

UNIVERSITY OF OKLAHOMA

GRADUATE COLLEGE

MUSEUM BIOMOLECULES: EXAMINING HUMAN PATHOGENS AND ASSESSING
DIFFERENTIAL COLLECTION PRESERVATION WITH DENTAL CALCULUS FROM
THE NATIONAL MUSEUM OF NATURAL HISTORY

A DISSERTATION

SUBMITTED TO THE GRADUATE FACULTY

in partial fulfillment of the requirements for the

Degree of

DOCTOR OF PHILOSOPHY

By

RITA MAVIS AUSTIN

Norman, Oklahoma

2020

MUSEUM BIOMOLECULES: EXAMINING HUMAN PATHOGENS AND ASSESSING
DIFFERENTIAL COLLECTION PRESERVATION WITH DENTAL CALCULUS FROM
THE NATIONAL MUSEUM OF NATURAL HISTORY

A DISSERTATION APPROVED BY THE
DEPARTMENT OF ANTHROPOLOGY

BY THE COMMITTEE CONSISTING OF

Dr. Courtney Hofman, Chair

Dr. Hayley Lanier

Dr. Cecil Lewis Jr.

Dr. Christina Warinner

Dr. Kimberly Marshall

© Copyright by RITA MAVIS AUSTIN 2020

All Rights Reserved.

Acknowledgments

I would like to express my sincere gratitude to everyone who has assisted me during my doctoral program. My committee members, in particular, have provided much needed guidance and support for the completion of this dissertation. The projects described herein would not have been possible without my chair, Dr. Courtney Hofman, who has patiently encouraged and taught me essential scientific and professional skills. The support and assurance provided by Dr. Christina Warinner and Dr. Cecil Lewis Jr. have propelled the dissertation (and me) forward. Thank you to Dr. Kimberly Marshall, who helped ground my research in broader, social frameworks. I also appreciate Dr. Hayley Lanier's involvement in ensuring that my writing and explanations are understood by audiences outside of anthropology. Dr. Sabrina Sholts, of the Smithsonian Institution, also positively contributed to my dissertation and helped me gain valuable museum and collection experience. I could not have participated or attended many conferences and workshops without the financial support provided by the Department of Anthropology and Graduate College. I thank all of my lab mates, including Nisha Patel, Dave Jacobson, Justin Lund, Tanvi Honap, Elizabeth Young, John Boyd, Allison Mann, Stephanie Schnorr, Krithivasan Sankaranarayanan, Cara Monroe, and Nihan D. Dagtas for patiently teaching me things. I also would like to thank all of the co-authors that will be a part of the published products resulting from this dissertation—the research would not have been possible without them, including Molly Zuckerman, Anne Stone, and Alexander Hübner. Finally, I thank Alvin Austin, who said “if it was easy, everyone would do it”, the Lashley and Walsh families for long-time help, support, and fun, Donna Frank for widening my perspectives about the world, and my mother, Joan Skopal, for continuous support, love, and inspiration throughout this process. And finally, I would like to broadly thank the individuals whose skeletal remains I analyzed and included in this dissertation.

Contents

Content	Pg
1 Introduction	10
2 Opinion: To Curate the Molecular Past, Museums Need a Carefully Considered Set of Best Practices	16
3 Curating Biomolecules: Archival and Molecular Legacies of Preservations Past Across Several Smithsonian's Biological Anthropology Collections	
3.1 Abstract	21
3.2 Introduction	21
3.3 Materials and Methods	25
3.3.1 Collection Archival Research	25
3.3.2 DNA Extraction	25
3.3.3 Illumina Library Preparation	26
3.3.4 Computation Methods	26
3.4 Results	27
3.4.1 Collection Histories and Treatment Summaries	27
Japan (AMM)	30
Egypt	30
Germany	31
Mongolia	31
Japan (IUCMT)	32
Switzerland	32
Terry Collection	33
3.4.2 Metagenomic Results	33
DNA Recovery	33
Microbial Community Authentication	35
Dental Calculus Microbial Community Comparisons	38
3.5 Discussion	43
3.6 Conclusions	48
3.7 Acknowledgments	48
4 Assessing Dental Calculus as a Source of Ancient Tuberculosis and Syphilis DNA from the Smithsonian's Terry Collection	
4.1 Abstract	49
4.2 Introduction	49
4.3 Methods	51
4.4 Results	55
4.4.1 Skeletal Diagnoses	55
4.4.2 Shotgun Sequencing	55
4.4.3 MTBC quantitative PCR	57
4.4.3 Hybridization Capture	58
4.5 Discussion	58
4.6 Conclusions	60

4.7 Acknowledgments	60
5 Dental Calculus Reveals Acute Bacterial Infections: A Case Study from the Smithsonian's Terry Collection	
5.1 Abstract	62
5.2 Introduction	62
5.3 Methods	64
5.4 Results	65
5.4.1 Skeletal Analysis	65
5.4.2 DNA Preservation and Authentication	66
5.4.3 Pathogen Identification	67
5.5 Discussion	69
5.5.1 The Substrate	69
5.5.2 The Individual	69
5.5.3 Recognizing Anatomies	72
5.6 Conclusions	73
5.7 Acknowledgments	74
6 Conclusion	75
7 References	80
Appendix A: Chapter 3 Supplementary Materials	94
Appendix B: Chapter 4 Supplementary Materials	115
Appendix C: Chapter 5 Supplementary Materials	148

List of Tables

3.1 Documented collection histories based on archival research on the accession, catalog cards, curator notes/ journal entries, and literature sources	29
5.1 Percent of reads mapping to pneumonia pathogen reference genome	68
A.1 Archival information referencing contextual and preservational insights for the analyzed individuals and collections	95
A.2 Sample removal criteria and information	104
A.3 sPLS-DA balance error rates (BER)	105
B.1 Tuberculosis indicator score criteria	122
B.2 Syphilis and tuberculosis indicator scores determined for each individual	122
B.3 Sample statistics when mapped to <i>M. tuberculosis</i> reconstruction	124
B.4 Sample statistics when mapped to <i>T. pallidum</i> reference genome	126
B.5 rpoB2 qPCR sample results	130
B.6 IS6110 qPCR results. All samples were run in duplicate	133
B.7 Capture samples mapped to <i>M. tuberculosis</i> reconstruction	136
B.8 Capture experiments MTBC lineage SNP Calling. Colors correspond to capture experiment 1 (blue) and 2 (green)	138
C.1 Musculoskeletal activity markers	159
C.2 Traumatic lesions and 'other' pathological lesions not directly associated with TB	160

C.3 TB indicators, following Dangvard Pedersen et al. (2019)	161
C.4 Observed oral pathologies	162
C.5 Oral microbial community composition of St.LI as classified by MetaPhlan2, sorted by descending percent abundance	163
C.6 Known causative agents of bacterial pneumonia with citations	164

List of Figures

2.1 Museum collection staff stand among anthropological collections at the Smithsonian's National Museum of Natural history	17
2.2 The graph shows variation in calculus preservation by collection, illustrating DNA yields (nanograms per milligrams)	19
3.1 Taphonomic stages affecting museum collections	23
3.2 Geographic locations of archaeological dental calculus included in this study	24
3.3 Normalized DNA (ng/mg) for each collection as measured by a Qubit fluorometer	34
3.4 SourceTracker analysis (A) and principal coordinate analysis (PCA) (B) of unweighted unifracs distances for estimated 16S rRNA gene identification	36
3.5 QIIME and MetaPhlAn2 alpha diversity and fragment length comparison	37
3.6 Microbial community differences between collection types	39
3.7 Microbial community differences between collection treatments	40
3.8 Microbial community differences without Egyptian samples...	41
3.9 Taxa contributing to collection discrimination by sPLS-DA	42
4.1 SourceTracker analysis of shotgun metagenomic reads sorted by increasing proportions of taxa	56
4.2 Proportion of unique, quality filtered reads mapped to genomes of <i>F. nucleatum</i> ...	57
5.1 Authentication of Dental Calculus Metagenomic DNA	67
A.1 Stacked bar plot of order level relative abundances as identified by QIIME and MetaPhlAn2	106
A.2 Unweighted 16S rRNA PCA analysis of contributing oral environmental sources	107
A.3 MapDamage terminal end damage patterns as mapped to <i>T. forsythia</i>	108
A.4 Taxa contributing to discriminating among collection types	109
A.5 Microbial community differences between collection treatments with single Egyptian category	110
A.6 Original analysis number of species level OTUs identified by collection with QIIME and MetaPhlAn2	111
A.7 Original microbial community differences among collection treatments	112
A.8 Original analysis taxa contributing to collection discrimination	113
A.9 Original microbial differences between collection types	113
B.1 Genome coverage and GC content plot of the top three samples (S16, T1, and T9) with unique reads mapping to <i>M. tuberculosis</i>	143

B.2 Genome coverage and GC content plot of the top three samples (T7, T2, and S6) with unique reads mapping to <i>T. pallidum</i>	144
B.3 SourceTracker analysis of shotgun metagenomic reads sorted by increasing proportions of taxa associated with modern subgingival plaque	145
B.4 Heatmap of HOPS identification results from dental calculus	145
B.5 Percent of shotgun metagenomic reads mapping to <i>T. pallidum</i> by sample	146
B.6 Percent of shotgun metagenomic reads mapping to <i>M. tuberculosis</i> genome reconstruction	147
C.1 Focal depression photos and radiograph	153
C.2 Anterior and posterior views of complete midline mandible fracture	154
C.3 Anterior view of the proximal left femur showing two clustered lytic pits	154
C.4 The right (A) and left (B) os coxae	155
C.5 Schematic indicating tooth and alveolar bone location of dental caries, antemortem and postmortem tooth loss, periodontal disease, gingivitis, and oral abscesses	156
C.6 Several vertebrae exhibit large pores or pits (solid arrows)	157
C.7 The cranial vault displays several pathological lesions, which may be associated with Tuberculosis	158
C.8 mapDamage plot of shotgun reads mapped to the <i>Mycobacterium tuberculosis</i> genome reconstruction	158
C.9 The right side of the mandible reveals a draining oral abscess	159

Abstract

Museum anthropology collections continue to provide insight into human experiences and disease in the past, especially in the era of biomolecular research. Molecular investigations of museum collections have increased dramatically since the recent high throughput sequencing revolution. However, basic biological questions regarding the preservation of museum biomolecules and the kind of information biomolecular substrates can provide need to be clarified. In this dissertation, four studies address the challenges to conducting molecular research with museum collections, in particular, the impact of unique acquisition and treatment histories of collections, and how museum biomolecules can be used to answer biological questions. Collections from the Smithsonian Institution's National Museum of Natural History (NMNH) Anthropology department are utilized in this dissertation to investigate: specific mechanisms and approaches to curating museum biomolecules (Chapter Two); the recovery of DNA among globally representative skeletal populations that have specific and unique acquisition and treatment histories (Chapter Three); the diagnostic potential of dental calculus to confirm COD of either syphilis or tuberculosis, using the Terry Collection (Chapter Four); and, finally, the end of life circumstance of a single individual using historic and collection contextualization, skeletal assessment, and oral microbiome molecular methods (Chapter Five).

Keywords

Museum biomolecules, ancient DNA, dental calculus, pathogens, archival research, skeletal assessment, differential preservation.

Chapter 1

Introduction

Museums act as meeting points of knowledge, research, materials, and the public. Distributing new ideas and understandings of the human experiences and the natural world is central to museum missions. More practically, museums also offer curated objects and remains, with museum records and relevant information to help elucidate human and natural histories.

Museum collection and specimens have been readily used to access biomolecules (e.g. DNA, RNA, proteins, isotopes) and use them as informative lines of evidence for understanding and reconstructing the past. Researchers seeking access and use of museum collections has increased over time, especially after the recent high throughput sequencing revolution which has allowed for more and cheaper data to be generated. Museums biomolecules' research utility is underscored by the continual inundation of destructive sampling requests to collections (Austin et al 2019, McLean et al 2016, Thomas 1994). With this increased demand for ancient materials, competition for rare and specific museum specimens among researchers has intensified, with some researchers suggesting that hoarding of museum and archaeological samples is being practiced by genetic laboratories in the United States and Europe (Lewis-Kraus 2019a, Makarewicz et al 2017). The asserted sample hoarding is due in part to the necessary screening that labs conducting biomolecular analyses need to conduct, where many samples must often be sorted to find those with exceptional preservation and pertinent biomolecular information. Necessary screening processes makes biomolecular analyses expensive and can readily use up limited museum and archaeological materials. Biomolecular analyses are also complicated by the recent realization that specimen age or provenience are not steadfast indicators of biomolecular potential or preservation (Antinick & Foran 2018, Kistler et al 2017).

Assumptions based on museum specimen age and provenience is also logically unsound when museum collection acquisition, preservation, and curation histories are also considered. Numerous conservation methods have been applied to museum collections to ensure the physical longevity of specimens over time (e.g. pesticides, cabinet fumigants, calcium hypochlorite, and formalin) (Anderson 1965, Eklund & Thomas 2010, Hangay & Dingley 1985, Hildebrand 1968, Hornaday & Holland 1891, McDonough et al 2018, Prendini et al 2002, Rowley 1925, Simmons

& Voss 2009). In combination with the original collecting intent (e.g. whether or not specific sections or regions, such as only crania, were sought after and obtained), the museum treatment and handling histories contribute to the current conditions and contents of museum collections (Austin et al 2019). Heavy handling and research being done on specific specimens and/or collections certainly affects museum collections on the molecular level.

Curating for museum biomolecules has been especially pertinent ever since the 1984 recovery of ancient DNA (aDNA) from the dried muscle of a salt preserved quagga (Higuchi et al 1984). Since this first study, ancient DNA researchers have attempted aDNA recovery from a number of substrates. Early use of museum specimens focused on soft tissue remains (Paabo 1985, Paabo et al 1988) followed by the amplification and isolation of aDNA from archaeological bones (Hagelberg et al 1989, Horai et al 1989). Optimizing and maximizing endogenous DNA yields from specimens continues to be an important area of innovation in the field; the proposal of new extraction protocols and authentication methods (Rohland et al 2015, Rohland & Hofreiter 2007), as well as alternative skeletal elements such as the petrous portion (Dabney et al 2013, Hansen et al 2017, Pinhasi et al 2015), are recent contributions to this line of inquiry.

However, there have been conflicting results provided in the literature regarding museum collection molecular recovery. In fresh tissue samples, various fixatives, pH, temperature, and duration of treatment conditions produce irreproducible DNA and RNA damage, including treatment induced DNA cross-linking and poly-A tail modification of RNA (Srinivasan et al 2002). Previous optimization work has suggested that DNA damage exclusively from museum conservation treatments to human and animal hard and soft tissues was indistinguishable from normal degradation, due to unaccounted factors (e.g. specimen age, input quantity of DNA, etc.) (Eklund & Thomas 2010, Lopez-Polin 2012). A recent study concluded that preservation techniques did not impact the DNA yield and quality from mammalian bone, claw, skin, and osteocrust samples because all samples yielded DNA sequence data (McDonough et al 2018). In this study, a tanned *Conepatus semistriatus* (striped hog-nosed skunk) skin sample produced the lowest endogenous DNA yield and shortest fragment size of all the assessed samples (McDonough et al 2018). Arsenic, which was a common storage pesticide (Anderson 1965, Hornaday & Holland 1891, Omstein 2010), may have been applied to the other samples, but the

impact on recovered DNA was not discussed (McDonough et al 2018). Ultimately, McDonough et al. (2018) only tested relative overall DNA preservation and did not systematically assess the impact of specific curation and conservation techniques on DNA preservation. Understanding how museum processing affects biomolecules is important for museum curators and managers who are tasked with weighing the expected scientific gain against the indisputable loss of priceless, timeless, and irreplaceable material.

Understanding and documenting museum collection biomolecular preservation differences is explored in this dissertation through the analysis of calcified dental plaque (dental calculus). One of the most molecularly rich human, skeletal substrates, dental calculus contains high-quality biomolecules, such as abundantly available microbial and pathogen molecules, and to a lesser extent dietary and human (host) debris, with very little burial environmental contamination (Adler et al 2013, De La Fuente et al 2013, Hendy et al 2018, Jeong et al 2018, Linossier et al 1995, Mackie et al 2017, Ozga et al 2016, Radini et al 2019, Warinner et al 2014b, Warinner et al 2015a, Wilkin et al 2020).

Dental calculus forms throughout life through the periodic calcification of dental plaque on tooth surfaces traps bacteria and microbes in a crystalline matrix. Oral fluids continue to nourish the outermost surfaces, allowing further plaque buildup and perpetuating the calculus formation cycle (Jin & Yip 2002, Lieverse 1999, Schroeder 1969). Host genetics and immune responses (Warinner et al 2014b), diet (e.g. high protein or carbohydrate) (Jin & Yip 2002, Lieverse 1999), cultural oral practices (Radini et al 2017), and mineral content of water (Lieverse et al 2007) all contribute to individually variable calculus formation. Factors contributing to differential dental calculus deposition in people have been identified, including salivation rate (Mandel 1974), diet (Adler et al 2013, Lieverse 1999), dental hygiene (Radini et al 2017), and individuals' genetics (Lieverse et al 2007). The influence of diet, whether high protein (Baer & White 1966) or high carbohydrate (Smith et al 1963), has been debated in the bioarchaeological literature, but is juxtaposed by dental literature (Akcalı & Lang 2018) that has emphasized the multifactorial development of dental calculus. Since the early 1960s, other contributing formation factors have also been speculated and discussed (Dawes 1970, Lieverse 1999, Radini et al 2017), such as accidental inhalation and using the mouth as a third hand. For example, across geographic

locations and times, archaeological evidence based on tooth wear and attrition, has been used as evidence of non-masticatory use of the mouth during cooking processes (e.g. roasting) and making cultural material (e.g. shredding plant fibers) (Buckley et al 2014, Hinton 1981, Molnar 2008), which may have also potentially contributed to or hindered calculus formation. Additionally, shifts in oral communities and periodontal disease, triggering host immune inflammatory reactions resulting in neutrophil recruitment to the mouth, can also contribute to host material being incorporated into dental calculus (Fuchs et al 2007, Mann et al 2018, Ryder 2010, Warinner et al 2014b).

In addition to aiding current understanding about biomolecular preservation impacts, anthropological questions concerning the health, diet, and population dynamics of past peoples have been made accessible through dental calculus recovered from human remains (Dewhirst et al 2010, Gismondi et al 2018, Harper & Armelagos 2013, Hendy et al 2016, Ozga et al 2016, Warinner et al 2014a, Wilkin et al 2020). Specifically related to deciphering past health conditions, Paleo-pathogenomics, or the assessment of pathogen evolution and origins using genomic data, has also recently emerged as a specialized subdiscipline of anthropological molecular research. From ancient human remains, pathogenomics has recovered several pathogens still affecting human populations today, such as tuberculosis (Bos et al 2014, Hendy et al 2016, Pepperell et al 2011), leprosy (Crespo et al 2017, Honap et al 2018, Schuenemann et al 2013) and the plague (Bos et al 2011). The incorporation of health indicating molecules into dental calculus has also been established (Warinner et al 2014b) and complements the recognized relationship between oral and systemic health in modern populations, specifically between heart and respiratory infections, and chronic conditions, where poor oral health is an indicator of and contributes to broader bodily systems (Cullinan et al 2009, Dietrich et al 2017, El Attar et al 2010, Sumi et al 2007, Warinner et al 2014b). However, the extent to which dental calculus can be used to reconstruct human health in the past remains unclear.

Reconstruction of past human health has largely relied on documentation and interpretation of skeletal markers. However, the limitations of skeletal collection (e.g. the osteological paradox) and the limited responses bone has to health stressors (e.g. bone loss or buildup) can confound molecular data interpretation (DeWitte & Stojanowski 2015, Ortner 2003). Material restrictions

inherent to skeletal collections can convolute specimen taphonomy (post-mortem changes) and differential morbidity, mortality, and frailty when doing population assessments (DeWitte & Stojanowski 2015). Additionally, the relatively slow reaction time of bone does not allow skeletal reactions to generate before death for many infectious diseases (e.g. pneumonia, influenza, cholera), due to their pathogen biology, pathogenicity, and virulence (Ortner 2003). Taken together, these can skew collection demographics and mortality estimates, and affect molecular analysis success (Austin et al 2019). Because of these factors, being able to diagnosis pathogens and health status of past peoples from skeletal remains using molecular techniques is especially meaningful. Here, dental calculus is tested as a diagnostic tool to identify causes of death (COD) from skeletal individuals with documented COD, specifically syphilis or tuberculosis, is analyzed.

Museum and biomolecular research ethics are also a consistent topic discussed throughout the dissertation. For bioarchaeologists working closely with or at museums, the preservation of human remains is of particular importance because the skeletal remains are the direct means to understand past peoples (Lambert & Walker 2018). Today, the respectful display, study, and treatment human remains in museum collections is directed by many organizational and institutional ethical guidelines and policies (Curtis 2003). The inappropriate acquirement and handling of many human remains that went on in the past, and the morality of keeping collections solely for and from a scientific perspective are emphasized by the significant legislative changes that have been made regarding museum skeletal collections, such as the National Museum of the American Indian Act (1989) and Native American Graves Protection and Repatriation Act (1990). Discussing museum collection histories is not uncommon, especially because skeletal collections are already commonly viewed as scientifically and culturally important for the “material memory” (Lambert & Walker 2018) and personalized interactions with the past they provide. Furthermore, skeletal remains also act as direct evidence of and connections to the individuals acquiring and handling them. Anthropologists have long recognized that the body is a point of social, political, and economic negotiation, that ultimately act as reflections of the world in which individuals live (Lock 1993). It is important to remember the skeletal remain providing information for molecular analysis as individuals. Furthermore, direct-to-consumer ancestry kits have made the connection between individuals and living

populations with the molecular past readily available, impacting how people relate to the past and understand themselves. This enhanced awareness of the molecular past has important implications for molecular research from human remains as increasingly personal and specific information is obtainable through molecular applications. As scientists move forward with research and using museum collections, ethical challenges related to the amount and kind of information molecular data provides will need to be navigated.

Incorporation of ethical considerations is demonstrated in a case study, where the holistic analysis of a specific individual from the Terry Collection provides insight into the living conditions, health history, and human experience. Integrating skeletal analysis, contextualizing literature, health documents, and dental calculus molecular assessment, circumstances about the days before death are reconstructed. Because knowledge gained about this individual is so personal and specific, ethical considerations compelled the redaction of this individual's museum identification and removal of sequences mapping to the human genome, specifically in an effort to maintain their genetic privacy.

Utilizing the Smithsonian Institution's (SI) National Museum of Natural History (NMNH) Anthropology skeletal collections, this dissertation encompasses biomolecular investigations to identify specific mechanisms and approaches to curating museum biomolecules (Chapter 2); to reconstruct collection acquisition and treatment histories and compare molecular recovery across several NMNH collections (Chapter 3); the NMNH Terry Collection is also used to test the diagnostic potential of dental calculus to confirm COD of either syphilis or tuberculosis, as well as test a new hybridization capture approach (Chapter 4); and, finally, to holistically present the end of life circumstance of a single individual using historic and collection contextualization, skeletal assessment, and oral microbiome molecular methods (Chapter 5).

Chapter 2

Opinion: To curate the molecular past, museums need a carefully considered set of best practices

Published January 2019 in PNAS.

Austin, R. M., Sholts, S. B., Williams, L., Kistler, L., & Hofman, C. A. (2019). Opinion: To curate the molecular past, museums need a carefully considered set of best practices. *Proceedings of the National Academy of Sciences*, 116(5), 1471-1474.



To curate the molecular past, museums need a carefully considered set of best practices

Rita M. Austin^{a,b,1}, Sabrina B. Sholts^{c,1,2}, LaShanda Williams^{d,e}, Logan Kistler^c, and Courtney A. Hofman^{a,b,2}

Biomolecular research has sparked a methodological revolution in the field of anthropology, and museums are now faced with the curatorial challenge of conserving and evaluating materials for these new methods. Since 2010, hundreds of genome-wide datasets from ancient human samples have been published, and thousands more have been generated; doubling the amount of data in the field of paleogenomics now

requires less time than the publication of a single article (1). Technological advances in laboratory and bioinformatic approaches have also led to a new era in paleomicrobiology, in which entire ancient microbial communities can be recovered from various substrates, such as calcified dental calculus (2). Increasingly, large-scale datasets for other biomolecules, such as proteins and metabolites, are generated and



Fig. 1. Museum collections staff stand among the anthropological collections at the Smithsonian's National Museum of Natural History. Museums are responsible for preserving collections and the information they contain for future generations. But in recent years, they've been given the increasingly challenging task of curating and conserving biomolecular data. Image credit: Chip Clark (Smithsonian Institution, Washington, DC).

^aLaboratories for Molecular Anthropology and Microbiome Research, Stephenson Research and Technology Center, Norman, OK, 73019;

^bDepartment of Anthropology, University of Oklahoma, Norman, OK 73019; ^cDepartment of Anthropology, National Museum of Natural History, Smithsonian Institution, Washington, DC 20560; ^dDepartment of Ecology, Evolution, and Natural Resources, School of Environmental and Biological Sciences, Rutgers, The State University of New Jersey, New Brunswick, NJ 08901-8551; and ^eDepartment of Anthropology, Center for Human Evolutionary Studies, New Brunswick, NJ 08901-1414

The authors declare no conflict of interest.

Published under the [PNAS license](#).

Any opinions, findings, conclusions, or recommendations expressed in this work are those of the authors and have not been endorsed by the National Academy of Sciences.

¹R.M.A. and S.B.S. contributed equally to this work.

²To whom correspondence may be addressed. Email: sholts@si.edu or courtney.hofman@ou.edu.

integrated into multi-omic understandings of the human past (2).

Anthropological collections are important for science and society for reasons that include their potential applications for biomolecular research (3, 4). Ethical issues are central in the changing circumstances for ancient biomolecular information. Potential negative impacts on indigenous communities (such as when paleogenomic studies contradict traditional histories and undermine territorial or repatriation claims) have magnified the importance of community-based practices (5). Competition among ancient-DNA laboratories for museum samples has been likened to “the Wild West” (6), and more strictly regulating access to samples has been advocated to prevent “hoarding” (7). As stewards of collections that contain ancient biomolecules (e.g., isotopes, proteins, DNA, and metabolites), museums play a critical role among stakeholders in biomolecular research and should be responsive to such concerns (Fig. 1).

Although laws concerning human remains in museum collections have had the positive effect of increasing communication between researchers and indigenous communities, they specify only a minimum course of action for community engagement that often falls short of what is ethically appropriate.

Museums are also being inundated with destructive-sampling requests, underscoring the notion that collections and their biomolecules are finite resources. Anthropological collections, including ethnographic, archaeobotanical, and zooarchaeological, as well as human remains, are a source of unique insights into the history of our species, and it is an ethical imperative to preserve this information for future generations (8). Tasked with ensuring the scientific and societal longevity of their collections, museums consider a variety of criteria in making sampling decisions. Representing a unique perspective and a rich specimen repository, the position of museums in protecting, preserving, and providing materials for research should be included in biomolecular discussions.

Symposia such as our recent one at the 2018 meeting of the Society for American Archaeology suggest that an open dialogue about museum practices and policies is beneficial for all stakeholders. By drawing on examples from the Department of Anthropology at the Smithsonian Institution’s National Museum of Natural History (NMNH), we hope to see more take part in this conversation. Researchers and curators need to better address the challenges of curating ancient biomolecules, and we should, as a professional community, work toward the development and dissemination of best practices.

Laws and Consultation

Although laws concerning human remains in museum collections have had the positive effect of increasing

communication between researchers and indigenous communities, they specify only a minimum course of action for community engagement that often falls short of what is ethically appropriate (9). Paleogenomic researchers, whose studies may destructively sample bone, teeth, dental calculus, hair, soft tissues, and/or feces, have been encouraged to consult with descendant communities at the earliest stage of project design (5). Museums can be instrumental in facilitating these efforts, especially in cases in which it is unclear how or with whom researchers should engage.

At the NMNH, which is subject to the National Museum of the American Indian Act—the Smithsonian-specific counterpart to the Native American Graves Protection and Repatriation Act—the Department of Anthropology’s Repatriation Office is represented on the destructive-sampling committee for requests that involve—or may involve—Native American remains. Culturally affiliated remains require consultation and support from the community. For remains that have not been culturally affiliated under NMNH’s repatriation process, there is a longstanding practice of encouraging researchers to consult with representatives of the likely affiliated tribes. Additionally, NMNH provides a letter from the Smithsonian Institution’s Native American Repatriation Review Committee encouraging researchers to consult with Native American communities, whether or not their research involves destructive sampling, and provides guidance of whom to contact in these communities.

Seeking Preservation

With each proposal for destructive sampling, a committee weighs the likely gain in scientific knowledge against the unquestionable loss of priceless and irreplaceable material. It is important to consider that one residue or substrate can be subjected to several different techniques to address very different questions. The petrous bone, for instance, has become the preferred sampling site for paleogenomic studies of ancient population history because of optimal preservation of endogenous, organismal DNA (10), but the bone can also be useful for stable isotope analysis as a supplement or proxy for teeth in reconstructing diet during early life (11), and it carries important morphological signals of population histories (12). If the entire inner ear is used for DNA extraction, this information is lost, and if sampling involves removal of a large part of the petrous bone, morphological information about sex and childhood disease may be lost as well (13). Likewise, well-preserved teeth can perform as well as petrous bones for paleogenomic research in some cases, but with a potential tradeoff in limiting their suitability for replicative or other types of analysis after destructive sampling (13).

Although the amount of sample needed for multi-omic analysis can be just a few milligrams, it may still represent a large proportion of the well-preserved biomolecules for that individual or even an entire collection. Our ongoing survey of NMNH’s human skeletal remains, consisting of a cross-regional sample from North America, Europe, and Africa and encompassing

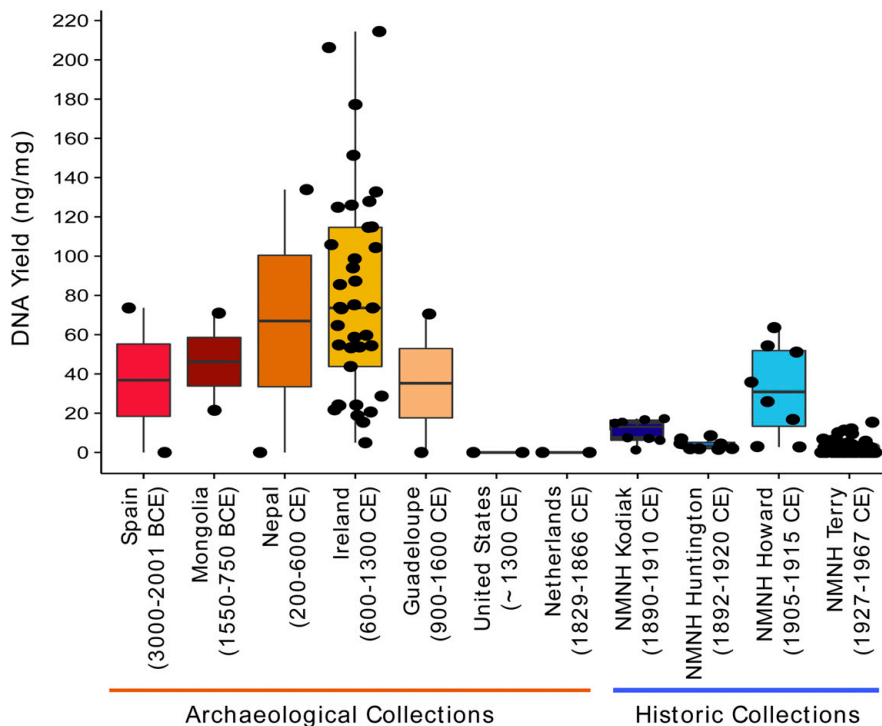


Fig. 2. The graph shows variation in calculus preservation by collection, illustrating DNA yields (nanograms per milligrams) for four 19th- to 20th-century NMNH anatomical collections and archaeological data (14) ranging from 3,000 BCE to the 19th century CE. All DNA samples were extracted with a widely used ancient DNA protocol (16) and quantified using a Qubit fluorometer. The DNA yields for archaeological collections were generally high with a more varied range of values whereas three of the four historic NMNH anatomical collections have consistently low amounts of DNA.

29 collections, highlights the disparate substrate availability within collections. Among 2,793 individuals, only about 11% have at least 5 mg of preserved dental calculus. Postmortem loss of calculus, due to intentional curatorial cleaning or accidental damage, was evident in roughly 50% of assessed individuals ($n = 1,552$). However, about 92% of individuals with calculus also have at least one intact petrous bone. Factors such as substrate availability and postmortem treatment are important to consider during study design, particularly for population-level investigations because collections may not contain sufficient dental calculus, petrous bone, or another substrate for representative analyses.

Moreover, knowledge about the curatorial conditions of collections, as well as historic and modern collection conservation practices (e.g., use of arsenic, glue, shellac, and pesticides) is important for assessing substrate-specific biomolecular yields and for interpreting the resulting data. Although dental calculus has proven to be a reservoir rich in endogenous human, microbial, pathogen, and dietary biomolecules, our analyses have shown considerable variation in DNA preservation between archaeological (14) and NMNH anatomical collections (Fig. 2). The latter collections are

more recent than the former (i.e., 100 versus several thousands of years old), yet yielded very low quantities of DNA, which may be because of postmortem preparation techniques, such as heat and chemical maceration with boiling water and benzene vapors.

As more is learned about factors of ancient DNA decay (15), it's becoming increasingly evident that assumptions should not be made about biomolecule preservation based solely on the age of remains or their provenience. Because the treatment of remains can cause unexpected taphonomic variation in ancient biomolecules, successful molecular recovery from one anatomical and/or archaeological locality does not guarantee comparable success from another. For this reason, researchers may encounter hurdles in obtaining approval for the destructive sampling of collections for which biomolecular preservation is unknown.

In the Department of Anthropology at NMNH, decisions about destructive sampling are based, in part, on the likelihood that the proposed analytical methods will yield the intended results and gain the most possible information from sampled collections. To safeguard these collections for studies in the near or distant future—when improved techniques may be

available—the destructive-sampling committee may approve a pilot test on a small subset of a larger request. The idea is to garner proof that biomolecular preservation is suitable for the proposed research before more extensive sampling can proceed.

Data Dissemination and Retention

A high likelihood of receiving analytical results from the requestor is a common criterion for destructive-sampling approval. Most museums have longstanding policies that require copies of documentation and publications from destructive sampling be kept as permanent records. However, many museums have yet to join with academic institutions, funding bodies, and scientific journals in adopting policies for proactive data archiving, maintenance, and sharing.

There is particular interest in paleogenomic data sharing for empirical investigations because of high experimental reliability and reproducibility standards (9) and museums have the ability to implement policies supporting the open science movement. For sampling requests to analyze nucleic acids (DNA, RNA, and epigenetic markers) from anthropological collections, NMNH has recently introduced requirements about the subsequent accessibility of raw data to ensure complete replicability of research and stable, open access to data deriving from collections. Requestors must now submit a Data Management Plan with their proposal that includes a specific strategy and timeline of data collection, management, backup, and release to field-standard genomic data repositories (https://naturalhistory.si.edu/sites/default/files/media/file/Anthro-SamplingPoliciesProcedures_0.pdf).

Like many institutions and organizations, museums are keen to adapt to changes in science and society. With respect to anthropological collections, museums everywhere face challenges to balance scientific interests, descendant concerns, and the need to preserve collections for future generations. Biomolecular techniques offer new avenues to understand the past, and curating specifically for biomolecules can increase their research applicability and continuing relevance. Although museums may differ in their views and policies toward destructive sampling, active discussion and consultation with stakeholders continue to be critical for preserving collections and developing innovative research partnerships. Museum collections are unique resources, and adjustments to these policies promise the sort of scientific and social benefits that only museums can provide.

Acknowledgments

We thank Siobain Duffy, Tanvi Honap, David Hunt, Dave Jacobson, Allison Mann, Susanna Sabin, Krithivasan Sankaranarayanan, Christina Warinner, and Molly Zuckerman for thought-provoking conversations; Nihan Kilic and Samuel Carreon for data-generation contributions; and William Billeck for thoughtful feedback on the manuscript. Finally, we thank all the participants from the 2018 Society for American Archaeology Biomolecules and Museums Collections: Challenges and Best Practices for Destructive Sampling session for contributing a range of perspectives related to museums and ancient biomolecules. This research was supported by grants from the Smithsonian Institution Graduate Fellowship Program; Department of Ecology, Evolution and Natural Resources, Rutgers, The State University of New Jersey; the Center for Human Evolutionary Studies, Rutgers, The State University of New Jersey; the University of Oklahoma; and the National Science Foundation (BCS-1643318 and BCS-1826220).

- 1 Reich D (2018) *Who We Are and How We Got Here: Ancient DNA and the New Science of the Human Past* (Oxford Univ Press, London).
- 2 Green EJ, Speller CF (2017) Novel substrates as sources of ancient DNA: Prospects and hurdles. *Genes (Basel)* 8:E180.
- 3 Sholtz SB, Bell JA, Rick TC (2016) Ecce homo: Science and society need anthropological collections. *Trends Ecol Evol* 31:580–583.
- 4 Hofman CA, Rick TC, Fleischer RC, Maldonado JE (2015) Conservation archaeogenomics: Ancient DNA and biodiversity in the Anthropocene. *Trends Ecol Evol* 30:540–549.
- 5 Bardill J, et al.; Summer internship for INdigenous peoples in Genomics (SING) Consortium (2018) Advancing the ethics of paleogenomics. *Science* 360:384–385.
- 6 Callaway E (August 11, 2017) Stop hoarding ancient bones, plead archaeologists. *Nature News*. Available at <https://www.nature.com/news/stop-hoarding-ancient-bones-plead-archaeologists-1.22445>. Accessed December 31, 2018.
- 7 Makarewicz C, Marom N, Bar-Oz G (2017) Palaeobiology: Ensure equal access to ancient DNA. *Nature* 548:158.
- 8 Walker PL (2008) Bioarchaeological ethics: A historical perspective on the value of human remains. *Biological Anthropology of the Human Skeleton*, eds Katzenberg MA, Saunders SR (Wiley-Liss, Hoboken, NJ), 2nd Ed, pp 3–40.
- 9 Anagnostou P, et al. (2015) When data sharing gets close to 100%: What human paleogenetics can teach the open science movement. *PLoS One* 10:e0121409.
- 10 Pinhasi R, et al. (2015) Optimal ancient DNA yields from the inner ear part of the human petrous bone. *PLoS One* 10:e0129102.
- 11 Järkov ML, Heinemeier J, Lynnerup N (2009) The petrous bone—a new sampling site for identifying early dietary patterns in stable isotopic studies. *Am J Phys Anthropol* 138:199–209.
- 12 Ponce de León MS, et al. (2018) Human bony labyrinth is an indicator of population history and dispersal from Africa. *Proc Natl Acad Sci USA* 115:4128–4133.
- 13 Hansen HB, et al. (2017) Comparing ancient DNA preservation in petrous bone and tooth cementum. *PLoS One* 12:e0170940.
- 14 Mann AE, et al. (2018) Differential preservation of endogenous human and microbial DNA in dental calculus and dentin. *Sci Rep* 8:9822.
- 15 Kistler L, Ware R, Smith O, Collins M, Allaby RG (2017) A new model for ancient DNA decay based on paleogenomic meta-analysis. *Nucleic Acids Res* 45:6310–6320.
- 16 Dabney J, et al. (2013) Complete mitochondrial genome sequence of a Middle Pleistocene cave bear reconstructed from ultrashort DNA fragments. *Proc Natl Acad Sci USA* 110:15758–15763.

Chapter 3

Metagenomic Analysis of Dental calculus Reveals Differential Preservation Trends Across Smithsonian's Biological Anthropology Collections

In preparation for submission to Philosophical Transactions of the Royal Society B.

3.1 Abstract

As repositories of finite resources, museums must make decisions about destructive sampling for biomolecular research, often with little or no information about expected biomolecule preservation. Additionally, the processes by which museum collections conserve and preserve collections add another post-mortem taphonomic phase to museum samples that can influence the recovery of biomolecules. To test differential museum collection preservation due to historic preparation and curation processes, metagenomic sequencing was conducted on dental calculus (n=45) from seven Smithsonian Institution (SI) Biological Anthropology museum collections. Explorations of archival documentation for each collection revealed unique acquisition and preservation histories that influence DNA preservation. Total DNA recovery from each collection varied between 0 and 38.21 ng/mg. Microbial community reconstructions also point to a global dental calculus signature, with most variation occurring at low abundances, as well as the differential affect taphonomic processes have on DNA fragmentation within museum collections. This study sets expectations for DNA preservation in dental calculus processed in particular ways that can contribute a data-driven approach to destructive sampling decision making. Ultimately, the various collection taphonomic histories and the subsequent impact to metagenomic results underscores the need for pilot studies before large scale destructive sampling of museum collections for genomic analysis.

3.2 Introduction

Museum collections of biological anthropology (e.g., human skeletal remains and preserved soft tissues) are invaluable sources of information about past events, processes, and people. Destructive sampling requests for these collections have dramatically increased since the advent of high throughput biomolecular analyses such as genomics and proteomics (Austin et al 2019, McLean et al 2016, Thomas 1994). Often many samples are required for screening to identify those with recoverable biomolecular information to address particular research questions.

Sample screening has become even more crucial as the age and provenience of remains are now understood to be unreliable indicators of ancient biomolecule preservation (Antinick & Foran 2018, Kistler et al 2017). This necessary screening process makes biomolecular analyses expensive in costs as well as in samples, which represent irreplaceable records of human history and ever-growing possibilities for scientific discovery.

Curation and conservation practices used by museums are important factors in ancient biomolecule preservation. The treatment of biological remains for and after museum accession effectively constitutes an additional taphonomic phase, in which physical and chemical modification may degrade biomolecules such as DNA and proteins (Figure 1) (LaMotta & Schiffer 2005). Most museum collections have been preserved over decades, if not centuries, with harsh chemicals, glues, and heat treatments (Caffell et al 2001, Eklund & Thomas 2010, Elder et al 1997, Mooney et al 1982, Storch 2003). Preservation techniques applied to collections have been highly variable, and may even be collection- or individual-specific (Goldberg 1996). Efforts to optimize and maximize endogenous DNA yields have resulted in new extraction protocols and authentication methods (Dabney et al 2013, Rohland et al 2015, Rohland & Hofreiter 2007), as well as alternative skeletal elements being used such as the petrous portion (Hansen et al 2017, Pinhasi et al 2015). In many labs, taphonomic processes are often only considered to the extent of whether or not a sample provides sufficient amounts of data for analysis. Methodological optimization endeavors have only systematically accounted for the specimen age, never for intervening process that may have affected the specimen. A 2018 study concluded that preservation techniques did not impact the DNA yield and quality from mammalian bone, claw, skin, and osteocrust samples because all samples yielded DNA sequence data (McDonough et al 2018). However, varied endogenous DNA yields and fragment sizes were reported (McDonough et al 2018) and only host DNA was targeted. Arsenic, which was a common storage pesticide (Anderson 1965, Hornaday & Holland 1891, Omstein 2010), may have been applied to some samples, but the impact on recovered DNA was not discussed (McDonough et al 2018), ultimately leaving the impact of specific curation and conservation techniques on DNA preservation undiscussed. While sample success is essential for all genomic research, understanding and anticipating how biomolecules are affected by different taphonomic processes, and how these alterations can bias downstream analysis are important for project

design and analyses. As ancient and historic biological remains continue to be used for biomolecular research, more information is needed about how museum taphonomic alterations (i.e. boiling, maceration, chemical treatments, etc.) impact biomolecular data and interpretation. For accurate and precise interpretation of biomolecular data, especially of health status in materials such as dental calculus or skeletal lesions, it is useful and important to know if and how museum collection biomolecules are differentially preserved.

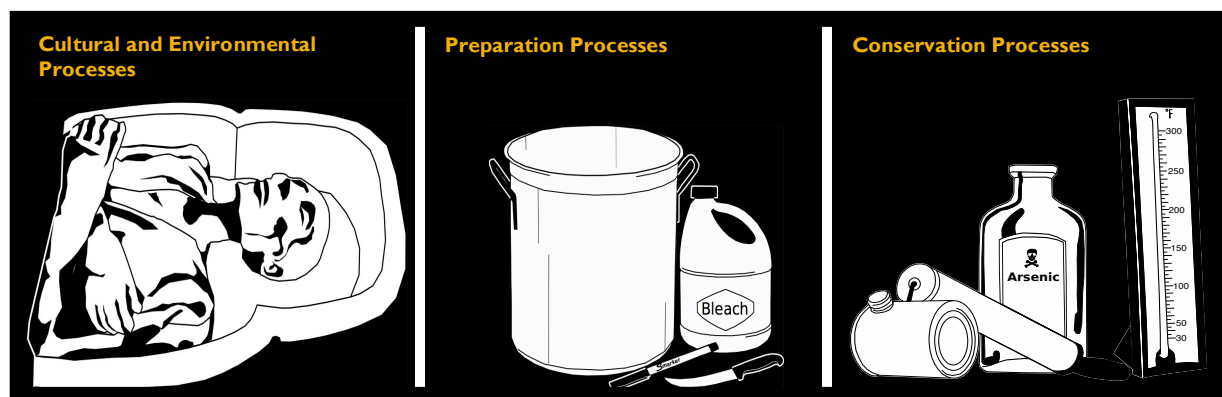


Figure 1. Taphonomic stages affecting museum collections. Cultural and environmental processes (left), which include burial practices and locality factors, are the first taphonomic phase to impact biomolecule preservation. Once brought to museums, processes (middle) to prepare the remains for storage and display are a second taphonomic phase. Lastly, conservation processes to maintain human remains over time can be considered a third taphonomic phase (right). Because detecting and characterizing molecular damage patterns is an integral part of ancient biomolecule authentication, information on how samples were collected, washed, preserved, and maintained is valuable for generating and interpreting high throughput biomolecular data.

In this study, we compared ancient biomolecule preservation in seven collections of biological anthropology at the Smithsonian’s National Museum of Natural History (NMNH) (Figure 2). Representing a wide geographic and temporal range of skeletal populations, these collections were selected for their variable taphonomic histories, much of which was reconstructed with museum archives and associated documentation. We utilized human dental calculus to explore molecular taphonomy, as dental calculus from ancient contexts has become an increasingly frequent sampling request due to its applications for oral microbiome research (Ozga et al 2016, Radini et al 2019, Warinner et al 2014b, Warinner et al 2015b). Oral microbiome communities have also been shown to be more stable than other human microbiome sites and across populations, making it possible to attribute metagenomic differences between collections to taphonomic impacts rather than expected population or human variation (Costello et al 2009,

Dewhirst et al 2010, Zhou et al 2013). Furthermore, high total DNA yields (Mann et al 2018) also make dental calculus a good source of DNA for testing collections for biomolecular, specifically ancient DNA (aDNA), preservation before destructively sampling other tissues. Our assessment of biomolecular preservation in dental calculus, a biofilm, also allowed other human tissues (bone and teeth) to be left intact.

We conducted metagenomic analysis on 45 dental calculus samples from 5 or 6 individuals from each collection (n=45) in order to assess aDNA preservation. Results were compared with published archaeological dental calculus and paired roots (Mann et al 2018), and modern dental calculus data (Velsko et al 2019) focusing on aDNA recovery, microbial community composition, and environmental contamination. This study provides the first metagenomic dataset of intra- and inter-collection preservation comparisons, providing insight into the biomolecular success of globally representative museum collections.

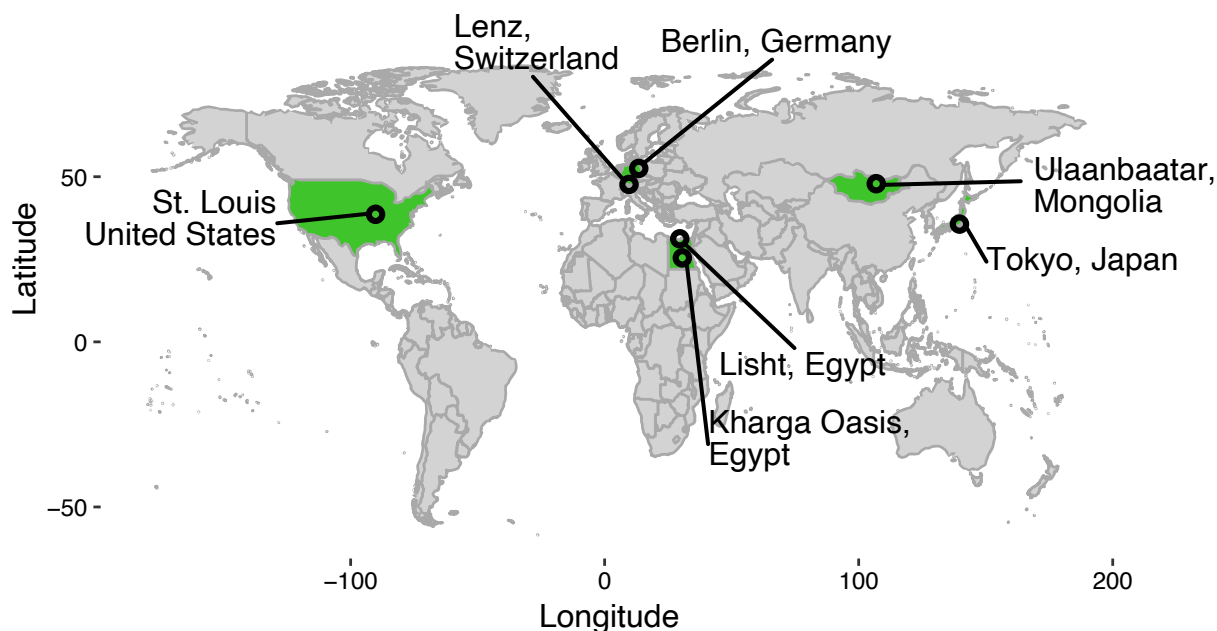


Figure 2. Geographic locations of archaeological dental calculus included in this study. Site locations (as specific as possible from collection records) and countries are denoted by circles and green shading, respectively.

3.3 Materials and Methods

3.3.1 Collection Archival Research

Our sample consisted of 45 dental calculus samples from 12 and 24th dynasty Egypt (n=10), 19th century Switzerland (n=6) and Japan AMM (Army Medical Museum; n= 5), and turn of the 20th century Japan IUCMT (Imperial University College of Medicine in Tokyo; n=6), Mongolia (n= 6), Germany (n= 6), and a random subsample from the Terry Collection (n=6) (Austin et al *in prep*) held at NMNH. These collections were accessioned into the Smithsonian Institution over an approximate 50-year period (1904-1950), to control for various environments, burial contexts, and preservation techniques impacting molecular preservation quality (Figure 2). In the time since accession, general departmental and pest control management strategies are assumed to have been applied across these collections.

Archival research was conducted on both the collections and individuals in this study to identify collection acquisition information regarding the contributing collector(s) (those that gathered or created the skeletal collections), location and cultures of origin, the institutions endorsing the assemblage, and documentation regarding applied preservation methods. Archival records were accessed at the National Anthropological Archives (NAA), in SI Accession Records, and from individual-specific SI catalog cards. Specific attention was paid to information that may have impacted aDNA preservation, such as environmental information, chemical treatments, and shipment conditions. Across Egypt, Mongolia, AMM, and Switzerland collections, we expected environmental and cultural processes to impact aDNA preservation due to originating from burial contexts. Limited cultural and environmental processes were expected for the Germany, IUCMT, and Terry collections because they were prepared for cadaver dissections and autopsies.

3.3.2 DNA Extraction

Ancient DNA extraction followed the protocol described in Ziesemer et al (2015) with only slight modifications. No UV irradiation was applied to samples before washing and digestion steps. Reaction blanks were processed in parallel to samples. Succinctly, between 1- 22 mg of dental calculus (6.1 mg on average) was washed with 0.5 M EDTA for 15mins before decalcification in an EDTA (Sigma) and 10% proteinase K (Qiagen) solution at room temperature for 72 hours. DNA was isolated via a modified Qiagen MinElute silica-column

based purification into a final 60 µl elution volume. Extraction elutions and blanks were quantified using a Qubit High Sensitivity fluorometer (Life Technologies).

3.3.3 Illumina Library Preparation

Across the global SI collections (excluding the Terry Collection), extracted DNA was built into dual indexed Illumina shotgun libraries following a modified single-tube preparation developed by Carøe et al (2017). Reaction blanks were processed in parallel to samples. Blunt-end repair was performed using 60 µl reactions with 50 µl of DNA extract for each sample, with incubations at 30 min at 20 C, immediately followed by 30 mins at 65 C. In the same tube, adapter ligation was completed in 77 µl reactions and incubated for 30 mins at 20 C and 10 mins at 65 C. An adapter fill-in step was then completed in a 91 µl reaction, followed by a 20 min incubation at 65 C and 20 mins at 80 C. All samples were then purified using Qiagen MinElute silica purification spin columns following Ozga et al (2018). Real-time quantitative PCR assay (qPCR, Lightcycler 480 Roche) targeting the IS7 and IS8 sequences in the universal Illumina adapters was used to quantify the libraries. Dual indices were added during sample amplification PCR in triplicate, 25 µl reactions using 4 µl template (12.5 µl of a 2x KAPA HiFi +Uracil master mix, 6 µl H₂O, 1 µl BSA (2.5 mg/ml), and 1 µl of each forward and reverse index (10 µl µM)). Libraries were amplified for 5 min at 98 C followed by 14–25 cycles of 20 seconds at 98 C, 15 seconds at 60 C, and 30 seconds at 72 C, followed by a final elongation step for 1 minute at 72 C. Agarose gel visualization was used to confirm amplification for each sample replicate. Amplified library replicates were then pooled and purified using SpeedBeeds and eluted in 30 µl EB (Rohland & Reich 2012). Following Qubit High Sensitivity fluorometer (Life Technologies) quantification, libraries were then diluted to 1 ng/µl for fragment visualization and quantification on a Fragment Analyzer. Samples were then pooled in equimolar ratios for a final purification on a PippinPrep with a 150-500 bp size range before being sent for 2x150 bp sequencing on an Illumina NovaSeq at the Oklahoma Medical Research Foundation (OMRF) in Oklahoma City

3.3.4 Computational Methods

Metagenomic sequences were processed with Adapter Removal v2 to remove Illumina adapters, merge paired-end sequences, and quality filter for minimum base quality and read length (20 and 25 respectively) (Schubert et al 2016). Cutadapt (Martin 2011) was applied to trimmed R1 files;

trimmed R2 were not included in analyses due to artificially generated sequences passing quality filtering. Analysis ready reads were then mapped to the Greengenes 16S rRNA database (DeSantis et al 2006) using the Burrows-Wheeler Aligner v0.7.12 (BWA) (Li & Durbin 2009). Samtools v1.5 (Li et al 2009) filtered unmapped and low quality reads, and removed PCR duplicates before converting files to fasta format. MapDamage v2.0.8-2 (Ginolhac et al 2011, Jónsson et al 2013) was employed to estimate the nick frequency, cytosine deamination rates, and read length patterns in order to authenticate quality filtered reads as ancient when BWA mapped to *Tannerella forsythia* (assembly ASM23821v1), a pathogenic oral microbe associated with periodontitis (Socransky & Haffajee 2005). To explore the microbial communities in dental calculus, additional comparative data sets from global archaeological calculus and paired root samples (Mann et al 2018), modern dental calculus samples from Spain (Velsko et al 2019), and museum dental calculus from the Terry Collection (Austin et al *in prep*) were included in all analyses. Only the trimmed and cutadapt R1, and trimmed and merged SI collection and comparative datasets files were included. Contributing bacterial sources were identified with SourceTracker v1.0.1 (Caporaso et al 2010) through UCLUST taxonomy assignment and visualized using unweighted unifracc distance matrix PCoA plots. MetaPhlAn2 (Truong et al 2015), previously identified for its accuracy in reconstructing taxonomy in simulated ancient bacterial communities (Velsko et al 2018), was used to generate metagenomic profiles. Sparse partial least squares-discriminant analysis (sPLS-DA) was performed to compare microbial communities by metadata categories (collection type, collection treatments) at the species level (as determined by MetaPhlAn2) using the mixOmics R package (Rohart et al 2017). Given classification categories for each sample, sPLS-DA looks for the maximum differences between categories, i.e. how sample categories can be discriminated. sPLS-DA analysis and methods follow Velsko et al. (2019). Data was visualized using R base packages and ggplot2 (R Core Team 2016, Wickham 2011).

3.4 Results

3.4.1 Collection Histories and Treatment Summaries

Like most museums, collections in this study were acquired and preserved over many years and in various ways (Vogel 2019). Until they were transported to the off-site Museum Support Center (Wilcox 1995), the physical anthropology collections were stored in wooden cabinets in

the National Museum of Natural History's attic and anthropology department halls (Goldberg 1996). Across anthropology collections, pests were managed through the application of fumigants and setting traps (Goldberg 1996). The building itself was constructed in 1910 so climate and temperature control was limited until 1960, when air conditioning was installed (Bellizzi 2014); these conditions exposed remains to the hot and humid summers as well as the cold and dry winters of Washington D.C. It is reasonable to assume that these conditions may have affected the collection biomolecules (Goldberg 1996).

Here we provide summaries of the documented histories of each collection from accession records, curator journals, and contextualizing literature sources (Table 1). These summaries encompass the broad taphonomic (LaMotta & Schiffer 2005) processes and circumstances that are fundamental to these collections. Critically, because they are museum collections, assessed remains have longer and more complex taphonomic histories.

Table 1. Summary of documented collection histories based on archival research on the accession, catalog cards, curator notes/ journal entries, and literature sources. Sorted by descending SI acquisition year.

Collection (n= 45)	Population Time Period	Collection Year	Taphonomic Processes		
			Cultural and Environmental	Collection Preparation	Collection Conservation
Japan: AMM (n= 5)	~1870 CE	~1870	Unspecified	Opportunistic acquisition by military medical personnel	Chemical fumigants, pesticides, seasonal changes
Egypt: Lisht (n= 5)	1990–1770 BCE [1]	1909	Mummification and pit burials	Minimal/ delicate handling	
Egypt: Kharga Oasis (n= 5)	500-600 CE [1]	1909	Mummification and bitumen	Heavy handling and fragmentation	
Germany (n=6)	1910 and 1911 CE	1910 and 1911	Autopsy/dissection	Autopsy and formalin injection; museum maceration	
Mongolia (n=6)	~1913 CE	1913	Surface burials	Water and potassium permanganate boiling	
Japan: IUCMT (n= 6)	~1920s CE	1923	Autopsy/dissection	Autopsy	
Switzerland (n= 6)	~1850 CE	1950	Ossuary internment	Private collection handling/ undocumented	
Terry Collection (n= 6)	1828-1967 [2]	1927-1967	Autopsy/dissection	Boiling and exposure to benzene vapors	

[1] (Irish 2006)

[2] (Hunt & Albanese 2005)

Japan (Army Medical Museum)

In the 19th and early 20th centuries, the Smithsonian actively exchanged collections with the Army Medical Museum (AMM), now known as the National Museum of Health and Medicine. AMM acquisition of remains was initiated by the 1862 Surgeon General proclamation, Circular No. 2, which solicited medical officers to collect and send “all specimens of morbid anatomy, surgical and foreign bodies removed, and such other matters as may prove of interest in the study of military medicine or surgery (Henry 1964).” From this directive, pathologies and traumas resulting from the Civil War and Westward Expansion, as well as overseas military operations, were amassed for medical study. Remains were also gathered to conduct racialized and classist comparisons (Juzda 2009, Redman 2012). For an extensive history of the AMM see (Henry 1964, Leikind 1954, Redman 2012). Individuals in this study were a part of larger collection exchanges between the AMM and the NMNH, acquired in either 1904 or 1910, but were collected in Japan by US military medical personnel around 1870. While documentation was required to accompany remains sent to the AMM, information is limited. Collectors, Jones M.D., Junghaus PhD., and Kidder, Surgeon had some metadata noted, ranging from the collection location, age and sex of the individual, and the person’s socioeconomic class (see Supplementary Table 1). Upon collection, conservation and preservation treatments applied to the analyzed individuals (n=5) is unknown, however is expected to be collector-specific, mirroring the collector-specific notes that are available.

Egypt

As part of the Egypt Expedition of the Metropolitan Museum of Art (New York City, NY), Aleš Hrdlička (NMNH Physical Anthropology curator, 1903-1942) worked under the direction of Mr. Albert M. Lythgoe at the Pyramids of Lisht and Mr. Winlock at the Kharga Oasis in 1908 and 1909 (1908, 1909a). During both field seasons, Hrdlička conducted field analyses of excavated human remains and hand-selected remains that would become a part of the NMNH collections. Across his international travels, Hrdlička would document his experiences in a personal journal. In journal remarks made during the Egypt Expedition(s), treatment of the remains is described; Lisht remains were commented to be “in [an] exceedingly fragile state” (December-March 1908-1909) at the time of their excavation, whereas sediment encrusted mummies were excavated at the Kharga Oasis (December-March 1908-1909). From Hrdlička’s personal notes, it seems as

though the Lisht remains were more delicately handled and packed for shipment, whereas many individuals from Kharga Oasis were treated more roughly, sometimes even with a saw during maceration efforts (December-March 1908-1909).

Germany

For the individuals in the Germany collection, Hrdlička was contacted by and entered a trade for anatomical specimens with Professor David von Hanseemann, a German pathologist and noted oncologist (1909b). Between 1909 and 1911, von Hanseemann provided the brains and crania of unclaimed individuals from the Rudolf-Virchow Hospital in Berlin in exchange for “Indian” crania from the NMNH (1909b). Per Hrdlička’s suggestion, intact brains and skulls were injected in situ with a formalin solution and sent whole to ensure the safe, undamaged arrival of the brains (1909b). The exact injection and shipment method (e.g. if additional preservatives, such as salts or charcoal, were also packed) is unknown, however, general anatomical dissection/autopsy conditions are assumed due to the individuals being prepared and originating from a hospital setting. Notes written by T. Dale Stewart (Physical Anthropology Curator, NMNH Director, and SI Secretary, 1942-1971) identified eight crania from the Germany collection that “had been placed in maceration about 1917 and was now in a dry state (Stewart 1942)”. There is no indication of which individuals Stewart is referring to, however, his note indicates the potential of the analyzed individuals to have been subjected to a prolonged period of dry heat.

Mongolia

In 1912 Hrdlička visited Mongolia to obtain remains for a collection. Staying just outside of the capital, Ulaanbaatar, at an old Russian military settlement (July-September 1912), Hrdlička hired three, local “Kossaks” (Cossacks) to gather surface remains in the surrounding area (1913). Skulls and body fragments were the primary elements that were scavenged due to cultural practice of sky burials (July-September 1912, MaMing et al 2016), which enabled immediate animal scavenging of remains. Notably, collected remains were boiled in water, manually cleaned, and soaked in a strong solution of “permanganate of potash” (potassium permanganate) due to the smell before shipment to the SI in wood boxes filled with charcoal (July-September 1912).

Japan (Imperial University College of Medicine in Tokyo)

By 1915 the NMNH skeletal collections were one of the largest and most extensively studied collections in the world. Trying to achieve complete global representation in the collections, Hrdlička contacted Professor Yoshiaki Koganei of the Imperial University College of Medicine, in Tokyo (IUCMT), Japan to ask advice on obtaining and/or arrange an exchange of dissected skeletal remains and associated brains (1923). Correspondences between Hrdlička and Koganei are not present between the initial contact letter in 1915 and the shipment invoice dated 1922, with communications likely stalling because of World War I and did not resume until Hrdlička traveled to Japan to personally arrange an exchange (1923). In his initial letter of contact, Hrdlička requests individuals' heads and brains to be injected with 5-10% formalin solution *in situ* (1923). As only "skeletons with skulls" are noted in accession records, it appears brains were not shipped in this exchange and had presumably (as there is no other indication) been skeletonized before shipment from Japan (1923). At this time in Japan, bodies for dissection were generally available through government sanctions directing bodies from prisons, poor houses, elderly homes, and university hospitals for medical school use (Cowdry 1920). Prior to dissection bodies were noted to be stored in "well-constructed tanks" in sanitary conditions, with refrigeration being reserved for thin sections (Cowdry 1920). There is no longer any record, in accession records or in Hrdlička's travel notes, of the specific dissection and skeletonization processes that were applied to the remains. However, with the establishment of the Tokyo Anthropological Society in 1884 and Koganei's contribution to the physical anthropology branch and being active in the Tokyo Imperial University Anthropological Institute (Cowdry 1920), a standardized, anthropologically-minded method is thought to have been applied to the individuals.

Switzerland

When moving from the U.S. to Switzerland in 1950, Dr. Adolph H. Schultz, a Swiss anthropologist known for his work with non-human primates, entered an exchange with T.D. Stewart of the Department of Anthropology. From Dr. Schultz's personal collection, the SI obtained skulls originating from the Platta and Lenz Swiss villages in the Canton Grisons estimated to originate during ~1850 (1950a). According to the collection history card, remains would have come from "neat stacks in ossuaries behind the village churches" (1950b). Prior to

and upon arrival at the SI, no documented conservation techniques are known to have been applied, effectively preserving the original, ossuary conditions they were kept in; although, because they were a part of a teaching professor 's private collection before being acquired by the SI, remains were assumed to have been handled and/or studied extensively.

Terry Collection

Assembled (1910-1967) the Terry Collection was cataloged, macerated, and stored under consistent and uniform protocols at the Washington University School of Medicine (WUSM) until its transfer to the Smithsonian Institution. The collection is known as one of the most complete and well documented anatomical collections in the world (Hunt & Albanese 2005). Information on individuals' true age, sex, ancestry, and cause of death are known from associated death records. Individuals in this collection were never buried, however, upon being received at the WUSM, individuals underwent manual maceration, boiling, drying, light degreasing by exposure to benzene vapors followed by a coating of a glue and formalin solution (for an in-depth account of the Terry Collection protocols and history see (Hunt & Albanese 2005, Trotter 1981)).

3.4.2 Metagenomic Results

DNA Recovery

Normalized DNA concentrations (ng/mg) varied across samples, collectors, and collections (Figure 3). A significant difference is observed when comparing results by who collected them (ANOVA F value= 14.63, $p < 0.05$) and between collections population/treatment method (ANOVA F value= 20.53, $p < 0.05$). However, Tukey HSD analysis reveals IUCMT/ Kogani to be the only collection/ collector significantly different from all others.

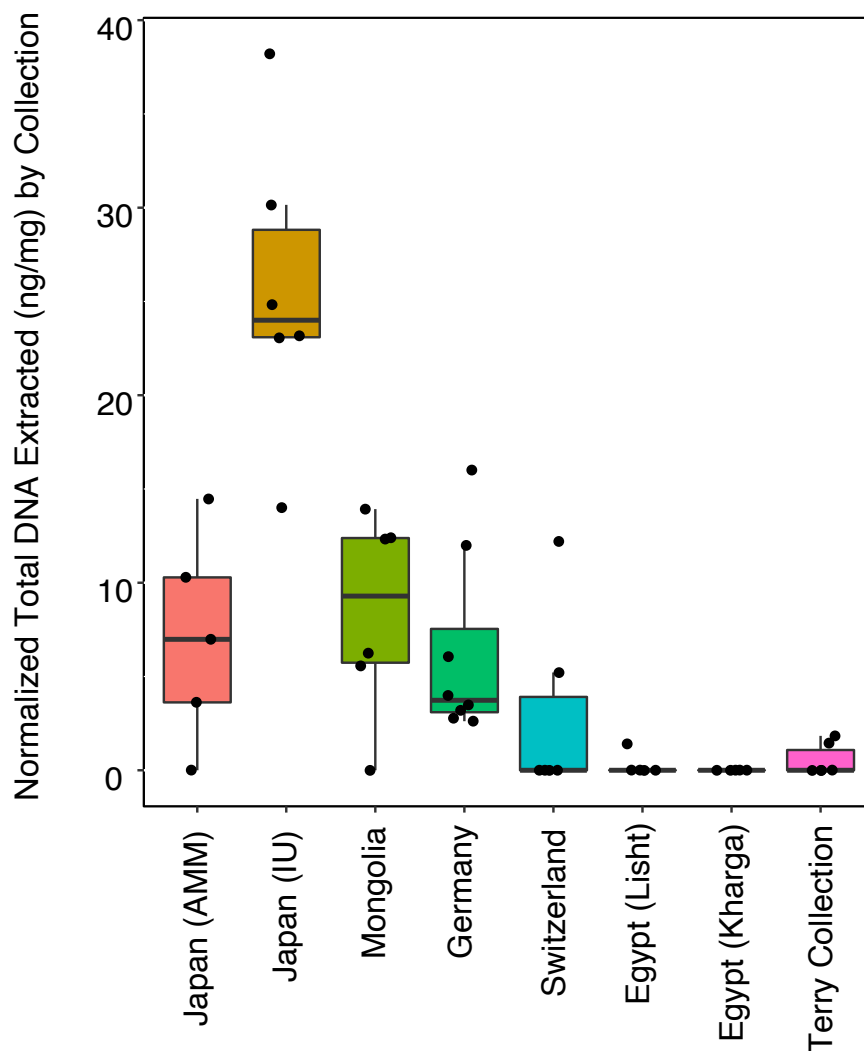


Figure 3. Normalized DNA (ng/mg) for each collection as measured by a Qubit fluorometer High Sensitivity Assay. Normalized DNA (ng/mg) yield of DNA extracts revealed significant differences among collecting individuals (ANOVA F value= 14.63, $p < 0.05$) and location/treatment method (ANOVA F value= 20.53, $p = < 0.05$). The Japan IUCMT collection showed significantly higher DNA concentration than all other collections.

Several samples ($n=6$) were removed from further analysis due to low quantity of reads following sequencing (<10 million) and quality filtering (<9 million) (Supplementary Table 2). To account for varying sampling depths, samples were rarefied to 2000 reads, after which three Egyptian samples were further removed from analyses. The removal of the three Egyptian samples is supported with MetaPhlan2 microbial classification assessment, where over 50% of identified taxa belonged to the order Burkholderiales, the only taxa identified in SI laboratory negatives (Supplementary Figure 1). Due to $>50\%$ Actinomycetales being identified in Terry

negatives, two more Egyptian samples, as well as one Terry sample, were removed from further analysis.

Microbial Community Authentication

To explore preservation of the oral microbial community in dental calculus, the proportion of each sample attributed to unique microbial environments was estimated through 16S rRNA SourceTracker analysis on remaining samples (n=39). Most samples show predominantly intact ancient calculus microbial signatures (Figure 4A). When assessing the differences between samples and collections with unweighted unifrac distances, all SI dental calculus and comparative oral sources (i.e. archaeological dental calculus and tooth roots, and modern dental calculus) (Mann et al 2018, Velsko et al 2019) cluster together and away from soil and gut microbial samples (Figure 4B). Unweighted unifrac distances of only oral sources (Supplementary Figure 2) generally distinguish between archaeological tooth roots and dental calculus samples. The Japan AMM and IUCMT collections, and the Switzerland collections show the most within collection clustering among oral sources. Skin source samples also cluster with oral samples indicating the microbial community similarities between these environments; it is thus, unsurprising that large proportions of skin are identified in some of the assessed oral samples (Figure 4A). Egyptian (average skin proportion= 39%), Mongolian (average skin proportion= 16 %), and Terry Collection individuals (average skin proportion= 22%) have the highest contributions attributed to skin sources.

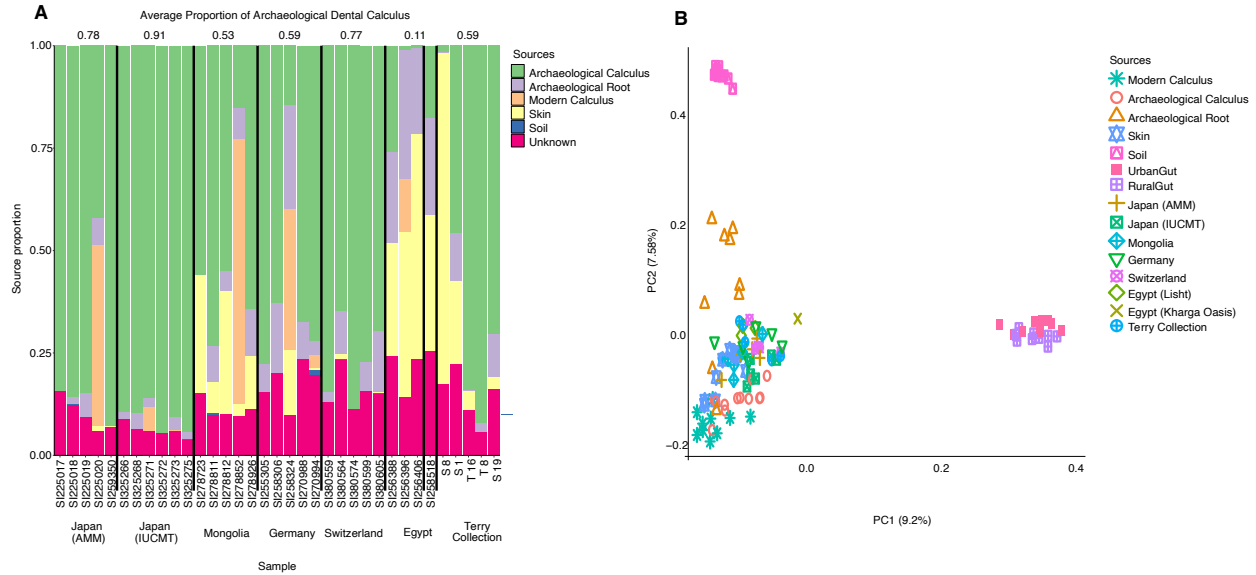


Figure 4. SourceTracker analysis (A) and principal coordinate analysis (PCA) (B) of unweighted unifracs distances for estimated 16S rRNA gene identification across museum collections. Excluding the Egyptian samples, SourceTracker analysis reports largely intact ancient dental calculus signatures. Oral samples (dental calculus and roots) cluster away from soil and gut microbial profiles, but overlap with skin communities.

Reads mapped to *T. forsythia* showed varying read lengths (ranging between 35 and 70bp), but with characteristic damage patterns of elevated C to T transitions at terminal ends (Supplementary Figure 3). Smoothness of plot lines relate to the number of reads contributing to the analysis. Terminal end damage is restricted to the first three base pairs for the Terry sample (T16) due to the application of partial uracil-DNA glycosylase (UDG) treatment to Terry samples (Austin et al *in prep*).

To investigate the microbial diversity within each sample, alpha diversity was estimated as observed OTUs at the species level with both 16s rRNA and Metaphlan2 datasets. QIIME 16S rRNA analysis identified many more OTUs than MetaPhlan2 (Figure 6). QIIME 16S rRNA OTU identification and fragment lengths show an inverse relationship (Figure 5A and B), with more identified species OTUs for museum collections even though fragment lengths are much shorter. Additionally, 16S rRNA fragment length distributions are significantly different across comparative datasets and museum collections (ANOVA F value= 60,518; $p < 0.05$). The only nonsignificant difference of fragment length distributions was seen between the Mongolian and AMM museum collections ($p = 0.07$, Tukey's HSD).

MetaPhlAn2 identification reflects an overall lower number of species (Figure 5C), with more variation between samples within museum collections. Interestingly, fragment lengths of MetaPhlAn2 communities have longer average fragment lengths than in the communities reconstructed by Qiime (Figure 5D). In general, archaeological and modern samples have higher numbers of identified species compared to the museum specimens. Like the 16S rRNA fragment distributions, MetaPhlAn2 fragment length distributions are also significantly different across comparative datasets and museum collections (ANOVA F value= 333,013; $p < 0.05$).

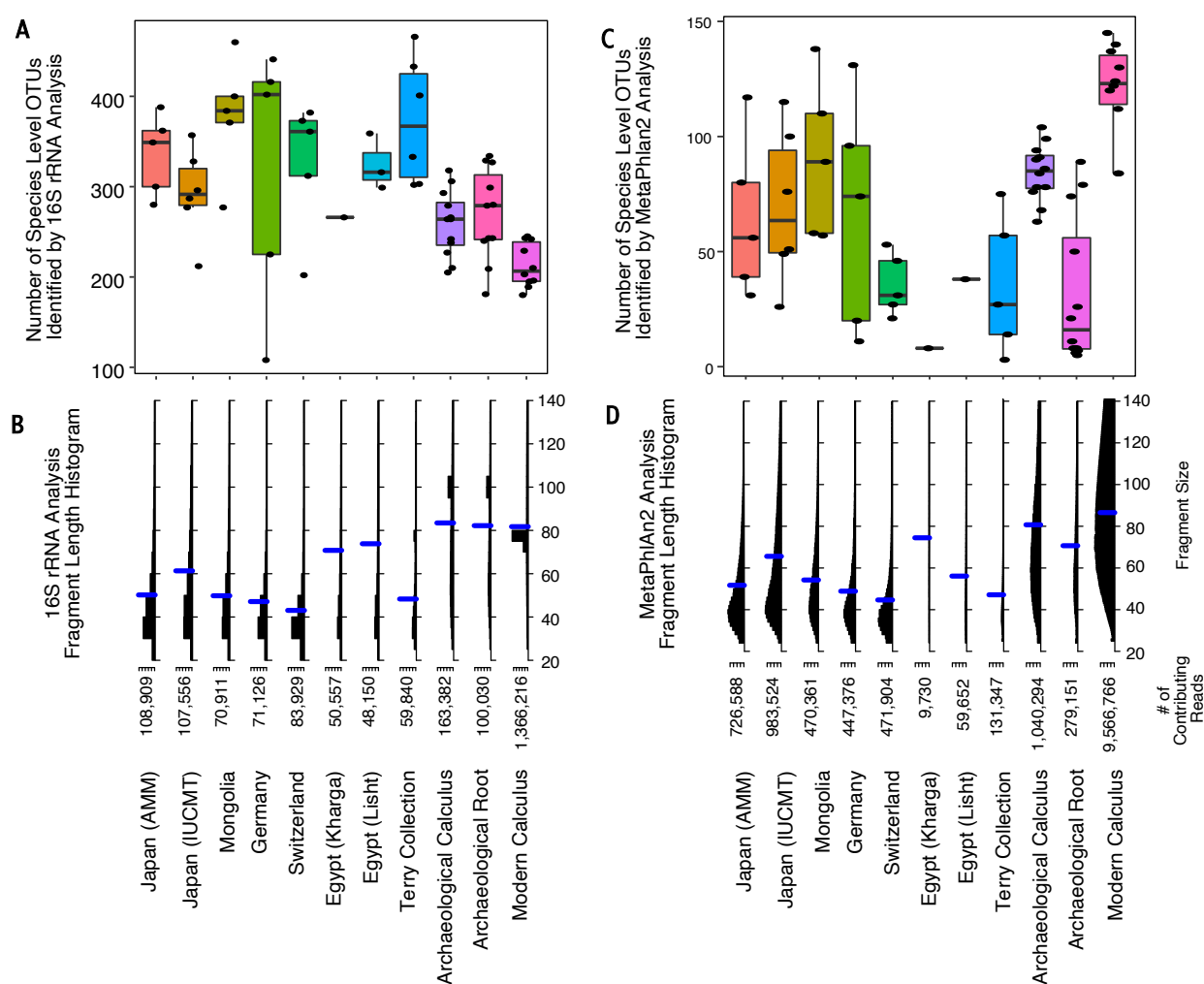


Figure 5. Number of species level OTUs identified by collection between QIIME 16S rRNA (A) and MetaPhlAn2 analyses (C). Fragment length distributions across collection samples' reads mapping to the 16S rRNA gene (B), and in MetaPhlAn2 (D). 16S rRNA histograms (B) are scaled from 0 to 50,000 for archaeological and museum collections, whereas modern dental calculus scaled from 0 to 400,000 due to the increased number of recovered molecules.

MetaPhlAn2 histograms (**D**) are scaled from 0 to 40,000 for archaeological and museum datasets, with modern calculus scaled from 0 to 80,000. Blue lines represent average fragment lengths for the collection.

Dental Calculus Microbial Community Comparisons

The role of collection type (i.e. archaeological, museum, and modern) and collection treatments in differentiating microbial communities was investigated using PCA and sPLS-DA on the Metaphlan2 community reconstructions (Figure 6). Differences among collection types and treatments would suggest that there are collection/treatment specific shared properties and conditions that are informative for biomolecular curation. Clustering of samples with laboratory negatives in the PCA reflects the depleted number of OTUs identified in samples and from different sample types (root vs calculus) rather than accounting for the types of microbes identified. However, sPLS-DA analysis differentiated museum and modern dental calculus samples, with very little overlap between 95% confidence interval ellipses (Figure 6B). Archaeological samples (root and calculus) distinctly cluster between modern and museum samples (Figure 6B). The balanced error rate (BER) was high (alpha greater than 0.05), which may be due to the low sample sizes of groups and/or high variability between samples within groups (Supplementary Table 3A). Overall, sPLS-DA analysis suggests that the microbial communities of archaeological samples are interstitial to modern and museum samples but are not distinguishable from either group. *Methanobrevibacter unclassified* is the most discriminating species between museum and modern dental calculus samples, where *Capnocytophaga gingivalis* distinguishes modern samples (Supplementary Figure 4).

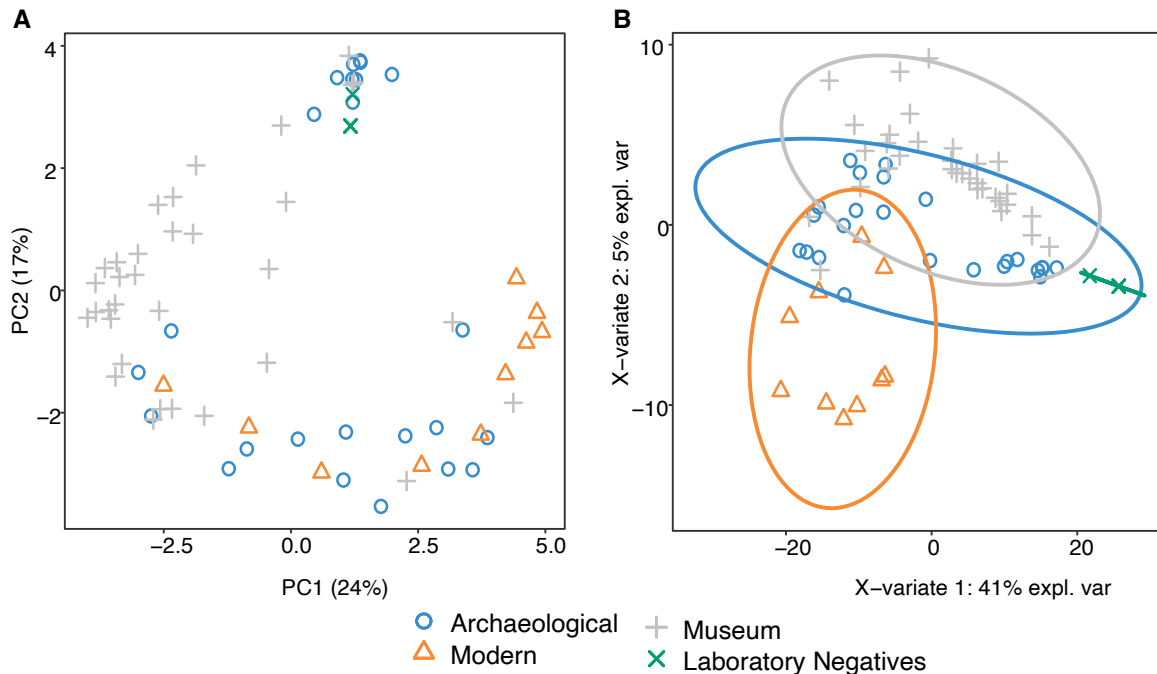


Figure 6. Microbial community differences between modern, archaeological, and museum dental calculus. Principal component analysis (PCA) (A) generated with the singular value decomposition (SVD) algorithm (Rohart et al 2017) shows archaeological and museum does not cluster distinctly. Modern dental calculus is seen in the lower portion of the plot but does not tightly cluster together. Several samples cluster with negative controls. Sparse partial least squares-discriminant analysis (sPLS-DA) (B) show microbial community differences between modern and museum dental calculus samples, with archaeological samples in between them. SI collection negative controls appear similar to archaeological and museum samples in the PCA analysis, however, are found within the 95% confidence interval ellipse of archaeological samples. Ellipses indicate 95% confidence intervals.

When samples are defined by collection rather than type (Figure 7A), two collections (Switzerland and IUCMT) and archaeological root samples appear to have internally consistent microbial communities while other collections show wider variation. Several Terry Collection samples cluster with archaeological root samples and laboratory negatives reflecting the depleted microbial communities recovered from the specific dental calculus samples and from the dental root sample type (Figure 7A). Modern dental calculus microbial profiles are relatively distinct from the museum and archaeological samples. As sPLS-DA (Figure 7B) generally discriminates the modern comparative dental calculus (Velsko et al 2019) from the museum and archaeological calculus, the analysis was also conducted without the modern calculus samples (Figure 7C). With the exception of one Terry Collection sample, collections and archaeological samples

heavily overlap, but more variation is seen among the Terry Collection, Germany, Mongolia, and AMM collections.

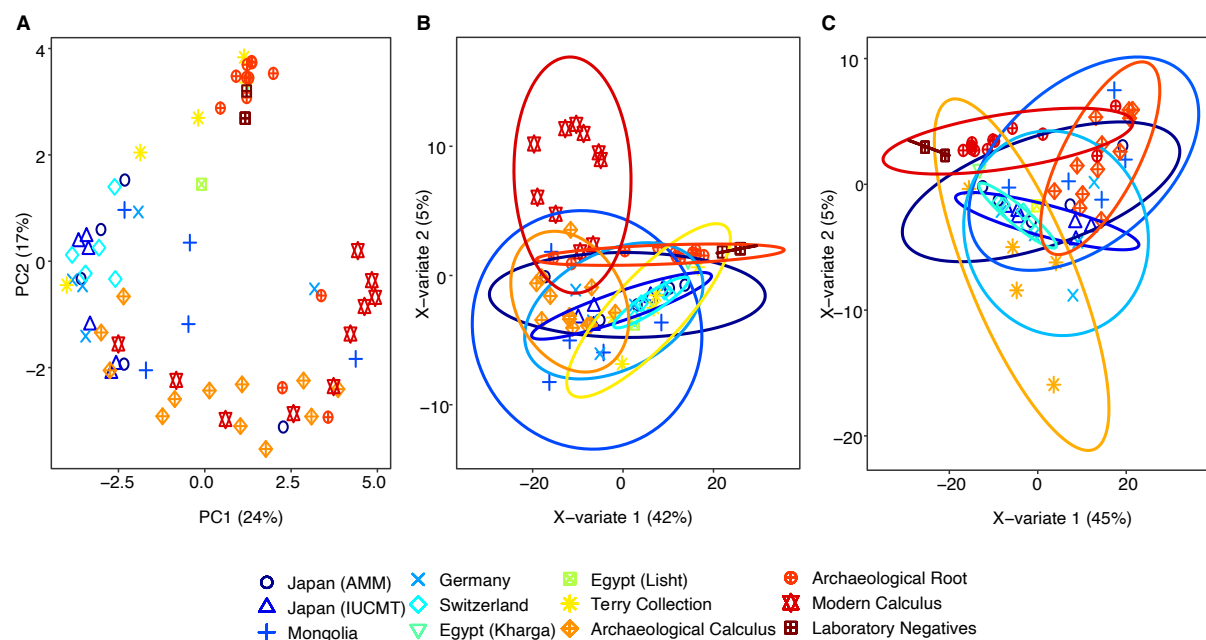


Figure 7. Microbial community differences between collection treatments. Principal component analysis (PCA) (A) shows slight grouping by collection. Sparse partial least squares-discriminant analysis (sPLS-DA) (B) indicate modern dental calculus microbial profiles to be more distinct from archaeological and museum dental calculus, which overlap in their ellipses (95% confidence intervals). Museum collection sPLS-DA species discrimination (C) was only minorly improved when modern dental calculus was removed.

When the remaining Kharga Oasis (n=1) and Lisht (n=1) samples are combined into a single category (“Egypt”) sPLS-DA discriminates modern dental calculus from archaeological and SI collection samples even more (Supplementary Figure 5A). However, the 95% confidence interval of modern samples continues to overlap with other collections. When modern comparative samples are removed (Supplementary Figure 5B) SI collections and archaeological samples are indistinguishable, with all confidence intervals overlapping.

When Egyptian samples are completely removed from analysis due to small sample size (n=2), sPLS-DA identifies less distinct taxa among museum dental calculus, and archaeological dental calculus and root samples but generally differentiates modern samples from SI and archaeological samples (Figure 8B). SPLS-DA continues to discriminate modern dental calculus

from all other samples. Even with the removal of modern comparative samples there are few distinguishing species among archaeological and SI collections microbial communities (Figure 8D). However, when laboratory negatives are also removed sPLS-DA distinguishes archaeological roots from SI and archaeological dental calculus (Figure 8E).

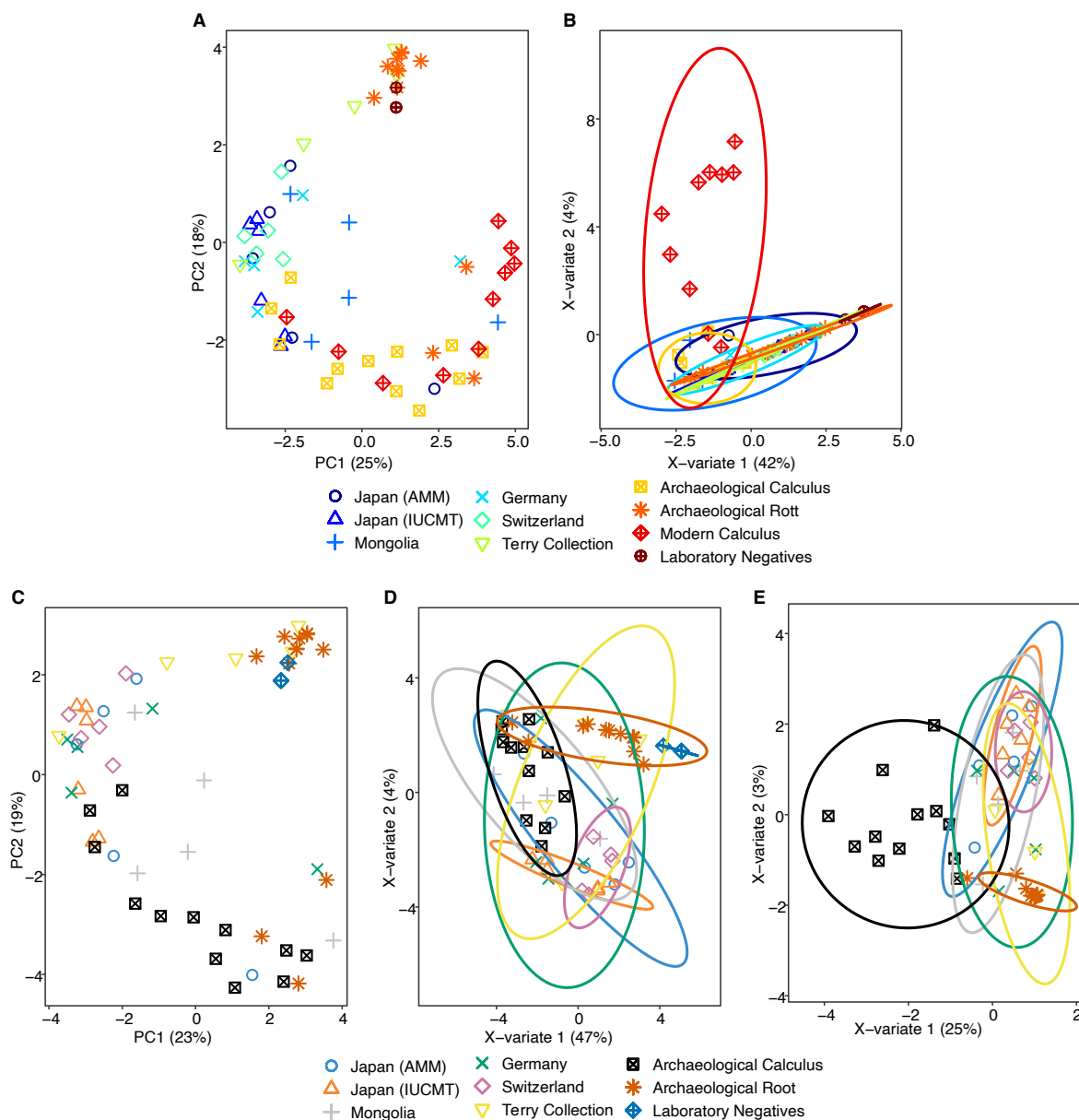


Figure 8. Microbial community differences between collection treatments without Egyptian samples (A and B), without Egyptian and modern samples (C and D), and without Egyptian, modern, and laboratory negatives (E). Principal component analyses (PCA) (A and C) mirror each other with slight grouping by collection. Sparse partial least squares-discriminant analysis (sPLS-DA) without Egyptian samples (B) show modern dental calculus microbial profiles to be more distinct from archaeological and museum samples, which overlap in their ellipses (95%

confidence intervals). Only the AMM and Mongolian collections, and archaeological dental calculus show any variation. When Egyptian and modern samples are removed sPLS-DA does not effectively differentiate among any of museum and archaeological samples (**D**). Archaeological dental calculus are distinguished from museum dental calculus and archaeological root samples using sPLS-DA, when Egyptian, modern, and laboratory negatives are excluded.

sPLS-DA identified *Streptococcus mutans* in the archaeological root microbial communities as the most distinguishing species among all collections and datasets (Figure 9A). The Kharga Oasis sample were also identified as contributing to collection discrimination. Removal of the Egyptian samples changes results to instead identify the AMM collection as contributing distinguishing species (Figure 9B). *Campylobacter showae* (a microbe associated with gingivitis and periodontitis (Lugonja et al 2016) is identified to be the most distinguishing species in the archaeological dental calculus samples when Egyptian and modern samples are excluded (Figure 9C). MetaPIAn2 analysis confirms *C. showae* to have the highest abundance in the archaeological dental calculus samples. High balanced error rate (BER) (alpha greater than 0.05; Supplementary Table 3B) observed for the sPLS-DA assessment also warrants caution for biologically interpreting the identified distinguishing species among data groups due to the small collection sample sizes and sample variability.

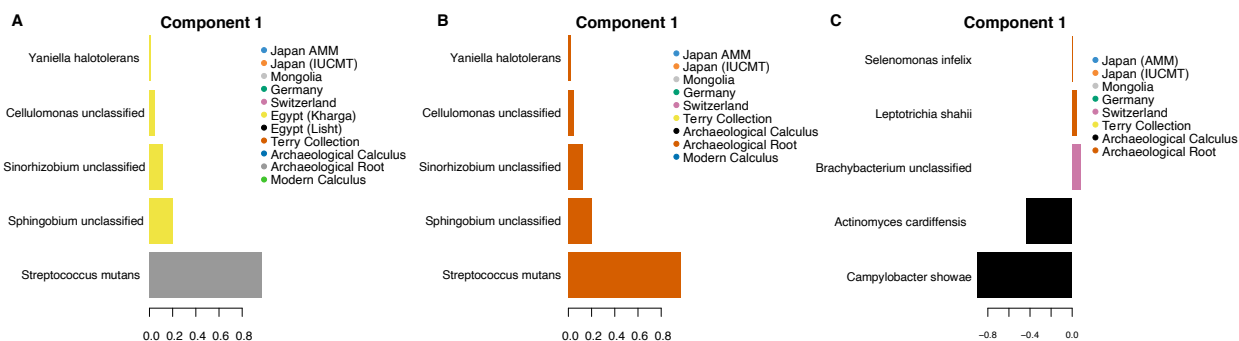


Figure 9. Taxa contributing to discriminating museum, archaeological, and modern samples on component 1 of the sPLS-DA. Archaeological root samples are consistently identified as contributing the most distinguishing species when Egyptian samples are included (**A**) compared to when they are removed (**B**). When modern and Egyptian samples are removed, archaeological calculus species contribute the most to distinguishing collections, followed to a much lesser extent by archaeological root and Switzerland species (**C**).

3.5 Discussion

Documenting Collection Histories

Archival research revealed unique, and sometimes, unexpected collection assemblage and treatment histories of the individuals in this study. While individuals in this study have been contextualized within their historic circumstances, the acquisition and trade of the analyzed human remains were under conditions that would be considered unethical today. Museum collectors, such as Hrdlička, were operating under established protocols that institutionalized and perpetuated colonial and racist ideals (Halperin 2007, Morris 1987, Redman 2012, Vogel 2019). Historic collecting methods have been discussed elsewhere (Halperin 2007, Humphrey 1973, Muller et al 2017, Watkins 2018), but often bodies, burials or remains of people of color, poor socioeconomic status, or without sociopolitical power were targeted for collection. This literal embodiment of social inequalities (Gravlee 2009, Kuzawa & Gravlee 2016) is reflected in the legacies of these collections.

Historic documentation provided very limited or no information on specific preservation details for even standard information such as provenience for some collections (i.e. AMM and Switzerland); historic documentation is often non-specific for recovery location(s) and use vague, often antiquated and racialized, categorization terms. While some collection histories were well-documented due to their widespread use in the field, others yielded unexpected findings. While anecdotally described as Bronze age, detailed analysis of Smithsonian records associated with the Mongolian individuals in this study suggests that they are rather more recent 20th century sky burials at the time of collection. AMS radiocarbon dating is ongoing to confirm sample age across collections (Catalog numbers: SI225017, SI325275, SI256388, SI258306, SI258518, SI258523, SI258519, SI278768, and SI380563). Importantly, historic contextualization of the excavations and sociocultural origins of collections helps depict a more complete understanding about the ways the dead were treated culturally and as part of museum assemblages.

The assessed collections reflect various cultural treatment of bodies following death from around the world and through time (Figure 1). For both Egyptian collections, individuals were mummified according to cultural practices (i.e. the intentional drying, wrapping, and burial of

human remains as part of religious burial rites) (Peck 1998). In addition to being interred in Egypt, a hot and dry environment, the mummification process was also intense, but effective in preserving the remains morphologically. Conversely, the Mongolian individuals underwent the sky burial cultural practice, where a body would be attached to a horse or cart which the rider would continue driving away from the city until the body “fell where it may” (July-September 1912). Remains were then left on the surface to decompose, or scavenged by animals (e.g. vultures and dogs) (July-September 1912). Individuals from Switzerland were sourced from village ossuaries (i.e. a mode of secondary internment of disarticulated skeletal remains stacked by skeletal elements) (1950b). Alongside decomposition mechanisms affecting remains directly upon death, cultural processes are commonly considered to be the initial processes contributing to specimen taphonomy.

Following initial taphonomic processes (i.e. cultural and environmental), are preparation and curation processes applied to remains upon becoming a part of collections. Post-mortem (i.e. anatomical) and collection treatments presented here vary widely as Smithsonian skeletal collections were prepared for specific goals. In the case of anatomical preparations of recent decedents, Hrdlička was interested in associated brains with their crania (i.e. Germany and Japan) and requested preservation methods in line with soft tissue preservation. With the Terry collection, methods focused on preserving complete skeletons to ensure the long-term availability of the collection for on-going anatomical research. For the Egyptian collections, Hrdlička primarily collected crania. It was not because the post crania were not found; human crania were just more interesting to the public and as scientific tools at the time, so Hrdlička selected for them (Hunt 2018). Between the Lisht and Kharga Oasis collections, Hrdlička handled each differently, with more care being given to the Lisht remains (December-March 1908-1909). The German and IUCMT collection were both subject to autopsy and anatomical preparation methods; even the exclusion of the IUCMT individuals from refrigeration and the formalin injection given to the German individuals likely contributed to taphonomic differences. Assemblage of individuals from their original burial contexts or acquirement directly after death to be a part of collections altered the remains through handling, preparation processes, and curation methods, continuing and adding new taphonomic layers to remains.

DNA Preservation

Collection preparation and preservation processes directly affected the amount of recoverable DNA. The Egyptian collections (range 0.00-1.41 ng/mg) and the Terry Collection (range 0.00-1.83 ng/mg) had the lowest DNA concentrations. The cultural practice of mummification and being interred in a hot, dry environment has made recovery of DNA challenging from Egyptian materials (Gilbert et al 2005, Khairat et al 2013, Marota et al 2002, Nerlich et al 1997, Paabo 1985, Schuenemann et al 2017). The collection documentation, low DNA yield, and high levels of skin contribution (39%) suggest that overall DNA preservation is poor in the Egyptian collections. Despite being older, Egyptian Lisht samples had more reads identified during 16S rRNA analysis than the Egyptian Kharga Oasis samples. This differential sequence recovery could be directly related to the delicate handling of these remains upon their excavation, whereas Kharga samples were handled with less care and are well studied today. Individuals in the Terry collection were boiled and exposed to benzene vapors which impacted both the quantity of DNA recovered and the DNA fragment length (average of 52 bp) when mapped to the common oral pathogen *T. forsythia*. In contrast, the IUCMT (Japan) collection yielded much higher DNA concentrations (range 5.21- 38.21 ng/mg), possibly reflecting the less consistent use of the collection for research, as SourceTracker estimated that it had very little skin contamination (0.25%). In general, postmortem preparation and museum taphonomic processes differentially impact the amount of total recoverable DNA across museum collections.

Despite variability in total DNA concentrations, source authentication of the microbial communities depicted largely intact ancient dental calculus communities. Cleaning processes applied to buried collections and comparative samples were successful in removing soil contaminants as all samples had minor or no soil contributing microbes identified in microbial communities. However, the regularity of research conducted on the Egyptian, Mongolian, and Terry collections (Hunt & Albanese 2005) is reflected in the observably higher proportions of skin associated microbes. Museums could minimize future skin contamination by asking those who handle human remains (or other material that would be of interest for genomic analysis) to wear gloves.

The global dataset of archaeological dental calculus samples from Mann et al (Mann et al 2018) are known to have been chosen for their good DNA preservation. Thus, their similarity to the modern dental calculus in fragment lengths and identified OTUs is not unexpected (Figure 5B). However, the fragment length distributions for museum collections are significantly different from one another as well as to the archaeological dental calculus, indicating that collection taphonomic histories impact aDNA fragmentation. While the significant differences seen across museum collections may specifically be impacted by individual sample variation, the overall mean fragment lengths indicate that some museum collections have shorter fragments as a result of the cultural/environmental (Egypt) and preparation (anatomical preparations) taphonomic phases.

Challenges in Ancient Metagenomics

Ancient metagenomics is faced with a challenge in conducting bioinformatic analyses as we are in the initial stages of understanding the impacts of using software tools developed for modern microbiome data on fragmented and degraded DNA. Here, we applied two commonly used metagenomic tools, QIIME for 16S rRNA and MetaPhlAn2. QIIME 16S rRNA analysis uses a single marker gene for taxonomic identification, while MetaPhlAn2 relies on series of genomic markers normalized for genome size (Truong et al 2015). Zeisemer et al (2015) has previously shown the necessity of shotgun metagenomic approaches for investigating ancient samples over the use of amplicon deep sequencing of 16S rRNA gene hypervariable regions, due to the taxonomic biases introduced by PCR on short fragment lengths. This study shows that measuring alpha diversity (number of observed species level OTUs) is also highly problematic for fragmented DNA.

Across collections and comparative data, QIIME identified more species than MetaPhlAn2. QIIME overestimates alpha diversity in ancient samples. This may be influenced by the fact that 16S rRNA sequences have been heavily studied and curated, resulting in high numbers of identified OTUs. Previous work has suggested that there are shifts in alpha diversity over time (Adler et al 2013). However, this may not be due to actual differences in microbial diversity in ancient and museum/archaeological samples, but rather, as Velesko et al (2019) propose, linked to the maturation stage of the dental calculus biofilm. Collections with very small average

fragment lengths (e.g. Germany, Switzerland, and Terry Collection) have some of the highest number of OTUs identified by QIIME. In contrast, MetaPIAn2 identifies fewer number of taxa in degraded samples compared to modern calculus. Using MetaPIAn2, Velsko et al (2019) found key metagenomic distinctions between modern dental plaque and calculus and indicated that these substrates can no longer be used interchangeably in analyses due to the positive impact of modern oral hygiene practices at limiting oral disease progression. Because of the impact of preservation on fragment lengths and alpha diversity estimates in ancient samples, alpha diversity should not be used for comparisons across time and space, instead multiple analyses and tools need to be applied for assessing ancient microbial community differences.

Assessment of microbial profile taxonomic differences using sPLS-DA verified expected differences between modern and ancient (archaeological and museum) dental calculus samples. The microbial community similarities seen between museum and archaeological samples reflect the similarity of the oral microbiome globally (Supplementary Figure 1), with some impact by taphonomic processes. Community reconstructions may appear more similar as a result of the variable number of species identified in these collections and samples overall. Low abundance taxa are especially helpful in discriminating among collections and datasets. For instance, all the species identified as maximizing group differences (Figure 10) are low abundance microbes. With the exception of *C. showae* and *A. cardiffensis* occurring in maximum relative abundances of 5.69% and 6.52%, respectively, all other distinguishing species have a maximum relative abundance below 0.2%. Importantly, the number and sample size of groups included in sPLS-DA altered the results. This is especially evident when the three collection types are analyzed versus when the collections are divided into the ten treatment categories; sPLS-DA illustrated more differences among the three collection types than when treatment histories are considered, where more homogenization is displayed. When increasing the number of analyzed groups it is important to also increase the within group sample sizes, as the sPLS-DA can estimate intergroup variation and content better.

To minimize collection impact, this study focused on 5-6 individuals from each collection, however additional samples could strengthen associations between collection and taphonomic processes. As more and more molecular data are generated from museum collections, returning

or publishing preservation data and results to museums as part of broad sampling agreements, which may already have data retention mechanisms in place, could also improve data driven decision making processes. Information about DNA recovery for specific samples and collections are invaluable for museum curators and other researchers for implementing museum policies (e.g. requiring researchers to wear gloves) and directing genomic research towards collections/samples that have previous worked.

3.6 Conclusions

Importantly, we have demonstrated that provenience alone cannot predict success of biomolecular analysis, even for a rich substrate like dental calculus. In order to effectively authenticate and anticipate accurate molecular results, preservation histories of museum collections need to be identified and incorporated into research designs prior to molecular analyses. With the exception of IUCMT, collection preservation revealed consistently low amounts of total DNA from the anatomical collections and well-handled material. Even so, reconstructed microbial communities and DNA damage profiles reflect intact ancient oral microbiome communities with expected degradation patterns for most collections.

Our findings have important implications for identifying taphonomic processes affecting DNA recovery and the applicability of museum collections for research questions dependent on molecular analyses. The subsequent impact of collection taphonomic histories on metagenomic results underscores the need for pilot studies on specific collections before large scale destructive sampling of museum collections for genomic analysis.

3.8 Acknowledgements

We would like to thank Katherine Crowe of the National Anthropology Archives for her assistance throughout the archival research. We also thank Dr. Kristina Zaranko for translating old scripted museum acquisition notes and letters in German, and Dr. David Hunt for contextualizing the collections within the NMNH over time.

Chapter 4

Smithsonian Institution's Terry Collection Reveals Insights into Ancient Tuberculosis and Syphilis Pathogen Detection Through Dental Calculus Metagenomics and Hybridization Capture

In preparation for submission to Philosophical Transactions of the Royal Society B.

4.1 Abstract

Two recognizable and revered diseases in humans are venereal syphilis and tuberculosis (TB), which have afflicted human populations for millennia. Recent metagenomic methodologies have made pathogens accessible from ancient human remains. Dental calculus, or calcified plaque, is an accessible biofilm from skeletal remains containing health informative biomolecules, potentially recording venereal syphilis and/or TB causes of death. To assess the potential of dental calculus to be a diagnostically informative substrate, 39 individuals from the Smithsonian Institution's Robert J. Terry Collection with known causes of death as either syphilis or TB were metagenomically tested via shotgun and capture approaches for the presence of *Treponema pallidum* and/or *Mycobacterium tuberculosis*. We demonstrate the recovery *T. pallidum* dental calculus to be problematic, whereas *M. tuberculosis* recovery appears more promising. While hybridization capture of *M. tuberculosis* did not recover lineage defining SNPs, the new capture approach presented was validated for a much lower cost than synthetic capture kits. However, confirming individuals' cause of death could not be achieved, which is thought to be as a direct result of the specific collection preparation that was applied. We anticipate our study to initiate further investigation into the diagnostic potential of dental calculus and the paleogenomic recovery of syphilis and tuberculosis from past populations.

4.2 Introduction

The origins and evolutionary histories of venereal syphilis and tuberculosis (TB) are of great interest to researchers and public health officials due to their significant cultural and health impacts on human populations (Barberis et al 2017, Bos et al 2014, Galagan 2014, Harper et al 2011, Prevention 2015, Sarbu et al 2014, Stone et al 2009, Šmajš et al 2012). Historically referred to as the "Great Pox" and "consumption", syphilis and TB (respectively) are communicable diseases with high, worldwide mortality rates (Barberis et al 2017, Organization

2019, Peterman & Kidd 2019, Sefton 2001, Sepkowitz 1996). Within the United States, the 2018 case rate (per 100,000 people) of syphilis was 35.3 (Prevention 2019a) and 2.8 for TB, but with up to 13 million people estimated to be infected with latent TB (Talwar et al 2019). Furthermore, antibiotic resistant TB is an increasingly important public health concern; of the 2018 TB cases reported in the US, approximately 9% were isoniazid drug resistant (INH TB), with 1.5% additional cases being drug resistant to both isoniazid and rifampin (multidrug-resistant TB) (Prevention 2019b).

Syphilis is a sexually transmitted infection caused by the *Treponema pallidum* bacterium. Untreated syphilis is an ulcerative disease that systemically affects the body over many years, and in many cases results in death (Aminov 2017, Musher 1991). Since the recognition (and heavy stigmatization) of syphilitic disease in Europe 500 years ago, regular debates about the New or Old World origin and antiquity of syphilis have raged (de Melo et al 2010, Harper et al 2011, Tampa et al 2014). With discussions claiming an Old World origin from probable cases of ancient syphilis based on skeletal evidence (Harper et al 2011). Genomic analysis of *T. pallidum* is notoriously difficult even for modern clinical samples, primarily due to the bacterium's inability to live and reproduce outside the human body (and purposefully infected rabbits) (Edmondson et al 2018). Positive molecular identification of *T. pallidum* from ancient remains has also been extremely difficult to do, and only recently achieved from skeletal lesions of three historic Mexico City human remains (Bouwman & Brown 2005, de Melo et al 2010, Meffray et al 2019, Schuenemann et al 2018).

Human infectious TB is caused by the *Mycobacterium tuberculosis* bacterium. TB remains a chronic infection seen around the world. Typically infecting the lungs first (but can infect any organ), symptoms and disease severity are dependent on the host's immune response, inoculum load, and bacterial strain, and is regularly fatal without proper treatment (Ortner 2003, Scriba et al 2017). Humans and *M. tuberculosis* have evolved and traveled with one another for thousands of years (Comas et al 2013, Wirth et al 2008). From this long history, molecular evidence from ancient remains suggests that there has been numerous exchanges of TB between humans and animals, particularly within American TB strains, with humans regularly transmitting the disease to animals and then being re-infected (Bos et al 2014).

Evolutionary understanding of how, when, and where syphilis and TB infects populations relies upon their positive/probable identification from temporally and spatially contextualized remains. Recently, dental calculus (calcified dental plaque) has emerged as a reservoir of high-quality human, microbial, pathogen, and dietary biomolecules (Adler et al 2013, De La Fuente et al 2013, Linossier et al 1995, Radini et al 2019, Warinner et al 2014b). Because dental calculus is a biofilm, in addition to its easy access and simple sampling procedures, its collection and testing is generally considered to be comparatively less destructive to human remains than the use of bone and/or teeth (Ozga et al 2016). Due to dental calculus being located in the mouth, oral pathogens have been the primary focus of disease-oriented research. However, the well-documented relationship between oral and systemic health (Cullinan et al 2009, Dietrich et al 2017, El Attar et al 2010, Sumi et al 2007), make dental calculus an optimal substrate to investigate pathogens that do not leave diagnostic skeletal lesions and affect the body systemically. However, diagnosis of systemic diseases using dental calculus remains untested.

In order to explore this potential, this study utilizes metagenomics to explore the recoverability of *T. pallidum* and *M. tuberculosis* (Comas et al 2010) from dental calculus using shotgun sequencing and targeted hybridization capture methods, alongside skeletal pathological assessment of individuals from the Smithsonian Institution's (SI) Robert J. Terry Anatomical Collection (Terry Collection). Systemic proliferation of *T. pallidum* during secondary and tertiary syphilis, and *M. tuberculosis* in pulmonary and systemic TB infections make these pathogens ideal to test the diagnostic potential of dental calculus for these diseases (Singh & Romanowski 1999). Additionally, the Terry Collection contains individuals with known syphilis and TB causes of death (COD), allowing us to ground truth our ability to diagnose perimortem diseases in museum collections and archaeological remains.

4.3 Materials and Methods

Detailed methodological information on skeletal assessments, as well as DNA shotgun, quantitative PCR, and capture procedures and analysis can be found in the supplementary material. Dental calculus samples from 39 individuals from the Smithsonian Institution's Terry Collection with COD as either, TB (n=20) or syphilis (n=19) were analyzed. More information

about the history, collection preparation, and broader impacts of the Terry Collection is provided in the supplementary material. Selected individuals were osteologically evaluated for dental health and general pathological reactions (Buikstra & Ubelaker 1994, Kerr 1988), ante- and perimortem systemic health, as well as disease-specific skeletal lesions (Harper et al 2011, Pedersen et al 2019). Diagnostic criteria based on several previous studies (Ortner 2003, Powers 2012, Santos & Roberts 2006, Steckel et al 2006) were synthesized and implemented in order to cohesively assess skeletal lesions consistent with tuberculosis infection (see Supplementary Table 1). A composite score of pathologic lesions for tuberculosis (Pedersen et al 2019, Zuckerman & Sandhu) and syphilis (Harper et al 2011) was determined for each individual as skeletal disease diagnosis estimate. Regardless of cause of death (COD), individuals were scored for both tuberculosis (TB) and syphilis skeletal diagnostic criteria.

Samples were extracted and prepared for sequencing at the Laboratories of Molecular Anthropology and Microbiome Research (LMAMR) in Norman, Oklahoma, USA, in a dedicated ancient laboratory. Extraction controls and reaction negatives were processed in parallel with all wet lab procedures. Extraction and sequencing preparations follow previously published descriptions in Mann et al (2018). Briefly, between 1 to 5.5mg of dental calculus was sampled and washed with 0.5M EDTA, prior to decalcification in a solution of (fresh) 0.5M EDTA and 10% proteinase K (Qiagen) at room temperature (Ozga et al 2016). A modified Qiagen MinElute silica-column based purification was performed to isolate DNA and then quantified using a Qubit High Sensitivity fluorometer (Life Technologies) (Ozga et al 2016). Samples were built into dual indexed Illumina shotgun libraries after a partial uracil-DNA-glycosylase (UDG) treatment (Rohland et al 2015). With the exception of Nebulization, NEBNext DNA Library Prep Master Set (New England Biolaboratories; E6076) kit was utilized, with samples purified with Qiagen MinElute silica spin column after each step (15 ul EB). Libraries were amplified in triplicate, pooled, and purified before final sample quantification, pooling, purification, and sequencing. Libraries were sequenced on an Illumina NextSeq (either 1x75 or 2x150; see Supplementary Tables 2 and 3) at the Max Planck Institute for the Science of Human History (Jena, Germany).

Sequencing adapters were removed, paired-end data was merged, and quality filtered for base quality and read length using Adapter Removal v2 (minimum quality 20 and read length 25) (Schubert et al 2016). Quality-filtered metagenomic reads were mapped to *T. pallidum* (str. Sea 81-4) and the ancestral *M. tuberculosis* (Comas et al 2010) reference genomes with Burrows-Wheeler Aligner v0.7.12 (Li & Durbin 2009). Unique, quality filtered shotgun reads mapping to both genomes were visualized using the BLAST Ring Image Generator (BRIG) program (Alikhan et al 2011). SourceTracker v1.0.1 profiles were estimated in order to identify potential contributing source environmental contexts (Caporaso et al 2010). Metagenomic profiles were generated using MetaPhlAn2 (Truong et al 2015). Shotgun data were also subjected to the pathogen screening tool HOPS (Heuristic Operations for Pathogen Screening) (Hübler et al 2019) as an additional assessment of pathogen profiles.

All sample extracts were screened for the presence of MTBC DNA by targeting the mycobacterial *rpoB* gene and the IS6110 multi-copy insertion element. The *rpoB* assay targets a sequence of the gene specific to the infectious MTBC, whereas the IS6110 assay targets MTBC species and other, specific closely related mycobacteria such as *M. kansasii* and *M. smegmatis* (Coros et al 2008, Harkins et al 2015, Thorne et al 2011). Ancient DNA extracts from bone samples from Peruvian mummies, which had previously yielded complete MTBC genomes (samples AD-54, -58, and -64) (Harkins et al 2015), as well as DNA from a modern MTBC strain H37Rv, were included as positive controls. Reaction conditions followed previously published protocols (Housman et al 2015). Detection of either *rpoB* or IS6110 indicates the presence of MTBC infectious species.

Six samples were selected for hybridization capture of *M. tuberculosis* based on the percentage of unique, quality filtered shotgun metagenomic reads mapped to the computationally reconstructed ancestral *Mycobacterium tuberculosis* reference genome (Comas et al 2010), the identification of MTBC-specific genome *rpoB* and/or IS6110 elements (Harkins et al 2015, Housman et al 2015, Thorne et al 2011) with quantitative PCR, and the proportion of 16s rRNA taxa belonging to an ancient dental calculus environment (as an estimate of environmental contamination).

Whole genome biotinylated probes for myBaits Hybridization Capture for Targeted NGS kit (manual v.4.01 April 2018) were generated directly from genomic *M. tuberculosis* DNA strains H37Rv, CDC1551, East African Indian 91_0079, and Indo-Oceanic T17X. Genomic *M. tuberculosis* strains were pooled in equimolar proportions (5 ug total) before being sent to Arbor Biosciences to be prepared into capture probes. Two capture experiments with six, total, dental calculus samples and one sample positive (either AD-54 or -64) (Harkins et al 2015) were conducted following manufacturer's instructions with slight modifications (see Supplementary Methods). The first capture experiment was used as an initial test to assess TB recovery. To confirm that experiment one results were not dependent on low sample input DNA volume (15ul) into the library preparation, a second capture experiment was conducted. When possible, new DNA extractions with higher extraction input weights as well as, larger library preparation volumes (50 ul template DNA) were used in the second experiment to maximize the amount of available DNA. Libraries in both capture experiments were hybridized at 65C for 12 hours, immediately followed by 60C for 36 hours in two, with two separate rounds of capture. After each round of capture, post-capture libraries were amplified in triplicate following the April 2018 myBaits capture protocol, using 5 ul of capture library for each replicate (15 ul total) and 15 ul NF water for a total replicate reaction volume of 50 ul. Importantly, 5.5 ul of biotinylated RNA baits were used for each sample and rounds of capture. Capture libraries were cleaned with the AMPure bead cleanup method, selecting for fragments longer than 150 bp, and resuspended in 30 ul of elution buffer. Capture experiment libraries were sequenced at the Oklahoma Medical Research Foundation Clinical Genomics Center in Oklahoma City, Oklahoma on a NextSeq using 150 bp paired-end read chemistry.

Capture sequencing data were analyzed by removing adapters, merging paired-end reads, and quality filtering for base quality and read length using Adapter Removal v2 (minimum quality 20 and read length 25) (Schubert et al 2016). Quality-filtered merged *M. tuberculosis* capture reads were mapped to the ancestral *M. tuberculosis* (Comas et al 2010) reference genome with Burrows-Wheeler Aligner v0.7.12 (aln seed length 1000 and maximum edit distance 0.1) (Li & Durbin 2009). Low quality reads and unmapped sequences were filtered, with the remaining reads being sorted before PCR duplicate and multiple mapping reads were removed and files

converted to fasta format using Samtools v1.5 (DeSantis et al 2006, Li et al 2009). Alignments were visually inspected for lineage informative SNPs identified by Bos et al (2014).

4.4 Results

4.4.1 Skeletal Diagnoses

Syphilis diagnostic categories ranged from 0 (non-specific lesions) to 5 (treponemal disease lesions present, observed on multiple elements of a single individual) (Harper et al 2011). Tuberculosis diagnostic categories ranged from 0 (non-specific lesions) to 3 (multiple ribs with periosteal reactions on visceral surfaces and destructive, lytic lesions observed on multiple thoracic and/or lumbar vertebrae) (Supplementary Table 1). Indicator scores for each individual are reported in Supplementary Table 2. Diagnostic criteria had 90% and 80% reliability for individuals with COD of tuberculosis and syphilis, respectively.

4.4.2 Shotgun Sequencing

An average of 16.8 million reads (range = 6.2 and 96 million) were recovered across all samples, with extraction and library blanks sequenced to an average of 3 million reads. Across all samples, the number of post-quality filtering reads that mapped to either the *T. pallidum* reference genome (GCA number 000604125.1) was a maximum of 127, and for the *M. tuberculosis* ancestral reference genome was a maximum of 4,415 (Comas et al 2010). Reads mapping to *M. tuberculosis* had more genome coverage (Supplementary Figure 1 and Supplementary Table 3) and lower clonality (2.7 to 22.5), contrasting to reads mapped to *T. pallidum* with poor genome coverage, high levels of read stacking, and high clonality (35.1 to 861.7) (Supplementary Figure 2 and Supplementary Table 4).

As *T. pallidum* was only recovered in minute amounts with high levels of read stacking and clonality, further analyses to identify *T. pallidum* were not carried out beyond shotgun screening. *T. pallidum* was not recovered in informative quantities from the Terry Collection to identify or confirm people were ill with venereal syphilis.

SourceTracker analysis successfully assigned recovered DNA to various sources, with all samples showing recognizable oral signatures (Figure 1 and Supplementary Figure 3). Consistent with the collection's lack of burial history, microbes from soil environments were recovered in negligible amounts. The proportion of skin-associated microbes varied across samples, potentially due to the long history of extensive handling and study of the Terry Collection individuals without gloves or regard for the biomolecules before such technology was available.

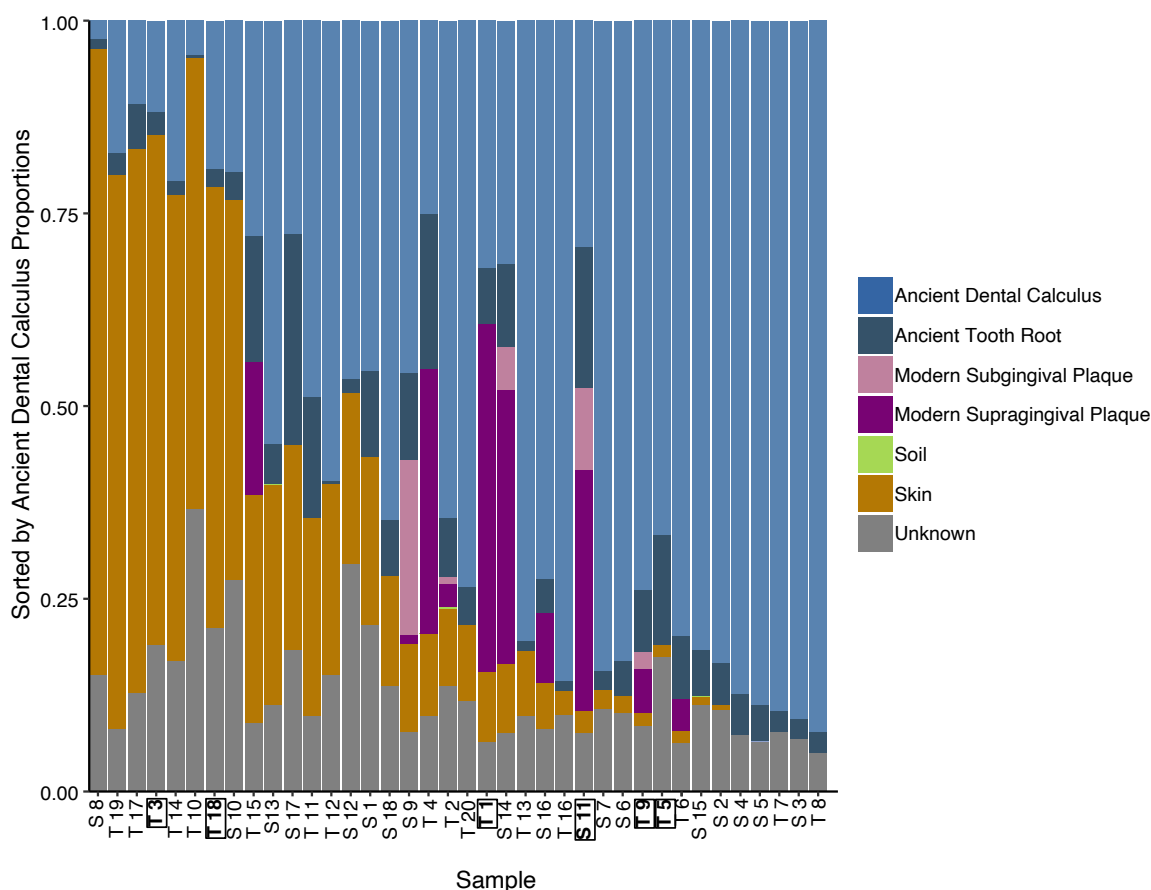


Figure 1. SourceTracker analysis of shotgun metagenomic reads sorted by increasing proportions of taxa associated with archaeological dental calculus samples (Mann et al 2018). Boxed and bolded sample labels indicate samples used in hybridization capture experiments.

HOPS results show several samples passing damage pattern and edit distance filters to genus *Treponema* (see red boxes in Supplementary Figure 4). However, upon further inspection of HOPS documentation, screening identified only *T. putidum* and *T. denticola* with greater than

80% surety across all samples for “*Treponema*”. Contrasting higher mapping rates to *M. tuberculosis*, HOPS results show little to no detection of MTBC reads across all samples.

All samples had higher overall mapping rates to *M. tuberculosis* than to *T. pallidum*, as well as to oral reference genomes of *Fusobacterium nucleatum* (a common, non-pathogenic oral microbe), and *Porphyromonas gingivalis* and *Tannerella forsyntia* (known oral pathogens) (Figure 2). These differential proportions of bacteria recoverability may also be a reflection of the general oral disease states observed across the assessed individuals.

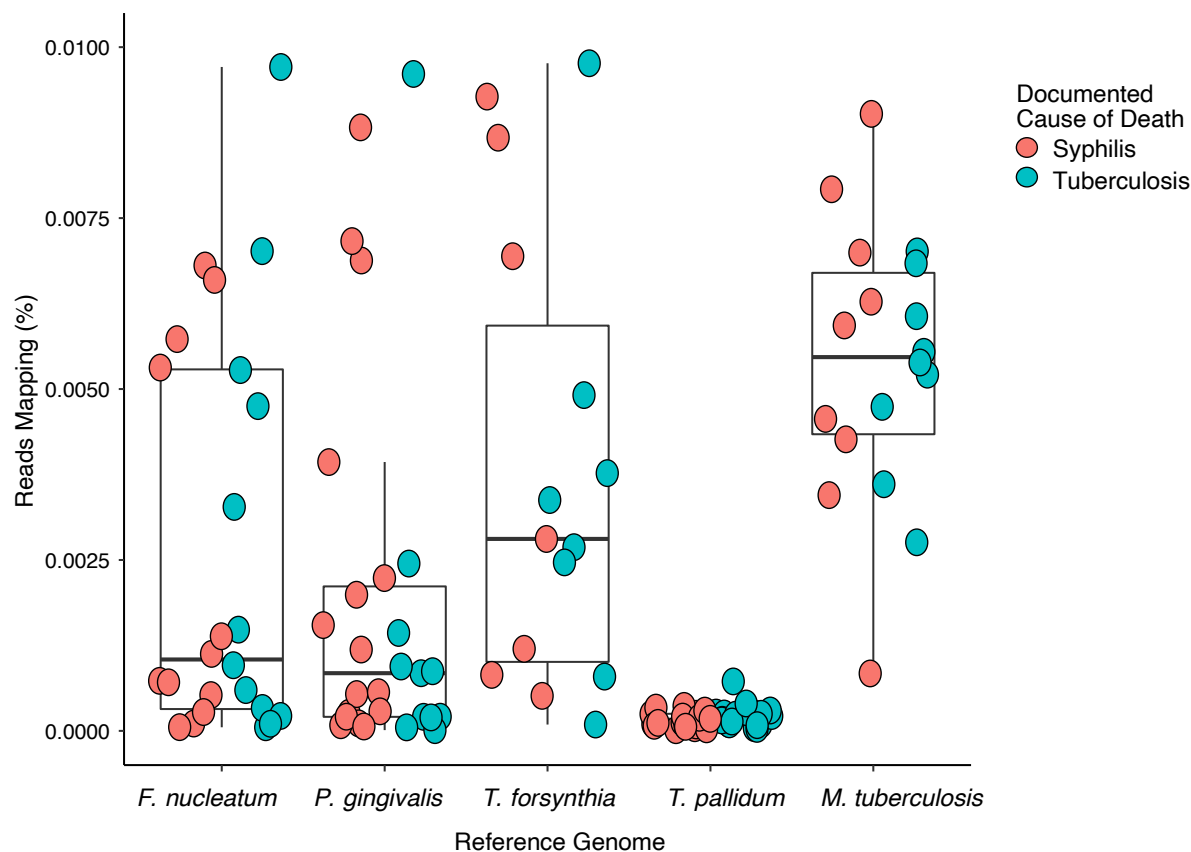


Figure 2. Proportion of unique, quality filtered reads mapped to genomes of *F. nucleatum* (GCA 000007325.1), *P. gingivalis* (GCA 000010505.1), *T. forsyntia* (GCA 000238215.1), *T. pallidum* (GCA 000604125.1), and *M. tuberculosis* (Comas et al 2010). In general, mapping rates to a given reference genome are low (less than 1%), however, *M. tuberculosis* proportions are visibly higher than common oral taxa. *T. pallidum* is nearly undetectable with this approach.

4.4.3 MTBC quantitative PCR

None of the samples tested positive for the *rpoB2* assay (Supplementary Table 5). Samples T1, T9, and T18 were positive for the IS6110 assay (Supplementary Table 6). TB sample positives

(Harkins et al (2015) samples AD-54, -58, and -64) and standards were confirmed the detection of IS6110 and *rpoB*, indicating successful experiments. Samples with positive calls for IS6110 were selected for *M. tuberculosis* hybridization capture.

4.4.4 Hybridization Capture

In-solution capture of *M. tuberculosis* using gene-nonspecific baits generated from modern TB strains was successful in reconstructing TB genomes for two previously published samples (Bos et al 2014, Harkins et al 2015); TB samples AD-54 and -64 yielded average genome coverages of 39X and 2.3X, respectively (Supplementary Tables 7 and 8). The implemented baitset is markedly different from those previously applied (Bos et al 2014) due to the inclusion of non-MTBC-specific bait sequences (e.g. *rpoB*, *gyrA*, *gyrB*, *katG*, and *mpt40* genes). In addition to replicating previously published results, the presented baitset is significantly cheaper to produce than other capture assays, further validating this approach.

Even so, less than 0.01% reads mapping to *M. tuberculosis* were recovered from these samples. With very high cluster factors (total tuberculosis mapped reads divided by the number of unique mapped tuberculosis reads; >564 from the first and >94 from the second), all reads were likely sequenced. Across both capture experiments, lineage defining single nucleotide polymorphisms (SNPs) were called for TB sample positives (AD-54 and -64), however, only one SNP position was observable for a single individual from each capture experiment (samples T1 and T3) (Supplementary Table 8). No change was observed for the samples between the ancestral *M. tuberculosis* reference genome and the lineage 4.3 base pair change (position 764995) from a cytosine to guanine nucleotide. Because the ancestral reference genome is attributed to TB lineage 4 and no SNP base pair variants were observed for the calculus samples, TB sequences may be lineage 4 TB strains. Overall, dental calculus capture experiments did recover putative DNA, but not enough was retrieved to complete genome reconstruction or identify specific lineages.

4.5 Discussion

Where TB was recovered from dental calculus using shotgun metagenomic methods, syphilis was not. All samples had low recoverability of *M. tuberculosis* and *T. pallidum*. While *M.*

tuberculosis was more reliably identified, the percent recovered across samples does not correspond to COD. *Treponema pallidum* reads were scarcely recovered, with high clonality and read stacking observed across samples. The metagenomic and COD juxtaposition is highlighted with a syphilis sample having the highest percent of reads mapping to *M. tuberculosis* but without any documented coinfection with TB (Figure 1 and Supplementary Figures 5 and 6). Documented co-infections do not correspond to read recovery rates for either disease. Neither, HOPS nor mapping analyses recovered *T. pallidum* in biologically meaningful amounts from these dental calculus samples. Diagnosing individuals with either TB or syphilis using shotgun metagenomics or the HOPS pipeline from Terry Collection dental calculus was not possible. Although, *Mycobacterium tuberculosis* is accessible in higher amounts than even common oral microbes and pathogens from dental calculus.

Tuberculosis-specific genetic screening and capture techniques did not provide more or better information corresponding to known incidents of TB COD. Supporting previous research, quantitative PCR successfully recovered MTBC-specific genomic elements from TB sample positives (Harkins et al 2015). MTBC IS6110 insertion was detectable from dental calculus, whereas the *rpoB2* gene was not. With <0.01% TB endogenous content recovered from both shotgun and capture methodologies, *M. tuberculosis* hybridization capture does not increase read or SNP recoverability from dental calculus. However, the presented capture approach is a much more affordable as well as valid option to apply in future research.

This study implemented myBaits capture probes developed from modern MTBC strains. *M. tuberculosis* lineage defining SNPs and genome reconstruction was successful for two previously tested Peruvian bone samples. Drastically reducing experiment expenses, the presented myBaits capture probes were generated from freely available TB DNA from various strains, as compared to expensive, commercially available options that use synthetic capture probes (e.g. Agilent). TB has possibly been recovered for the Terry Collection individuals, however more work is needed to model whether this is a result of mis-mapping of other *Mycobacterium* DNA in conserved regions of the genome due to the short read lengths in Terry individuals. Future research should test dental calculus in individuals with known TB and that were not preserved with intensive boiling. Furthermore, viable oral microbiome signatures were successfully

estimated across the majority of samples revealing the accessibility of microbial community and broader systemic health information. This is particularly meaningful because of the really short DNA fragments recovered across individuals, which may have resulted directly from the skeletonization/preservation process that was applied, as well as the long history of intensive research that has been done with the Terry Collection. Dental calculus sourced from other collections and/or individuals, whom were skeletonized and preserved through different methods and have known disease histories, could yield more and longer TB DNA sequences.

4.6 Conclusions

Our findings have important implications for understanding ancient DNA pathogen recoverability and the utility of dental calculus to molecularly diagnose TB and syphilis in ancient remains; *T. pallidum* was not recovered using shotgun metagenomics; *M. tuberculosis* may have been recovered from dental calculus using shotgun metagenomic methods; and hybridization capture recovered sparse amounts of putative TB DNA from dental calculus, although, not enough to determine specific MTBC lineages. While we were not able to successfully identify individuals, who suffered from TB or syphilis based on metagenomic results from dental calculus, the negative results presented here may stem from the intense skeletal preparation methods (e.g. benzene vapors) and research agendas applied to the Terry Collection. These findings emphasize the need for better understanding of differential DNA recovery across museum and archaeological collections (Austin et al 2019), especially for careful consideration of destructive sampling of museum and/or archaeological remains.

4.7 Acknowledgements

We thank Sabrina Sholts and David Hunt for thought-provoking conversations, support, and access to the Terry Collection at the National Museum of Natural History. We thank Andrea Eller for constructive and thoughtful feedback, Stephanie L. Schnorr and Jiawu Xu for assistance with labwork. We thank the individuals whose genetic and skeletal data were used in this study. Genomic DNA from *M. tuberculosis* strains H37Rv (NR-48669), CDC1551 (NR-48981), East African Indian 91_0079 (NR-44095), and Indo-Oceanic T17X (NR-44096) was obtained through BEI Resources, NIAID, NIH. We thank all the individuals whose genetic and skeletal data was used in this study (Tuberculosis Individuals 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17,

18, 19, and 20, as well as Syphilis Individuals 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, and 19). We extend our gratitude towards them as individuals with unique histories in their own right, as well as (regular) contributors to scientific and anthropological knowledge through being a part of the NMNH Terry Collection. The deidentification from the long-held Terry numbers applied here was done in an effort to maintain these individual's genetic privacy, and offer bodily and individual respect posthumously in recognition of the remains as individuals. Providing authorship or individually thanking skeletal individuals should not be seen as a solution to dealing with the nuances and convoluted histories of anatomical collections, rather our actions are meant to be a tool/mechanism to recognize these individuals as research contributors. Anatomical collections can no longer be thought of or treated as neutral within scientific contexts.

Chapter 5

Dental Calculus Reveals Acute Bacterial Infections: A Case Study from the Smithsonian's Terry Collection

In preparation for submission to the New England Journal of Medicine.

5.1 Abstract

Metagenomic approaches have allowed researchers to explore pathogen evolutionary histories from a new perspective, which has previously relied on historical, archaeological, and skeletal identification of disease. Mineralized dental plaque, or dental calculus, from skeletal remains of past populations is a recognized source of authentic ancient biomolecules, which has revealed complete microbial communities, health indicators, and dietary pathways from dental calculus elucidating past human lifestyles and experiences from a molecular perspective. Through the integration of skeletal, documentary, contemporary literature, and dental calculus metagenomic evidence, a cohesive account about an individual's life and death is presented here. Originating from St. Louis and a part of the Smithsonian Institution's Robert J. Terry Anatomical Collection, the examined individual has a documented cause of death of lobar pneumonia. Results reveal direct identification of pneumonia, which until now has been archaeologically invisible and obscured in historical records. The presence of *Acinetobacter*, a hospital acquired pathogen, is also identified. More broadly, an acute infection immediately preceding death is identified through dental calculus metagenomics. Skeletally, the individual displays chronic oral infections, as well as skeletal lesions indicative of a probable tuberculosis infection, confirming systemic ill-health. Documentary and newspaper evidence contextualize end-of-life circumstances for this individual in a segregated hospital and as being a part of an anatomical collection.

5.2 Introduction

Study of the history and evolution of infectious disease using historical, archaeological, and skeletal evidence is limited by the biases and incompleteness of documentary evidence (Mitchell 2017), biomolecular preservation, and the limited range of skeletal responses (i.e. bone growth and loss or altered development) that can be indicative of disease states (Ortner 2003). These factors are exacerbated by pathogen biology, pathogenicity, virulence, and the relatively slow reaction time of bone, whereby many infectious diseases (e.g. pneumonia, influenza, cholera) do

not generate skeletal reactions before death occurs. These material restrictions can convolute specimen taphonomy (post-mortem changes) and differential morbidity, mortality, and frailty in past individuals (DeWitte & Stojanowski 2015). Reference skeletal collections (e.g., anatomical collections) are further compounded by sampling bias (DeWitte & Stojanowski 2015, Komar & Grivas 2008, Zuckerman 2018), towards the inclusion of unique, distinctive, or advanced disease presentations in remains. Taken together, these can skew collection demographics and mortality estimates, and affect molecular analysis success (Austin et al 2019). However, with the emergence of ancient pathogen genomics, infectious diseases are regularly being identified in ancient contexts (Devault et al 2014, Spyrou et al 2019). Advances in molecular methods and bioinformatic approaches allow researchers to explore pathogen evolutionary histories, and can ultimately play a role in how we respond to current and future disease epidemics.

While teeth and skeletal material displaying pathological lesions have regularly been targeted for ancient pathogen research, dental calculus (calcified dental plaque) serves as a viable alternative, making human, microbial, pathogen, and dietary biomolecules accessible, with limited burial environment contamination (Adler et al 2013, De La Fuente et al 2013, Linossier et al 1995, Radini et al 2019, Warinner et al 2014b). Because dental plaque calcifies periodically throughout life, dental calculus has been considered a long-term record of an individual's oral health and pathogens that were either ingested or inhaled (Fuchs et al 2007, Jin & Yip 2002, Schroeder & Shanley 1969).

Here, we present a case study revealing end-of-life conditions for an individual from the Smithsonian Institution's Robert J. Terry Anatomical Collection (Terry Collection). The cause of death (COD) for individuals in the Terry Collection is documented in death certificates, providing an exceptional opportunity to conduct pathogen identification from dental calculus. According to the death certificate, the presented St. Louis individual (St.LI) was a 23-year-old, African American male, who died of lobar pneumonia in St. Louis, Missouri in the 1930s. In this study, metagenomic analysis of dental calculus, coupled with analyses of skeletal and contemporary records are integrated to form a cohesive narrative about the life and death of St.LI.

Pneumonia represented a major source of mortality in past populations, particularly in the pre-industrial and early-industrial periods (Gage 2005). However, it does not generate distinctive skeletal lesions (Ortner 2003), and is often misclassified in historical records (Gage 2005), confounding reconstructions of its impacts on public health. Hospital- and community-acquired pneumonia can be caused by a wide array of bacterial organisms including *Streptococcus pneumoniae*, *Haemophilus influenza*, *Moraxella catarrhalis*, *Staphylococcus aureus*, *Klebsiella pneumoniae*, *Acinetobacter* species, and *Pseudomonas aeruginosa*. Because pneumonic agents can be transmitted through inhalation or aspiration, regular contact between the respiratory tract and the oral mucosa create a high likelihood of their incorporation into dental calculus, thus potentially increasing their accessibility for molecular analysis.

For the first time, this study presents pneumonia as an investigable disease in historical contexts via integrating skeletal and metagenomic analyses. This study also refutes a long-held assumption that the data obtained from dental calculus always reveals cumulative information of a person's life (Warinner et al 2015a); we show that it is possible to recover pathogens associated with acute, terminal diseases from dental calculus, allowing the reconstruction of end-of-life circumstances.

5.3 Methods

Skeletal and pathological inventories of St.LI were conducted following established standards (Buikstra & Ubelaker 1994, Kerr 1988, Ortner 2003). Skeletal indicators of tuberculosis (TB) (Pedersen et al 2019) were also assessed due to the regular observation of periosteal and lytic reactions across St.LI's skeleton.

Ancient DNA was extracted from St.LI's dental calculus following established protocols (Ziesemer et al 2015) and built into partial uracil-DNA-glycosylase (UDG) treated, shotgun Illumina sequencing libraries (Meyer & Kircher 2010). The program, AdapterRemoval (Schubert et al 2016), prepared quality filtered sequences for assessing the preserved microbial community through MapDamage (Jonsson et al 2013) DNA preservation metrics, 16S rRNA bacterial taxonomic composition using MetaPhlan2 (Truong et al 2015) and identification of contamination source profiles via SourceTracker (Knights et al 2011). Sequences were mapped

to reference genomes of known bacterial causes of pneumonia and the human genome (hg19) using BWA and Bowtie2 (Langmead & Salzberg 2012, Li & Durbin 2009). PanPhlan (Scholz et al 2016) identified pneumonia strain based on the gene composition across all available complete genomes from the National Center for Biotechnology Information Genome database (September 2018). The genetic presence of tuberculosis was also tested by mapping reads to the computationally reconstructed ancestral *Mycobacterium tuberculosis* reference genome (Comas et al 2010).

5.4 Results

5.4.1 Skeletal analysis

St.LI has markedly robust post-cranial skeletal elements, with consistently enlarged and rugose muscle attachment sites, as well as musculoskeletal activity markers on the pubis and clavicles indicative of localized muscles trauma (Supplementary Table 1). Several non-metric traits, including metopism, and a depression on the anterior aspect of the right femoral head and neck were observed (Supplementary Table 2; Supplementary Figure 1). The depression has rounded margins indicating a long-standing presence, with minimal associated remodeling of the underlying cortical bone seen on radiographs (Supplementary Figure 1). A review of the published literature has not enabled identification of the cause of this depression (e.g. activity related). St.LI also exhibits a remodeling, complete midline mandible fracture (Supplementary Table 2; Supplementary Figure 2). Well remodeled, healing periosteal reactions (sub-periosteal bone deposition) are present on multiple post-cranial skeletal elements (femora, tibia, fibulae). Clustered pits or pores, with rounded margins indicative of healing, are also present on multiple elements (scapulae, radii). Areas of lytic pitting or pores, some with evidence of healing and others active at time of death (sharp margins), are present on multiple elements (left femur, right radius, sternum) and observable radiographically (Supplementary Table 2; Supplementary Figure 3 & 4). Additionally, St.LI exhibits multiple oral pathologies, including dental caries, antemortem tooth loss, oral abscesses, and evidence of gingivitis and both acute and latent periodontal disease (PD) (Supplementary Table 4; Supplementary Figure 5).

St.LI also displays multiple skeletal indicators of a chronic and/or episodic TB infection. Many (66%) diagnostically valuable indicators, with high sensitivity and specificity to TB (Pedersen et

al 2019) were observed, including bone proliferation and small clusters of lytic pits across multiple vertebrae, and deep lytic cavities and pits on the bones of the pelvis (Supplementary Table 3; Supplementary Figure 4 & 6). Less indicative TB diagnostic markers, (Pedersen et al 2019) specifically bone proliferation on a single rib and lytic pitting on multiple elements (vertebrae, humeri) were also observed, with varying degrees of activity and/or healing (Supplementary Table 3; Supplementary Figure 4 & 6). Further, radiographs reveal diffuse bone loss in the cranial vault. Macroscopically, this loss is reflected in well-remodeled areas of bone lysis and proliferation on the ecto- and endocranial surfaces, respectively (Supplementary Figure 7); these lesions are likely also associated with TB (Ortner 2003).

5.4.2 DNA Preservation and Authentication

Shotgun metagenome sequencing revealed 0.013% of reads mapped to the human genome; with short read lengths (62 bp average $\pm$ 23.2) and terminal base damage (averaging 5'C→T: 0.002, 3'G→A: 0.002) patterns characteristic of ancient DNA (Figure 1A) (Rohland et al 2015). Further analysis on recovered human DNA was not conducted to ensure genetic confidentiality. The microbial community profile for STL generated through taxonomic analysis of 16S rRNA reads (see methods) was compared to profiles generated from numerous sources including modern human skin, supra- and sub-lingival plaque, and soil, and ancient dental calculus and tooth root samples (Mann et al 2018) . The microbial community in St.LI more closely resembles other ancient tooth root microbiomes (Mann et al 2018) (Figure 1B). Expectantly, the microbial community is distinct from soil microbial communities, as the Terry Collection individuals were never buried (Hunt & Albanese 2005). These results verify the recovered DNA as representative of an oral context.

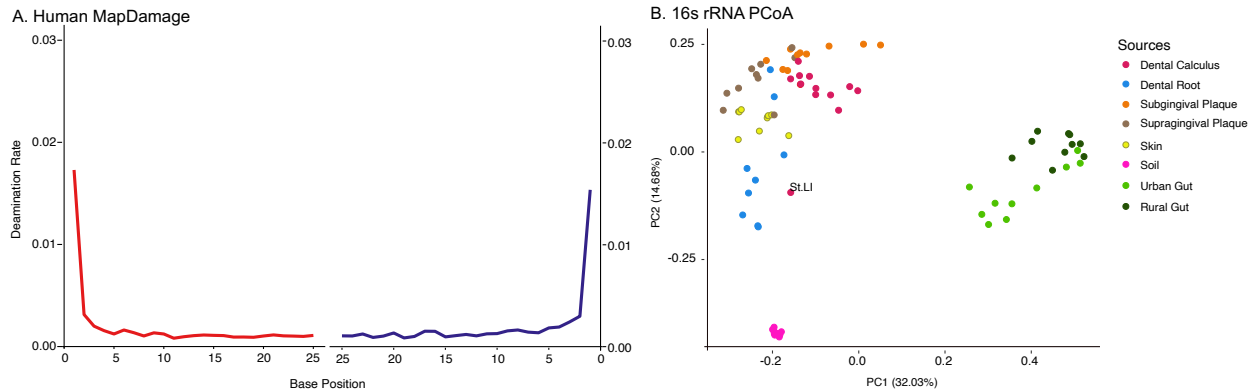


Figure 1. Authentication of Dental Calculus Metagenomic DNA (A) MapDamage plot of quality-filtered reads mapped to the human genome. These data display characteristic ancient terminal end read damage treated with partial UDG treatment. (B) Principle coordinate analysis of weighted unifracs distances for the 16s rRNA gene. St.LI Individual calculus clusters with ancient tooth roots, and away from soil and gut profiles. The position of the comparative skin sources (palm swabs) between dental root and plaque sources may be due to frequent hand-mouth interactions innate to humans (Oh et al 2016).

5.4.3 Pathogen Identification

The top ten microbial species observed in St.LI's microbial profile are either primarily known as, or can be, opportunistic or nosocomial pathogens that are not characteristic to a “normal” human oral microbiome (Supplementary Table 5) (Batra et al 2016, Bisharat et al 2012, Brooke 2012, Kim et al 2012, Mezzatesta et al 2012, Visca et al 2011, Wyres & Holt 2016). *Acinetobacter* bacteria, a well-known genus of nosocomial pathogens, dominates (~46%) the recovered microbial community composition (Munoz-Price & Weinstein 2008). Relative abundance of observed oral taxa using MetaPhlan2 corroborates the detection of *Klebsiella pneumoniae*, with it being the third most abundant bacterial species detected, accounting for an estimated 16% of the microbial community (Supplementary Table 5).

To identify a potential causal organism for the lobar pneumonia COD, quality filtered reads were mapped to genomes of known bacterial causes of pneumonia (Table 1). Over 669 times more reads mapped to *K. pneumoniae* than other potential causative agents. Reads mapping and genome coverage to the ancestral *M. tuberculosis* reference genome reconstruction were very limited (0.004% and 0.06%, respectively). Authentication with MapDamage (Supplementary Figure 8) or *M. tuberculosis* lineage identification was not possible with this small amount of mapped data.

Table 1. Percent of reads mapping to pneumonia pathogen reference genome, sorted by descending percent of reads mapping (see Supplementary Table 6 for references).

Pathogen	Percent of the Total Reads		
	Avg. Read Length (bp)	Mapping to Pathogen Reference	Mean genome coverage
<i>Klebsiella pneumoniae</i>	67.8	9.262	11.155x
<i>Staphylococcus aureus</i>	68.0	0.014	0.033x
<i>Coxiella burnetii</i>	49.5	0.012	0.029x
<i>Chlamydia psittaci</i>	45.7	0.003	0.011x
<i>Streptococcus pneumoniae</i>	38.0	0.002	0.003x
<i>Haemophilus influenzae</i>	40.1	0.002	0.004x
<i>Moraxella catarrhalis</i>	42.5	0.002	0.004x
<i>Neisseria meningitis</i>	36.5	0.002	0.003x
<i>Chlamydophila pneumoniae</i>	43.5	0.002	0.008x
<i>Mycoplasma pneumoniae</i>	44.5	0.002	0.009x
<i>Legionella pneumophila</i>	42.1	0.001	0.001x
<i>Streptococcus pyogenes</i>	37.5	0.0003	0.0006x

Outside of hospital settings, *Klebsiella* is a known skin microbe, particularly of the hands (Casewell & Phillips 1977, Oh et al 2016), whereas *Acinetobacter* species are found ubiquitously in the environment (Munoz-Price & Weinstein 2008). While these microbes are commonly found, the present study took methodological and computational approaches to guarantee the authenticity of reads, ensuring analyzed reads originated from St.LI's oral cavity; namely, an EDTA wash to remove outer contaminants before DNA extraction, including experimental control negatives, evaluating mapped reads for characteristic terminal-end damage, homoplasy and strain identification. Strain identification was accomplished by comparing pangenome markers to determine strain level similarities (Scholz et al 2016). *K. pneumoniae* strain 12208 shared the highest percent gene identity (92.1%) to the *K. pneumoniae* that infected St.LI.

5.5 Discussion

5.5.1 *The Substrate*

The amount and rapidity of calculus formation is influenced by diet, genetic variation, oral hygiene practices, age, sex, and mineral levels (calcium and phosphate) of the blood and saliva (White 1997). While dental calculus may form within 2-week periods (Radini et al 2017, White 1991), there is considerable variation between individuals, with people being described as “heavy” and “light” calculus formers. Several studies estimate that 0.10-0.15% of calculus by dry mass is deposited, daily (Akcalı & Lang 2018). While we cannot assess whether St.LI exhibited “heavy” or “light” calculus deposition, the unusual microbial composition indicates the consistent exposure of the mouth and dental calculus to *Acinetobacter* and *Klebsiella* bacteria. The oral community clustering more closely with tooth roots further emphasizes the St.LI’s dysbiotic oral health. These analyses enable direct identification of pneumonia, a major source of mortality in ancient and historic populations, but which has been archaeologically invisible and obscured in historical records. More broadly, results reveal direct identification of an acute infection immediately preceding death via dental calculus, a ubiquitous human skeletal substrate (Warinner et al 2015a).

5.5.2 *The Individual*

In addition to the acute infection responsible for their death, St.LI suffered long-term illness suggestive of a compromised immune system. Due to varied stages of lysis activity and healing at time of death, it remains unclear in what order St.LI developed the chronic infections (oral lesions and those associated with tuberculosis) and the subsequently expressed immunological response. Chronic systemic infections alter levels of systemic inflammatory responses mechanisms, potentially dysregulating and elevating inflammatory cytokine expression (Guyot-Revol et al 2006, Shaddox et al 2010). This is seen during the active phase of TB infections, where cytokine levels (e.g. tumor necrosis factor alpha (TNF α)) are elevated (Van Crevel et al 2002), which are also responsible for generating skeletal lesions (Ortner 2003). The healed mandibular fracture, activity markers, and other pathological lesions (Supplementary Table 3) indicate that St.LI also experienced chronic and/or episodic stressors throughout early adulthood, which may have resulted in prominent bodily stress levels (McEwen & Wingfield 2003). The observed periodontal disease and periosteal reactions would have also impacted St.LI’s bodily

stress response, with respective outcomes of hyper-inflammatory and immunological suppression responses. These processes and oral diseases would have opened exchanges among St.LI's immune and circulatory system to reactive bone and inflamed gum tissues, creating substantial opportunity for comorbidities to become systemic (Falcao & Bullón 2019, Scannapieco et al 2003, Straub 2011), corroborating the link between oral bacteria, and coronary (Dietrich et al 2017) and respiratory health (El Attar et al 2010, Sumi et al 2007) risks.

Molecular results corroborate the skeletal evidence. The top taxa identified through metagenomic analysis, *Acinetobacter pittii* and *A. junii*, are strictly aerobic, recognized nosocomial microbes (Wong et al 2017). As oral abscesses are likely to be primarily comprised of strict anaerobes (Rusinovci et al 2018, Shweta 2013) the presence of *Acinetobacter* suggests it is associated with the perimortem condition of St.LI, and active infection that potentially contributed to the recorded lobar pneumonia COD. The resemblance of this community to other tooth root microbial communities (Figure 1B) may be a reflection of having sampled calculus on the abscessed mandibular, right first molar (Supplementary Figure 5 & 9). If so, the dental calculus reflects a depreciated microbial community and the continued, unhindered development of dental calculus despite the presence of an abscess. Microbes would have freely accessed nearby host circulatory and lymphatic vessels due to concurrent local inflammatory responses. The strain level identification and genome coverage of the recovered *K. pneumoniae*, and *Acinetobacter* species genome and phylogenetic reconstruction exemplifies the depth of community imbalance observed in St.LI. The abnormal and nosocomial oral microbial community, combined with observed skeletal lesions, suggest elevated coinfection susceptibility from an incapacitated immune system.

K. pneumoniae infections progress and present rapidly (Ko et al 2002, Paczosa & Mecsas 2016). In conjunction with high morbidity associated with any *K. pneumoniae* infection, the chronic oral infections and inflammatory response, probable TB infection, and recovered nosocomial microbes would have likely suppressed St.LI's ability to effectively combat the pneumonic infection, regardless of his young age (Marais et al 2013). If St.LI did not, or was not able to, seek medical attention promptly, their condition would have quickly deteriorated. According to their death certificate, St.LI was admitted to St. Louis' only public (segregated) hospital for

African-Americans, which was notoriously under-staffed, cramped, and under-funded, with consistently poor patient outcomes (1925, 1928c, 1929, Brundidge 1932, Fisher 1933). Further, no effective antibiotics were available at this time to treat pneumatic infections. St.LI's pneumonia and systemic infection would have resulted in his lungs filling with pus, ultimately causing death through suffocation at the end of his 11-day hospital admission.

In conjunction with death records and predicted disease pathogenesis and immunopathology, these skeletal and microbial data provide insight into aspects of St.LI lived experiences as a young, African American laborer during a time of great social division and change in the U.S. and St. Louis. Race, like other modes of social inequality, is a culturally constructed framework of structural disenfranchisement that has both social (e.g., disparities in access to medical resources) and biological (e.g., health disparities) consequences (Kuzawa & Gravlee 2016). Race's biological consequences occur in part because the particular biological and social conditions produced by racism (i.e. structural discrimination, interpersonal discrimination) become embodied (Gengo 2014, Gravlee 2009, Krieger 2014). Chronic physiological stressors observed in the microbiome and skeleton of St.LI, and other Terry Collection individuals (De la Cova 2011, de la Cova 2019), can be associated with the environment of structural violence and racialized terror that exist(ed) for African-American communities, as well as the Jim Crow American South that people fled, in the early 20th century (Gengo 2014). Amongst other sources, establishment of the National Association of the Advancement of Colored People in 1909, federally mandated segregation in 1913, the continuation of the Great Migration, race riots and Jim Crow laws across the country, the Harlem Renaissance, Prohibition, World War I, and the Great Depression nationally punctuated the life of St.LI. Locally, St.LI would have felt the impacts of when St. Louis passed the country's first residential segregation ordinance in 1916 (invalidated in 1918), and the 1917 St. Louis Race Riots. Further, St. Louis African American communities endured structural discrimination, such as crowded living conditions, poor sanitation, and inadequate access to health care (De la Cova 2011, de la Cova 2019).

In fact, racial health disparities in St. Louis came to a vote in 1923. A portion of a civic improvement bond went to build a new African American City Hospital (O'Neil 2014) (later named the Homer G. Phillips Hospital) due to the squalid conditions and lack of space in the

existing, public African American hospital. However, construction of the new hospital was not initiated until 1932 and not opened until 1937 due to city discord over where the new hospital would be built. The debate centered on whether the hospital should be independent or associated with the established white hospital (1929, O’Neil 2014). The delay and institutional mismanagement caused the public African American City Hospital to fall further into filth and disrepair throughout this period; by 1928 a New York hospital review stated: “sanitary conditions for the animals at Forest Park Zoo [In St. Louis] are better than those at the Negro hospital” (1928b). These conditions contributed to poor patient health outcomes (1932), such as that experienced by St.LI, and other Terry individuals, while also exacerbating negative health effects (1928a, Fisher 1933) of the poor biosocial conditions in St. Louis.

5.5.3 Recognizing Anatomies

The increasingly personal information and narratives obtainable from skeletal remains with the use of modern, interdisciplinary techniques reveal parts of St.LI’s lived experiences and death.. While for some communities, portraits like the one presented here may be valuable remembrances and an opportunity to tell untold histories, we recognize that this view is not universal. Especially, since it is well documented that historic anatomical collections were often dubiously sourced by taking unclaimed remains, specifically of poor and minority groups (de la Cova 2019, Jones & Whitaker 2012, Zimmer 2018, Zuckerman 2018). This is in large part because dissection was highly stigmatized, and dissection—especially bodies within reference communities—functioned as a form of structural violence, enforcement of postmortem social inequality, and punishment for marginalized communities (Hildebrandt 2008, Muller et al 2017), enforcing social inequalities experienced by living populations (Garment et al 2007, Jones & Whitaker 2012, Nystrom 2014). Social inequalities have led to many individuals being incorporated into skeletal collections like the Terry Collection and allow for their continual use in scientific endeavors. As a result of this history, discussions about the ethics of skeletal and genetic analysis of skeletal collections (Bardill et al 2018, Callaway 2017, Makarewicz et al 2017), recognizing the immense academic and medical contributions from collection individuals, and having a shared remembrance of the people in these reference collections is increasingly important.

In fact, anatomical and medical reference collections, including the Terry, have been fundamental to medical and forensic methods and procedures, including the development of sex, age, and ancestry estimation methods, and diagnostic criteria for pathological conditions (Humphrey 1973, Hunt & Albanese 2005, Zuckerman 2018). For reference collections, which may include individuals with living descendants, being cognizant of personally identifiable information (PII) and of anatomical personhood, especially in the age of genetic testing, needs to be a priority for medical and academic communities, and the wider public interested in genetic testing and genealogy. Active recognition of anatomical collections' social history, and the lived experiences of individuals within them, as well as how information gained from these collections affects living people, should be evaluated during project design (De la Cova 2011, Watkins 2018, Zuckerman 2018). Importantly, we are not advocating for anatomical collections or publications to specify names of skeletal individuals; rather, to recognize the individuals in reference collections as participants (Nystrom 2014) that provide broader social, biological, and historical understandings of the past and present. In doing so, the temporal and ideological gaps between researchers and the analyzed individuals can be bridged, literally showing how the past informs the present.

5.6 Conclusions

The present study reveals an incredible breadth and depth of information about the lived experiences and end-of-life circumstances for St.LI by considering multiple lines of evidence. St.LI has also revealed that dental calculus, once thought to be a cumulative substrate, can also incorporate short-term infectious diseases (e.g., pneumonia), potentially making numerous ancient to early-industrial health conditions (Omran 1971) microbially accessible. *Klebsiella pneumoniae* and *Acinetobacter* genomes have also been reconstructed for the first time from ancient remains and as direct contributors to COD. All the evidence culminates to depict a young, immunocompromised individual whose (ill)health experiences and death reflect their embodied biosocial conditions of the early 20th century U.S.

It is incumbent upon researchers using these collections to recognize the processes of inequality that led to the assemblage of many anatomical collections and to actively discuss how those systems continue to be enacted through scientific research (Smith 2013). A driving question

needs to be continually asked throughout the research process: for whom does the scientific knowledge being generated benefit? This is especially poignant when working with skeletal collections because the individuals being assessed can no longer benefit from the information they reveal. Normalizing the identification and the discussion of social inequalities deriving from colonial and racist paradigms is critical for accessible and meaningful scientific knowledge production.

5.7 Acknowledgments

We thank Sabrina Sholts and David Hunt for thought-provoking conversations, support, and access to the Terry Collection at the National Museum of Natural History; and Lesley Rankin-Hill, Tassie Hirschfeld, and Michael Blakey for thoughtful conversations about working with the Terry Collection and contextualizing St. Louis in the early 1900s. This research was supported by grants to the University of Oklahoma and Mississippi State University by the National Science Foundation (NSF BCS-1643318). Rita M. Austin is supported by a Smithsonian Pre-doctoral Research Fellowship.

Chapter 6

Conclusions

Museum collections are an undeniable resource to humanity for their ability to inform scientists and the public about past human experiences (Sholts et al 2016). With the destructive sampling of ancient human remains for molecular analysis, new and valuable information has been gained about past populations. Beyond population histories, insights into the evolution and variation of the human microbiome (Warinner 2016), dietary practices (Hendy et al 2018, Warinner et al 2014a), and pathogen evolution (Devault et al 2014, Hendy et al 2016, Vågene et al 2018, Warinner et al 2014b) have shaped the future of paleogenomics and paleoproteomics. However, the advances in molecular technologies have resulted in the dramatic increase of destructive sampling requests for anthropological and skeletal collections, making baseline knowledge of molecular taphonomy critical for optimizing research design and ensuring the longevity of museum materials (Austin et al 2019, McLean et al 2016, Thomas 1994).

Through the assessment of several Smithsonian Institution National Museum of Natural History (NMNH) skeletal collections', this dissertation integrated multiple lines of evidence and approaches, including archival documentation, skeletal assessments, and metagenomic analyses on dental calculus to provide applicable information future research optimization and to better inform decisions made by curators. The presented projects provide a dynamic understanding of museum collections as repositories of biological and historic processes. Specifically, archival research revealed collections' the variation in museum taphonomic histories. Population/collection-level skeletal survey for dental calculus quantified the presence or absence of the biofilm for both ante and postmortem contexts, showing biases in museum collections as a result of museum taphonomies. Metagenomic analysis of dental calculus ancient DNA (aDNA) from global populations demonstrate the impact of museum taphonomic processes on aDNA recovery and DNA fragment lengths, and the overlap in microbial profile taxa and taxa abundances between archaeological and museum samples. In assessing individuals from the Terry Collection, the shotgun and capture sequencing identified the presence of *Mycobacterium tuberculosis* in dental calculus. However, *Treponema pallidum* (causative agent of venereal syphilis) could not be recovered from dental calculus using shotgun metagenomics. Through a

case study approach, pathogen aDNA recovered from a single individual revealed acute, systemic infections with *Acinetobacter* species and identified a likely agent (*Klebsiella pneumoniae* str. 12208) in the lobar pneumonia cause of death. In-depth skeletal assessment of the case study individual revealed systemic and morphological health indicators of a probable case of tuberculosis infection and oral diseases. With these findings, this dissertation has highlighted the utility of museum collections for interdisciplinary and multi-methodological research.

The interdisciplinary and holistic perspectives of this dissertation emphasize their necessity and advantages for museum and ancient metagenomic research; various lines of evidence contribute to a more complete biological and historically contextualized understanding of the past. For instance, if only skeletal and metagenomic assessments had been conducted for the case study the broader historic context (e.g. segregation) and St. Louis-specific socioeconomic processes (e.g. decrepit hospital conditions) from exploring contemporary literature acting on the individual would have been missed.

In Chapter Two, overarching considerations and policies intrinsic to conducting biomolecular research using museum collections are explored, with particular attention given to working with human skeletal remains. A survey of several NMNH human skeletal collections revealed material and molecular preservation biases related to the artificial absence of skeletal elements within collections, as evidenced by the approximately 50% loss of collection dental calculus due to handling, accidental loss and/or curatorial cleaning. Additionally, this work showed that historic curation and preservation processes impacted total ancient DNA recovery, where museum samples showed consistently lower DNA yields than archaeological samples. Best practices related to accessing collections, community consultation, maintenance of collections for longevity, and knowledge dissemination and retention are also presented. These findings and suggestions will hopefully influence future destructive sampling decisions made by researchers and curators alike.

Chapter Three tested microbial community aDNA preservation across several NMNH skeletal collections. Archival research illuminated the acquisition and conservation histories of globally

representative collections. Metagenomic aDNA preservation within dental calculus reveals lower amounts of total aDNA from museum collections, but with authentic recoverable ancient dental calculus microbial signatures. Pragmatically, these results illustrate the collection-specific information and approach that is needed when conducting museum research due to differential taphonomic histories and molecular preservation. This information can also assist museum curators as they assess the molecular potential of specific collections in relation to the potential for the collection to address a set of research questions. With these molecular and archival overviews of collection taphonomies research and on-going collection curation can be optimized.

When focusing on the Robert J. Terry Anatomical Collection, Chapter Four explores the potential of dental calculus to be a metagenomic diagnostic tool to identify syphilis (*T. pallidum*) and tuberculosis (TB; *M. tuberculosis*) in past populations. These data suggest that *T. pallidum* likely cannot be recovered from dental calculus through shotgun metagenomics. *M. tuberculosis* DNA may be recovered from dental calculus with a shotgun metagenomic approach, however, the extremely short fragment lengths of the Terry Collection limited analysis with this approach. Using a novel in-solution hybridization capture approach, *M. tuberculosis* DNA was enriched, however lineage defining SNPs (Bos et al 2014) were not identified in Terry Collection samples. Importantly, Chapter Four results are impacted by the collection-specific preservation processes (i.e. benzene vapors) that were applied to Terry Collection individuals. Low aDNA recovery and very short fragment lengths seen from Terry Collection samples in Chapter Three confirm the overall molecular influence collection preparations processes on Terry Collection biomolecules. *M. tuberculosis* pathogen recovery from dental calculus presented here demonstrate hopeful expectations for future research investigating TB from dental calculus when applied to collections less impacted by taphonomic processes. More targeted investigations are required for *T. pallidum* research. In being able to recover pathogens from dental calculus and confirm diagnosis disease postmortem, the evolutionary histories of these pathogens can be investigated more widely and can potentially influence how we understand and treat diseases today.

Uniquely, because of the associated Terry Collection documentation, skeletal preservation, and the first application of metagenomic methods on the collection, Chapter Five presents a holistic case study revealing and contextualizing the life and death of a single individual. Skeletal

assessment revealed life-long development and immunological reactions of oral and systemic disease, and metagenomic analysis of the dental calculus indicated a dysbiotic oral microbiome. Amazingly, the reconstructed dental calculus microbial community confirmed the cause of death of lobar pneumonia, identified a potential contributing strain (*Klebsiella pneumonia* strain 12208), and revealed a probable hospital-acquired infection with *Acinetobacter* species. These results are impactful due to the recovery of information about an acute infection, as well as identification of pneumonia, which has been obscured in historic documentation and skeletally invisible as it does not leave indicative skeletal lesions (Mitchell 2017, Ortner 2003).

Museum skeletal collections are a direct reflection of the decision-making processes of bioanthropologists of the past. Future applications of the findings described here will hopefully inform other ancient metagenomic researchers using museum collections, as well as museum curators, about museum research and destructive sampling considerations. Molecular and methodological potential of specific collections based on their preservation history and metagenomic results is meaningful information for museum curators and managers who are tasked with weighing the expected scientific gain against the indisputable loss of collection materials. This information is critical in a time of increased sampling requests as well as increased competition within the ancient biomolecule field (Lewis-Kraus 2019b, Makarewicz et al 2017).

Results presented here also provide critical information about dental calculus as a biologically informative substrate for ancient metagenomic research. Metagenomic assessments of dental calculus have found specific and reliable dietary molecules (Hendy et al 2018, Warinner et al 2014a), helping to reconstruct dietary pathways, identify source species (e.g. goats, cows, etc.), and understand the spread of dairy subsistence (Wilkin et al 2020). In conjunction to the various types of molecules found in dental calculus (e.g. phytoliths, proteins, and DNA) and the easy accessibility of the substrate on human remains, targeting other molecules commonly found in dental calculus makes sense. However, from the presented data, which molecules can be targeted for diagnostic purposes, is dependent on sample/collection preservation and the taxa being investigated (*M. tuberculosis*, *T. pallidum*, or *K. pneumoniae*). The recovery of a skeletally invisible and acute disease like pneumonia is especially exciting. This work is impactful for

investigating diseases which do not leave skeletal markers due to their short disease period and the limited bone reactions that are possible (Ortner 2003).

The challenges associated with conducting research on the dead in this dissertation emphasize the need for on-going ethical discussions. This is particularly true for individuals in anatomical collections, such as those in the Terry Collection, who may have lived and died in the recent past. Innovative responses within scientific literature are also called for, as more personal information (e.g. names, genetic material) may be accessible through research on museum and anatomical individuals; making the past more and more personal (De La Cova 2010, De la Cova 2011, de la Cova 2019, Muller et al 2017, Watkins 2018). Through the personalized analyses and research that are now being conducted, narratives describing the disenfranchisement, exploitation, and social inequalities past peoples were subjected to are readily accessible and explored in scientific articles. However, in exposing past social inequalities, reflection on and normalizing discussion of how social inequalities persist needs to be brought to the forefront of scientific proceedings. Anatomical collections, while previously considered scientifically neutral, are not. New acknowledgements (e.g. thanking study individuals for their contribution) and regulations (e.g. requiring ethic statements regarding collection histories in publications) are needed. Engagement with potential descendant communities, regardless of their status under the Native American Graves Protection and Repatriation Act (1990), should also be incorporated and considered during research using museum and anatomical collections. While past wrongs cannot be undone or righted, current and future research can strive to not perpetuate or ignore social inequalities.

7 References

1908. Accession Record 49258. ed. S Institution. Washington, D.C.: National Museum of Natural History
- 1909a. Accession Record 50330. ed. S Institution. Washington, D.C.: National Museum of Natural History
- 1909b. Accession Record 50142. ed. S Institution. Washington, D.C.: National Museum of Natural History
1913. Accession Record 54928. ed. S Institution. Washington, D.C.: National Museum of Natural History
1923. Accession Record 070271. ed. S Institution. Washington, D.C.: National Museum of Natural History
1925. HOSPITAL IN ST. LOUIS IS MISMANAGED: FORMER RESIDENT DOCTORS MAKE CHARGES. *The Chicago Defender (National Edition) (1921-1967)*:
- 1928a. CALLS ST. LOUIS HOSPITAL UNFIT FOR PATIENTS: ST. LOUISANS SEE NEED OF NEW HOSPITAL. *The Chicago Defender (National Edition) (1921-1967)*:
- 1928b. Medics Face Scathing Critic Of Lax Conditions In ST. Louis Hospital: "In Grip O' Politic Intrigue" New York Physician Declares City Hospital No. 2 Should Be Taken Out Of Partisan Politics. *The Pittsburgh Courier (1911-1950)*:
- 1928c. ONE FRUIT OF SEGREGATION. *The Pittsburgh Courier (1911-1950)*:
1929. CITIZENS OF ST. LOUIS FIGHT FOR A DECENT HOSPITAL: COMMITTEE ASKS CITY TO BUY PORO COLLEGE PROPERTY. *New Journal and Guide (1916-2003)*:
1932. Patients Packed Bed-to-Bed at Negro Hospital: One Bath Tub for 149 Patients-- Contagious Cases Not Isolated. *St. Louis Star and Times*:
- 1950a. Accession Record 188059. ed. S Institution. Washington, D.C.: National Museum of Natural History
- 1950b. History. In *Catalog Card 188059*, ed. S Institution. Washington, D.C.: National Museum of Natural History
- December-March 1908-1909. "From My Journeys" 1898 -1938. In *Ales Hrdlicka Papers*. National Anthropology Archives: Smithsonian Institution
- July-September 1912. "From my Journeys" 1898-1938. In *Ales Hrdlicka Papers*, pp. 35, 44. National Anthropological Archives: Smithsonian Institution
- Adler CJ, Dobney K, Weyrich LS, Kaidonis J, Walker AW, et al. 2013. Sequencing ancient calcified dental plaque shows changes in oral microbiota with dietary shifts of the Neolithic and Industrial revolutions. *Nat Genet* 45: 450-5, 55e1
- Akcalı A, Lang NP. 2018. Dental calculus: the calcified biofilm and its role in disease development. *Periodontology* 2000 76: 109-15
- Alikhan N-F, Petty NK, Zakour NLB, Beatson SA. 2011. BLAST Ring Image Generator (BRIG): simple prokaryote genome comparisons. *BMC genomics* 12: 402
- Aminov R. 2017. History of antimicrobial drug discovery: Major classes and health impact. *Biochemical pharmacology* 133: 4-19
- Anderson RM. 1965. *Methods of collecting and preserving vertebrate animals*. Department of the Secretary of State.
- Antinick TC, Foran DR. 2018. Intra- and Inter-Element Variability in Mitochondrial and Nuclear DNA from Fresh and Environmentally Exposed Skeletal Remains. *J Forensic Sci*

- Austin RM, Sholts SB, Williams L, Kistler L, Hofman CA. 2019. Opinion: To curate the molecular past, museums need a carefully considered set of best practices. *Proceedings of the National Academy of Sciences* 116: 1471-74
- Baer PN, White CL. 1966. Studies on experimental calculus formation in the rat. IX. The effect of varying the protein and fat content of the diet on calculus deposition and alveolar bone loss. *J Periodontol* 37: 113-7
- Barberis I, Bragazzi N, Galluzzo L, Martini M. 2017. The history of tuberculosis: from the first historical records to the isolation of Koch's bacillus. *Journal of preventive medicine and hygiene* 58: E9
- Bardill J, Bader AC, Garrison NA, Bolnick DA, Raff JA, et al. 2018. Advancing the ethics of paleogenomics. *Science* 360: 384-85
- Batra P, Mathur P, Misra MC. 2016. *Aeromonas* spp.: an emerging nosocomial pathogen. *Journal of laboratory physicians* 8: 1
- Bellizzi C. 2014. How SI Staff Beat the Heat. In *Smithsonian Collections Blog*, ed. Esposito. si-siris.blogspot.com
- Bisharat N, Gorlachev T, Keness Y. 2012. 10-Years hospital experience in *Pseudomonas stutzeri* and literature. *Open Infect Dis J* 6: 21-4
- Bos KI, Harkins KM, Herbig A, Coscolla M, Weber N, et al. 2014. Pre-Columbian mycobacterial genomes reveal seals as a source of New World human tuberculosis. *Nature* 514: 494-7
- Bos KI, Schuenemann VJ, Golding GB, Burbano HA, Waglechner N, et al. 2011. A draft genome of *Yersinia pestis* from victims of the Black Death. *Nature* 478: 506-10
- Bouwman AS, Brown TA. 2005. The limits of biomolecular palaeopathology: ancient DNA cannot be used to study venereal syphilis. *Journal of Archaeological Science* 32: 703-13
- Brooke JS. 2012. *Stenotrophomonas maltophilia*: an emerging global opportunistic pathogen. *Clinical microbiology reviews* 25: 2-41
- Brundidge HT. 1932. Patients Packed Bed-To-Bed At Negro Hospital: One Bath Tub for 149 Patients—Contagious Cases Not Isolated. *The St. Louis Star and Times*:
- Buckley S, Usai D, Jakob T, Radini A, Hardy K. 2014. Dental Calculus Reveals Unique Insights into Food Items, Cooking and Plant Processing in Prehistoric Central Sudan. *Plos One* 9: e100808
- Buikstra J, Ubelaker D. 1994. Standards for Recording Human Remains Arkansas. *Archaeological Survey Research Series*
- Caffell AC, Roberts CA, Janaway RC, Wilson AS. 2001. Pressures on osteological collections: the importance of damage limitation: Archaeopress
- Callaway E. 2017. Stop hoarding ancient bones, plead archaeologists. *Nature*
- Caporaso JG, Kuczynski J, Stombaugh J, Bittinger K, Bushman FD, et al. 2010. QIIME allows analysis of high-throughput community sequencing data. *Nature methods* 7: 335
- Carøe C, Gopalakrishnan S, Vinner L, Mak SS, Sinding MHS, et al. 2017. Single-tube library preparation for degraded DNA. *Methods in Ecology and Evolution*
- Casewell M, Phillips I. 1977. Hands as route of transmission for *Klebsiella* species. *Br Med J* 2: 1315-17
- Cohen S, Tyrrell DA, Smith AP. 1991. Psychological stress and susceptibility to the common cold. *New England journal of medicine* 325: 606-12
- Comas I, Chakravarti J, Small PM, Galagan J, Niemann S, et al. 2010. Human T cell epitopes of *Mycobacterium tuberculosis* are evolutionarily hyperconserved. *Nature genetics* 42: 498

- Comas I, Coscolla M, Luo T, Borrell S, Holt KE, et al. 2013. Out-of-Africa migration and Neolithic coexpansion of *Mycobacterium tuberculosis* with modern humans. *Nature genetics* 45: 1176
- Coros A, DeConno E, Derbyshire KM. 2008. IS6110, a *Mycobacterium tuberculosis* complex-specific insertion sequence, is also present in the genome of *Mycobacterium smegmatis*, suggestive of lateral gene transfer among mycobacterial species. *Journal of bacteriology* 190: 3408-10
- Corruccini RS. 1974. An examination of the meaning of cranial discrete traits for human skeletal biological studies. *American Journal of Physical Anthropology* 40: 425-45
- Costello EK, Lauber CL, Hamady M, Fierer N, Gordon JI, Knight R. 2009. Bacterial community variation in human body habitats across space and time. *Science* 326: 1694-7
- Cowdry EV. 1920. Anatomy in Japan. *The Anatomical Record* 18: 67-95
- Crespo FA, Klaes CK, Switala AE, DeWitte SN. 2017. Do leprosy and tuberculosis generate a systemic inflammatory shift? Setting the ground for a new dialogue between experimental immunology and bioarchaeology. *American journal of physical anthropology* 162: 143-56
- Cullinan M, Ford P, Seymour G. 2009. Periodontal disease and systemic health: current status. *Australian dental journal* 54: S62-S69
- Curtis NG. 2003. Human remains: the sacred, museums and archaeology. *Public Archaeology* 3: 21-32
- Dabney J, Meyer M, Pääbo S. 2013. Ancient DNA damage. *Cold Spring Harbor perspectives in biology* 5: a012567
- Dangvard Pedersen D, Milner GR, Kolmos HJ, Boldsen JL. 2019. The association between skeletal lesions and tuberculosis diagnosis using a probabilistic approach. *International journal of paleopathology*
- Dawes C. 1970. Effects of diet on salivary secretion and composition. *J Dent Res* 49: 1263-73
- De La Cova C. 2010. Cultural patterns of trauma among 19th-century-born males in cadaver collections. *American anthropologist* 112: 589-606
- De la Cova C. 2011. Race, health, and disease in 19th-century-born males. *American journal of physical anthropology* 144: 526-37
- de la Cova C. 2019. Marginalized bodies and the construction of the Robert J. Terry anatomical skeletal collection: A promised land lost In *Bioarchaeology of Marginalized People*, pp. 133-55: Elsevier
- De La Fuente C, Flores S, Moraga M. 2013. DNA from human ancient bacteria: a novel source of genetic evidence from archaeological dental calculus. *Archaeometry* 55: 767-78
- de Melo FL, de Mello JCM, Fraga AM, Nunes K, Eggers S. 2010. Syphilis at the crossroad of phylogenetics and paleopathology. *PLoS neglected tropical diseases* 4: e575
- DeSantis TZ, Hugenholtz P, Larsen N, Rojas M, Brodie EL, et al. 2006. Greengenes, a chimera-checked 16S rRNA gene database and workbench compatible with ARB. *Applied and environmental microbiology* 72: 5069-72
- Devault AM, Golding GB, Waglechner N, Enk JM, Kuch M, et al. 2014. Second-pandemic strain of *Vibrio cholerae* from the Philadelphia cholera outbreak of 1849. *New England Journal of Medicine* 370: 334-40
- Dewhirst FE, Chen T, Izard J, Paster BJ, Tanner AC, et al. 2010. The human oral microbiome. *J Bacteriol* 192: 5002-17

- DeWitte SN, Stojanowski CM. 2015. The osteological paradox 20 years later: past perspectives, future directions. *Journal of Archaeological Research* 23: 397-450
- Dietrich T, Webb I, Stenhouse L, Pattni A, Ready D, et al. 2017. Evidence summary: the relationship between oral and cardiovascular disease. *British dental journal* 222: 381
- Edmondson DG, Hu B, Norris SJ. 2018. Long-term in vitro culture of the syphilis spirochete *Treponema pallidum* subsp. *pallidum*. *MBio* 9: e01153-18
- Eklund JA, Thomas MG. 2010. Assessing the effects of conservation treatments on short sequences of DNA in vitro. *Journal of Archaeological Science* 37: 2831-41
- El Attar M, Zaghloul M, El Menoufy H. 2010. Role of periodontitis in hospital-acquired pneumonia. *EMHJ-Eastern Mediterranean Health Journal*, 16 (5), 563-569, 2010
- Elder A, Madsen S, Brown G, Herbel C, Collins C, et al. 1997. Adhesives and consolidants in geological and paleontological conservation: A wall chart. *SPNHC Leaflets* 1: 1-4
- Falcao A, Bullón P. 2019. A review of the influence of periodontal treatment in systemic diseases. *Periodontology* 2000 79: 117-28
- Fisher R. 1933. THREE PATIENTS CROWDED INTO 2 HOSPITAL BEDS: ST. LOUIS HOSPITAL IS DECLARED UNFIT; NEW ONE SOUGHT. *New Journal and Guide (1916-2003)*:
- Fuchs TA, Abed U, Goosmann C, Hurwitz R, Schulze I, et al. 2007. Novel cell death program leads to neutrophil extracellular traps. *The Journal of cell biology* 176: 231-41
- Gafni M. 2018. Here's the 'open-source' genealogy DNA website that helped crack the Golden State Killer case. *Mercury News* 26
- Gage TB. 2005. Are modern environments really bad for us?: revisiting the demographic and epidemiologic transitions. *American Journal of Physical Anthropology: The Official Publication of the American Association of Physical Anthropologists* 128: 96-117
- Galagan JE. 2014. Genomic insights into tuberculosis. *Nature Reviews Genetics* 15: 307-20
- Garment A, Lederer S, Rogers N, Boulton L. 2007. Let the dead teach the living: The rise of body bequeathal in 20th-century America. *Academic Medicine* 82: 1000-05
- Gengo RG. 2014. "Race Becomes Biology": Co-occurring Oral and Systemic Disease as Embodiment of Structural Violence in an American Skeletal Sample.
- Gilbert MTP, Barnes I, Collins MJ, Smith C, Eklund J, et al. 2005. Long-term survival of ancient DNA in Egypt: Response to Zink and Nerlich (2003). *American journal of physical anthropology* 128: 110-14
- Ginolhac A, Rasmussen M, Gilbert MTP, Willerslev E, Orlando L. 2011. mapDamage: testing for damage patterns in ancient DNA sequences. *Bioinformatics* 27: 2153-55
- Gismondi A, D'Agostino A, Canuti L, Di Marco G, Martínez-Labarga C, et al. 2018. Dental calculus reveals diet habits and medicinal plant use in the Early Medieval Italian population of Colonna. *Journal of Archaeological Science: Reports* 20: 556-64
- Goldberg L. 1996. A history of pest control measures in the anthropology collections, National Museum of Natural History, Smithsonian Institution. *Journal of the American Institute for Conservation* 35: 23-43
- Gravlee CC. 2009. How race becomes biology: embodiment of social inequality. *American journal of physical anthropology* 139: 47-57
- Guyot-Revol V, Innes JA, Hackforth S, Hinks T, Lalvani A. 2006. Regulatory T cells are expanded in blood and disease sites in patients with tuberculosis. *American journal of respiratory and critical care medicine* 173: 803-10
- Hagelberg E, Sykes B, Hedges R. 1989. Ancient bone DNA amplified. *Nature* 342: 485

- Halperin EC. 2007. The poor, the black, and the marginalized as the source of cadavers in United States anatomical education. *Clinical Anatomy: The Official Journal of the American Association of Clinical Anatomists and the British Association of Clinical Anatomists* 20: 489-95
- Hangay G, Dingley M. 1985. Biological museum methods
- Hansen HB, Damgaard PB, Margaryan A, Stenderup J, Lynnerup N, et al. 2017. Comparing Ancient DNA Preservation in Petrous Bone and Tooth Cementum. *PLoS One* 12: e0170940
- Harkins KM, Buikstra JE, Campbell T, Bos KI, Johnson ED, et al. 2015. Screening ancient tuberculosis with qPCR: challenges and opportunities. *Philosophical Transactions of the Royal Society B: Biological Sciences* 370: 20130622
- Harper KN, Armelagos GJ. 2013. Genomics, the origins of agriculture, and our changing microbe-scape: Time to revisit some old tales and tell some new ones. *American journal of physical anthropology* 152: 135-52
- Harper KN, Zuckerman MK, Harper ML, Kingston JD, Armelagos GJ. 2011. The origin and antiquity of syphilis revisited: An Appraisal of Old World pre-Columbian evidence for treponemal infection. *American journal of physical anthropology* 146: 99-133
- Hendy J, Collins M, Teoh KY, Ashford DA, Thomas-Oates J, et al. 2016. The challenge of identifying tuberculosis proteins in archaeological tissues. *Journal of Archaeological Science* 66: 146-53
- Hendy J, Warinner C, Bouwman A, Collins MJ, Fiddyment S, et al. 2018. Proteomic evidence of dietary sources in ancient dental calculus. *Proceedings of the Royal Society B* 285
- Henry RS. 1964. *The Armed Forces Institute of Pathology: Its First Century, 1862-1962*. Office of the Surgeon General, Department of the Army.
- Higuchi R, Bowman B, Freiberger M, Ryder OA, Wilson AC. 1984. DNA sequences from the quagga, an extinct member of the horse family. *Nature* 312: 282-4
- Hildebrand M. 1968. *Anatomical preparations*. Univ of California Press.
- Hildebrandt S. 2008. Capital punishment and anatomy: history and ethics of an ongoing association. *Clinical Anatomy* 21: 5-14
- Hinton RJ. 1981. Form and patterning of anterior tooth wear among aboriginal human groups. *American Journal of Physical Anthropology* 54: 555-64
- Honap TP, Pfister L-A, Housman G, Mills S, Tarara RP, et al. 2018. Mycobacterium leprae genomes from naturally infected nonhuman primates. *PLoS neglected tropical diseases* 12
- Horai S, Hayasaka K, Murayama K, Wate N, KOIKE H, Nakai N. 1989. DNA amplification from ancient human skeletal remains and their sequence analysis. *Proceedings of the Japan Academy, series B* 65: 229-33
- Hornaday WT, Holland WJ. 1891. *Taxidermy and zoological collecting: a complete handbook for the amateur taxidermist, collector, osteologist, museum-builder, sportsman, and traveller*. C. Scribner's sons.
- Housman G, Malukiewicz J, Boere V, Grativol AD, Pereira LCM, et al. 2015. Validation of qPCR methods for the detection of Mycobacterium in New World animal reservoirs. *PLoS neglected tropical diseases* 9: e0004198
- Humphrey DC. 1973. Dissection and discrimination: the social origins of cadavers in America, 1760-1915. *Bulletin of the New York Academy of Medicine* 49: 819

- Hunt DR, Albanese J. 2005. History and demographic composition of the Robert J. Terry anatomical collection. *American Journal of Physical Anthropology: The Official Publication of the American Association of Physical Anthropologists* 127: 406-17
- Hübner R, Key FM, Warinner C, Bos KI, Krause J, Herbig A. 2019. HOPS: Automated detection and authentication of pathogen DNA in archaeological remains. *Genome Biology* 20: 1-13
- Irish JD. 2006. Who were the ancient Egyptians? Dental affinities among Neolithic through postdynastic peoples. *American Journal of Physical Anthropology: The Official Publication of the American Association of Physical Anthropologists* 129: 529-43
- Jeong C, Wilkin S, Amgalantugs T, Bouwman AS, Taylor WTT, et al. 2018. Bronze Age population dynamics and the rise of dairy pastoralism on the eastern Eurasian steppe. *Proceedings of the National Academy of Sciences* 115: E11248-E55
- Jin Y, Yip H-K. 2002. Supragingival calculus: formation and control. *Critical reviews in oral biology & medicine* 13: 426-41
- Jones DG, Whitaker MI. 2012. Anatomy's use of unclaimed bodies: reasons against continued dependence on an ethically dubious practice. *Clinical Anatomy* 25: 246-54
- Jonsson H, Ginolhac A, Schubert M, Johnson PL, Orlando L. 2013. mapDamage2.0: fast approximate Bayesian estimates of ancient DNA damage parameters. *Bioinformatics* 29: 1682-4
- Juzda E. 2009. Skulls, science, and the spoils of war: craniological studies at the United States Army Medical Museum, 1868–1900. *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences* 40: 156-67
- Jónsson H, Ginolhac A, Schubert M, Johnson PL, Orlando L. 2013. mapDamage2. 0: fast approximate Bayesian estimates of ancient DNA damage parameters. *Bioinformatics* 29: 1682-84
- Kerr N. 1988. A method of assessing periodontal status in archaeologically derived skeletal material. *Journal of Paleopathology* 2: 67-78
- Khairat R, Ball M, Chang C-CH, Bianucci R, Nerlich AG, et al. 2013. First insights into the metagenome of Egyptian mummies using next-generation sequencing. *Journal of applied genetics* 54: 309-25
- Kim SE, Park S-H, Park HB, Park K-H, Kim S-H, et al. 2012. Nosocomial *Pseudomonas putida* bacteremia: high rates of carbapenem resistance and mortality. *Chonnam medical journal* 48: 91-95
- Kistler L, Ware R, Smith O, Collins M, Allaby RG. 2017. A new model for ancient DNA decay based on paleogenomic meta-analysis. *Nucleic Acids Res* 45: 6310-20
- Knights D, Kuczynski J, Charlson ES, Zaneveld J, Mozer MC, et al. 2011. Bayesian community-wide culture-independent microbial source tracking. *Nat Methods* 8: 761-3
- Ko W-C, Paterson DL, Sagnimeni AJ, Hansen DS, Von Gottberg A, et al. 2002. Community-acquired *Klebsiella pneumoniae* bacteremia: global differences in clinical patterns. *Emerging infectious diseases* 8: 160
- Komar DA, Grivas C. 2008. Manufactured populations: what do contemporary reference skeletal collections represent? A comparative study using the Maxwell Museum documented collection. *American Journal of Physical Anthropology: The Official Publication of the American Association of Physical Anthropologists* 137: 224-33
- Krieger N. 2014. Discrimination and health inequities. In *Social Epidemiology*, ed. L Berkman, I Kawachi, M Glymour, pp. 63-125. New York: Oxford University Press

- Kuzawa CW, Gravlee CC. 2016. Beyond genetic race: biocultural insights into the causes of racial health disparities. *New Directions in Biocultural Anthropology*: 89-106
- Lambert PM, Walker PL. 2018. BIOARCHAEOLOGICAL ETHICS: PERSPECTIVES ON THE USE AND VALUE OF HUMAN REMAINS IN SCIENTIFIC RESEARCH. *Biological Anthropology of the Human Skeleton*: 1-42
- LaMotta V, Schiffer M. 2005. Archaeological formation processes. *Archaeology: The Key Concepts*: 121-7
- Langmead B, Salzberg SL. 2012. Fast gapped-read alignment with Bowtie 2. *Nat Methods* 9: 357-9
- Leikind MC. 1954. Army Medical Museum and Armed Forces Institute of Pathology in Historical Perspective. *The Scientific Monthly* 79: 71-78
- Lewis-Kraus G. 2019a. Is Ancient DNA Research Revealing New Truths — or Falling Into Old Traps? *New York Times*:
- Lewis-Kraus G. 2019b. Is ancient DNA research revealing new truths—or falling into old traps. *New York Times Magazine* 17
- Li H, Durbin R. 2009. Fast and accurate short read alignment with Burrows–Wheeler transform. *bioinformatics* 25: 1754-60
- Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, et al. 2009. The sequence alignment/map format and SAMtools. *Bioinformatics* 25: 2078-79
- Lieverse AR. 1999. Diet and the aetiology of dental calculus. *International Journal of Osteoarchaeology* 9: 219-32
- Lieverse AR, Link DW, Bazaliiskiy VI, Goriunova OI, Weber AW. 2007. Dental health indicators of hunter–gatherer adaptation and cultural change in Siberia's Cis-Baikal. *American Journal of Physical Anthropology* 134: 323-39
- Linossier A, Gajardo M, Olavarria J. 1995. Paleomicrobiological study in dental calculus: *Streptococcus mutans*. *Scanning microscopy* 10: 1005-13; discussion 14
- Lock M. 1993. Cultivating the body: Anthropology and epistemologies of bodily practice and knowledge. *Annual review of anthropology* 22: 133-55
- Lopez-Polin L. 2012. Possible interferences of some conservation treatments with subsequent studies on fossil bones: A conservator's overview. *Quaternary International* 275: 120-27
- Lugonja B, Yeo L, Milward MR, Smith D, Dietrich T, et al. 2016. Periodontitis prevalence and serum antibody reactivity to periodontal bacteria in primary Sjögren's syndrome: a pilot study. *Journal of Clinical Periodontology* 43: 26-33
- Mackie M, Hendy J, Lowe AD, Sperduti A, Holst M, et al. 2017. Preservation of the metaproteome: variability of protein preservation in ancient dental calculus. *Sci Technol Archaeol Res* 3: 74-86
- Makarewicz C, Marom N, Bar-Oz G. 2017. Ensure equal access to ancient DNA. *Nature* 548: 158
- MaMing R, Lee L, Yang X, Buzzard P. 2016. Vultures and sky burials on the Qinghai-Tibet Plateau. *Vulture News* 71: 22-35
- Mandel ID. 1974. Biochemical aspects of calculus formation: II. Comparative Studies of Saliva in Heavy and Light Calculus Formers. *Journal of periodontal research* 9: 211-21
- Mann AE, Sabin S, Ziesemer K, Vagene AJ, Schroeder H, et al. 2018. Differential preservation of endogenous human and microbial DNA in dental calculus and dentin. *Sci Rep* 8: 9822

- Marais BJ, Loennroth K, Lawn SD, Migliori GB, Mwaba P, et al. 2013. Tuberculosis comorbidity with communicable and non-communicable diseases: integrating health services and control efforts. *The Lancet infectious diseases* 13: 436-48
- Marota I, Basile C, Ubaldi M, Rollo F. 2002. DNA decay rate in papyri and human remains from Egyptian archaeological sites. *American Journal of Physical Anthropology: The Official Publication of the American Association of Physical Anthropologists* 117: 310-18
- Martin M. 2011. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet. journal* 17: 10-12
- McDonough MM, Parker LD, Rotzel McInerney N, Campana MG, Maldonado JE. 2018. Performance of commonly requested destructive museum samples for mammalian genomic studies. *Journal of Mammalogy*: gyy080-gyy80
- McEwen BS, Wingfield JC. 2003. The concept of allostasis in biology and biomedicine. *Hormones and behavior* 43: 2-15
- McLean BS, Bell KC, Dunnum JL, Abrahamson B, Colella JP, et al. 2016. Natural history collections-based research: progress, promise, and best practices. *J Mammal* 97: 287-97
- Meffray A, Perrin M, Richier A, Schmitt A, Ardagna Y, Biagini P. 2019. Molecular detection of *Treponema pallidum* subspecies *pallidum* in 150-year-old foetal remains, southeastern France. *Journal of medical microbiology* 68: 761-69
- Meyer M, Kircher M. 2010. Illumina sequencing library preparation for highly multiplexed target capture and sequencing. *Cold Spring Harbor Protocols* 2010: pdb. prot5448
- Mezzatesta ML, Gona F, Stefani S. 2012. Enterobacter cloacae complex: clinical impact and emerging antibiotic resistance. *Future microbiology* 7: 887-902
- Mitchell PD. 2017. Improving the use of historical written sources in paleopathology. *International journal of paleopathology* 19: 88-95
- Molnar P. 2008. Dental wear and oral pathology: possible evidence and consequences of habitual use of teeth in a Swedish Neolithic sample. *American Journal of Physical Anthropology: The Official Publication of the American Association of Physical Anthropologists* 136: 423-31
- Mooney MP, Kraus EM, Bardach J, Snodgrass JJ. 1982. Skull preparation using the enzyme-active detergent technique. *The Anatomical Record* 202: 125-29
- Morris AG. 1987. The reflection of the collector: San and Khoi skeletons in museum collections. *The South African Archaeological Bulletin*: 12-22
- Muller JL, Pearlstein KE, de la Cova C. 2017. Dissection and documented skeletal collections: Embodiments of legalized inequality In *The bioarchaeology of dissection and autopsy in the United States*, pp. 185-201: Springer
- Munoz-Price LS, Weinstein RA. 2008. Acinetobacter infection. *New England Journal of Medicine* 358: 1271-81
- Musher DM. 1991. Syphilis, neurosyphilis, penicillin, and AIDS. *Journal of Infectious Diseases* 163: 1201-06
- Nerlich AG, Haas CJ, Zink A, Szeimies U, Hagedorn HG. 1997. Molecular evidence for tuberculosis in an ancient Egyptian mummy. *The Lancet* 350: 1404
- Nystrom KC. 2014. The Bioarchaeology of Structural Violence and Dissection in the 19th-Century United States. *American Anthropologist* 116: 765-79
- Oh J, Byrd AL, Park M, Kong HH, Segre JA, Program NCS. 2016. Temporal stability of the human skin microbiome. *Cell* 165: 854-66

- Omran AR. 1971. The epidemiologic transition. A theory of the epidemiology of population change. *Milbank Mem Fund Q* 49: 509-38
- Omstein L. 2010. Poisonous heritage: pesticides in museum collections.
- Organization WH. 2019. Global tuberculosis report 2019 *C C B Y- N C- S A 3 . 0 I G O*
- Ortner DJ. 2003. *Identification of pathological conditions in human skeletal remains*. Academic Press.
- Ozga AT, Nieves-Colon MA, Honap TP, Sankaranarayanan K, Hofman CA, et al. 2016. Successful enrichment and recovery of whole mitochondrial genomes from ancient human dental calculus. *Am J Phys Anthropol* 160: 220-8
- O'Neil T. 2014. Look Back 250 • Crusading lawyer becomes namesake of new hospital for black St. Louis. In *St. Louis Post-Dispatch*
- Paabo S. 1985. Preservation of DNA in ancient Egyptian mummies. *Journal of Archaeological Science* 12: 411-17
- Paabo S, Gifford JA, Wilson AC. 1988. Mitochondrial DNA sequences from a 7000-year old brain. *Nucleic Acids Res* 16: 9775-87
- Paczosa MK, Meccas J. 2016. *Klebsiella pneumoniae*: going on the offense with a strong defense. *Microbiol. Mol. Biol. Rev.* 80: 629-61
- Padro CJ, Sanders VM. *Seminars in immunology* 2014, 26: 357-68. Elsevier.
- Peck WH. 1998. *Mummies of ancient Egypt*. na.
- Pedersen DD, Milner GR, Kolmos HJ, Boldsen JL. 2019. The association between skeletal lesions and tuberculosis diagnosis using a probabilistic approach. *International journal of paleopathology*
- Pepperell CS, Granka JM, Alexander DC, Behr MA, Chui L, et al. 2011. Dispersal of *Mycobacterium tuberculosis* via the Canadian fur trade. *Proc Natl Acad Sci U S A* 108: 6526-31
- Peterman TA, Kidd SE. 2019. Trends in Deaths due to Syphilis, United States, 1968—2015. *Sexually transmitted diseases* 46: 37
- Pinhasi R, Fernandes D, Sirak K, Novak M, Connell S, et al. 2015. Optimal ancient DNA yields from the inner ear part of the human petrous bone. *PLoS One* 10: e0129102
- Powers N. 2012. Human osteology method statement. *London: Museum of London*
- Prendini L, Hanner R, DeSalle R. 2002. Obtaining, storing and archiving specimens and tissue samples for use in molecular studies In *Techniques in molecular systematics and evolution*, pp. 176-248: Springer
- Prevention CfDca. 2015. U.S. Public Health Service Syphilis Study at Tuskegee. ed. CfDca Prevention. <https://www.cdc.gov/tuskegee/>
- Prevention CfDca. 2019a. Sexually Transmitted Disease Surveillance 2018. Atlanta, GA: U.S. Department of Health and Human Services
- Prevention CfDca. 2019b. Trends in Tuberculosis, 2018. ed. DoT Elimination
- R Core Team. 2016. R: A Language and Environment for Statistical Computing. Vienna, Austria
- Radini A, Nikita E, Buckley S, Copeland L, Hardy K. 2017. Beyond food: The multiple pathways for inclusion of materials into ancient dental calculus. *Am J Phys Anthropol* 162 Suppl 63: 71-83
- Radini A, Tromp M, Beach A, Tong E, Speller C, et al. 2019. Medieval women's early involvement in manuscript production suggested by lapis lazuli identification in dental calculus. *Science Advances* 5: eaau7126

- Redman SJ. 2012. *Human Remains and the Construction of Race and History, 1897-1945*. UC Berkeley
- Rohart F, Gautier B, Singh A, Lê Cao K-A. 2017. mixOmics: An R package for ‘omics feature selection and multiple data integration. *PLoS computational biology* 13: e1005752
- Rohland N, Harney E, Mallick S, Nordenfelt S, Reich D. 2015. Partial uracil–DNA–glycosylase treatment for screening of ancient DNA. *Phil. Trans. R. Soc. B* 370: 20130624
- Rohland N, Hofreiter M. 2007. Ancient DNA extraction from bones and teeth. *Nat Protoc* 2: 1756-62
- Rohland N, Reich D. 2012. Cost-effective, high-throughput DNA sequencing libraries for multiplexed target capture. *Genome research* 22: 939-46
- Rowley J. 1925. *Taxidermy and museum exhibition*. USA.
- Rusinovci S, Sejdini M, Salihu S, Stubljär D, Haliti N. 2018. Bacteria That Cause Dentoalveolar Abscesses after Failed Endodontic Treatments: A Pilot Study. *Journal of International Dental and Medical Research* 11: 823-29
- Ryder MI. 2010. Comparison of neutrophil functions in aggressive and chronic periodontitis. *Periodontol* 2000 53: 124-37
- Sabin S, Herbig A, Vagene AJ, Ahlström T, Bozovic G, et al. 2019. A seventeenth-century Mycobacterium tuberculosis genome supports a Neolithic emergence of the Mycobacterium tuberculosis complex. *BioRxiv*
- Santos AL, Roberts CA. 2006. Anatomy of a serial killer: differential diagnosis of tuberculosis based on rib lesions of adult individuals from the Coimbra Identified Skeletal Collection, Portugal. *American Journal of Physical Anthropology: The Official Publication of the American Association of Physical Anthropologists* 130: 38-49
- Sarbu I, Matei C, Benea V, Georgescu S. 2014. Brief history of syphilis. *Journal of medicine and life* 7: 4
- Scannapieco FA, Bush RB, Paju S. 2003. Associations between periodontal disease and risk for nosocomial bacterial pneumonia and chronic obstructive pulmonary disease. A systematic review. *Annals of periodontology* 8: 54-69
- Scholz M, Ward DV, Pasolli E, Tolio T, Zolfo M, et al. 2016. Strain-level microbial epidemiology and population genomics from shotgun metagenomics. *Nature methods* 13: 435
- Schroeder HE. 1969. Formation and inhibition of dental calculus. *J Periodontol* 40: 643-6
- Schroeder HE, Shanley D. 1969. Formation and inhibition of dental calculus. *Journal of Periodontology-Periodontics* 40: 643-46
- Schubert M, Lindgreen S, Orlando L. 2016. AdapterRemoval v2: rapid adapter trimming, identification, and read merging. *BMC Res Notes* 9: 88
- Schuenemann VJ, Lankapalli AK, Barquera R, Nelson EA, Hernandez DI, et al. 2018. Historic Treponema pallidum genomes from Colonial Mexico retrieved from archaeological remains. *PLoS neglected tropical diseases* 12
- Schuenemann VJ, Peltzer A, Welte B, Van Pelt WP, Molak M, et al. 2017. Ancient Egyptian mummy genomes suggest an increase of Sub-Saharan African ancestry in post-Roman periods. *Nature communications* 8: 15694
- Schuenemann VJ, Singh P, Mendum TA, Krause-Kyora B, Jager G, et al. 2013. Genome-wide comparison of medieval and modern Mycobacterium leprae. *Science* 341: 179-83

- Scriba TJ, Penn-Nicholson A, Shankar S, Hraha T, Thompson EG, et al. 2017. Sequential inflammatory processes define human progression from M. tuberculosis infection to tuberculosis disease. *PLoS pathogens* 13
- Sefton A. 2001. The Great Pox that was... syphilis. *Journal of applied microbiology* 91: 592-96
- Sepkowitz KA. 1996. How contagious is tuberculosis? *Clinical Infectious Diseases* 23: 954-62
- Shaddox L, Wiedey J, Bimstein E, Magnuson I, Clare-Salzler M, et al. 2010. Hyper-responsive phenotype in localized aggressive periodontitis. *Journal of dental research* 89: 143-48
- Sholts SB, Bell JA, Rick TC. 2016. Ecce homo: Science and society need anthropological collections. *Trends Ecol Evol* 31: 580-83
- Shweta S. 2013. Dental abscess: A microbiological review. *Dental research journal* 10: 585
- Simmons NB, Voss RS. 2009. Collection, preparation, and fixation of specimens and tissues. *Ecological and Behavioral Methods for the Study of Bats*. Johns Hopkins University Press, Baltimore
- Singh AE, Romanowski B. 1999. Syphilis: review with emphasis on clinical, epidemiologic, and some biologic features. *Clinical microbiology reviews* 12: 187-209
- Smith LT. 2013. *Decolonizing methodologies: Research and indigenous peoples*. Zed Books Ltd.
- Smith LW, Baer PN, King CT, White CL. 1963. Studies on Experimental Calculus Formation in the Rat III. Calculus Formation as Influenced by High Fat, High Carbohydrate Diets, and Sucrose in the Drinking Water. *Journal of Periodontology* 34: 327-29
- Socransky SS, Haffajee AD. 2005. Periodontal microbial ecology. *Periodontol 2000* 38: 135-87
- Spyrou MA, Bos KI, Herbig A, Krause J. 2019. Ancient pathogen genomics as an emerging tool for infectious disease research. *Nature Reviews Genetics*: 1
- Srinivasan M, Sedmak D, Jewell S. 2002. Effect of fixatives and tissue processing on the content and integrity of nucleic acids. *Am J Pathol* 161: 1961-71
- Steckel RH, Larsen CS, Sciulli PW, Walker PL. 2006. Data collection codebook. *The Global History of Health Project* 2006: 1-41
- Stewart T. 1942. Notes on skulls from Germany sent at various times by Prof. Hansemann. In *Catalog Card 051505*, ed. S Institution. Washington, D.C.: National Museum of Natural History
- Stone AC, Wilbur AK, Buikstra JE, Roberts CA. 2009. Tuberculosis and leprosy in perspective. *Am J Phys Anthropol* 140 Suppl 49: 66-94
- Storch PS. 2003. Field and laboratory methods for handling osseous materials. Available from: *Minnesota Historical Society, St. Paul* http://www.mnhs.org/preserve/conservation/reports/osseous_materials.pdf (accessed 1 Aug 05)
- Straub RH. 2011. Concepts of evolutionary medicine and energy regulation contribute to the etiology of systemic chronic inflammatory diseases. Elsevier
- Sumi Y, Miura H, Michiwaki Y, Nagaosa S, Nagaya M. 2007. Colonization of dental plaque by respiratory pathogens in dependent elderly. *Archives of gerontology and geriatrics* 44: 119-24
- Talwar A, Tsang C, Price S, Pratt R, Walker W, et al. 2019. Tuberculosis — United States, 2018. *MMWR Morb Mortal Wkly Rep* 68
- Tampa M, Sarbu I, Matei C, Benea V, Georgescu S. 2014. Brief history of syphilis. *Journal of medicine and life* 7: 4
- Thomas R. 1994. Analysis of DNA from natural history museum collections In *Molecular ecology and evolution: approaches and applications*, pp. 311-21: Springer

- Thorne N, Borrell S, Evans J, Magee J, de Viedma DG, et al. 2011. IS6110-based global phylogeny of *Mycobacterium tuberculosis*. *Infection, Genetics and evolution* 11: 132-38
- Trotter M. 1981. Robert J. Terry, 1871–1966. *American journal of physical anthropology* 56: 503-08
- Truong DT, Franzosa EA, Tickle TL, Scholz M, Weingart G, et al. 2015. MetaPhlAn2 for enhanced metagenomic taxonomic profiling. *Nature methods* 12: 902
- Van Crevel R, Ottenhoff TH, Van Der Meer JW. 2002. Innate immunity to *Mycobacterium tuberculosis*. *Clinical microbiology reviews* 15: 294-309
- Velsko IM, Frantz LA, Herbig A, Larson G, Warinner C. 2018. Selection of appropriate metagenome taxonomic classifiers for ancient microbiome research. *MSystems* 3: e00080-18
- Velsko IM, Yates JAF, Aron F, Hagan RW, Frantz LA, et al. 2019. Microbial differences between dental plaque and historic dental calculus are related to oral biofilm maturation stage. *Microbiome* 7: 102
- Visca P, Seifert H, Towner KJ. 2011. *Acinetobacter* infection—an emerging threat to human health. *IUBMB life* 63: 1048-54
- Vogel G. 2019. Natural history museums face their own past. American Association for the Advancement of Science
- Vågene ÅJ, Herbig A, Campana MG, García NMR, Warinner C, et al. 2018. *Salmonella enterica* genomes from victims of a major sixteenth-century epidemic in Mexico. *Nature ecology & evolution* 2: 520
- Warinner C. 2016. Dental Calculus and the Evolution of the Human Oral Microbiome. *J Calif Dent Assoc* 44: 411-20
- Warinner C, Hendy J, Speller C, Cappellini E, Fischer R, et al. 2014a. Direct evidence of milk consumption from ancient human dental calculus. *Sci Rep* 4: 7104
- Warinner C, Rodrigues JF, Vyas R, Trachsel C, Shved N, et al. 2014b. Pathogens and host immunity in the ancient human oral cavity. *Nat Genet* 46: 336-44
- Warinner C, Speller C, Collins MJ. 2015a. A new era in palaeomicrobiology: prospects for ancient dental calculus as a long-term record of the human oral microbiome. *Philos Trans R Soc Lond B Biol Sci* 370: 20130376
- Warinner C, Speller C, Collins MJ, Lewis CM, Jr. 2015b. Ancient human microbiomes. *J Hum Evol* 79: 125-36
- Watkins R. 2018. Anatomical collections as the anthropological other: Some considerations In *Bioarchaeological Analyses and Bodies*, pp. 27-47: Springer
- White DJ. 1991. Processes contributing to the formation of dental calculus. *Biofouling* 4: 209-18
- White DJ. 1997. Dental calculus: recent insights into occurrence, formation, prevention, removal and oral health effects of supragingival and subgingival deposits. *European journal of oral sciences* 105: 508-22
- Wickham H. 2011. ggplot2. *Wiley Interdisciplinary Reviews: Computational Statistics* 3: 180-85
- Wilcox UV. 1995. Detached storage: the Smithsonian Institution's Museum Support Center. *Museum International* 47: 18-22
- Wilkin S, Miller AV, Taylor WT, Miller BK, Hagan RW, et al. 2020. Dairy pastoralism sustained eastern Eurasian steppe populations for 5,000 years. *Nature ecology & evolution* 4: 346-55
- Wirth T, Hildebrand F, Allix-Béguec C, Wölbeling F, Kubica T, et al. 2008. Origin, spread and demography of the *Mycobacterium tuberculosis* complex. *PLoS pathogens* 4

- Wong D, Nielsen TB, Bonomo RA, Pantapalangkoor P, Luna B, Spellberg B. 2017. Clinical and pathophysiological overview of Acinetobacter infections: a century of challenges. *Clinical microbiology reviews* 30: 409-47
- Wyres KL, Holt KE. 2016. Klebsiella pneumoniae population genomics and antimicrobial-resistant clones. *Trends in microbiology* 24: 944-56
- Zhou Y, Gao H, Mihindukulasuriya KA, La Rosa PS, Wylie KM, et al. 2013. Biogeography of the ecosystems of the healthy human body. *Genome biology* 14: R1
- Ziesemer KA, Mann AE, Sankaranarayanan K, Schroeder H, Ozga AT, et al. 2015. Intrinsic challenges in ancient microbiome reconstruction using 16S rRNA gene amplification. *Scientific Reports* 5: 16498
- Zimmer AN. 2018. More Than the Sum Total of Their Parts: Restoring Identity by Recombining a Skeletal Collection with Its Texts In *Bioarchaeological Analyses and Bodies*, pp. 49-69: Springer
- Zuckerman M. In Press. Identity and Health: Exploring Relationships between Health, Disease, and Identity in Past Populations. In *Identity Revisited: The Bioarchaeology of Identity in the Americas and Beyond*, ed. C Stojanowski, K Knudson. Gainesville, FL: University of Florida Press
- Zuckerman MK. 2018. Recovering the lived body from bodies of evidence: Interrogation of diagnostic criteria and parameters for disease ecology reconstructed from skeletons within anatomical and medical anatomical collections In *Bioarchaeological Analyses and Bodies*, pp. 89-107: Springer
- Zuckerman MK, Sandhu KA. Translating between paleopathology and paleoepidemiology: Further incorporation of evidence-based diagnostic criteria for chronic infectious disease into standardized skeletal data recording systems. *American Journal of Physical Anthropology*
- Šmajš D, Norris SJ, Weinstock GM. 2012. Genetic diversity in Treponema pallidum: implications for pathogenesis, evolution and molecular diagnostics of syphilis and yaws. *Infection, Genetics and Evolution* 12: 191-202

Appendix A: Chapter 3

Supplementary Materials: Metagenomic Analysis of Dental Calculus Reveals Differential Preservation Trends Across Smithsonian's Biological Anthropology Collections

Multiple Microbial Community Analyses

Microbial community analyses were conducted twice due to the late discovery of artificial base pair runs in the SI collections' R1 and R2 sequences after quality filtering and adapter trimming reads sequenced (excluding the Terry Collection), which were all ran on the same sequencing lane. Terry Collection, archaeological(Mann et al 2018) , and modern samples (Velsko et al 2019) did not have this issue, which were each sequenced independently. Additionally, the poly base pair runs were retained following high (30) base quality filtering. This suggests the base pair runs are a sequencing artifact. Due to low merging rates, the trimmed R1, R2, and merged sequencing files were originally included in all analyses. By including reads and the R2 files containing the artificial bases, fragment length distributions were altered when analyzed with MetaPhlAn2, resulting numerous long fragments (100+bp) being detected in the original analyses (Supplementary Figure 6D).

However, once the poly base pair runs at the end of trimmed R1 reads and in the trimmed R2 file were identified, analysis was redone to only include trimmed R1 files that had poly base pair runs removed. To remove artificial base pairs cutadapt (Martin 2011) was applied to R1 reads before following the computational steps described in the main text. Importantly, only the trimmed and cutadapt R1, and trimmed and merged files were included in the reanalysis because trimmed R2 files mostly contained artificial sequences. All subsequent processing and analysis steps remained the same across analyses.

Inaccurate sPLS-DA Results

Prior to finding the artificial poly base pair runs in trimmed R1 and R2 files, all MetaPhlAn2 taxonomic assessment were run with trimmed R1 and R2, and trimmed and merged files. Sparse partial least squares discriminant analysis (sPLS-DA) was run with the resulting species level taxonomic table, which had some differing results than those presented in the main text. When

all SI collections and comparative datasets are included, sPLS-DA continues to differentiate modern dental calculus samples away from collections and archaeological samples (Supplementary Figure 7B). However, with the removal of modern dental calculus sPLS-DA shows archaeological root samples differentiating from archaeological dental calculus and SI collections (Supplementary Figure 7C). This contrasts the accurate analysis presented in the main text where archaeological root and dental calculus continue to cluster with SI collections (Figure 7C).

sPLS-DA also identifies more and different species contributing the most to distinguishing datasets (Figure 10A and Supplementary Figure 8A). Analysis also shows archaeological root to distinguish among collections, where archaeological dental calculus was originally identified as contributing the most distinct communities. However, the Kharga Oasis sample is consistently identified as contributing species that distinguish among species in both the original (Supplementary Figure 8A) and updated (Figure 10A) analyses. When the Kharga Oasis and modern samples are removed, sPLS-DA instead identifies archaeological calculus, followed by the AMM collection, to contribute the most to distinguishing groups (Supplementary Figure 8B). While consistent in identifying archaeological calculus as the highest contributor, sPLS-DA then finds the Switzerland collection as the next group contributing the most to the updated analysis (Figure 10C).

The most dramatic alteration to results was seen when collection types were assessed. In the original sPLS-DA analysis only the modern dental calculus was notably different from the archaeological and SI collections (Supplementary Figure 9B). In the updated results (Figure 7B), SI collections are much more distinct from archaeological samples, which are distinctly found in between SI collections and modern dental calculus. These differences between analyses emphasize the necessity for the updated analysis described in the main text.

Supplementary Tables

Supplementary Table 1. Archival information referencing contextual and preservational insights for the analyzed individuals and collections.

SI catalog#	Locality (as given with remains)	Collector	Catalog Card Notes	Accession Notes (information pertains to all collection individuals)	NAA quotes
P256388-0	Egypt: Upper Egypt, Lisht: Near Matanieh (Al Mataniyah)	Hrdlicka: Egyptain Expedition	Locality: Rock-tombs; People: Egyptian XX-XXV Dyn; Remarks: Female. Damaged. Almost complete skeleton	"The skeletal remains we are getting date from about 2700 BC, and in part from a period between 1100 and 700 BC"-Hrdlicka (January 1909 letter, Accession Record 50330)	"All the bones from these excavations, and there are great many though mostly damaged, are brought to me to the "laboratory", where I examine everything, and make selection for the National Museum. But the bones are in exceedingly fragile state, so that the skulls have to be packed into individual little boxes which then go into a large case, yet even thus there is much apprehension as to their fate" ("From My Journeys" 1898 - 1938, December-March 1908-1909).
P256393-0	Egypt: Upper Egypt, Lisht: Near Matanieh (Al Mataniyah)	Hrdlicka: Egyptain Expedition	Locality: Rock-tombs; People: Egyptian XX-XXV Dyn; Remarks: Female. Almost complete skeleton		Superficial burials were noted to likely be of "poor people", while the pits were seen as unique and probably belonging to separate familie. With the Lisht cemetery containing officials and
P256396-0	Egypt: Upper Egypt, Lisht: Near Matanieh (Al Mataniyah)	Hrdlicka: Egyptain Expedition	Locality: Rock-tombs; People: Egyptian XX-XXV Dyn; Remarks: Female.		

P256406-0	Egypt: Upper Egypt, Lisht: Near Matanieh (Al Mataniyah)	Hrdlicka: Egyptain Expedition	Locality: Rock- tombs; People: Egyptian XX-XXV Dyn; Remarks: Female. Almost complete skeleton		"richer people" ("From My Journeys" 1898 -1938, December-March 1908- 1909).
P256560-0	Egypt: Upper Egypt, Lisht: Near Matanieh (Al Mataniyah)	Hrdlicka: Egyptain Expedition	Locality: Rock- tombs; People: Egyptian XX-XXV Dyn; Remarks: Adult Female. Almost complete skeleton	"...for as tomb and pit after pit are opened they are found plundered, which means that all of commercial value has been taken and the rest destroyed or badly damaged. This plundering has been practiced ... from very early times, and has been carried out to almost incredible extremes"- Hrdlicka (January 1909 letter, Accession Record 50330)	"Another mummy. They are encrusted with sediments that have already turned more or less into stone, also with gravelly salt, and are often as dry as old parchment and as tough. Now and then the rocky crust must be chiseled and hammered off, and the body dismembered, to be made ready for maceration, with a saw" ("From My Journeys" 1898 -1938, December-March 1908- 1909).
P258519-0	Egypt: Upper Egypt: Kharga Oasis: El Baguat	Hrdlicka: Egyptain Expedition	Locality: El Baguat; People: Egyptian; Remarks: Male. Almost complete skeleton		
P258537-0	Egypt: Upper Egypt: Kharga Oasis: El Baguat	Hrdlicka: Egyptain Expedition	Locality: El Baguat; People: Egyptian; Remarks: Female		
P258518-0	Egypt: Upper Egypt: Kharga Oasis: El Baguat	Hrdlicka: Egyptain Expedition	Locality: El Baguat; People: Egyptian Mummy; Remarks: Male		

P258521-0	Egypt: Upper Egypt: Kharga Oasis: El Baguat	Hrdlicka: Egyptain Expedition	Locality: El Baguat; People: Egyptian Mummy; Remarks: NA		
P258523-0	Egypt: Upper Egypt: Kharga Oasis: El Baguat	Hrdlicka: Egyptain Expedition	Locality: El Baguat; People: Egyptian Mummy; Remarks: Female		
P255305-0	Germany: Berlin	Prof David Paul Von Hansemann	Locality: Berlin, Germany; People: German; Remarks: Male. Age 53 years. Died of pneumonia	"I am in the position to provide you with anatomical things that would be of interest to you, not anthropological things"- Prof. Hansemann, translated from German by Kristina Zaranko (June 10, 1908 letter, Accession Record 50142)	NA
P255309-0	Germany: Berlin	Prof David Paul Von Hansemann	Locality: Berlin, Germany; People: German; Remarks: Male. 48 years old. Died of chronic bronchitis		
P258306-0	Germany: Berlin ▪	Prof David Paul Von Hansemann	Locality: Berlin, Germany; People: German (?Pole); Remarks: Male. 43 years. Died of nephritis		
P258324-0	Germany: Berlin ▪	Prof David Paul Von Hansemann	Locality: Berlin, Germany; People: German; Remarks:	"If the brain is injected with ten per cent formalin and the head sent as a whole, is very easy to	

			Male. 46 years. Died of tuberculosis	removed the brain when the specimen arrives in the laboratory, and have it in perfect condition"- Hrdlicka (July 3, 1908 letter, Accession Record 50142)	
P270988-0	Germany: Berlin	Prof David Paul Von Hansemann	Locality: Berlin, Germany; People: German; Remarks: Adult not identified as to age + sex		
P270994-0	Germany: Berlin	Prof David Paul Von Hansemann	Locality: Berlin, Germany; People: German; Remarks: Adult not identified as to age + sex		
P325266-0	Japan: Honshu: Tokyo Prefecture Tokyo	Prof Yoshihiyo Koganei	Locality: Tokyo, Japan; People: Japanese; Remarks: Adult. Male. 27 years old	"What I would like most would be heads from dissection room, Japanese, Chinese, or Koreans, in which the brains had been hardened by an injection, soon after death, of 5 or 10% formalin."- Hrdlicka (April 9 1915 letter, Accession Record 70271)	NA
P325268-0	Japan: Honshu: Tokyo Prefecture Tokyo	Prof Yoshihiyo Koganei	Locality: Tokyo, Japan; People: Japanese; Remarks: Adult. Male. 40 1/3 yrs		
P325271-0	Japan: Honshu: Tokyo Prefecture Tokyo	Prof Yoshihiyo Koganei	Locality: Tokyo, Japan; People: Japanese; Remarks: Adult. Female. "mandible and zygoma cleaned with naphtha to remove tape adhesive left by researcher. 2/96"		

P325272-0	Japan: Honshu: Tokyo Prefecture Tokyo	Prof Yoshihiyo Koganei	Locality: Tokyo, Japan; People: Japanese; Remarks: Adult. Male. "mandible and zygoma cleaned with naptha to remove tape adhesive left by researcher. 2/96"		
P325273-0	Japan: Honshu: Tokyo Prefecture Tokyo	Prof Yoshihiyo Koganei	Locality: Tokyo, Japan; People: Japanese; Remarks: Adult. Male		
P325275-0	Japan: Honshu: Tokyo Prefecture Tokyo	Prof Yoshihiyo Koganei	Locality: Tokyo, Japan; People: Japanese; Remarks: Adult. Male		
P225017-0	Japan: Honshu: Hyogo (Hiogo) Prefecture	W.H. Jones, M.D.	Locality: Hiogo, Japan; People: Japanese; Remarks: See papers: "Japanese skull taken from a cemetery in Hiogo, Japan, by Dr. W.H. Jones	NA	NA
P225018-0	Japan: Honshu: Tokyo	Prof L. H. Junghaus	Locality: Jeddo; People: Japanese; Remarks: Skull of educated Japanese		"I forward to you for your Institution two skulls, prepared by him during his residence in Jeddo. The white

	Prefecture Tokyo		man from Jeddo. Prepared by Prof. L.H. Junghaus, for the Smithsonian Institution in 1872.		one belonged to a well educated Japanese gentleman. (act.c.35). The other cut in two pieces and not bleached, is the skull of a Japanese criminal, 34 years of age" - Dr. Frikeisen (Frikeisen, 1872)
P225019-0	Japan: Honshu: Tokyo Prefecture Tokyo	Prof L. H. Junghaus	Locality: Jeddo, Japan; People: Japanese; Remarks: Skull of a Japanese criminal. 34 years of age. Prepared by Prof. L.H. Junghaus, for the Smithsonian Institution in 1872		
P225020-0	Japan: Honshu: Tokyo/Kana gawa Prefecture	J.H. Kidder, Surgeon, MSN	Locality: Between Yokohama+ Jeddo, Japan; People: Japanese; Remarks: Skull of Japanese criminal beheaded at execution grounds b/w Yokohama & Jeddo in 1869. Skulls belonged to the ninsok or coalie class- markedly different from the higher classes. Male		NA
P259350-0	Japan: Honshu: Tokyo/Kana gawa Prefecture	J.H. Kidder, Surgeon, MSN	Locality: Japan, on road bet. Yeddo and Yokohama; People: Japanese (Ninsok class); Remarks:		NA

			Male. A thief executed in 1869		
P278723-0	Mongolia: Ulan Bator (Urga)	Ales Hrdlicka	Locality: Urga; People: Mongolian; Remarks: Male	NA	"They throw the bodies of the dead out on the hills or into gorges near settlements or camps, to be devoured by their dogs, wolves, vultures and ravens, and what remains of the bones is left to whiten the soil" ("From my Journeys" 1898-1938, July-September 1912).
P278811-0	Mongolia: Ulan Bator (Urga)	Ales Hrdlicka	Locality: Urga; People: Mongolian; Remarks: Male		"At last reach Urga...in the outskirts, rather foul air in places--learn it is the smell of what remains of the bodies that are brought out, whenever there is death, for "burial by animals"" ("From my Journeys" 1898-1938, July-September 1912).
P278812-0	Mongolia: Ulan Bator (Urga)	Ales Hrdlicka	Locality: Urga; People: Mongolian; Remarks: Female		"But many of the specimens found were more or less recent. Quite a few had still the brains in. And many such smelled unendurably almost, except perhaps to one used to dissection and autopsies" ("From my Journeys" 1898-1938, July-September 1912).
P278852-0	Mongolia: Ulan Bator (Urga)	Ales Hrdlicka	Locality: Urga; People: Mongolian; Remarks: Male		

P278875-0	Mongolia: Ulan Bator (Urga)	Ales Hrdlicka	Locality: Urga; People: Mongolian; Remarks: Male		"The boiling and cleaning,..., proceeded ceaselessly until the whole precious material was in good condition. There were a total of 213 skulls...They were ranged on planks in the shed - but they still smelled. So a good amount of permanganate of potash was obtained and they were steeped in a strong solution--yet on drying they continued smelling...a lot of small boxes and packed the specimens in charcoal...some months afterwards the boxes were opened in [NMNH]- and the skulls still smelled. It was the brains" ("From my Journeys" 1898-1938, July-September 1912).
P278926-0	Mongolia: Ulan Bator (Urga)	Ales Hrdlicka	Locality: Urga; People: Mongolian; Remarks: Female		
P380559-0	Switzerland: Graubunden Caton (Grisons): Albula D: Lantch (Lenz)	Adolph. H. Schultz	Locality: Switzerland: Lenz, Canton Grisons; People: White; Remarks: Male	"In my private collection there is a series of 64 skulls from the Canton Grisons in the Eastern Swiss Alps; they all come from the villiage churches of Platta and Lenz, where they had been house since at least 1850." - Schultz (September 5, 1950 letter, Accession Record 188059)	NA
P380564-0	Switzerland: Graubunden Caton (Grisons): Albula D:	Adolph. H. Schultz	Locality: Switzerland: Lenz, Canton Grisons; People: White; Remarks: Male		

	Lantch (Lenz)				
P380574-0	Switzerland: Graubunden Caton (Grisons): Albula D: Lantch (Lenz)	Adolph. H. Schultz	Locality: Switzerland: Lenz, Canton Grisons; People: White; Remarks: Male		
P380584-0	Switzerland: Graubunden Caton (Grisons): Albula D: Lantch (Lenz)	Adolph. H. Schultz	Locality: Switzerland: Lenz, Canton Grisons; People: White; Remarks: Female		
P380599-0	Switzerland: Graubunden Caton (Grisons): Albula D: Lantch (Lenz)	Adolph. H. Schultz	Locality: Switzerland: Lenz, Canton Grisons; People: White; Remarks: Female		
P380605-0	Switzerland: Graubunden Caton (Grisons): Albula D: Lantch (Lenz)	Adolph. H. Schultz	Locality: Switzerland: Lenz, Canton Grisons; People: White; Remarks: Female?		

Supplementary Table 2. Sample removal criteria and values. Sorted by descending sample ID.

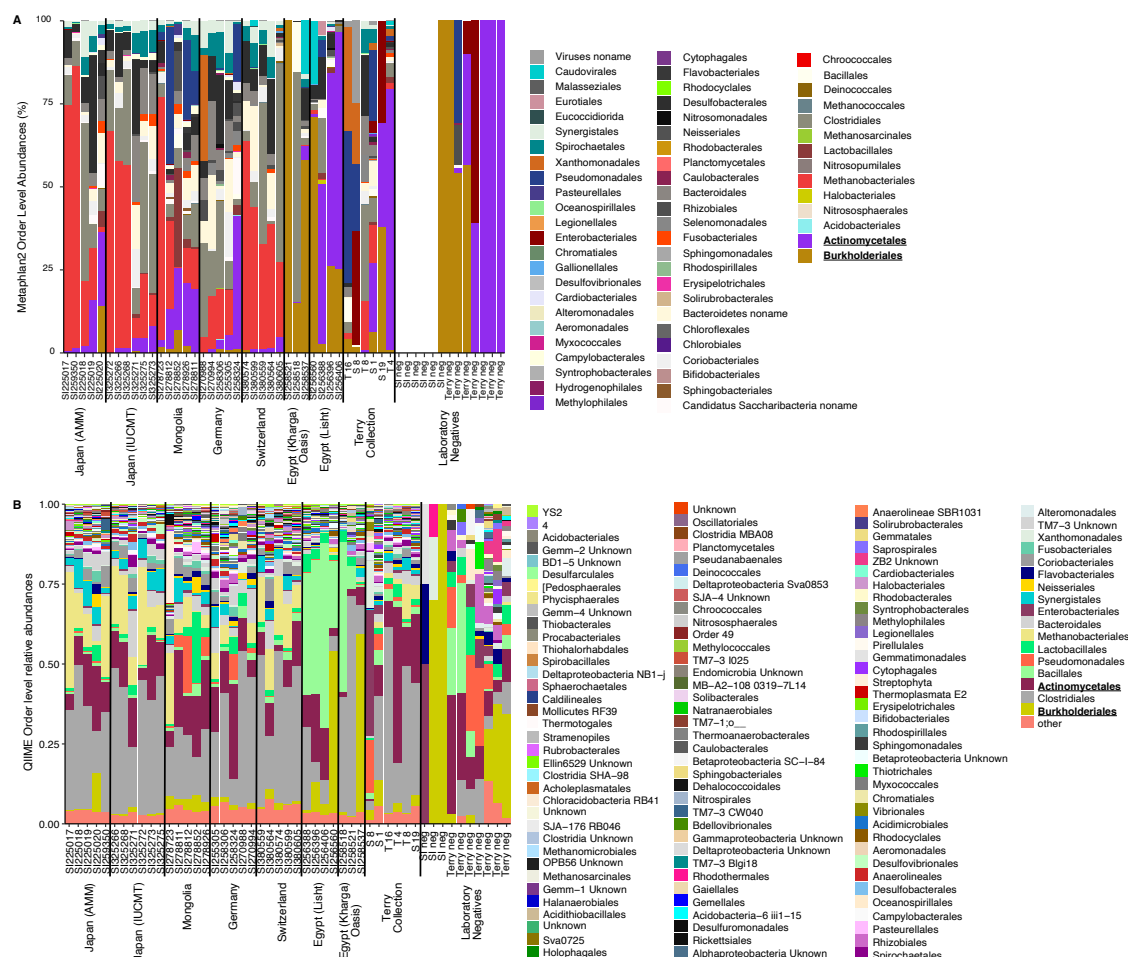
Sample ID	Collection	Total Reads	Retained Reads	Avg Read Length	% Burkholderiales	% Actinomycetales
SI225017	AMM	16,351,128	16,304,219	44.71	0.00	0.71
SI225018	AMM	23,467,427	24,824,677	51.01	0.00	2.20
SI225019	AMM	17,036,386	17,230,103	50.33	0.00	16.05
SI225020	AMM	18,926,703	16,971,689	58.03	0.25	22.17
SI255305	Germany	29,017,880	22,405,453	46.29	0.00	4.71
SI255309	Germany	35,721,877	6,764,245	46.52	0.00	1.46
SI256388	Egypt	43,186,802	32,069,264	55.63	0.00	48.44
SI256393	Egypt	15,288,592	8,326,073	64.97	89.54	0.21
SI256396	Egypt	17,993,084	14,407,378	42.27	18.20	58.33
SI256406	Egypt	20,819,639	17,702,593	76.64	25.53	71.07
SI256560	Egypt	20,632,049	14,649,802	58.03	70.26	0.71
SI258306	Germany	23,823,171	20,663,211	46.39	1.32	2.73
SI258324	Germany	33,367,600	35,895,358	48.38	0.00	39.84
SI258518	Egypt	21,888,756	20,228,505	73.58	15.08	0.16
SI258519	Egypt	11,938,361	7,788,171	60.5	0.00	0.00
SI258521	Egypt	23,328,806	14,682,030	84.18	100.00	0.00
SI258523	Egypt	1,851,437	1,161,868	44.4	44.01	0.77
SI258537	Egypt	19,105,813	9,941,052	49.55	57.70	4.46
SI259350	AMM	18,238,401	18,239,274	60.92	0.16	1.33
SI270988	Germany	17,813,451	15,243,386	50.63	0.00	0.00
SI270994	Germany	18,171,180	16,483,118	49.6	0.00	1.15
SI278723	Mongolia	21,158,753	12,977,835	57.3	0.00	2.09
SI278811	Mongolia	19,987,732	9,211,007	51.33	0.00	19.19
SI278812	Mongolia	16,036,693	9,027,202	49.77	0.00	12.23
SI278852	Mongolia	27,669,996	17,965,331	56.07	0.00	18.71
SI278875	Mongolia	8,504,579	6,351,896	59.73	0.00	0.43
SI278926	Mongolia	14,280,926	13,422,940	49.54	0.00	18.91
SI325266	Japan	28,413,936	26,581,496	61.44	0.03	1.54
SI325268	Japan	16,939,250	17,350,879	70.54	0.00	1.57
SI325271	Japan	19,788,261	20,790,810	55.12	0.00	4.04
SI325272	Japan	16,617,378	16,978,573	80.6	0.00	0.09

SI325273	Japan	19,619,675	14,678,972	67.13	0.00	8.04
SI325275	Japan	15,140,231	13,947,910	77.1	0.05	4.63
SI380559	Swiss	20,698,069	20,463,957	41.09	0.00	1.35
SI380564	Swiss	26,801,745	17,132,071	45.99	0.00	1.58
SI380574	Swiss	20,565,357	12,257,079	45.99	0.86	0.18
SI380584	Swiss	10,060,628	3,594,758	47.57	0.00	5.20
SI380599	Swiss	34,964,667	25,841,206	43.45	0.00	1.07
SI380605	Swiss	22,584,618	18,510,637	41.54	0.00	4.78
T 4	Terry	10,126,316	9,149,612	41.02	2.91	78.69
S 19	Terry	8,288,014	6,461,832	73.95	0.00	31.46
T16	Terry	10,541,297	10,143,352	73.76	0.00	0.00
T 8	Terry	9,226,004	8,896,910	46.49	0.00	0.79
S 1	Terry	9,941,418	9,073,034	48.57	6.22	20.65
S 8	Terry	12,031,853	11,831,703	74.36	0.00	0.22
SI neg1	neg	250,523	43,609	53.15	0.00	0.00
SI neg2	neg	1,326,331	84,497	61.94	100.00	0.00
SI neg3	neg	253,850	85,719	33.57	0.00	0.00
SI neg4	neg	971,558	47,101	53.16	0.00	0.00
SI neg5	neg	377,953	72,012	30.14	0.00	0.00
T neg1	neg	497908	163068	118.05	0.00	100.00
T neg2	neg	386449	152368	114.84	0.00	100.00
T neg3	neg	373696	95122	79.72	0.00	100.00
T neg4	neg	4796495	4736576	74.76	0.00	39.09
T neg5	neg	6493195	6294988	74.99	0.00	0.00
T neg6	neg	8472534	7810350	61.45	54.27	1.43
T neg7	neg	789571	596214	71.75	100.00	0.00
T neg8	neg	2212946	734429	45.68	56.76	33.47

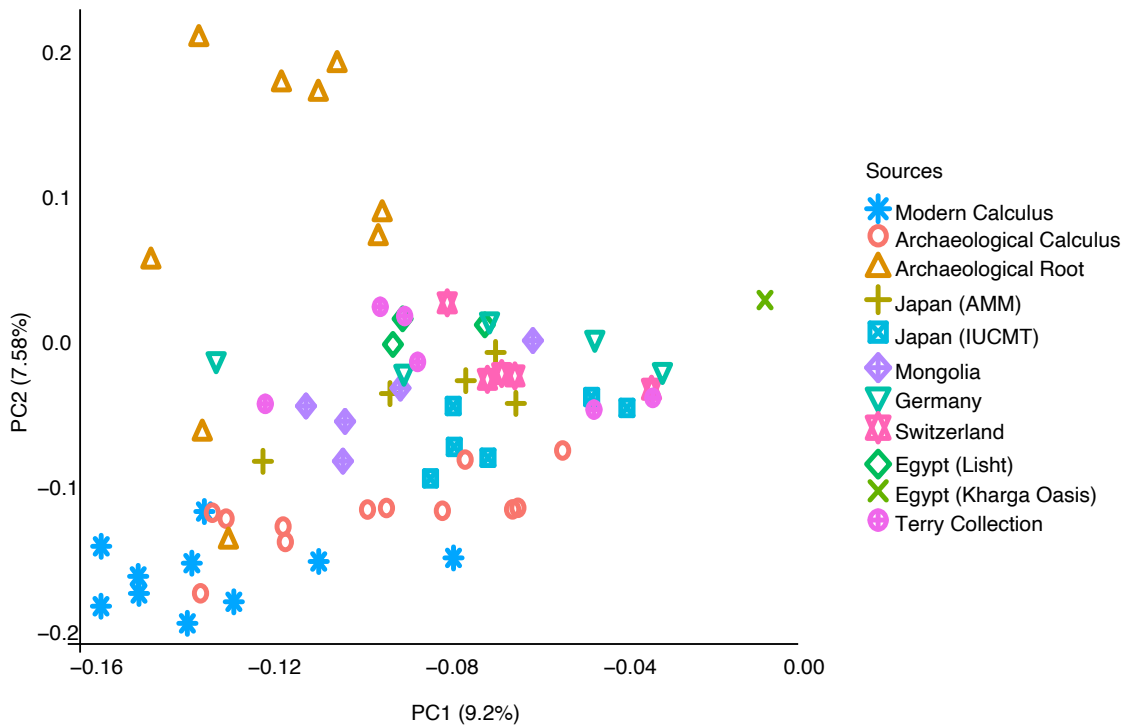
Supplementary Table 3. sPLS-DA balance error rates (BER).

	component	maximum distance error rate
A. BER when comparing collection type (i.e. modern, archaeological, and museum status)	1	0.8168333
	2	0.7646667
	3	0.744
	4	0.6953333
	5	0.6596667
B. BER when comparing museum collections and archaeological samples.	1	0.7797917
	2	0.7358333
	3	0.6816667
	4	0.62
	5	0.6452083

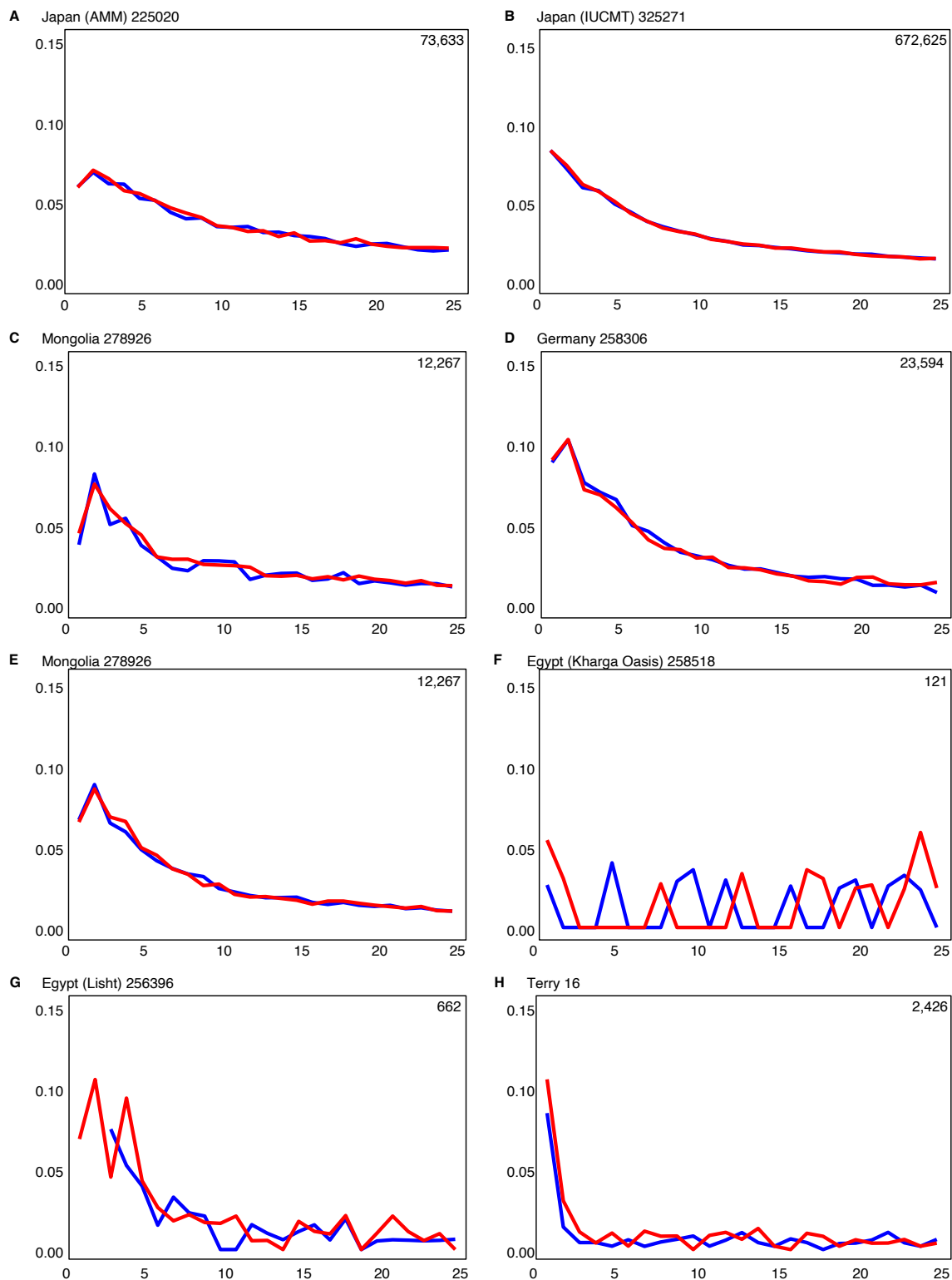
Supplementary Figures



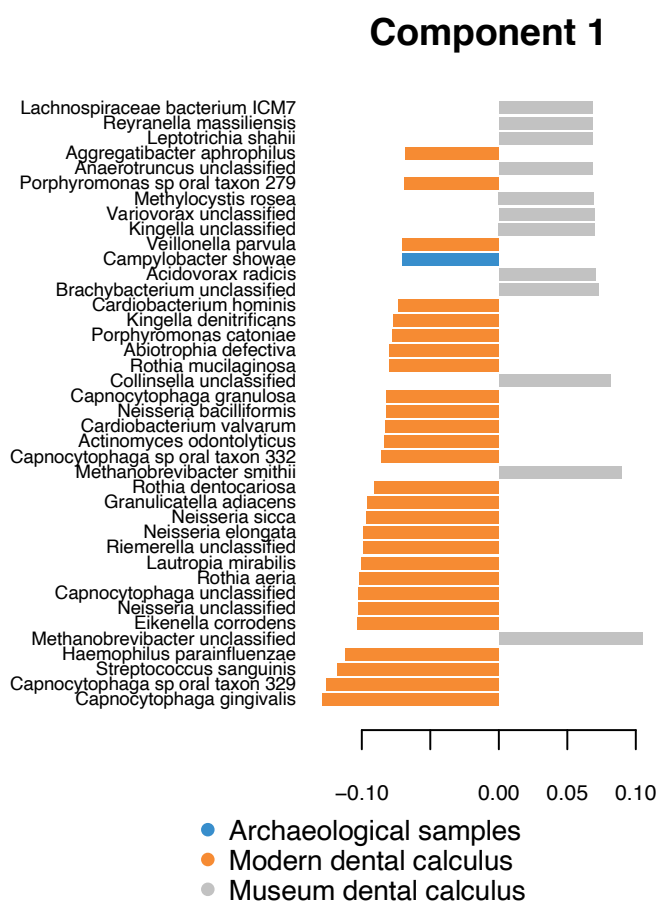
Supplementary Figure 1. Stacked bar plot displaying percent relative abundance of taxa at the order level as classified by Metaphlan2 (A) and QIIME 16S rRNA (B) analyses across SI collections. Laboratory negatives revealed Burkholderiales and Actinomycetales bacteria as primary contaminants; samples with greater than 50% identified Burkholderiales or Actinomycetales (bolded in legend) reads were removed from further analysis. This also illustrates the variable diversity seen within and between collection as lower taxonomic levels are classified.



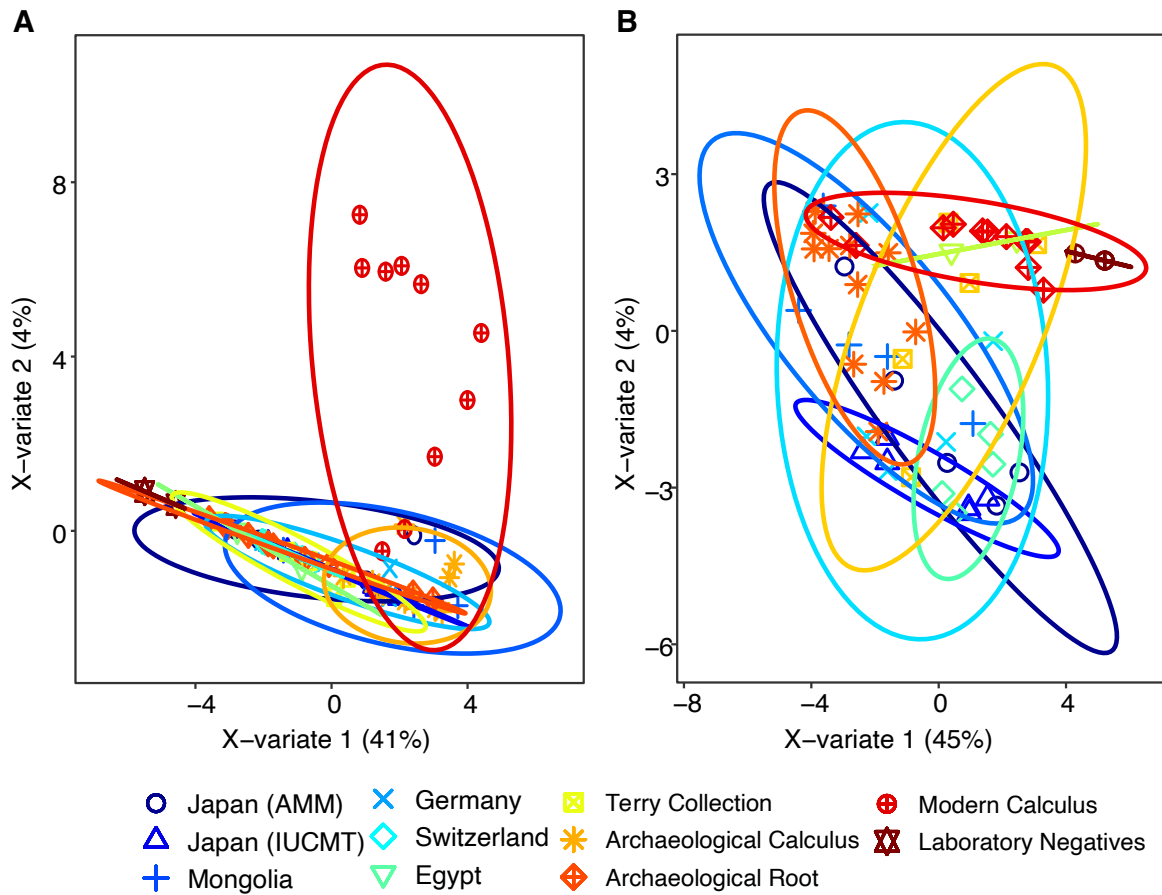
Supplementary Figure 2. Unweighted 16S rRNA PCA analysis of contributing oral environment sources. AMM, IUCMT, and Switzerland collection show the most within collection clustering.



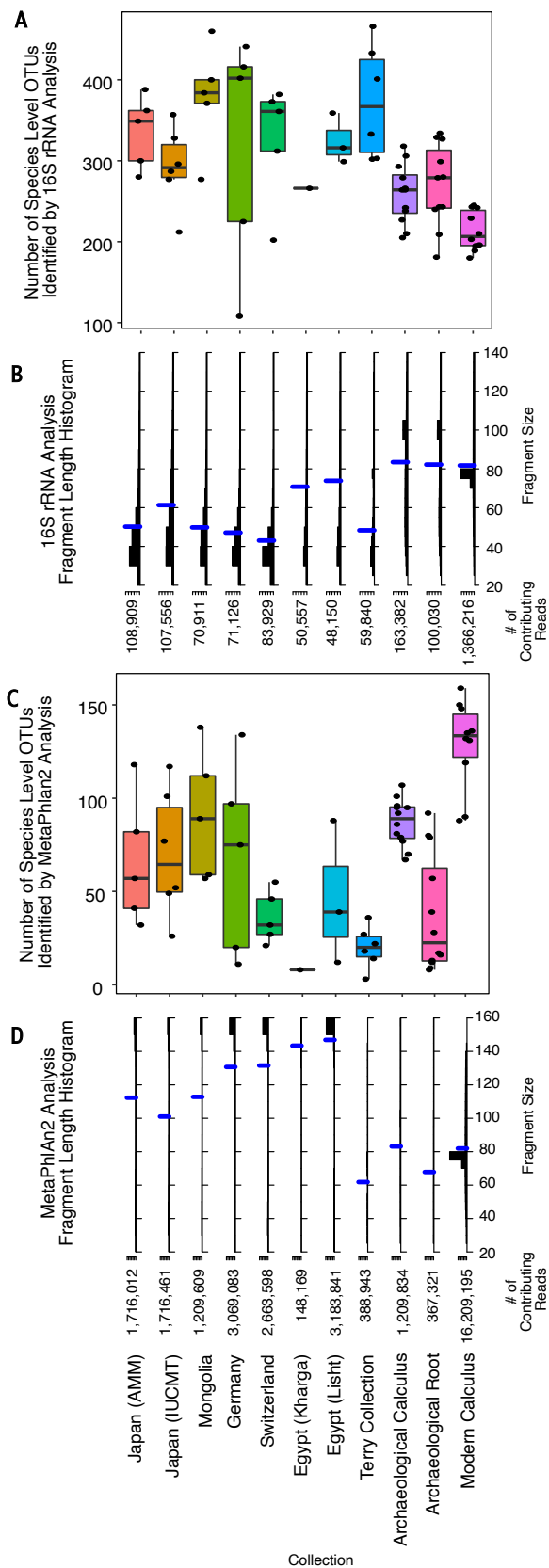
Supplementary Figure 3. MapDamage terminal end damage patterns as mapped to *Tannerella forsythia*, with number of reads mapping indicated in the upper right corners. Cytosine to thymine terminal end transitions are seen in 5 prime (blue) and 3 prime (red) strands. The eight samples shown are representative samples from each SI collection.



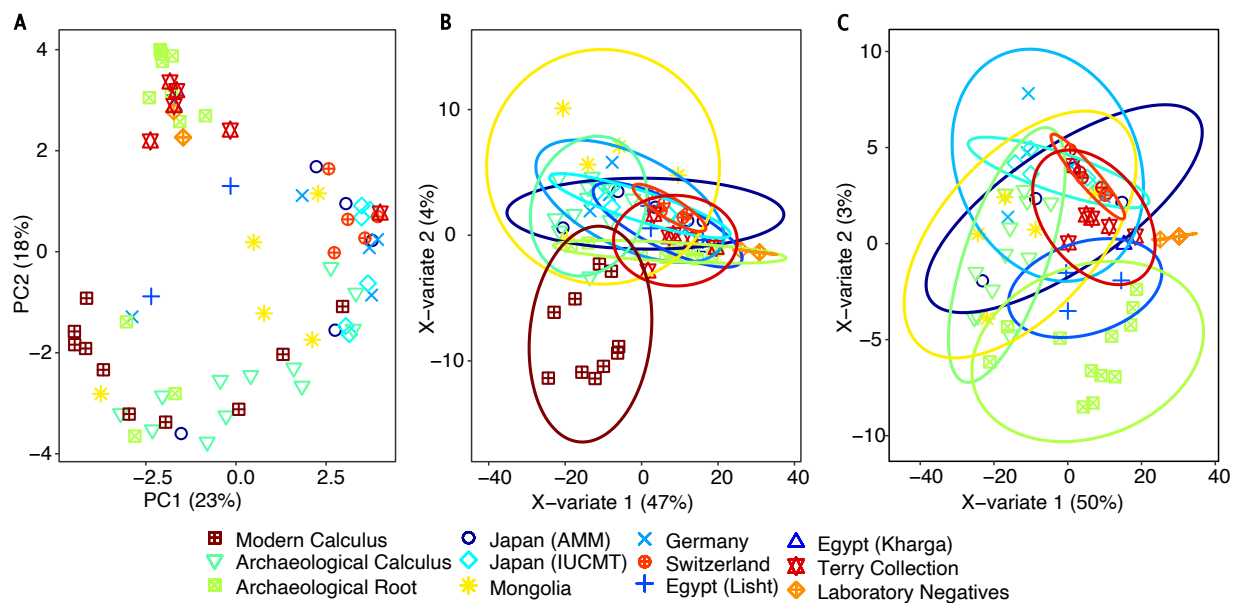
Supplementary Figure 4. Top 40 taxa contributing to discriminating among archaeological, modern, and museum samples along component 1 of sPLS-DA. Modern and museum dental calculus microbial species contribute the most to distinguishing collection types.



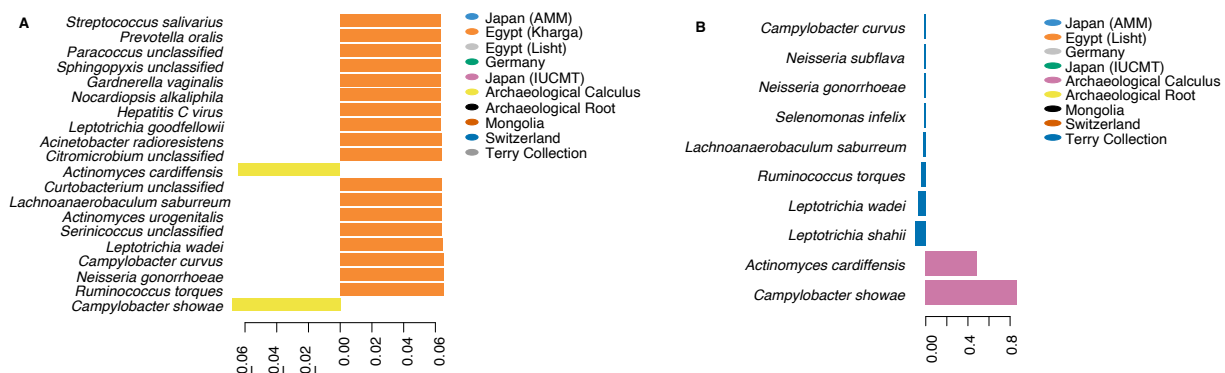
Supplementary Figure 5. Microbial community differences between collection treatments with Kharga Oasis and Lisht collections combined into single category (Egypt). Sparse partial least squares-discriminant analysis (sPLS-DA) (**A**) show modern dental calculus microbial profiles to be more distinct from archaeological and museum dental calculus. Archaeological and SI collections overlap in their ellipses (95% confidence intervals) and show limited variation among them. Museum and archaeological collection sPLS-DA (**B**) continues to show heavy overlap of microbial community content when modern samples are removed.



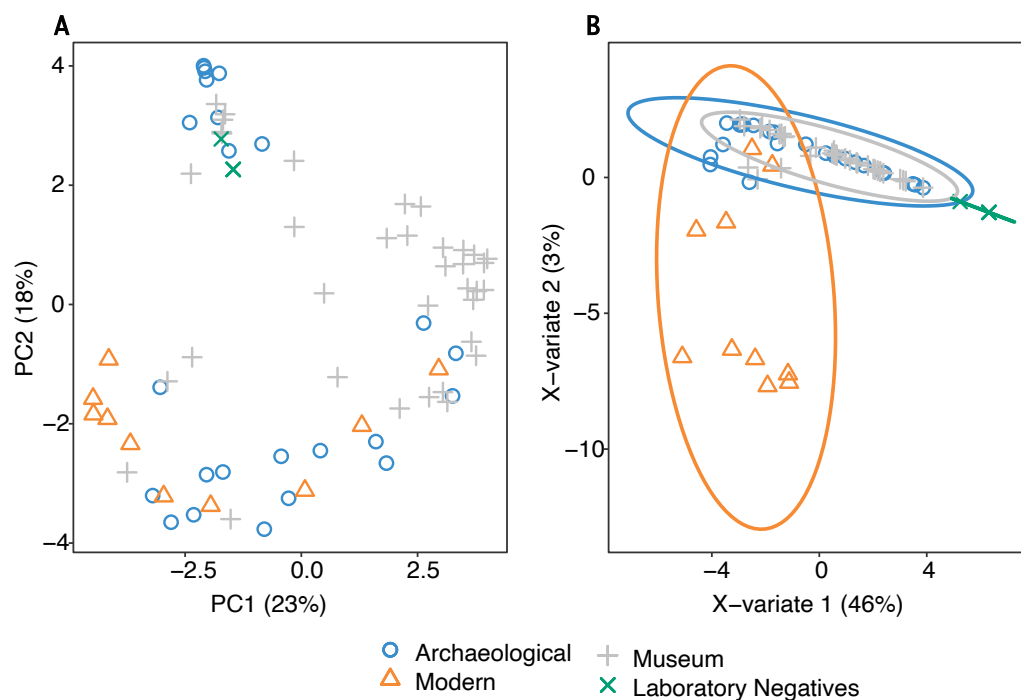
Supplementary Figure 6. Original analysis (i.e. before identifying the artificial base pairs) number of species level OTUs identified by collection between QIIME 16S rRNA (A) and MetaPhlAn2 analyses (C). Fragment length distributions across collection samples' reads mapping to the 16S rRNA gene (B), and in MetaPhlAn2 (D). 16S rRNA histograms (B) are scaled from 0 to 50,000 for archaeological data and museum collections, whereas modern dental calculus scaled from 0 to 400,000 due to the increased number of recovered molecules. MetaPhlAn2 histograms (D) are all scaled from 0 to 3,000,000. Blue lines represent average fragment lengths for the collection.



Supplementary Figure 7. Original microbial community differences among collection treatments. Prior to the identification of poly base pair tails in the global SI collection samples, principal component analysis (PCA) (A) shows slight grouping by collection. Sparse partial least squares-discriminant analysis (sPLS-DA) (B) indicate modern dental calculus microbial profiles to be more distinct from archaeological and museum dental calculus, which overlap in their ellipses (95% confidence intervals). Museum collection sPLS-DA species discrimination (C) was only minorly improved when modern dental calculus was removed. Archaeological root microbial profiles generally cluster away from museum and archaeological dental calculus communities.



Supplementary Figure 8. Prior to the identification of artificial poly base pairs in SI collection sequences, sPLS-DA analysis identified Archaeological dental calculus and the Kharga Oasis samples with the most distinguishing species (A). sPLS-DA identified species contributing to distinguishing archaeological and museum samples, with the removal of the Egyptian Kharga Oasis sample (B), where species from the AMM communities discriminate samples the most following archaeological dental calculus.



Supplementary Figure 9. Prior to the identification of artificial poly base pairs in SI collection sequences, microbial community differences between modern, archaeological, and museum dental calculus. Principal component analysis (PCA) (A) generated with the singular value decomposition (SVD) algorithm (Rohart et al 2017) shows archaeological and museum does not cluster distinctly. Modern dental calculus is seen in the lower portion of the plot but does not tightly cluster together. Several samples cluster with negative controls. Sparse partial least squares-discriminant analysis (sPLS-DA) (B) show general microbial community differences between modern dental calculus samples and overlapping archaeological and museum dental

calculus samples. SI collection negative controls appear similar to archaeological and museum samples in the PCA analysis, however, are distinguishable based on species identified (e.g. *Burkholderia*). Ellipses indicate 95% confidence intervals.

Appendix B: Chapter 4

Supplementary Materials: Smithsonian Institution's Terry Collection Reveals Insights into Ancient Tuberculosis and Syphilis Pathogen Detection Through Dental Calculus Metagenomics and Hybridization Capture

Materials and Methods

Terry Collection Overview

The Terry Collection was assembled (1910-1967), cataloged, macerated, and stored under consistent and uniform protocols at the Washington University School of Medicine (WUSM) until its transfer to the Smithsonian Institution. The collection is known as one of the most complete and well documented anatomical collections in the world (Hunt & Albanese 2005). Information on individuals' true age, sex, ancestry, and cause of death are known from associated death records. Individuals in this collection were never buried, however, upon being received at the WUSM, individuals underwent manual maceration, boiling, drying, light degreasing by exposure to benzene vapors followed by the application of a glue and formalin solution coating (for an in-depth account of the Terry Collection protocols and history see (Hunt & Albanese 2005, Trotter 1981)).

People Assessed from Terry Collection

Between 1910 and 1941 at the Washington University School of Medicine, Robert J. Terry amassed the majority of the human remains that would become the Robert J. Terry Collection. Many of the individuals incorporated into the collection by Terry, himself, would have been impacted by, or even a part of, the Great Migration, with many racial and socioeconomic dynamics also affecting people and communities (de la Cova 2019). Mildred Trotter, Terry's colleague and an anatomy instructor, continued to expand the collection until her 1967 retirement (de la Cova 2019). Trotter's collecting helped balance the collection's sex demographics with the inclusion of more females (Hunt & Albanese 2005). Even so, the Terry Collection persists as an example of demographic biases seen across anatomical collections, which exemplify the structural and legislative violence that was allowed and accepted to take place upon deceased bodies from marginalized communities (de la Cova 2019, Muller et al 2017).

Until the 1955 Missouri Willed Body Law, Terry Collection individuals and/or their families were legally, and undoubtedly, heavily influenced by societal and economic pressures. Socially unequal stressors would have directly affected where human remains went and what was done to them. Once taken into custody, the anatomical treatment and final destination of an individual's body was likely unknown to family and community members (De la Cova 2011). It is important to recognize this history and remember the people as posthumous contributors to this study (and so many others). Without them this study would not have been possible.

Trotter's collection methods also shifted towards incorporating more individuals who elected to donate their bodies due to the socio-medical legislative shift (Trotter 1981). It is because of the recent establishment of the Terry Collection (1910-1967), the recent molecular revolution, and widespread use of genetic ancestry testing for personal and criminal investigations (Gafni 2018), that we use study-specific identifiers, and removed reads mapping to the human genome from analysis. These precautions were done in the sole effort to maintain these individuals' genetic privacy, while also offering posthumous bodily and individual respect in recognition of the individuals. In line with the Terry Collection's emblematic position within the scientific community, data precautions and ethical statements made here represent the fact that anatomical collections can no longer be considered or treated/used neutrally within scientific contexts. Furthermore, our actions are meant to be a tool/mechanism to recognize these individuals as contributors and should not be seen as the solution to dealing with the nuanced and tumultuous histories anatomical collections contain.

Skeletal Analyses

Selected individuals were evaluated for dental health and general pathological reactions (Buikstra & Ubelaker 1994, Kerr 1988), ante- and peri-mortem systemic health, as well as disease-specific skeletal lesions (Harper et al 2011, Pedersen et al 2019). Diagnostic criteria based on several previous studies (Ortner 2003, Powers 2012, Santos & Roberts 2006, Steckel et al 2006) were synthesized and implemented in order to cohesively assess skeletal lesions consistent with tuberculosis infection (see Supplementary Table 1). A composite score of pathologic lesions for tuberculosis (Pedersen et al 2019, Zuckerman & Sandhu) and syphilis (Harper et al 2011) was determined for each individual as skeletal disease diagnosis estimate.

Regardless of cause of death (COD), individuals were scored for both tuberculosis (TB) and syphilis skeletal diagnostic criteria.

Molecular Analyses

Sampling

Subsampling of dental calculus was conducted at the National Museum of Natural History's (NMNH) Museum Support Center in the Biological Anthropology Collections, located in Suitland, Maryland. In accordance with previously studies (Warinner et al 2014b) , a bleach sterilized dental scalar and nitrile gloves were used to scarp dental calculus directly from the individual into a sterile 1.5 ml tube.

All laboratory work and analyses were performed at the Laboratories for Molecular Anthropology and Microbiome Research (LMAMR) at the University of Oklahoma in Norman, Oklahoma. Samples were processed for sequencing in the LMAMR Ancient DNA Laboratory, a dedicated, six-chambered clean room, abiding by established contamination control workflows, including physical separation from other laboratories performing molecular research, and unidirectional work and positive air flows to avoid cross-contamination. Researchers were outfitted with full-body Tyvek suit, mask, and gloves to protect the sample from investigator contamination. Non-template extraction controls and reagent blanks were also processed in parallel to identify contamination.

DNA Extraction

With slight modifications, DNA extraction followed the protocol described in Ziesemer et al. (Ziesemer et al 2015). No UV irradiation was applied to samples before washing and digestion steps. Briefly, between 1- 5.5 mg of dental calculus was washed with 0.5 M EDTA for 15 mins before being allowed to decalcify in a solution of (fresh) 0.5M EDTA and 10% proteinase K (Qiagen) at room temperature on a nutator for 72 hours. A modified Qiagen MinElute silica-column based purification was performed to isolate DNA into a final 60 ul elution volume. The extraction elution was then quantified using a Qubit High Sensitivity fluorometer (Life Technologies).

Library Preparation

Prior to library construction, partial uracil-DNA-glycosylase (UDG) treatment was applied in order to ensure the longest possible, authentic read lengths (Rohland et al 2015). Partial UDG treatment achieves this by eliminating interior nucleotide damage that would otherwise cause polymerase enzymes to detach, while also preserving terminal nucleotide damage for later DNA authentication (Rohland et al 2015). 30 ul of extract was used as input for partial UDG treatment. Following the partial UDG treatment and enzyme inactivation, samples were treated with the NEBNext DNA Library Prep Master Set (E6076) with blunt-end modified Illumina adapters. NEBNext DNA Library Prep manufacturer's instructions were followed with slight modifications. Following the partial UDG treatment, sequence end-repair was completed by only adding the NEBNext End Repair enzyme to the partial UDG reaction solution and completed with a 30 minute, 20°C incubation. Other End Repair buffers were not added because reaction reagents and conditions were created during the UDG treatment. Samples were then purified with a Qiagen MinElute silica spin column and eluted in 15 ul EB. Illumina adapters were then ligated to each sample, followed by a 15-minute incubation at 20°C. Before entering the Adapter Fill-In step, samples were cleaned with a Qiagen MinElute silica spin column and eluted in 15 ul EB. A 20 minute, 37°C incubation completed adapter fill-in before a final Qiagen MinElute silica spin column purification and 50 ul elution. Libraries were quantified using a qPCR KAPA Hifi Uracil+ enzyme assay (Lightcycler 480 Roche) to determine appropriate number of cycles for the indexing PCR.

Indexing

Sample libraries were dual-indexed in triplicate, in 25 ul PCR reactions, with 4 ul library input, 12.5 ul KAPA U + master mix, 6 ul H₂O, 1 ul of 2.5 mg/ml BSA, and 0.75 ul unique forward and reverse barcodes (10 uM). Based on the qPCR fluorescence values (C_q), libraries were binned into similar cycling groups for amplification, with four cycles added to each group to ensure DNA fragment abundance. Thermocycling conditions were 5 min at 98°C followed by 14 cycles of 98°C (20s), 60°C (15s), and 72°C (30s), followed by a final elongation step 72°C (1m). Amplified PCR replicates were then pooled and purified with a Qiagen MinElute silica spin column in 30 ul EB.

Shotgun Screening Sequencing

Following Fragment Analyzer quantification, sampling pooling, and PippinPrep purification to remove adapter dimers, libraries were sent for sequencing at the Max Planck Institute for the Science of Human History (Jena, Germany) on two Illumina NextSeqs using single-end 75 bp or paired-end 150 bp read chemistry. Between 6.2 and 12 million reads were recovered, with extraction and library blanks sequenced to a depth of 0.4-8 million reads.

rpoB and IS6110 Quantitative PCR

Extracts (2 ul per reaction) were amplified using TaqMan FastStartII no UNG (10 ul per reaction; Thermo Fisher Scientific) with a real-time quantitative PCR assay (qPCR, Lightcycler 480 Roche). Reaction recipes and cycling conditions followed previously published instructions (Harkins et al 2015).

Hybridization Tuberculosis Capture.

Manufacturer's instructions were followed for the myBaits Hybridization Capture for Targeted NGS kit (manual v.4.01 April 2018) for two capture experiments. As positive controls, bone samples from Peruvian mummies, which had previously yielded complete MTBC genomes (samples AD-54 and -64) (Harkins et al 2015) were processed in parallel with the presented dental calculus samples. Importantly, 5.5 ul of biotinylated RNA baits were added to each sample and allowed to hybridize for 48 hours (65C for 12 hours and 60C for 36 hours). Post-capture libraries from the each round of tuberculosis capture were amplified in triplicate following the April 2018 myBaits capture protocol (5 ul of capture library total for each replicate, and adding 15 ul NF water for a total reaction volume of 50 ul). Clean capture libraries were resuspended in 30 ul of elution buffer.

Post-capture libraries were quantified using the KAPA Library Quantification Kit, standards, and analysis template (v4.14) to estimate cycle numbers for optimal post-capture libraries amplification. Triplicate quantification PCR reactions consisted of 5 ul of NF water, 12.5 ul of a 2X KAPA HiFi HotStart Ready Mix, 1 ul BSA, 0.75 ul each of Illumina IS5 and IS6 indexing primers (10 uM), and 1 of SYTO 9 green fluorescent nucleic acid stain (25 ul reactions). PCR

reaction condition were: 2 minutes at 95C, followed by 40 cycles of 20 seconds at 98C, 15 seconds at 60C, 30 seconds at 72C, with a final 1 minute elongation step at 72 C and melt curve cycle (95C for 10 seconds, 65C for 1 minute, and 97 for 1 second). Following the KAPA Library Quantification Kit analysis template, the available concentration (ng) of each library was determined. Samples were then grouped by cycle number for triplicate amplification. Post capture amplification consisted of 5 ul of library template, 12.5 ul of a 2X KAPA HiFi HotStart Ready Mix, 5 ul NF water, and 2.5 ul each of a forward and reverse indexing primer (10 uM). Capture libraries were amplified following 2 minutes at 98C, followed by 10–12 cycles of 20 seconds at 98C, 30 seconds at 60C, 20 seconds at 72C, with a final 5 minute extension step at 72 C. Post-capture amplified libraries were isolated from the capture beads (that were included during the PCR) by pelleting them on an MPC until clear, keeping the supernatant. Pooled sample replicates were cleaned with AMPure bead cleanup method.

The clean, enriched libraries were quantified using a Fragment Analyzer to determine sample concentration (nM). Based on sample concentration, equimolar ratios of each were pooled together, and purified using the PippinPrep to remove adapter dimers. Capture experiments were sequenced on a NextSeq using 150 bp paired-end read chemistry at the Oklahoma Medical Research Foundation Clinical Genomics Center in Oklahoma City, Oklahoma.

Metagenomic Data Analysis

Shotgun and capture metagenomic reads were processed with the following steps: adapter removal, paired-end data was merged, and quality filtered for base quality (minimum 20) and read length (25 bp minimum) using Adapter Removal v2 (Schubert et al 2016).

Analysis ready reads were then mapped to the 16S rRNA Greengenes database using bowtie2 v1.1.2 (retaining no unaligned reads; --no-unal flag) (DeSantis et al 2006). Samtools v1.5 filtered unmapped and low quality reads, sorted, removed PCR duplicate and multiple mapping reads before converting files to fasta format (Li et al 2009). Metagenomic profiles were analyzed and generated using MetaPhlAn2 (Truong et al 2015). In order to identify potential contributing sources SourceTracker v1.0.1 was used (Caporaso et al 2010). Weighted and unweighted

UniFrac distance matrixes were generated in QIIME v1.9.1 and visualized in R as PCoA plots using ggplot2 (Caporaso et al 2010, Wickham 2011).

To authenticate retained reads as ancient, mapDamage v2.0.8-2 was applied to reads mapping to the human genome (hg19) using Bowtie v1.1.2 (Langmead & Salzberg 2012). mapDamage estimated the nick frequency, cytosine deamination rates, and read length patterns (Jónsson et al 2013) .

The top three samples with the highest percent unique, quality filtered reads mapping to *M. tuberculosis* (Comas et al 2010) and *T. pallidum* (str. Sea 81-4) reference genomes were also visualized for the genomes' breadth and depth of read coverage using BRIG (BLAST Ring Image Generator) (Alikhan et al 2011). Reads were mapped to the reference genome with an e value threshold of 1e-10 and a minimum percent identity of 95%. Coverage was calculated across a 100bp window. Blue spikes in the coverage graphs indicate those areas with read stacking in the genome (more than one standard deviation from the mean coverage) (Supplementary Figures 1 and 2).

The bioinformatics tool HOPS (Heuristic Operations for Pathogen Screening) (Hübler et al 2019) was also ran for comparison. The HOPS reference database consisted of complete reference genomes of the species listed in Supplementary Figure 3, as available from the National Center for Biotechnology Information (NCBI) Refseq (December 2016). Quality filtered and merged reads were then 'locally' aligned to bacterial references, particularly between highly conserved segments of the reference genomes.

Supplementary Tables

Supplementary Table 1. Tuberculosis indicator score criteria.

Score	Criteria
0	Lesions consistent with a non-tuberculosis process (e.g., taphonomic process, non-infectious etiology, etc.)
1	Periosteal reactions on visceral surfaces of rib on a single rib OR predominantly destructive, lytic lesions observable on one lumbar or thoracic vertebrae (e.g., vertebral collapse)
2	Periosteal reactions on visceral surfaces of rib on multiple ribs OR predominantly destructive, lytic lesions observable on more than one lumbar or thoracic vertebrae (e.g., vertebral collapse)
3	Periosteal reactions on visceral surfaces of rib on multiple ribs and predominantly destructive, lytic lesions observable on more than one lumbar or thoracic vertebrae (e.g., vertebral collapse)

Supplementary Table 2. Tuberculosis and venereal syphilis indicator scores for each of the assessed individuals. Study identification codes denote which disease each individual is documented to have died from (S: venereal syphilis; T: tuberculosis).

Study ID	Infection/Co-infection: TB diagnostic score (0-3) (see Supplementary Table 1)	Infection/ Co-infection: syphilis diagnostic score (0-5) (see (Harper et al 2011))
S1	2	1
S10	0	0
S11	0	0
S12	0	1
S13	0	3
S15	0	1
S16	0	1
S17	0	1
S18	0	0
S19	0	1
S2	0	1
S3	2	1
S4	0	2
S5	2	3
S6	0	1
S7	1	3
S8	0	1
S9	2	1

T1	2	1
T10	3	1
T11	0	1
T12	1	0
T13	2	0
T14	3	1
T15	2	1
T16	2	1
T17	2	1
T18	1	1
T19	2	1
T2	3	1
T20	0	0
T3	2	0
T4	2	1
T5	2	1
T6	3	1
T7	2	1
T8	2	1
T9	2	3

Supplementary Table 3. Sample statistics when mapped to *M. tuberculosis* reconstruction.

Sample	Medically Diagnosed Pathogen	SE or PE sequencing	Total Reads	Analysis-ready reads (trimmed and merged)	Proportion of reads mapped	Unique Q37 mapped reads	Average length of mapped reads	% mapped DNA
T20	T	SE	9674263	8029676	0.0005	1,058	39.6474	0.010936234
T19	T	PE	10765422	9718642	0.0002	755	37.5338	0.007013195
T18	T	PE	10192831	9776418	0.0004	1,532	37.3336	0.015030172
T17	T	SE	8639763	7979479	0.0002	479	38.1608	0.005544134
T16	T	SE	10541297	10244153	0.0002	721	38.4258	0.006839766
T15	T	PE	10395922	9665867	0.0008	2,200	38.9445	0.021162144
T14	T	SE	10421456	7622933	0.0002	543	40.6703	0.005210404
T13	T	SE	7434592	6990102	0.0002	451	38.5211	0.006066237
T12	T	SE	8563068	7965366	0.0001	309	40.1197	0.003608520
T11	T	PE	72232292	56356336	0.0006	1,993	40.2524	0.002759154
T10	T	SE	9556608	7093547	0.0004	515	42.9398	0.005388941
T9	T	PE	10366877	9729140	0.0008	2,342	36.5004	0.022591182
T8	T	PE	9226004	8896910	0.0004	1,139	39.8464	0.012345540
T7	T	PE	8696577	8307841	0.0004	1,212	37.0553	0.013936518
T6	T	SE	8380223	8052947	0.0006	1,582	39.9475	0.018877779
T5	T	PE	92029333	77766893	0.0012	4,363	33.0383	0.004740880
T4	T	PE	10126316	9149612	0.0008	2,053	36.7842	0.020273908

T3	T	PE	9589718	9099396	0.0003	1,057	35.8798	0.011022222
T2	T	PE	8962934	8201534	0.0005	1,358	35.2121	0.015151289
T1	T	SE	7085101	6833872	0.0008	1,858	40.8111	0.026224044
S19	S	SE	8288014	7880191	0.0003	748	34.881	0.009025081
S18	S	SE	7232609	5968925	0.0003	506	40.8972	0.006996092
S17	S	SE	8160682	6332373	0.0002	484	39.8264	0.005930877
S16	S	PE	10150353	9545083	.0011	2824	38.1254	0.027821693
S15	S	PE	9391030	8974921	0.0005	1,487	36.074	0.015834259
S14	S	SE	9974471	9610087	0.0007	1,996	39.0085	0.020011086
S13	S	SE	8166313	7409275	0.0002	647	38.7094	0.007922792
S12	S	SE	10720154	8268521	0.0006	673	43.9822	0.006277895
S11	S	PE	96739684	40834098	0.0013	4,415	41.559	0.004563794
S10	S	PE	84769767	30715800	0.0006	713	36.2258	0.000841102
S9	S	SE	6624608	5938166	0.0006	1,389	37.6703	0.020967278
S8	S	SE	12031853	8817035	0.0002	415	41.8699	0.003449178
S7	S	SE	7642968	7476903	0.0003	890	36.0697	0.011644691
S6	S	PE	8999548	8382949	0.0005	1,491	34.7089	0.016567499
S5	S	SE	6276063	6131565	0.0004	941	37.4867	0.014993476
S4	S	SE	6643968	6503027	0.0005	1,091	37.4592	0.016420910
S3	S	SE	7612937	7262676	0.0003	795	40.3987	0.010442750
S2	S	SE	6676754	6512938	0.0004	925	37.0205	0.013854037
S1	S	PE	9941418	9073034	0.0001	424	35.4693	0.004264985
neg3	N	PE	8472534	7810350	0	44	42.3864	0.000519325

neg2	N	PE	2212946	734429	0.0002	11	36.4545	0.000497075
neg1	N	PE	789571	596214	0	2	50.5	0.000253302
libneg3	N	SE	6493195	772547	0	3	42	0.000046202
libneg2	N	SE	4796495	217287	0.0001	1	63	0.000020849
libneg	N	PE	373696	95122	0.0002	-	There were no unique mapped reads	0.000000000
Eneg2	N	PE	386449	152368	0.0001	4	61.75	0.001035065
Eneg1	N	PE	497908	163068	0.0001	4	60.5	0.000803361

T: Tuberculosis; **S:** Syphilis

SE: Single end; **PE:** Paired end

Supplementary Table 4. Sample statistics when mapped to *T. pallidum* reference genome.

Sample	Medically Diagnosed Pathogen	SE or PE sequencing	Total Reads	Analysis-ready reads (trimmed and merged)	Proportion of reads mapped	Unique Q37 mapped reads	Average length of mapped reads	% mapped DNA
T20	T	SE	9,674,263	8029676	0.83	26	32.8462	0.0002687543
T19	T	PE	10,765,422	9718642	0.902765	27	34.0741	0.0002508030
T18	T	PE	10,192,831	9776418	0.959146	26	34.5769	0.0002550812
T17	T	SE	8,639,763	7979479	0.9235	9	33.5556	0.0001041695
T16	T	SE	10,541,297	10244153	0.9718	23	34.6957	0.0002181895
T15	T	PE	10,395,922	9665867	0.929775	10	36.5	0.0000961916
T14	T	SE	10421456	7622933	0.7314	17	35.2353	0.0001631250

T13	T	SE	7,434,592	6990102	0.9402	7	35.2857	0.0000941545
T12	T	SE	8,563,068	7965366	0.9302	17	34.1176	0.0001985270
T11	T	PE	72,232,292	56356336	0.7802	111	35.3694	0.0001536709
T10	T	SE	9,556,608	7093547	0.7422	4	35	0.0000418559
T9	T	PE	10,366,877	9729140	0.938483	31	31	0.0002990293
T8	T	PE	9,226,004	8896910	0.96433	22	35.6364	0.0002384564
T7	T	PE	8,696,577	8307841	0.9553	63	34.3333	0.0007244230
T6	T	SE	8,380,223	8052947	0.9609	23	36.3478	0.0002744557
T5	T	PE	92,029,333	77766893	0.845	127	31.0236	0.0001379995
T4	T	PE	10,126,316	9149612	0.903548	3	33.3333	0.0000296258
T3	T	PE	9,589,718	9099396	0.94887	24	36	0.0002502680
T2	T	PE	8,962,934	8201534	0.91505	36	30.8333	0.0004016542
T1	T	SE	7,085,101	6833872	0.9645	6	33.1667	0.0000846847
S19	S	SE	8,288,014	7880191	0.9507	2	32	0.0000241312
S18	S	SE	7,232,609	5968925	0.8252	14	33.6429	0.0001935678
S17	S	SE	8,160,682	6332373	0.7759	5	33	0.0000612694
S16	S	PE	10,150,353	9545083	0.94037	12	31.5	0.0001182225
S15	S	PE	9,391,030	8974921	0.955691	28	33.5357	0.0002981569
S14	S	SE	9,974,471	9610087	0.9634	25	34.84	0.0002506399
S13	S	SE	8,166,313	7409275	0.9072	10	39.4	0.0001224543
S12	S	SE	10,720,154	8268521	0.7713	19	34.3684	0.0001772363
S11	S	PE	96,739,684	40834098	0.4221	38	37.5789	0.0000392807
S10	S	PE	84,769,767	30715800	0.3623	9	33.1111	0.0000106170

S9	S	SE	6,624,608	5938166	0.8963	5	37.2	0.0000754762
S8	S	SE	12,031,853	8817035	0.7328	12	33.25	0.0000997353
S7	S	SE	7,642,968	7476903	0.9782	19	35.0526	0.0002485945
S6	S	PE	8,999,548	8382949	0.931486	33	31.6364	0.0003666851
S5	S	SE	6,276,063	6131565	0.9769	5	33.8	0.0000796678
S4	S	SE	6,643,968	6503027	0.9787	9	32.7778	0.0001354612
S3	S	SE	7,612,937	7262676	0.9539	26	42.6923	0.0003415239
S2	S	SE	6,676,754	6512938	0.9754	14	33	0.0002096827
S1	S	PE	9,941,418	9073034	0.91265	8	30.375	0.0000804714
neg3	N	PE	8,472,534	7810350	0.921843	1	31	0.0000118028
neg2	N	PE	2,212,946	734429	0.331878	0	no unique mapped reads	0.0000000000
neg1	N	PE	789,571	596214	0.755111	0	no unique mapped reads	0.0000000000
libneg3	N	SE	6,493,195	772547	0.1189	0	no unique mapped reads	0.0000000000
libneg2	N	SE	4,796,495	217287	0.0453	0	no unique mapped reads	0.0000000000

libneg1	N	PE	373,696	95122	0.2545	0	no unique mapped reads	0.0000000000
Eneg2	N	PE	386,449	152368	0.3942	0	no unique mapped reads	0.0000000000
Eneg1	N	PE	497,908	163068	0.3275	0	no unique mapped reads	0.0000000000

T: Tuberculosis; **S:** Syphilis

SE: Single end; **PE:** Paired end

Supplementary Table 5. rpoB2 qPCR sample results. All samples were ran in duplicate.

Sample Type	Sample Name	Replicate	Cq	Call	Cq Mean	Cq Error	EPF
Calculus	T1	1	-	Negative	-	-	0.04
Calculus	T1	2	-	Negative	-	-	0.03
Calculus	T10	1	-	Negative	-	-	0.04
Calculus	T10	2	-	Negative	-	-	0.05
Calculus	T11	1	-	Negative	-	-	0.03
Calculus	T11	2	-	Negative	-	-	0.01
Calculus	T12	1	-	Negative	-	-	0.02
Calculus	T12	2	-	Negative	-	-	0.02
Calculus	T13	1	-	Negative	-	-	0.01
Calculus	T13	2	-	Negative	-	-	0.02
Calculus	T14	1	-	Negative	-	-	0.01
Calculus	T14	2	-	Negative	-	-	0.03
Calculus	T15	1	-	Negative	-	-	0.02
Calculus	T15	2	-	Negative	-	-	0.03
Calculus	T16	1	-	Negative	-	-	0.05
Calculus	T16	2	-	Negative	-	-	0.03
Calculus	T17	1	-	Negative	-	-	0.05
Calculus	T17	2	-	Negative	-	-	0.06
Calculus	T18	1	-	Negative	-	-	0.03
Calculus	T18	2	-	Negative	-	-	0.05
Calculus	T19	1	-	Negative	-	-	0.01
Calculus	T19	2	-	Negative	-	-	0.02
Calculus	T2	1	-	Negative	-	-	0.01
Calculus	T2	2	-	Negative	-	-	0.01
Calculus	T20	1	-	Negative	-	-	0.05
Calculus	T20	2	-	Negative	-	-	0.07
Calculus	T3	1	-	Negative	-	-	0.01
Calculus	T3	2	-	Negative	-	-	0.01
Calculus	T4	1	-	Negative	-	-	0.03

Calculus	T4	2	-	Negative	-	-	0.06
Calculus	T5	1	-	Negative	-	-	0.03
Calculus	T5	2	-	Negative	-	-	0.02
Calculus	T6	1	-	Negative	-	-	0.03
Calculus	T6	2	-	Negative	-	-	0.02
Calculus	T7	1	-	Negative	-	-	0.05
Calculus	T7	2	-	Negative	-	-	0.03
Calculus	T8	1	-	Negative	-	-	0.02
Calculus	T8	2	-	Negative	-	-	0
Calculus	T9	1	-	Negative	-	-	0.03
Calculus	T9	2	-	Negative	-	-	0.03
Calculus	S11	1	-	Negative	-	-	0
Calculus	S11	2	-	Negative	-	-	0.01
Positive	Std 6 (0.00005 ng/ul)	1	31.46	Positive	31.46	0	4.17
Positive	Std 6 (0.00005 ng/ul)	2	31.69	Positive	31.69	0	4.19
Positive	Std 6 (0.00005 ng/ul)	1	34.85	Positive	34.85	0	3.63
Positive	Std 6 (0.00005 ng/ul)	2	34.21	Positive	34.21	0	3.75
Positive	Std 4 (0.005 ng/ul)	1	27.34	Positive	27.34	0	4.95
Positive	Std 4 (0.005 ng/ul)	2	27.09	Positive	27.09	0	4.81
Positive	Std 4 (0.005 ng/ul)	1	24.8	Positive	24.8	0	4.78
Positive	Std 4 (0.005 ng/ul)	2	24.86	Positive	24.86	0	4.68
Positive	Std 2 (0.5ng/ul)	1	17.82	Positive	17.82	0	5.1
Positive	Std 2 (0.5ng/ul)	2	17.73	Positive	17.73	0	5.14
Positive	Std 2 (0.5ng/ul)	1	18.56	Positive	18.56	0	5.67
Positive	Std 2 (0.5ng/ul)	2	18.86	Positive	18.86	0	5.6
Negative	Neg6	1	-	Negative	-	-	0.01

Negative	Neg6	2	-	Negative	-	-	0
Negative	Neg5	1	-	Negative	-	-	0.02
Negative	Neg5	2	-	Negative	-	-	0.02
Negative	Neg4	1	-	Negative	-	-	0.04
Negative	Neg4	2	-	Negative	-	-	0.04
Negative	Neg3	1	-	Negative	-	-	0
Negative	Neg3	2	-	Negative	-	-	0.01
Negative	Neg2	1	-	Negative	-	-	0.03
Negative	Neg2	2	-	Negative	-	-	0.02
Negative	Neg1	1	-	Negative	-	-	0.01
Negative	Neg1	2	-	Negative	-	-	0.06
qPCR neg	Neg	1	-	Negative	-	-	0
qPCR neg	Neg	2	-	Negative	-	-	0
qPCR neg	Neg	1	-	Negative	-	-	0.02
Bone+	AD64	1	40.88	Positive	40.88	0	2.04
Bone+	AD64	2	-	Negative	-	-	0.01
Bone+	AD58	1	39.36	Positive	39.36	0	2.48
Bone+	AD58	2	36.9	Positive	36.9	0	2.64
Bone+	AD54	1	30.88	Positive	30.88	0	4.23
Bone+	AD54	2	31.39	Positive	31.39	0	4.2
Bone+	AD54	1	30.02	Positive	30.02	0	4.27
Bone+	AD54	2	30.06	Positive	30.06	0	4.04

Supplementary Table 6. IS6110 qPCR results. All samples were ran in duplicate.

Sample Type	Sample Name	Replicate	Cq	Call	Cq Mean	Cq Error	EPF
Calculus	T9	1	-	Negative	-	-	0.03
Calculus	T9	2	36.45	Positive	36.45	0	3.68
Calculus	T8	1	-	Negative	-	-	0
Calculus	T8	2	-	Negative	-	-	0.02
Calculus	T7	1	-	Negative	-	-	0.03
Calculus	T7	2	-	Negative	-	-	0.03
Calculus	T6	1	-	Negative	-	-	0.02
Calculus	T6	2	-	Negative	-	-	0.02
Calculus	T5	1	-	Negative	-	-	0.02
Calculus	T5	2	-	Negative	-	-	0.03
Calculus	T4	1	-	Negative	-	-	0
Calculus	T4	2	-	Negative	-	-	0
Calculus	T3	1	-	Negative	-	-	0
Calculus	T3	2	-	Negative	-	-	0.01
Calculus	T20	1	-	Negative	-	-	0
Calculus	T20	2	-	Negative	-	-	0.01
Calculus	T2	1	-	Negative	-	-	0.03
Calculus	T2	2	-	Negative	-	-	0.03
Calculus	T19	1	-	Negative	-	-	0.01
Calculus	T19	2	-	Negative	-	-	0
Calculus	T18	1	38.57	Positive	38.57	0	3.01
Calculus	T18	2	36.21	Positive	36.21	0	3.65
Calculus	T17	1	-	Negative	-	-	0.01
Calculus	T17	2	-	Negative	-	-	0
Calculus	T16	1	-	Negative	-	-	0.01
Calculus	T16	2	-	Negative	-	-	0.02
Calculus	T15	1	-	Negative	-	-	0.01
Calculus	T15	2	-	Negative	-	-	0.03
Calculus	T14	1	-	Negative	-	-	0

Calculus	T14	2	-	Negative	-	-	0.04
Calculus	T13	1	-	Negative	-	-	0.03
Calculus	T13	2	-	Negative	-	-	0
Calculus	T12	1	-	Negative	-	-	0.01
Calculus	T12	2	-	Negative	-	-	0
Calculus	T11	1	-	Negative	-	-	0.04
Calculus	T11	2	-	Negative	-	-	0.01
Calculus	T10	1	-	Negative	-	-	0.02
Calculus	T10	2	-	Negative	-	-	0
Calculus	T1	1	42.23	Positive	42.23	0	1.77
Calculus	T1	2	35.42	Positive	35.42	0	3.95
Calculus	S11	1	-	Negative	-	-	0
Calculus	S11	2	-	Negative	-	-	0
Positive	Std 6 (0.00005 ng/ul)	1	30.41	Positive	30.41	0	5.57
Positive	Std 6 (0.00005 ng/ul)	2	30.35	Positive	30.35	0	5.66
Positive	Std 4 (0.005 ng/ul)	1	23.29	Positive	23.29	0	6.98
Positive	Std 4 (0.005 ng/ul)	2	23.93	Positive	23.93	0	6.99
Positive	Std 3 (0.05ng/ul)	1	12.06	Positive	12.06	0	8.11
Positive	Std 2 (0.5ng/ul)	1	13.33	Positive	13.33	0	9.17
Positive	Std 2 (0.5ng/ul)	2	13.38	Positive	13.38	0	9.13
Positive	Std 2 (0.5ng/ul)	1	13.63	Positive	13.63	0	8.05
Positive	Std 2 (0.5ng/ul)	2	13.88	Positive	13.88	0	8.11
Positive	Std 1 (5ng/ul)	1	9.75	Positive	9.75	0	9.55
Positive	Std 1 (5ng/ul)	2	9.73	Positive	9.73	0	9.51
Negative	Neg6	1	-	Negative	-	-	0
Negative	Neg6	2	-	Negative	-	-	0.02
Negative	Neg5	1	-	Negative	-	-	0.01
Negative	Neg5	2	-	Negative	-	-	0.01
Negative	Neg4	1	-	Negative	-	-	0.01

Negative	Neg4	2	-	Negative	-	-	0.02
Negative	Neg3	1	-	Negative	-	-	0.01
Negative	Neg3	2	-	Negative	-	-	0.02
Negative	Neg2	1	-	Negative	-	-	0.02
Negative	Neg2	2	-	Negative	-	-	0.03
Negative	Neg1	1	-	Negative	-	-	0
Negative	Neg1	2	-	Negative	-	-	0.03
qPCR neg	Neg	1	-	Negative	-	-	0
qPCR neg	Neg	2	-	Negative	-	-	0.02
qPCR neg	Neg	1	35.23	Positive	35.23	0	4.45
Bone+	AD64	1	36.01	Positive	36.01	0	0.86
Bone+	AD64	2	43.68	Positive	43.68	0	0.34
Bone+	AD58	1	42	Positive	42	0	0.45
Bone+	AD58	2	39.81	Positive	39.81	0	0.47
Bone+	AD54	1	28.35	Positive	28.35	0	1.56
Bone+	AD54	2	28.71	Positive	28.71	0	1.62
Bone+	AD54	1	28.22	Positive	28.22	0	1.22
Bone+	AD54	2	28.36	Positive	28.36	0	1.32

Supplementary Table 7. Capture samples mapped to *M. tuberculosis* reconstruction. Colored by experiment, with blue being capture experiment 1 and green being capture experiment 2.

Sample	Total Reads	Analysis-ready reads	Proportion of reads kept after trimming and merging	Unique Q37 mapped reads	Average length of mapped reads	Mean coverage	Std dev of mean coverage	% reference covered $\geq 1X$	% reference covered $\geq 5X$	Cluster factor	% endogenous DNA	# of homozygous SNPs	# of heterozygous SNPs
AD54A	28565744	27337761	0.96	2726439	62.72	38.76X	27.49X	96.51%	91.37%	9.2	9.54	924	171
LibNeg15	347302	282569	0.81	44	48.27	0.00X	0.02X	0.04%	0%	2019.27	0.013	0	0
LibNeg50	73721	67235	0.91	101	49.70	0.00X	0.04X	0.11%	0%	450.52	0.14	0	0
T9	18550188	15370913	0.82	3,241	49.11	0.04X	1.33X	0.43%	0.11%	969.60	0.012	25	90
T1	24508357	23213515	0.95	19,262	51.7	0.23X	2.53X	12.56%	0.28%	592.65	0.08	93	189

TAD64	15271247	9549339	0.63	187327	54.01	2.29X	2.84X	73.40%	15.12%	32.65	1.23	113	93
TLneg	58610	713	0.01	0	no unique mapped reads	N/A	N/A	N/A	N/A	N/A	0.00	N/A	N/A
TEneg	93378	40705	0.44	12	50.75	0.00X	0.03X	0%	0%	16.58	0.01	0	0
S11	8281633	6206833	0.75	81	62.88	0.00X	0.08X	0.03%	0.01%	1406.88	0.00	2	6
T5	93530094	67648301	0.72	137	63.64	0.00X	0.16X	0.05%	0.01%	7624.06	0.00	0	9
T18	162796769	119581980	0.73	725	61.27	0.01X	0.52X	0.24%	0.04%	6402.42	0.00	4	33
T3	64834272	54915975	0.85	1202	44.72	0.01X	0.66X	0.16%	0.05%	154.28	0.00	9	29
T9	7515949	5518550	0.73	1784	48.54	0.01X	0.89X	0.20%	0.07%	211.82	0.02	12	47
T1	24168968	11954670	0.5	2141	70.5	0.03X	1.54X	0.36%	0.08%	94.83	0.01	42	59

Capture Experiment 1

Capture Experiment 2

Supplementary Table 8. Capture experiments MTBC lineage SNP Calling. Colors correspond to capture experiment 1 (blue) and 2 (green).

[illegible]

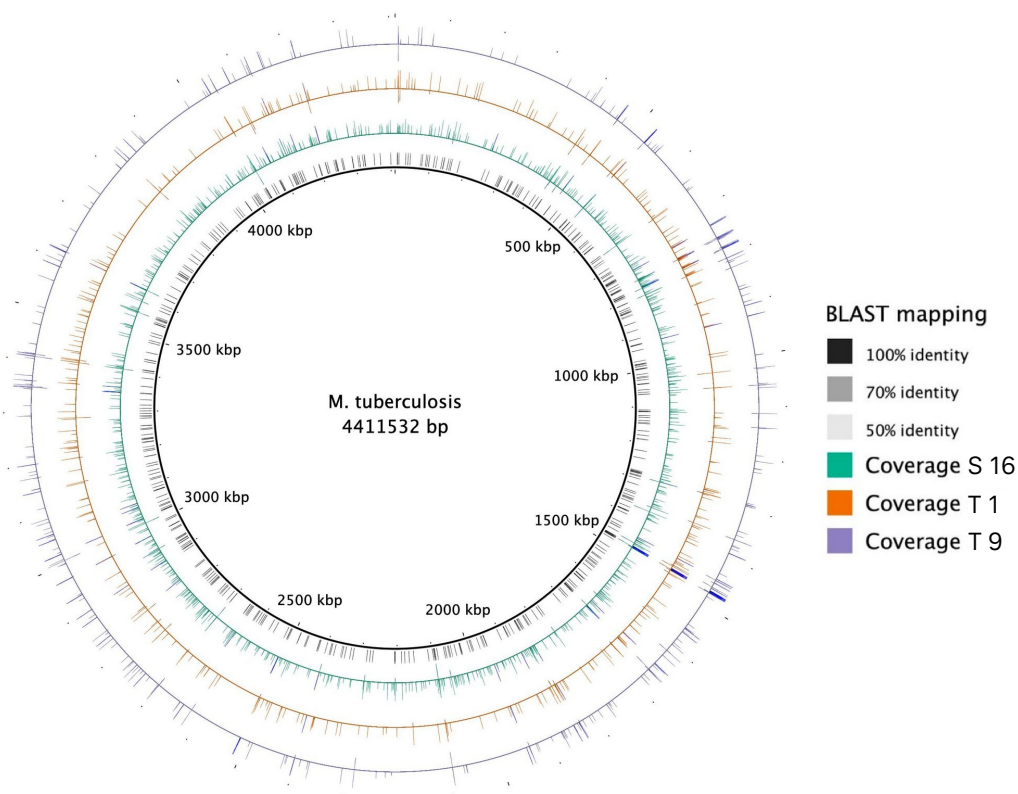
4.5	4.4.2	4.4.1.2	4.4.1.1	4.4.1	4.4	4.3.4.2.1	4.3.4.2	4.3.4.1	4.3.4	4.3.3
1719757	4246508	2694560	355181	4151558	4307886	1502120	1132368	4398141	3977226	403364
G	G	G	G	G	G	C	C	G	G	G
G/T	G/A	G/C	G/A	G/A	G/A	C/A	C/T	G/A	G/A	G/A
0/0:100:53:53:0:0%:1E0:41:0:28:25:0:0	0/0:44:24:24:24:0:0%:1E0:40:0:8:16:0:0	0/0:63:33:33:33:0:0%:1E0:40:0:17:16:0:0	0/0:24:19:18:17:1:5.56%:5E-1:41:41:7:10:0:1	0/0:100:53:53:53:0:0%:1E0:41:0:27:26:0:0	0/0:44:23:23:23:0:0%:1E0:41:0:10:13:0:0	0/0:59:32:32:32:0:0%:1E0:40:0:18:14:0:0	0/0:40:22:22:22:0:0%:1E0:40:0:10:12:0:0	0/0:55:30:30:30:0:0%:1E0:40:0:14:16:0:0	0/0:85:45:45:45:0:0%:1E0:40:0:18:27:0:0	0/0:40:24:21:21:0:0%:1E0:41:0:8:13:0:0
N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
0/0:14:7:7:7:0:0%:1E0:41:0:4:3:0:0	N/A	N/A	N/A	0/0:14:7:7:7:0:0%:1E0:41:0:6:1:0:0	N/A	N/A	N/A	N/A	N/A	N/A
N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A

5	4.9	4.8	4.7	4.6.2.2	4.6.2.1	4.6.2	4.6.1.2	4.6.1.1	4.6.1	4.6
1799921	1759252	3836739	4249732	2875883	3570528	4125058	1501468	874787	4260268	3466426
C	T	G	C	C	C	G	G	G	G	G
C/A	G/T	G/A	C/G	C/T	C/G	G/C	G/C	G/A	G/C	G/A
0/0:14:8:8: 8:0:0%:1E0 :41:0:3:5:0: 0	0/0:74:40:4 0:40:0:0%: 1E0:40:0:2 2:18:0:0	0/0:70:37:3 7:37:0:0%: 1E0:40:0:1 8:19:0:0	0/0:47:27:2 6:26:0:0%: 1E0:40:0:9: 17:0:0	0/0:156:84: 84:84:0:0%: :1E0:40:0:4 2:42:0:0	0/0:74:40:3 9:39:0:0%: 1E0:40:0:1 5:24:0:0	0/0:22:11:1 1:11:0:0%: 1E0:40:0:3: 8:0:0	0/0:36:20:2 0:20:0:0%: 1E0:40:0:9: 11:0:0	0/0:70:39:3 8:38:0:0%: 1E0:41:0:1 9:19:0:0	0/0:37:19:1 9:19:0:0%: 1E0:41:0:1 1:8:0:0	0/0:111:60: 60:60:0:0%: :1E0:40:0:3 2:28:0:0
N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A

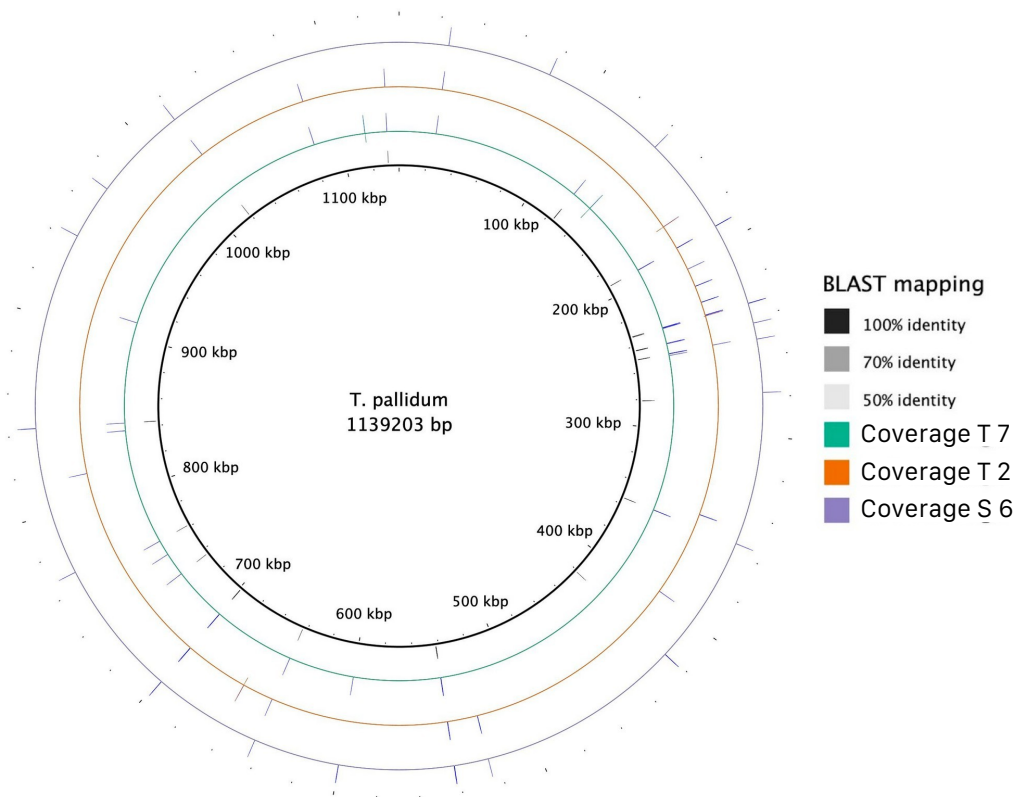
6	1816587	C	C/G	0/0:18:9:9:9:0:0%:1E0:40:0:4:5:0:0	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
7	1137518	G	G/A	0/0:22:11:11:11:0:0%:1E0:40:0:6:5:0:0	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A

Capture Experiment 1
Capture Experiment 2

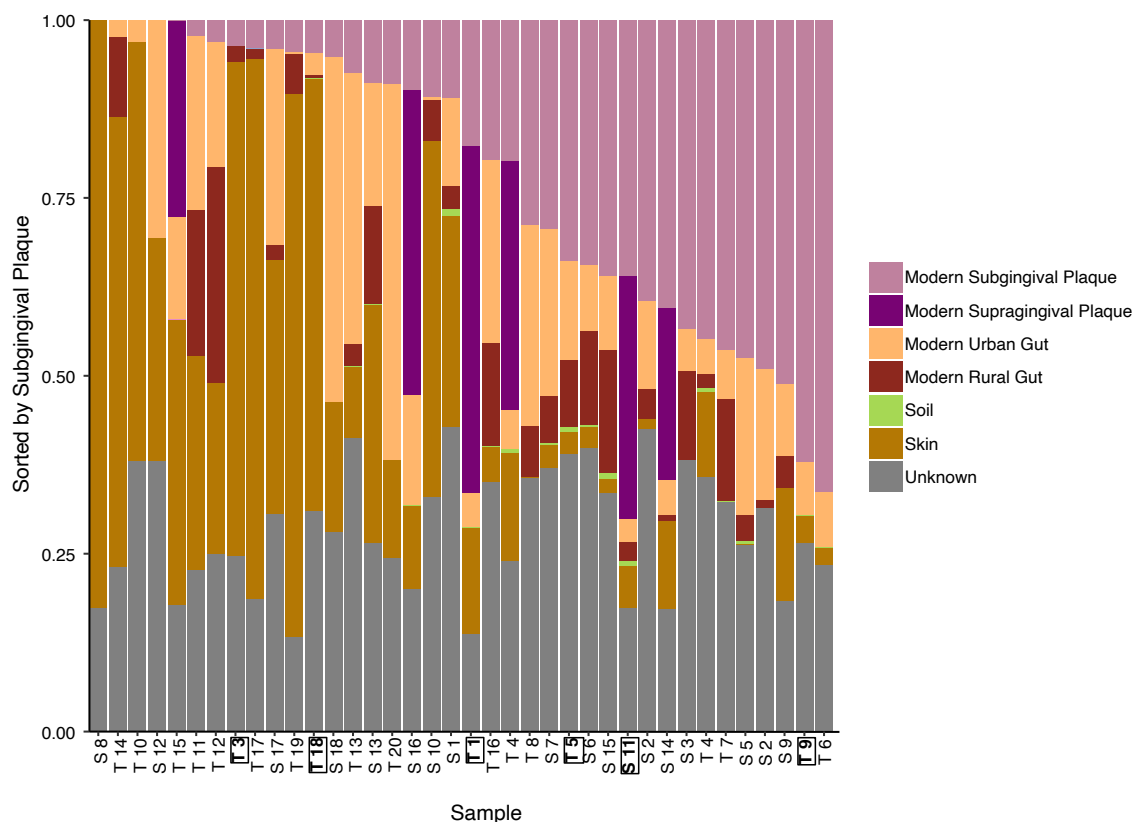
Supplementary Figures



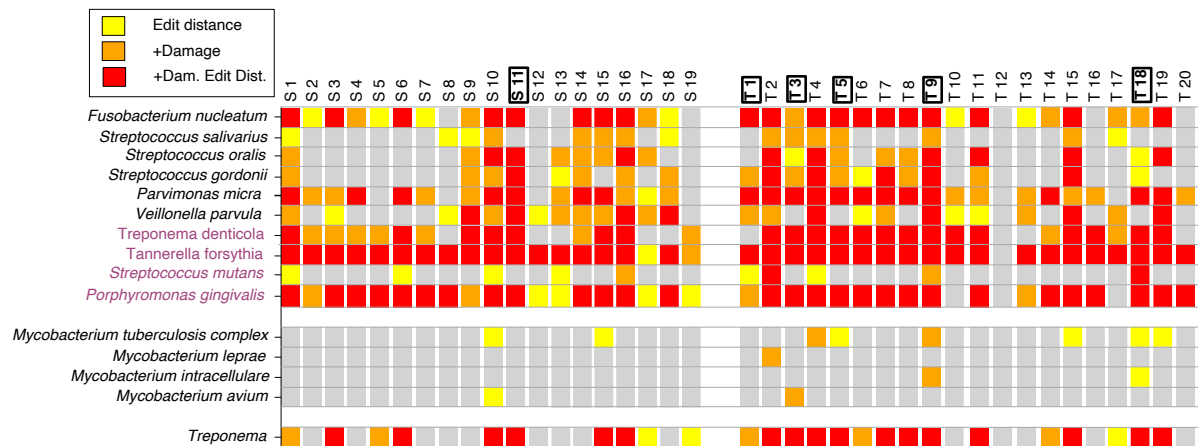
Supplementary Figure 1. Genome coverage and GC content plot of the top three samples (S16, T1, and T9) with unique reads mapping to *M. tuberculosis* reconstructed genome (Comas et al 2010). Generated using the BRIG (Alikhan et al 2011), tick marks indicate reads and their position on the reference genome (red) with the percent identity (shades of red) indicating read similarity to reference genome as determined in BLAST results. The next three rings are sample histograms (green, orange, and purple) showing genome coverage, calculated as the average over a rolling base-pair window (1000), at particular locations of the *M. tuberculosis* genome. Genome regions with coverage more than one standard deviation from the mean coverage are represented as blue spikes. The inner-most, black ring shows the GC content of the reference genome. A good portion of the reference genome was covered with high percent identities (red ticks).



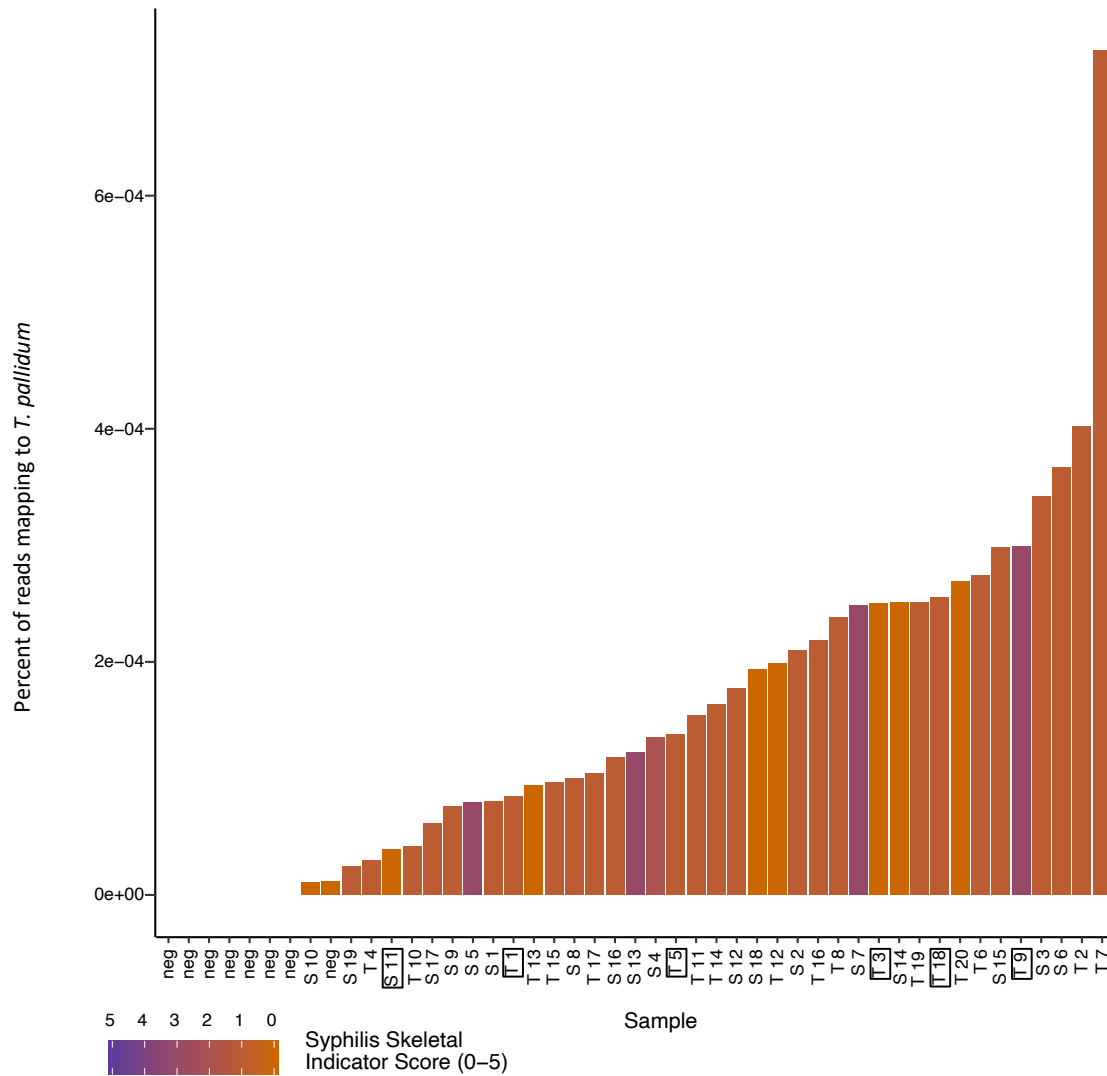
Supplementary Figure 2. Genome coverage and GC content plot of the top three samples (T7, T2, and S6) with unique reads mapping to *T. pallidum*. Generated using the BRIG (Alikhan et al 2011), the inner-most circle is the reference genome (*T. pallidum* subsp. *pallidum* str. Sea 81-4). Tick marks indicate reads and their position on the reference genome (red) with the percent identity (shades of red) indicating read similarity to reference genome as determined in BLAST results. The next three rings are sample histograms (green, orange, and purple) showing genome coverage, calculated as the average over a rolling base-pair window (1000), at particular locations of the *T. pallidum* genome. Genome regions with coverage more than one standard deviation from the mean coverage are represented as blue spikes. The inner-most, black ring shows the GC content of the reference genome. While reads were recovered across the *T. pallidum* genome, only a very small proportion of the total genome was retrieved. Additionally, high read stacking (blue ticks) is seen across samples without any visible correlation to GC content.



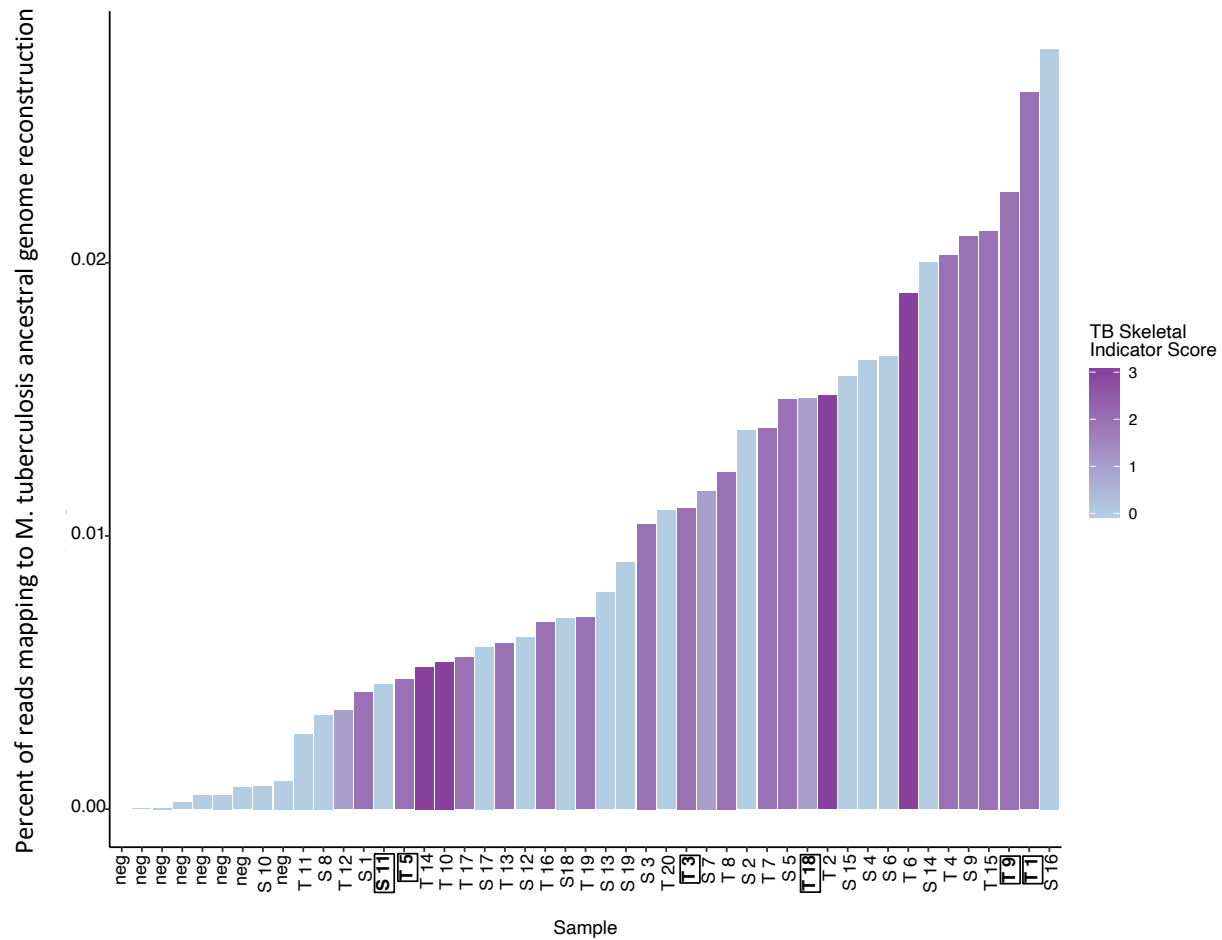
Supplementary Figure 3. SourceTracker analysis of shotgun metagenomic reads sorted by increasing proportions of taxa associated with modern subgingival plaque. Boxed and bolded sample names indicate samples used in hybridization capture experiments. In only using modern community profiles for comparison samples show lower proportions of taxa related to oral communities.



Supplementary Figure 4. Heatmap of HOPS identification results from dental calculus. Color indicates surety of detection, with red having both, detectable damage patterns and edit distance. Bolded and boxed samples were subjected to *M. tuberculosis* hybridization capture. Purple font species are common oral pathogens. *Mycobacterium* species were not detected across samples. When screening for various *Treponema* species, only *T. putidum* and *T. denticola* were detected across samples; *T. pallidum* was not detected.



Supplementary Figure 5. Percent of shotgun metagenomic reads mapping to *T. pallidum* by sample. Boxed and bolded sample names indicate samples selected for *M. tuberculosis* hybridization capture. There is not an apparent relationship between indicator scores of skeletal syphilis and the percentage of sequencing reads mapping to *T. pallidum* from the assessed dental calculus.



Supplementary Figure 6. Percent of shotgun metagenomic reads mapping to *M. tuberculosis* genome reconstruction (Comas et al 2010) by sample, colored by scored TB skeletal diagnostic criteria. Bolded and Boxed samples indicate those selected for *M. tuberculosis* hybridization capture. There is not an apparent relationship between skeletal indicator scores and the percentage of sequencing reads mapping to *M. tuberculosis* from the assessed dental calculus.

Appendix C: Chapter 5

Supplementary Materials: Dental Calculus Reveals Acute Bacterial Infections: A Case Study from the Smithsonian's Terry Collection

Supplementary Information

Terry Collection

With the intent of representing human variation, the Robert J. Terry Anatomical Collection (Terry) was collected (1910-1967), cataloged, macerated, and stored under well-established and uniform protocols, making the collection one of the most complete anatomical collections in the world (de la Cova 2019, Hunt & Albanese 2005, Trotter 1981). Upon release to the University of Washington at St. Louis Medical School, individuals would be skeletonized through manual maceration, boiling in water, drying, degreasing by exposure to benzene vapors, and then coated with a glue and formalin solution (Hunt & Albanese 2005, Trotter 1981). In addition, individuals' chronological age, sex, ethnicity, and cause of death are known thanks to the available anatomical preparation records and death certificates.

It is now recognized that the Terry Collection is primarily comprised of poor, African-Americans, and were often unskilled laborers (Corruccini 1974, De la Cova 2011) who arrived in Missouri as part of the early 20th Great Migration (De la Cova 2011). Individuals identified as Black in the Terry manifest greater frequencies of periodontal disease and periosteal reactions, which are indicative of elevated local and systemic inflammatory burdens, as well as weapon-related trauma, attributed to inter-personal violence, than do White individuals in the collection (De La Cova 2010, De la Cova 2011, Gengo 2014). The collection therefore represent many aspects of the embodiment of racial structural and interpersonal discrimination in demographic and epidemiological characteristics (Zuckerman In Press) in the contemporary U.S.

Skeletal Analysis

Skeletal and pathological inventories were performed on St.L Individual following established standards (Buikstra & Ubelaker 1994, Dangvard Pedersen et al 2019, Kerr 1988, Ortner 2003). These included macroscopic assessment of oral health and pathological lesions such as those associated with infectious diseases (e.g., tuberculosis). Radiographs of femora and cranial vault

pathologies were taken to visualize the underlying trabecular structure and the extent of bone remodeling.

Sampling

Subsampling was conducted at the National Museum of Natural History's (NMNH) Museum Support Center in the Biological Anthropology Collections, located in Suitland, Maryland by CAH. In accordance with previously studies (Warinner et al 2014b), a bleach sterilized dental scalar and nitrile gloves were used to scrap dental calculus directly from the individual into a sterile 1.5ml tube.

DNA Extraction and Library Preparation

All laboratory methods and analyses were performed at the Laboratories for Molecular Anthropology and Microbiome Research (LMAMR) at the University of Oklahoma in Norman, Oklahoma. Dental calculus was processed by RMA for extraction and library construction in the LMAMR Ancient DNA Laboratory, a dedicated, six-chambered clean room, abiding by established contamination control workflows, including physical separation from other laboratories performing molecular research, unidirectional work, and positive air flows to avoid cross-contamination. Throughout, RMA was outfitted with full-body Tyvek suit, mask, and gloves to protect the sample from investigator contamination. Non-template extraction controls and reagent blanks were also processed in parallel with the sample.

With slight modifications, DNA extraction followed previously established protocols (Ziesemer et al 2015). No UV irradiation was applied to samples before washing and digestion steps. Briefly, 4.9mg of dental calculus was washed with 0.5M EDTA for 15mins before being allowed to decalcify in a solution of (fresh) EDTA and 10% proteinase K (Qiagen) at room temperature on a nutator for 72 hours. A modified Qiagen MinElute silica-column based purification was performed to isolate DNA into a final 60ul elution volume. The extraction elution was then quantified using a Qubit High Sensitivity fluorometer (Life Technologies), yielding 0.176ng/ul of extracted DNA.

Prior to library construction, partial uracil-DNA-glycosylase (UDG) treatment was applied in order to ensure the longest possible, authentic read lengths (Rohland et al 2015). Partial UDG treatment achieves this by eliminating interior nucleotide damage that would otherwise cause polymerase enzymes to detach, while also preserving terminal nucleotide damage for later DNA authentication (Rohland et al 2015). Half of the extract (30 μ l) was treated with partial UDG and built into dual indexed Illumina shotgun libraries using the NEBNext DNA Library Prep Master Set (E6076). With the exception of Nebulization, NEBNext manufacturer's instructions were only slightly modified. Blunt-end repair immediately followed partial UDG treatment by adding only the NEBNext End Repair enzyme followed by a 30-minute incubation at 20°C. Other End Repair buffers were not added because reaction reagents and conditions were already present from the partial UDG treatment. Samples were then purified with a Qiagen MinElute silica spin column and eluted in 15 μ l EB. Illumina adapters were then ligated to each sample, followed by a 15-minute incubation at 20°C. Before entering the Adapter Fill-In step, samples were cleaned with a Qiagen MinElute silica spin column and eluted in 15 μ l EB. Adapter Fill-In was incubated for 20 minutes at 37°C and diluted with PCR grade water for a final volume of 50 μ l.

Library quantification was achieved with a real-time quantitative PCR assay (qPCR, Lightcycler 480 Roche), using primers *IS7_short_amp.P5* and *IS8_short_amp.P7* 3 to quantify ng/ μ l. qPCR conditions were 0.3 μ M of forward and reverse primers, 1X KAPA Hifi Uracil+ ReadyMix PCR Kit (KapaBiosystems), 0.1 mg/mL bovine serum albumin, 8 μ M SYBRgreen fluorescent dye (Invitrogen), and 1 μ l template DNA in a 25 μ l final qPCR volume. A 95°C initial denaturation (5 min) followed by 45 cycles of 98°C (20s), 50°C (15s), and 72°C (30s) was performed.

Sample libraries were dual-indexed in triplicate, in 25 μ l PCR reactions, with 4 μ l library input, 12.5 μ l KAPA U + master mix, 6 μ l H₂O, 1 μ l of 2.5 mg/ml BSA, and 0.75 μ l unique forward and reverse barcodes (10 μ M). Thermocycling conditions were 5 min at 98°C followed by 14 cycles (based on the qPCR fluorescence values) of 98°C (20s), 60°C (15s), and 72°C (30s), followed by a final elongation step 72°C (1m). Amplified PCR replicates were then pooled and purified with a Qiagen MinElute silica spin column in 30 μ l EB. Extraction and library controls were processed simultaneously.

Sequencing

Following Fragment Analyzer quantification, sampling pooling, and Pippin Prep purification to remove adapter dimers, libraries were sent for sequencing on an Illumina NextSeq at the Max Planck Institute for the Science of Human History (Jena, Germany) using 75bp paired-end read chemistry.

Shotgun Metagenomic Data Analysis

Shotgun metagenomic reads were quality filtered before downstream analysis. Adapters were removed, paired-end data was merged, and quality filtered for base quality (at least 20 bp) and read length (30 bp minimum) using AdapterRemoval v2 (Schubert et al 2016). Analysis ready reads were then mapped to the Greengenes database and human genome (hg19) using the Burrows-Wheeler Aligner v0.7.12 (BWA) and bowtie2 v1.1.2, respectively (DeSantis et al 2006, Langmead & Salzberg 2012, Li & Durbin 2009). Samtools v1.5 (Li et al 2009) further filtered unmapped and low quality reads, sorted, removed PCR duplicate and multiple mapping reads before converting files to fasta format. Potential contributing sources were calculated based the bacterial and archaeal taxonomic identification using a closed reference taxonomic assignment in Qiime v1.9.1 (Caporaso et al 2010) and SourceTracker v1.0.1 (Knights et al 2011). Weighted and unweighted UniFrac distance matrixes were generated in QIIME v1.9.1 (Caporaso et al 2010) and visualized in R (R Core Team 2016) as PCoA plots using ggplot2 (Caporaso et al 2010, Wickham 2011). Metagenomic profiles were also analyzed and generated using MetaPhlAn2 (Truong et al 2015). PanPlan v1.2.2 (Scholz et al 2016) was used for strain identification based on the pangenome composition of reference genomes. To authenticate reads as ancient, mapDamage v2.0.8-2 (Jonsson et al 2013) was employed to estimate the nick frequency, cytosine deamination rates, and read length patterns, which were then visualized in R using the ggplot2 (Wickham 2011). Extraction and library controls averaged 1.5 million reads and less than 12,000 unique filtered reads mapping to the human genome, as opposed to the 131,585 reads mapping from St.L Individual.

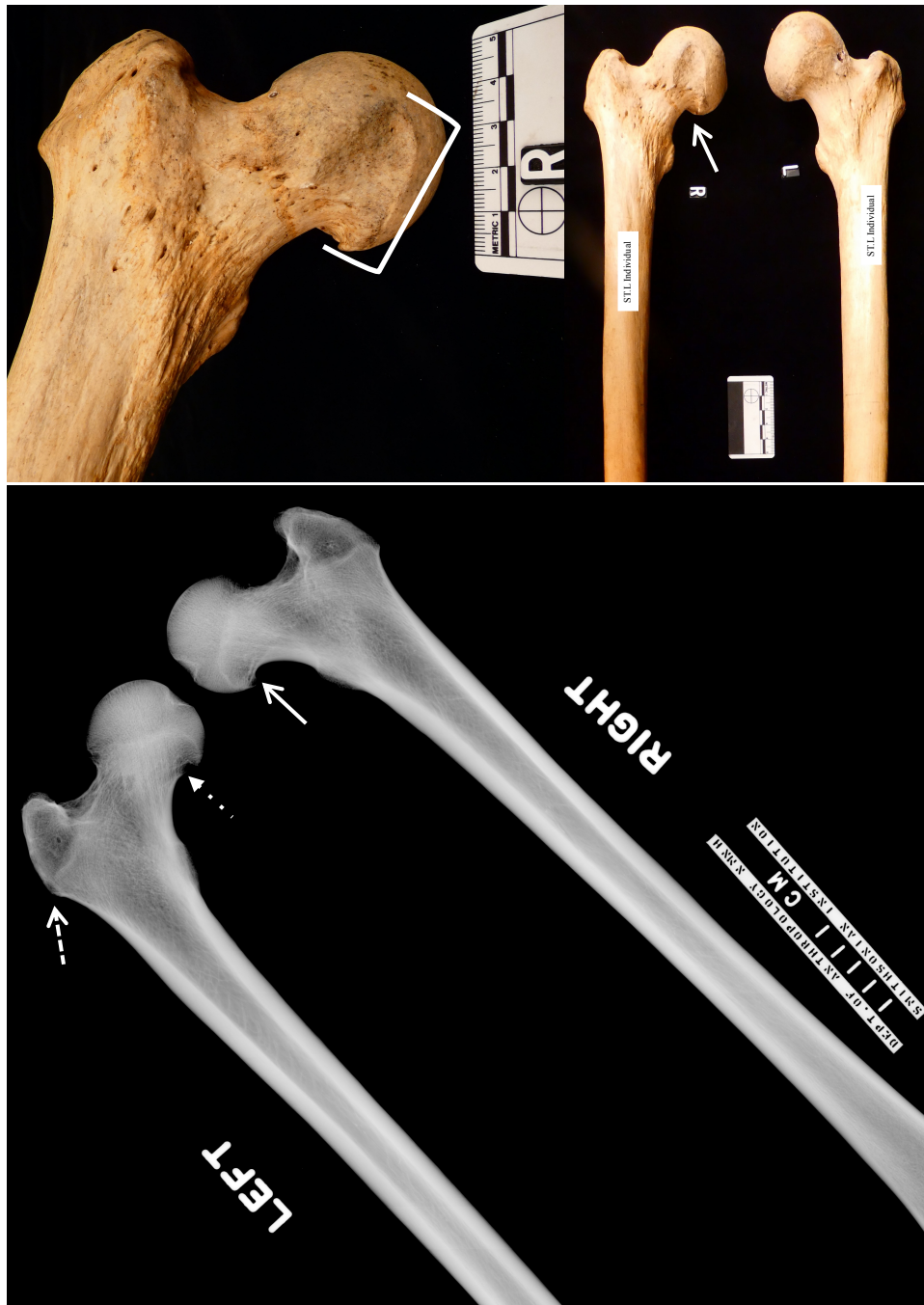
Supplementary Results

Because of the deep personal history information gained during this study we boldly position St.L Individual as an active participant in this research.

A holistic skeletal, oral microbiome assessment, and documentary assessment bared a complex and personal history. In addition to the acute infection responsible for their death, St.L Individual experienced co-morbid chronic disease conditions: localized oral infection and probable tuberculosis infection. Physiological stress (allostatic) load can cause immunological suppression (Padro & Sanders 2014), particularly of the cellular immune response (CIR) against other pathogens (Cohen et al 1991). This relationship systemically affects the body, since a hyper-inflammatory state can cause both localized and systemic inflammatory tissue damage and an elevated inflammatory response to other chronic infections, such as those causal to dental caries and periodontal disease (Crespo et al 2017). Bi-directional relationships can also occur between oral pathologies and systemic inflammatory conditions; strong localized inflammatory responses, such as those in periodontal disease, can become systemic through spillover of inflammatory cytokines and other active immune cells. In some cases, this can lead to chronic inflammatory systemic diseases (Straub 2011).

10,084,925 million reads were recovered, with blanks sequenced to a depth of 0.7-8 million reads. The extensive genetic data gained through this study has also led us to retain the anonymity, both in identification and genetically, of St.L Individual.

Supplementary Figures



Supplementary Figure 1. (A) Focal depression with clear margins on inferolateral portion of the anterior aspect of right femoral head, extending onto the anterior portion of the femoral neck (c. 23 mm in maximum diameter). (B) The trait (arrow) is not bilaterally symmetrical. (C) Radiographs of the left and right femora also show limited reorganization of the trabecular bone underlying the depression on the right femur (solid arrow). The left femoral head also shows osteolytic bone resorption occurring on the inferior aspect (dotted arrow) and a lytic lesion within the trabecular bone (dashed arrow), that is not observable on the bone cortex.



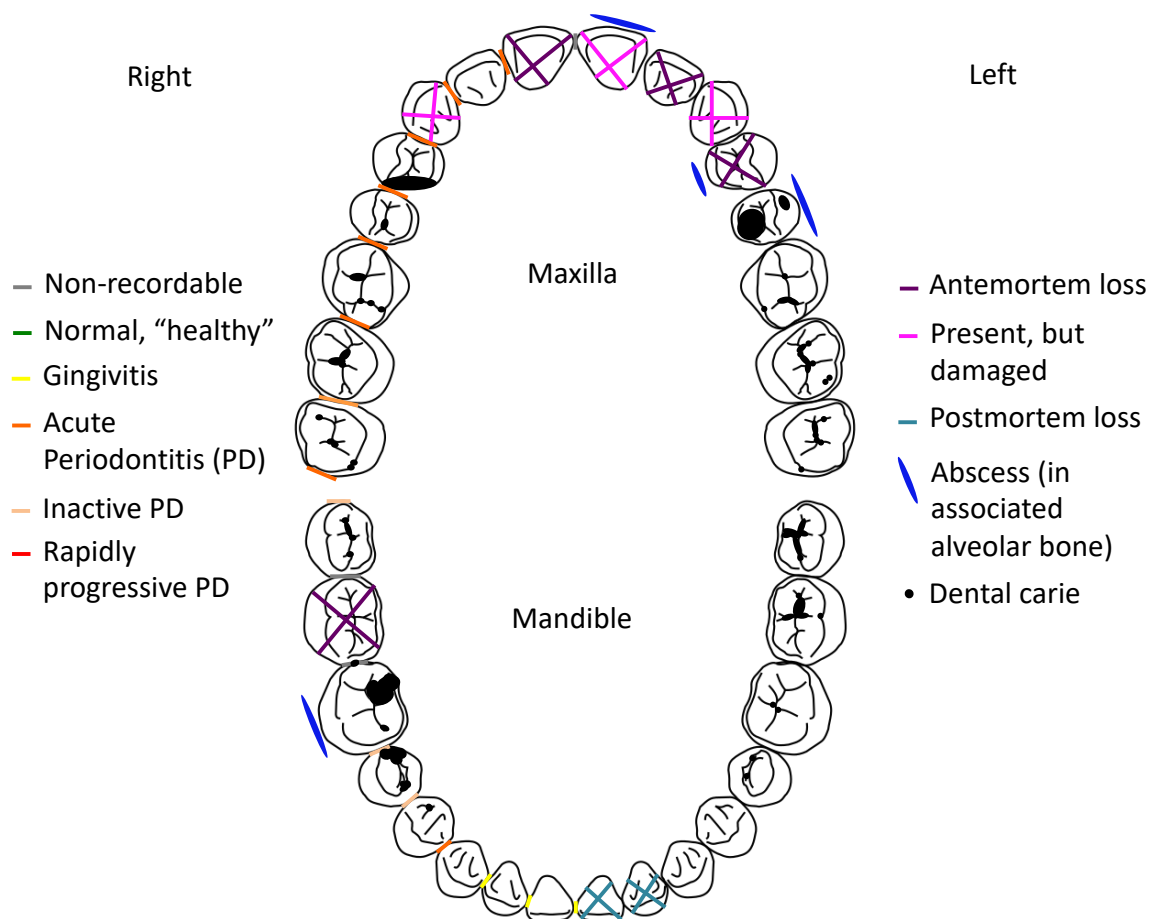
Supplementary Figure 2. (A) Anterior and (B) posterior views of complete midline mandible fracture. The calcified hematoma (solid arrows) and probable fracture line (dashed arrow) are indicated.



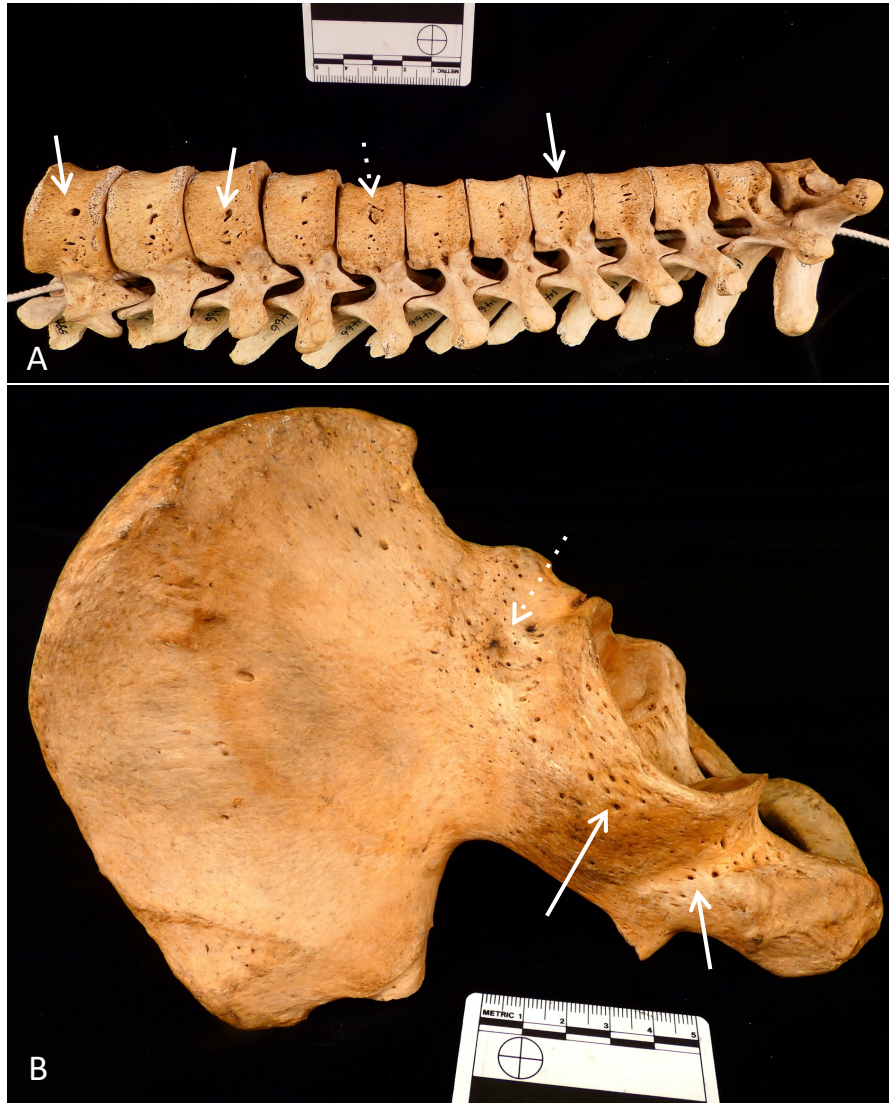
Supplementary Figure 3. Anterior view of the proximal left femur showing two clustered lytic pits (inferior pit: c. 23 mm in maximum diameter; superior pit: c. 10 mm in maximum diameter). The margins are sharp, meaning that the lesion was likely active at time of death, with irregular trabecular bone within the base of the pits.



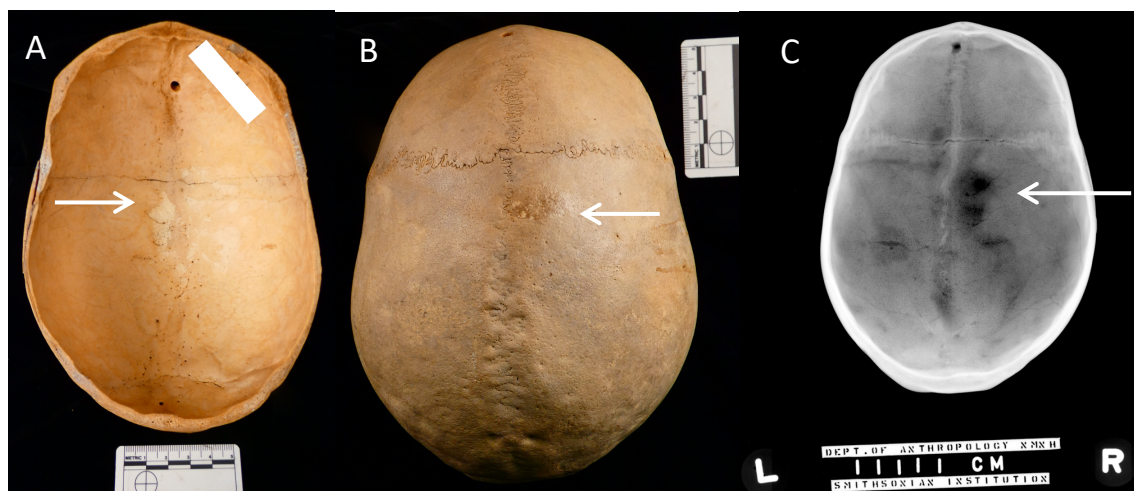
Supplementary Figure 4. The right (A) and left (B) os coxae. Clustered pits, with rounded margins indicative of healing, are observable on medial portion of the pubic crest (dashed circles). The arrows indicate the presence of multiple deep cavities with lytic pitting in the acetabulae. The borders of the pits exhibit woven bone growth with dense trabeculae. The lesions on the right (A) os coxa have sharp margins, indicating that they were likely active at time of death, while those on the left (B) have rounded margins, which is indicative of healing. These lesions correspond to the ACE2 indicator, which has high diagnostic value, for tuberculosis infection (Pedersen et al 2019).



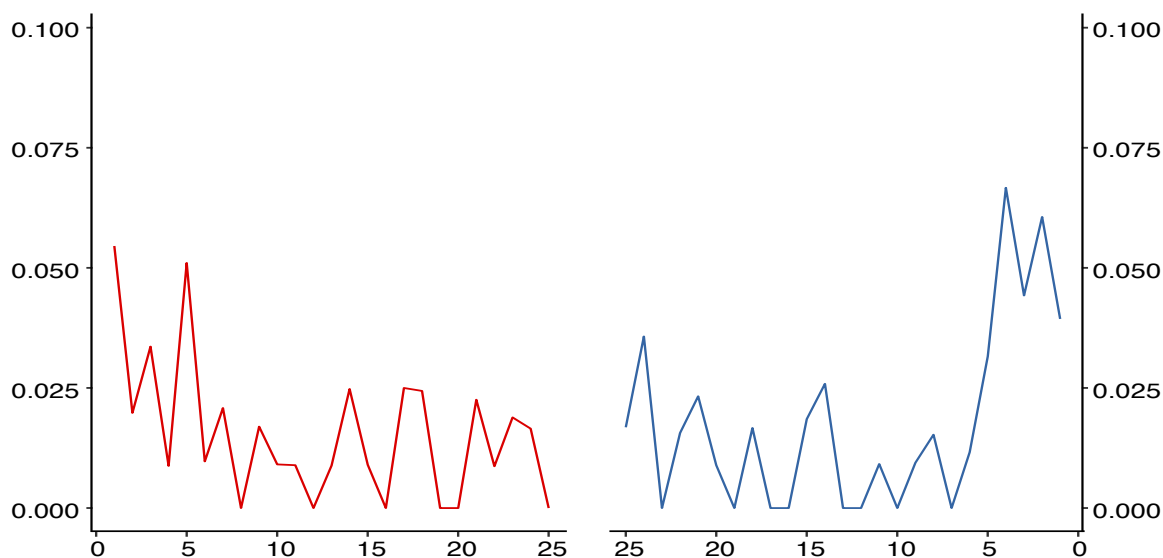
Supplementary Figure 5. Schematic indicating tooth and alveolar bone location of dental caries, antemortem and postmortem tooth loss, periodontal disease, gingivitis, and oral abscesses.



Supplementary Figure 6. (A) Several vertebrae exhibit large pores or pits (solid arrows). Several thoracic vertebrae, specifically T1 to T12, have irregular to circular shaped pits (measuring ≥ 3 mm in maximum diameter) with rounded margins, indicative of healing. Several lumbar vertebrae also have clusters (≥ 3) of large pits with irregular to circular shapes and rounded margins, specifically L1 to L5. Several of these have irregular trabecular bone at the base (dotted arrow) and correspond to the VEN1 indicator for tuberculosis lesions (Dangvard Pedersen et al 2019). (B) Multiple large, roughly circular pits, are present on the right os coxae, with several measuring ≥ 3 mm in diameter. These lesions correspond to the BOD indicator for tuberculosis (Dangvard Pedersen et al 2019), which has high diagnostic value. Most have rounded margins, indicative of healing (solid arrows), though one lytic foci has sharp margins, and was likely active at time of death (dotted arrow).



Supplemental Figure 7. The cranial vault displays several pathological lesions, which may be associated with Tuberculosis. (A) The endocranial surface exhibits an area of well-remodeled, healing reactive bone located along the mid-point of the sagittal suture (white arrow). Deidentification made with white box. (B) The ectocranial surface shows an area of diffuse, well-remodeled, healing lytic destruction located along the mid-point of the sagittal suture (white arrow). (C) On the radiograph of the cranial vault, diffuse bone loss (radiolucent areas) is evident in several areas of the cranial vault, alongside the sagittal suture and on the parietal bosses. The largest area of loss is indicated (white arrow).



Supplemental Figure 8. mapDamage plot of shotgun reads mapped to the *Mycobacterium tuberculosis* genome reconstruction (Comas et al 2010). The erratic pattern illustrates the uncertainty to authenticate the very small quantity of 402 unique reads mapping to *M. tuberculosis*. *M. tuberculosis* DNA is more commonly recovered in historical and ancient samples near active lesions and tooth pulp (Harkins et al 2015, Sabin et al 2019). More work is needed to discern whether tuberculosis can be recovered from dental calculus.



Supplementary Figure 9. The right side of the mandible reveals a draining oral abscess (white arrow) associated with the 1st molar. The sharp margins of this lesion indicate that it was active at time of death.

Supplementary Tables

Supplementary Table 1. Musculoskeletal activity markers

Skeletal element	Description
Pubis	Clustered pitting on the medial portion of the pubic crest/ tubercle, which originates the rectus abdominus muscle, likely representing muscle micro-trauma.
Clavicle	Area of focal bone loss, likely representing muscle microtrauma, on inferior aspect of the medial portion of the left clavicle (c. 5.2 mm in maximum diameter). Area encompasses an attachment site for part of the pectoralis major muscle.

Supplementary Table 2. Traumatic lesions and ‘other’ pathological lesions not directly associated with TB, following established descriptions and criteria for pathologies and trauma.

Skeletal element	Pathological and traumatic lesions
Left femur	Two clustered lytic pits (larger one: c. 23 mm in diameter; smaller: c. 10 mm in diameter) on anterior aspect of the femoral neck, anterior aspect. Sharp edged lesion, likely active at time of death, irregular trabecular bone within the base of the pits, no evidence of reactive bone.
Right radius	Lytic area on lateral aspect of the radial head (c. 1.5 mm in diameter) No evidence of healing, reactive bone.
Mandible	Well-remodeled fracture, confined within the mandibular symphysis, with ossified hematoma.
Right and Left radii	Clustered pits on the radial necks.
Sternum	Clustered pitting (c. ≥ 3 mm in diameter) on the dorsal aspect of the manubrium. No evidence of remodeling/ healing.
	Clustered pitting (c. ≥ 3 mm in diameter) on areas (c. 11 mm in maximum length) surrounding the right and left costal notches on the sternal body.
Femora	Right femur: Areas of well remodeled periosteal reactions on the medial, lateral, and posterior aspects of the diaphysis (71 mm – 23.5 mm in maximum length).
	Left femur: Areas of well remodeled periosteal reactions on the medial, lateral, and posterior aspects of the diaphysis (159 mm – 73 mm in maximum length).
Tibia	Areas of well remodeled periosteal reactions on the anterior aspect of the diaphysis of the right tibia (c. 12 mm in maximum length).
Scapulae	Right scapula: Clustered pitting on the acromion process (c. 5 mm in maximum length). No evidence of remodeling/ healing. Also, diffuse bone loss, with cortical thinning affecting the infra- and supraspinous fossae.
	Left scapula: Clustered pitting on the acromion process (c. 10 mm in maximum length). No evidence of remodeling/ healing. Also, diffuse bone loss, with cortical thinning affecting the infra- and supraspinous fossae.
Fibula	Right fibula: Area of well remodeled periosteal reaction on the anterolateral aspect of the diaphysis of the right fibula (c. 98 mm in maximum length).
	Left fibula: Area of well remodeled periosteal reaction on the posterolateral aspect of the diaphysis of the right fibula (c. 156 mm in maximum length).

Supplementary Table 3. TB indicators, following Dangvard Pedersen et al. (2019). High diagnostic value indicators are in grey.

Indicator	Score	Description
RIB1	0 (R), 1 (L)	Left rib 8 exhibits rough texture covering c. 5 cm
RIB2	0 (R), 0 (L)	No lesions related to TB present
VER	0 (R), 0 (L)	No lesions related to TB present
VEN1	1 (thoracic), 1 (lumbar)	Thoracic vertebrae: Multiple (≥ 3) large pits with irregular to circular shapes and rounded to sharp margins measuring ≥ 3 mm in diameter are present on T1-T12.
		T6 and 12 also exhibit larger openings (T6: c. 20 mm in diameter; T12: c. 8 mm in diameter) with irregular trabecular bone at the base.
		Lumbar vertebrae: Clusters (≥ 3) large pits with irregular to circular shapes and rounded to sharp margins measuring ≥ 3 mm in diameter are present on L1-L5. L1 also exhibits larger openings (L1: c. 4 mm in diameter) with irregular trabecular bone at the base.
VEN2	1 (thoracic), 1 (lumbar)	Thoracic vertebrae: T7-T9 exhibit bone proliferation on c. 50% of the ventral surfaces.
		Lumbar vertebrae: L1-4 exhibit bone proliferation on c. 50% of the ventral surfaces.
BOD	0 (R), 0 (L)	No bone changes related to TB are present.
ACE1	0 (R), 0 (L)	No bone changes related to TB are present.
ACE2	1 (R), 1 (L)	Right: multiple deep cavities with lytic pitting present on $>50\%$ of the area. Borders exhibit woven bone with dense trabeculae.
		Left: multiple deep cavities, the largest with a diameter of c. 20 mm, with lytic pitting present on $>50\%$ of the area. Borders exhibit woven bone with dense trabeculae.
PF1	0 (R), 0 (L)	No lesions related to TB are present.
PF2	0 (R), 0 (L)	No lesions related to TB present
ILI	0 (R), 1 (L)	Clustered pits present on the left iliac auricular surface, two of which are ≥ 3 mm in diameter.
DF	0 (R), 0 (L)	No lesions related to TB present
PT1	0 (R), 0 (L)	No lesions related to TB present
PT2	0 (R), 0 (L)	No lesions related to TB present
DH1	1 (R), 1 (L)	Right: Clustered pits/pores cover (c. 40 mm in max. height) of the lateral epicondyle and the medial epicondyle (c. 37 mm in max height). Pits also extend onto the anterior aspect.
		Left: Clustered pits/pores cover (c. 39 mm in max. height) of the lateral aspect and the medial aspect (c. 62 mm in max height) of the epicondyles.

DH2	1 (R), 1 (L)	Right: Lytic area on lateral, posterior margin of the trochlea (c. 6 mm in diameter). No evidence of healing, reactive bone.
		Lytic area on medial, posterior, inferior margin of the trochlea (c. 6 mm in diameter). No evidence of healing, reactive bone.
		Left: Lytic area on the posterior articular surface, medial margin of the trochlea (c. 5 mm in diameter). No evidence of healing, reactive bone.
		Lytic area on the lateral margin of the trochlea (c. 3 mm in diameter). No evidence of healing, reactive bone.
		Lytic area on the anterior margin of the trochlea (c. 7 mm in diameter). No evidence of healing, reactive bone.
		Lytic area on the medial, posterior, inferior margin of the trochlea (c. 4 mm in diameter). No evidence of healing, reactive bone.
PR	0 (R), 0 (L)	No lesions related to TB present.
PU	0 (R), 0 (L)	No lesions related to TB present.

Supplementary Table 4. Observed oral pathologies.

Oral pathology	Teeth Affected	% Affected	
Antemortem tooth loss	<i>Maxillary:</i> Left lateral incisor (10); Left canine (12); Right central incisor (8).	% <i>Observable teeth affected by pathology</i>	9.30%
	<i>Mandibular:</i> Right 2 nd molar (31)		
Dental caries	<i>Maxillary :</i> Right 1 st – 3 rd molars (1-3), right 1 st – 2 nd premolars (4-5). Left 2 nd premolar (13); left 1 st – 3 rd molars (14-16).	% <i>Observable teeth affected by pathology</i>	66%
	<i>Mandibular:</i> Right 3 rd and 1 st molars (30, 32); right 1 st -2 nd premolars (28, 29). Left 1 st - 2 nd premolars (20, 21); left 1 st – 3 rd molars (17-19)	% <i>Observable teeth affected by ≥1 caries</i>	11%
Abscesses	<i>Maxillary:</i> Left central incisor (9), labial abscess; left lateral incisor (10), labial abscess; canine (11): labial abscess; left 1 st premolar (12): lingual abscess; left 2 nd premolar (13): buccal abscess.	% <i>Observable teeth (associated alveolar bone) affected by pathology</i>	28%

	<i>Mandibular:</i> Left canine (22): labial abscess; left lateral incisor (23): labial and lingual abscess; left central incisor (24): labial and lingual abscess. Right 1 st molar (30): buccal.					
	% observable Interproximal (IP) spaces	% IP spaces w/ healthy bone (1)	% IP spaces w/ gingivitis (2)	% IP spaces w/ acute PD (3)	% IP spaces w/ inactive PD (4)	% IP spaces w/ rapidly progressive PD (5)
Periodontal disease (PD)	77%	0%	21%	50%	26%	0%

- (#): Tooth number, following Universal Tooth Numbering System/ American System.
- AMTL % affected: percentage of teeth missing with resorbing or fully resorbed alveolus, of all teeth present for observation and all teeth missing, but with intact/ observable associated alveolar bone (codes: 1-2, 4-5, 7 in dentition inventory (Table 2)15; does not include teeth that are present but unobservable (not erupted, in crypt) or missing with no associated alveolar bone).
- Caries: % affected: number of caries present/ number of teeth present for observation (codes: 1-2; 7 in dentition inventory (Table 2)15); excludes all teeth missing/ not present for observation).
- Abscesses: % affected includes portion of alveolar bone associated with each tooth that was present for observation. Abscesses were also recorded for teeth that were missing, either postmortem or antemortem.
- Periodontal disease (PD): scored following Kerr (1988)14 (0-5). Scores were pooled for maxilla and mandible. Periodontal disease was scored (in the same number of possible locations) on the left maxilla and mandible when the right was not present for observation.

Supplementary Table 5. Oral microbial community composition of St.LI as classified by MetaPhlan2, sorted by descending percent abundance.

Microbe	% Abundance
<i>Acinetobacter junii</i>	45.97653
<i>Acinetobacter pittii calcoaceticus nosocomialis</i>	16.6801
<i>Klebsiella pneumoniae</i>	16.0708
<i>Enterobacter cloacae</i>	9.8593
<i>Acinetobacter radioresistens</i>	2.34663
<i>Acinetobacter baumannii</i>	2.32759
<i>Pseudomonas putida</i>	1.49636
<i>Aeromonas unclassified</i>	0.99223
<i>Acinetobacter bereziniae</i>	0.65878

<i>Pseudomonas stutzeri</i>	0.62717
<i>Aeromonas caviae</i>	0.52446
<i>Stenotrophomonas maltophilia</i>	0.48665
<i>Escherichia unclassified</i>	0.44604
<i>Comamonas unclassified</i>	0.33937
<i>Lysinibacillus fusiformis</i>	0.30276
<i>Bacillus cereus thuringiensis</i>	0.17745
<i>Delftia acidovorans</i>	0.16344
<i>Escherichia coli</i>	0.14663
<i>Enterococcus faecalis</i>	0.13167
<i>Kocuria rhizophila</i>	0.06303
<i>Eubacterium saphenum</i>	0.05009
<i>Delftia unclassified</i>	0.04754
<i>Citrobacter unclassified</i>	0.03969
<i>Comamonas testosteroni</i>	0.03261
<i>Citrobacter freundii</i>	0.00898
<i>Parvimonas unclassified</i>	0.0041

Supplementary Table 6. Known causative agents of bacterial pneumonia with citations.

Causative agent	Source
<i>Streptococcus pneumoniae</i>	Weiser JN, Ferreira DM, Paton JC. Streptococcus pneumoniae: transmission, colonization and invasion. Nature Reviews Microbiology 2018;16:355.
<i>Staphylococcus aureus</i>	Aliberti S, Reyes LF, Faverio P, et al. Global initiative for meticillin-resistant Staphylococcus aureus pneumonia (GLIMP): an international, observational cohort study. The Lancet Infectious Diseases 2016;16:1364-76.
<i>Streptococcus pyogenes</i>	Tamayo E, Montes M, Vicente D, Pérez-Trallero E. Streptococcus pyogenes pneumonia in adults: clinical presentation and molecular characterization of isolates 2006-2015. PLoS One 2016;11:e0152640.
<i>Klebsiella pneumoniae</i>	Wyres KL, Holt KE. Klebsiella pneumoniae population genomics and antimicrobial-resistant clones. Trends in microbiology 2016;24:944-56.
<i>Haemophilus influenzae</i>	Everett ED, Rahm AE, Adaniya R, Stevens DL, McNitt TR. Haemophilus influenzae pneumonia in adults. Jama 1977;238:319-21.

<i>Moraxella catarrhalis</i>	Verduin CM, Hol C, Fleer A, van Dijk H, van Belkum A. <i>Moraxella catarrhalis</i> : from emerging to established pathogen. <i>Clinical microbiology reviews</i> 2002;15:125-44.
<i>Neisseria meningitis</i>	Cohn AC, MacNeil JR, Harrison LH, et al. Changes in <i>Neisseria meningitidis</i> disease epidemiology in the United States, 1998–2007: implications for prevention of meningococcal disease. <i>Clinical Infectious Diseases</i> 2010;50:184-91.
<i>Legionella pneumophila</i>	Falcó V, de Sevilla TF, Alegre J, Ferrer A, Vázquez JMM. <i>Legionella pneumophila</i> : a cause of severe community-acquired pneumonia. <i>Chest</i> 1991;100:1007-11.
<i>Chlamydophila pneumoniae</i>	Grayston JT, Aldous MB, Easton A, et al. Evidence that <i>Chlamydia pneumoniae</i> causes pneumonia and bronchitis. <i>Journal of Infectious Diseases</i> 1993;168:1231-5.
<i>Mycoplasma pneumoniae</i>	Mansel JK, Rosenow EC, Smith TF, Martin JW. <i>Mycoplasma pneumoniae</i> pneumonia. <i>Chest</i> 1989;95:639-46.
<i>Chlamydia psittaci</i>	Oldach DW, Gaydos CA, Mundy LM, Quinn TC. Rapid diagnosis of <i>Chlamydia psittaci</i> pneumonia. <i>Clinical infectious diseases</i> 1993;17:338-43.
<i>Coxiella burnetii</i>	Marrie TJ. <i>Coxiella burnetii</i> (Q fever) pneumonia. <i>Clinical infectious diseases</i> 1995;21:S253-S64.