDG	6567102	93.98	2097	0.49	7.52	0.02	0.82	0.39	0.18	0.13	0.10	0.08	0.05	0.03	0.03	0.03	43.32	39	60.15
18907	EB 8.8.18	UDG	7841726	92.44	2059	0.44	8.20	0.02	0.80	0.40	0.18	0.13	0.10	0.08	0.04	0.04	0.03	0.02	43.34	40	60.22
18908	LB 8.10.18	UDG	1951666	96.15	1515	0.74	4.33	0.01	0.71	0.24	0.13	0.09	0.07	0.06	0.04	0.04	0.02	0.02	42.95	39	59.96
18909	LB 8.16.18	UDG	6158456	96.39	1262	0.79	18.19	0.01	0.64	0.22	0.11	0.07	0.06	0.05	0.03	0.05	0.01	0.03	42.40	38	59.53
Positive Control																					
EAD001.A01	-	UDG	32988014	89.252	2543284	19.971	1.366	27.0578	28.5206	96.63	96.11	95.48	94.67	93.71	0.0143	0.0016	0.0142	0.0016	46.93	45	63.94

Supplementary Table 4. Mapping statistics for the MTBC captured UDG treated libraries. These data were mapped to the human reference (Hg19). It should be noted that MPI-SHH ID's are showing the last five digits, as all external libraries are given the prefix XXX001.A… (XXX001.A18873).

MPI Pandora ID	ASU ID	Library Type	# of Raw sequencing reads	% Merged reads	Mapped reads after RMDup	Endogenous DNA QF (%)	Cluster factor	Mean coverage	std. dev. coverage	Coverage >= 1X in %	Coverage >= 2X in %	Coverage >= 3X in %	Coverage >= 4X in %	Coverage >= 5X in %	DMG 1st Base 3'	DMG 2nd Base 3'	DMG 1st Base 5'	DMG 2nd Base 5'	average fragment length	median fragment length	GC content in %
18873	353.3	UDG	22990494	98.57	74587	1.01	1.47	0.74	16.07	7.92	3.20	2.19	1.72	1.43	0.02	0.02	0.02	0.02	43.75	42	60.77
18874	354.1	UDG	24326394	99.28	157351	3.66	2.58	1.54	28.20	10.02	6.11	4.58	3.69	3.10	0.02	0.02	0.02	0.02	43.16	41	61.19
18875	355.1	UDG	25869886	99.35	107029	2.37	2.75	1.09	28.39	4.57	2.16	1.73	1.49	1.34	0.02	0.03	0.02	0.03	45.05	42	60.14
18876	358.2	UDG	25038084	98.86	127243	1.40	1.32	1.31	18.25	12.13	7.63	5.82	4.76	4.05	0.02	0.02	0.02	0.02	45.33	43	61.34
18877	361.2	UDG	22779122	98.91	101090	1.31	1.42	1.02	18.07	9.07	5.36	4.03	3.26	2.78	0.02	0.02	0.02	0.02	44.69	42	60.85
18878	362	UDG	24264322	97.99	121366	1.58	1.53	1.19	19.60	10.33	6.07	4.49	3.64	3.08	0.02	0.02	0.02	0.02	43.42	41	60.63
18879	393.2	UDG	28337868	98.74	111339	1.72	2.08	1.13	26.61	5.73	3.15	2.37	1.96	1.70	0.02	0.02	0.02	0.02	44.93	42	60.7
18880	398.2	UDG	22608316	98.35	34823	0.48	1.51	0.36	10.60	4.90	1.65	1.10	0.87	0.72	0.02	0.03	0.02	0.03	45.54	43	60.21
18881	427.3	UDG	23252946	97.84	221907	2.69	1.38	2.47	21.79	61.28	31.76	15.96	9.01	6.15	0.01	0.01	0.01	0.01	49.20	44	62.75
18882	427.4	UDG	25244208	98.83	176741	2.11	1.44	1.82	23.34	21.06	8.42	6.03	5.02	4.38	0.02	0.02	0.02	0.02	45.43	42	61.65
18883	429.1	UDG	24197828	99.18	454305	6.35	1.64	4.95	26.78	87.49	75.09	62.02	49.54	38.48	0.01	0.01	0.01	0.01	48.04	44	62.21
18884	429.2	UDG	22855792	99.00	100822	1.95	2.15	1.11	29.41	19.07	4.85	2.44	1.79	1.47	0.02	0.02	0.02	0.02	48.76	44	60.59
18885	429.3	UDG	27954202	98.81	148302	2.43	2.23	1.64	29.13	42.59	15.64	6.56	3.74	2.75	0.02	0.02	0.02	0.02	48.86	44	61.62
18886	430.2	UDG	18940414	98.60	55123	1.64	2.65	0.58	18.38	1.48	1.04	0.86	0.76	0.68	0.03	0.04	0.03	0.04	46.15	43	59.13
18887	483	UDG	24670736	95.68	51619	2.29	5.40	0.57	19.72	5.23	1.91	1.27	0.99	0.83	0.02	0.03	0.03	0.04	48.89	45	60.01
18888	488	UDG	26927666	98.49	80907	1.22	1.85	0.77	15.79	6.17	3.36	2.46	2.00	1.69	0.02	0.03	0.02	0.03	42.21	40	60.35
18889	489	UDG	23878152	98.51	64138	1.00	1.64	0.60	12.23	5.81	3.05	2.23	1.77	1.49	0.02	0.03	0.02	0.03	41.54	39	60.32
18890	491	UDG	22151648	99.15	36693	2.04	5.81	0.37	15.51	8.60	1.11	0.50	0.39	0.34	0.03	0.04	0.04	0.04	44.16	41	58.99
18891	492	UDG	28312858	99.28	226427	7.13	4.31	2.31	25.05	75.20	51.22	31.52	17.82	9.36	0.01	0.01	0.01	0.01	44.96	42	61.67
18892	495	UDG	24416656	98.84	812035	9.47	1.30	7.70	10.82	91.28	84.30	76.86	69.26	61.79	0.00	0.00	0.00	0.00	41.84	40	61.14
18893	499	UDG	29942282	93.18	53919	0.61	1.71	0.61	17.56	6.43	2.92	1.97	1.53	1.26	0.02	0.03	0.02	0.03	49.64	45	59.98
18894	500	UDG	24024250	95.96	119558	1.57	1.56	1.24	24.71	7.26	4.02	3.08	2.57	2.24	0.01	0.02	0.01	0.02	45.75	43	60.63
18895	502	UDG	25975156	95.77	48434	0.56	1.46	0.52	14.09	5.24	2.62	1.81	1.40	1.16	0.02	0.03	0.02	0.02	47.30	43	60.23
18896	505	UDG	22630316	92.94	35468	0.46	1.46	0.41	11.60	4.62	2.32	1.57	1.18	0.97	0.02	0.03	0.02	0.03	51.31	45	60.02
18897	506	UDG	23062290	95.86	135135	1.61	1.32	1.44	21.10	10.27	6.66	5.26	4.42	3.85	0.02	0.02	0.02	0.02	46.86	43	61.01
18898	507	UDG	27495152	99.03	69257	0.77	1.45	0.72	14.70	7.69	4.19	3.01	2.36	1.97	0.02	0.02	0.02	0.03	45.65	43	60.8
18899	508.2	UDG	32586100	98.85	133312	1.21	1.44	1.38	21.37	9.53	6.47	5.16	4.34	3.78	0.02	0.03	0.02	0.03	45.66	43	61.54
18900	508.3	UDG	26138118	97.39	88713	1.06	1.54	0.98	23.01	5.91	3.36	2.49	1.99	1.67	0.02	0.02	0.02	0.02	48.86	45	61.22

Supplementary Table 5. Estimations of MTBC reads and evaluation of mycobacterial background contamination of samples using MALT analysis. The yellow highlight indicates the four genomes reconstructed thus far. Genomes 18883/429.1, 18891,492, and 18892/495 show remarkable values of 99-100% MTBC reads, suggesting little to no environmental mycobacteria background. Genome 18881/427.3 has an estimated 63% of MTBC reads suggesting a significant amount of contamination from environmental mycobacteria. It should be noted that MPI-SHH ID's are showing the last five digits, as all external libraries are given the prefix XXX001.A… (XXX001.A18873).

MPI-SHH ID: UDG-libraries	ASU sample ID	Elements sampled	Total Assigned Reads in MALT 95ID	Total Summarized Reads Mycobacteria	Total Assigned Reads Mycobacteria	Total non-MTBC Myocbacterial Reads	Total Summarized Reads MTBC	Total Assigned Reads MTBC	Upper Boundary Calculation Estimation of MTBC Reads (%)	Lower Boundary Calculation Estimation of MTBC reads (%)	Middle Calculation Estimation of MTBC reads (%)
18873	**353.3**	vertebra	827869	25320	14433	19785	5535	4262	0.79	0.22	0.22
18874	**354.1**	rib	615838	14826	9389	12541	2285	1585	0.79	0.15	0.16
18875	**355.1**	mandibular right 2nd molar	845210	9836	5461	7896	1940	1661	0.75	0.20	0.20
18876	**358.2**	pelvis	383994	61876	39717	52348	9528	6475	0.80	0.15	0.16
18877	**361.2**	vertebra	527895	51871	34809	44638	7233	5043	0.81	0.14	0.14
18878	**362**	pelvis	448041	40713	27400	34927	5786	4055	0.82	0.14	0.14
18879	**393.2**	vertebra	954528	14568	9262	12294	2274	1648	0.79	0.16	0.16
18880	**398.2**	vertebra	439060	12204	5573	8027	4177	3601	0.80	0.34	0.35
18881	**427.3**	vertebra	1117229	30295	16012	21304	8991	8582	0.83	0.30	0.30
18882	**427.4**	vertebra	896005	97700	34284	42765	54935	53190	0.91	0.56	0.57
18883	**429.1**	vertebra	1208995	255203	6207	7689	247514	243328	0.99	0.97	0.98
18884	**429.2**	Rib	1344858	7769	2650	3309	4460	4460	0.92	0.57	0.58
18885	**429.3**	Rib	1520618	19425	4105	5372	14053	13823	0.93	0.72	0.73
18886	**430.2**	calculus	1268785	7197	4243	5237	1960	1639	0.86	0.27	0.28
18887	**483**	vertebra	1148326	8382	3139	4442	3940	3773	0.84	0.47	0.47
18888	**488**	vertebra	395898	16483	10499	13839	2644	1959	0.80	0.16	0.16
18889	**489**	vertebra	392160	17502	11705	14855	2647	1945	0.82	0.15	0.15
18890	**491**	vertebra	3063760	16259	1531	1531	14728	14728	1.00	0.91	0.91
18891	**492**	maxillary right 2nd molar	3003423	277787	3263	3263	274524	271507	1.00	0.99	1.00
18892	**495**	vertebra	4491665	1071598	11252	11252	1060346	1049781	1.00	0.99	1.00
18893	**499**	vertebra	286321	15107	9292	12614	2493	1909	0.78	0.17	0.17
18894	**500**	sacrum	784709	48354	28633	39894	8460	6109	0.77	0.17	0.18
18895	**502**	metacarpal	335493	14625	9222	12300	2325	1646	0.79	0.16	0.16
18896	**505**	Thoracic centrum (unfused)	421028	21228	13700	18168	3060	2104	0.79	0.14	0.15
18897	**506**	Thoracic centrum (unfused)	501901	61738	42413	54070	7668	5554	0.81	0.12	0.13
18898	**507**	rib head	333561	26078	17535	22186	3892	2944	0.82	0.15	0.15
18899	**508.2**	proximal pedal phalanx	413100	50168	26118	39022	11146	8005	0.74	0.22	0.22
18900	**508.3**	rib	555252	35383	17649	27534	7849	5293	0.72	0.22	0.22



Supplementary Figure 1. Location of pre-colonial TB samples included in this project. Map provided by Kelly E. Blevins and constructed using Google Maps.