THE UNIVERSITY of ADELAIDE

Ancient DNA studies of dental calculus

Andrew G. Farrer

Australian Centre for Ancient DNA
School of Biological Sciences
Faculty of Sciences
University of Adelaide
AUSTRALIA

Supervisors: Prof. Alan Cooper,
Dr Laura S. Weyrich, A. Prof. Neville Gully,
and Prof. Keith Dobney (University of Liverpool, UK).

Thesis submitted in fulfilment of the requirements for the degree of Doctor of Philosophy

October 2016

Thesis Declaration

I certify that this work contains no material which has been accepted for the award of any other degree or diploma in my name, in any university or other tertiary institution and, to the best of my knowledge and belief, contains no material previously published or written by another person, except where due reference has been made in the text. In addition, I certify that no part of this work will, in the future, be used in a submission in my name for any other degree or diploma in any university or other tertiary institution without the prior approval of the University of Adelaide and, where applicable, any partner institution responsible for the joint-award of this degree.

I give consent to this copy of my thesis, when deposited in the University Library, being made available for loan and photocopying, subject to the provisions of the Copyright Act 1968.

I acknowledge that copyright of published works contained within this thesis (as listed on page iii) resides with the copyright holder(s) of those works.

I also give permission for the digital version of my thesis to be made available on the web, via the University's digital research repository, the Library Search, and also through web search engines, unless permission has been granted by the University to restrict access for a period of time.

12/10/2016

Andrew G. Farrer Date

“But what if in future one should tell such people that there are living more animals in the unclean matter on the teeth in one's mouth than there are men in a whole Kingdom?”

Antoni van Leeuwenhoek, “Letter to the Royal Society”
28th December 1683.

Translation: Digitale Bibliotheek voor de Nederlandse Letteren, 1952

Table of Contents

Thesis Abstract Page i

Publications Page iii

Acknowledgements Page vii

Author's Note Page ix

Introduction Page 1
- *Overview* *3*
- *Working with ancient DNA* *5*
- *Human microbiota* *11*
- *Modern, industrialised human microbiota are distinct from the ancestral state* *18*
- *Ancient DNA analysis of human microbiota* *21*
- *Great Britain allows fine scale analysis of microbiota* *24*
- *Communication benefits science and society* *30*
- *Thesis overview* *32*

Chapter I Page 49
Ancient bacteria from dental calculus as an archaeological and anthropological tool
- *Statement of Authorship* *51*
- *Abstract* *55*
- *Introduction* *57*
- *The human microbiota* *59*
- *Ancient DNA analysis of dental calculus* *62*
- *Considerations when using ancient DNA* *66*
- *Using the oral microbiota to unravel past human behaviour* *70*
- *Conclusions* *75*

Chapter II Page 91
Effectiveness of decontamination protocols when analysing ancient DNA preserved in dental calculus.
- *Statement of Authorship* *93*
- *Abstract* *97*
- *Introduction* *99*
- *Methods* *101*
- *Results* *106*
- *Discussion* *114*

Chapter III Page 127
Diet driven differentiation of oral microbiota in ancient Britain
- *Statement of Authorship* *129*
- *Abstract* *133*
- *Introduction* *135*
- *Results* *138*
- *Discussion* *150*
- *Methods* *155*

Chapter IV Page 165
Biological and cultural drivers of oral microbiota in Medieval and Post-Medieval London
- *Statement of Authorship* *167*
- *Abstract* *173*
- *Introduction* *175*
- *Methods* *177*
- *Results* *182*
- *Discussion* *192*

Chapter V Page 205
Incorporating science communication into standard group research practice
- *Statement of Authorship* *207*
- *Abstract* *211*
- *Introduction* *213*
- *1. Planning and building a science communication programme* *214*
- *2. Producing and promoting content* *215*
- *3. Evaluating success* *217*
- *Conclusion* *220*

Discussion Page 225
- *Overall summary and significance* *227*
- *Oral microbiota are defined by individualistic factors* *230*
- *Changes through time* *236*
- *Working with ancient DNA* *239*
- *Future work* *243*
- *Science communication* *250*
- *Conclusion* *253*

Appendix I Page 263
Chapter III: Supplementary Information

Appendix II Page 281
Chapter IV: Supplementary Information

Appendix III Page 317
Reconstructing Neandertal behavior, diet, and disease using ancient DNA from dental calculus (Weyrich *et al*. 2016)

Appendix IV Page 357
Ethical approval for analysis of human dental calculus (H-2012-108)

Thesis Abstract

Over the last 10 years, the human microbiota has been identified as a major force in human health and disease. Microbiota are the bacterial communities that live on the internal and external surfaces of the body, and comprise ~50% of the total cell count of a human individual. Recent studies have indicted the role of cultural and environmental factors on shaping these bacterial communities, including diet, interaction with people and animals, and medical treatments. However, the majority of microbiota studies are in modern human populations or animal models. Consequently, there is limited knowledge on the diversity of microbiota in the past, and how this diversity has been altered through time.

Ancient DNA analyses of the oral microbiota preserved in dental calculus (calcified dental plaque) offer a way to examine historical microbiota composition. Thus, microbiota alterations through time can be mapped, and, with use of detailed archaeological and historical records, the cultural and environmental factors that trigger change identified. Further, elucidating fine scale population structure may be possible due to the rapid response of microbiota to changing environments. Results from ancient DNA studies are critical in understanding historical microbiota composition and population substructure, examining how microbiota change and adapt through time, defining the health status of historical and modern populations, and indicating routes of investigation for medical manipulation of microbiota in disease prevention.

This thesis provides the most detailed analysis of historical microbiota to date, complemented with comprehensive metadata. Initially, I explored methods to minimize the impact of environmental contamination on analyses of ancient dental calculus collected from museums and archaeological sites. This allowed me to identify the optimum decontamination protocol and prepare over 250 British dental calculus samples from the Pre-Roman period to the Early Victorian period (~ 2,000 years). Utilizing high-throughput shotgun sequencing, I identified distinct, unique bacterial community

structures present throughout history that are driven by the diets of different socio-economic classes and are not evident in the modern oral cavity. I then focused on an ~800-year period of London history (1066 – 1853) and identified the first associations of microbiota and disease in an ancient population. Ultimately, these studies alter our understanding of the modern oral microbiota and help define and calibrate ancient microbiota analysis as a powerful new tool for studying human history.

Finally, I provided a framework for science communication within a research group that provides benefits and training to each member. I highlighted and demonstrated that science communication is a powerful tool for informing and engaging the public, and can provide direct research benefits. Moving forward, this framework should be utilized to disseminate research results to peers and the public with the aim to stimulate collaborations, and inform and engage public in new understandings of human biology, history, and medicine.

Publications

Journals

Weyrich *et al.* 2016 Reconstructing Neandertal behavior, diet, and disease using ancient DNA from dental calculus, Nature, 2016
- Included in this thesis (Appendix III)

Cooper, A.; Weyrich, L.S.; Farrer, A.G. 2015, The relevance of ancient DNA to contemporary disease, Pathology, 47, S28-S28

Weyrich, L.S.; Dixit, S.; Farrer, A.G.; Cooper, A.J.; Cooper, A.J. 2015, The skin microbiome: Associations between altered microbial communities and disease, Australasian Journal of Dermatology, 56, 4, 268-274

Conference Items

Farrer, A.G.; Dobney, K., Weyrich, L.S.; Cooper, A. 2016, "Unappreciated ecosystem diversity of oral microbiota detected in ancient Britain", 7th Annual International Symposium for Biomolecular Archaeology (ISBA7), Oxford, United Kingdom

Farrer, A.G.; Lekis, M.; Weyrich, L.S.; Cooper, A. 2016, "Training and Inspiring University Students: Incorporating Science Communication into Research Groups", Festival of Learning and Teaching, University of Adelaide, South Australia

Farrer, A.G.; Weyrich, L.S.; Cooper, A. 2015, "Ancient DNA studies provide a basis for medical research", Australian Society for Medical Research (ASMR) SA Scientific Meeting, South Australia

Grants and Funding

2015: $1,250 – Royal Society of South Australia Small Research Grants Scheme: "The changing function of the microbiota throughout Medieval and Post-Medieval London, U.K."

2014: $15,000 – L. F. and D. Denholm Scholarship, University of Adelaide

Community engagement

"Extracting DNA: separating the wheat from the chaff"
- Blog Post: Explaining the basis of DNA, ACAD *August 2016*

“ScienceAlive! Career Ambassador”
- Event: Speaking to high school students about science careers and university, National Science Week *August 2016*

"The life on us: Fossils of a microscopic world"
- Article: Outreach to high schools on why ancient dental calculus can be used to study impacts of changing culture, eScience Magazine *May 2016*

"LabARTory"
- Blog Post: Artistic elements of an ancient DNA lab, ACAD *December 2015*

"Thank you to our generous supporters"
- Video: Non-speaking role to thank University supporters, specifically scholarship donor Mrs Denholm *December 2015*

"Science says! - Adelaide"
- Event: Panel show organised by Science Nation, hosted by RiAus *November 2015*

"Revealing a skeleton's history by cleaning its teeth"
- Article: Letter to Mrs. Denholm in the Lumen (University of Adelaide Magazine) *Winter 2015*

"Andrew Farrer – Ancient DNA and the life on us"
- Interview: Discussing the 3 Minute Thesis competition and PhD research, Radio Adelaide *October 2015*

" Ancient DNA and the life on us"
- Talk: TransTasman 3 Minute Thesis: Final 10 of the Asia-Pacific competition, University of Queensland *October 2015*

"Ancient DNA and the life on us"
- Talk: Winner and People’s Choice, University of Adelaide *September 2015*

"OAGR: a brand new genome database"
- Blog Post: Showcasing a new ancient DNA database, Co-author: Jimmy Breen, ACAD *August 2015*

"The Good, the Bad, and the Ugly: Bacteria in your Body"
- Children's University, extra-curricular science for primary school children, Co-presenter: Laura Weyrich *July 2015*

"Cleaning the teeth of skeletons: Ancient DNA studies of the oral microbiota"
- Talk: ADOHTA Professional Development Evening *July 2015*

"Cleaning the teeth of skeletons: Ancient DNA studies of dental calculus"
- Talk: Community presentation, Boroughbridge, UK *March 2015*

"I hope you remembered to shower"
- Blog Post: Insight into sterile ancient DNA protocols, ACAD *February 2014*

Acknowledgements

This all began among the collections of the Natural History Museum, London, where I first met Alan Cooper and Laura Weyrich (and apparently didn't alarm them enough that they quietly disappeared back to Australia and stopped returning my emails). In the three years since that meeting I have amassed a long list of people who need thanking for their guidance and support. Firstly, to my supervisory team, Alan Cooper, Laura S. Weyrich, Neville Gully, and Keith Dobney, who provided stimulating discussion and challenges, listened to, supported, and helped develop my ideas, and have allowed me to develop professionally and personally. Particular thanks to Alan Cooper, who supported my application, pushed my scientific thinking, and supported my interest in science communication.

The Australian Centre for Ancient DNA has been an incredibly supportive and enjoyable place to work and learn. To all members who have shared time with me in the offices and labs, thank you for making the opposite side of the world feel like home. The dental calculus team, which has expanded dramatically since I started three years ago, it has been rewarding to work with you all, and to enjoy a few social events together as well.

Special thanks go to Laura Weyrich and Maria Lekis. Laura, you were a supervisor and a friend throughout. Seemingly endless patience and always able to restore my confidence, it would not have been the same without you. Maria, academia is a bubble world, and time spent with you has made me aware of the world outside the lab and the importance of working as a team to accomplish the most. I will keep "Finder, Minder, Grinder" in mind wherever I end up. To both of you, your friendship has been and is the highlight of Australia.

To friends and family. I've been terrible at keeping in touch but have enjoyed every catch-up and visit. Finally, to my girlfriend: Cadda you were unbelievably supportive for so long. You put up with my flaws and my dark moments, and were there to cheer me up. For this, I will be forever grateful.

Author's Note

"I would like to introduce you all to Georgina. Georgina, as I've called her, lived in Medieval London during some of the worst disease epidemics to ever strike the city, beginning with the Black Death in 1348. Sadly, she died before her 26th birthday, and was buried in an abbey cemetery close to the Tower of London. But today, I have recruited her, and over 200 other men and women from the last one and a half thousand years of British history, for a medical study. A study to look at how their life experiences impacted the bacteria that lived on their bodies."

I opened my 3-minute thesis presentation with these lines. They serve to impress upon the audience, and myself, that beyond the literature, protocols, contamination, bioinformatics, taxonomy, and functional pathways; that behind the samples, this is a human story.

Image credits:
Top – Museum of London, UK.
"Georgina", who provided a dental calculus sample during my visit to the Museum, March 2015.
Bottom –University of Queensland, Australia.
Presenting my research at the TransTasman 3MT final, October 2015.

Introduction

Overview

Recently acquired knowledge of the microbial communities living on and in the human body has been referred to as the beginning of a revolution in the understanding of health and disease (1). There are ~ 7 x 10^{13} microorganisms living on the surfaces of the human body, almost outnumbering the human cells present (2). These communities of microbes are 'microbiota' (3), and they play critical roles within the body, performing up to 10,000,000 functions, many of which the human body is incapable of completing (4). The combined genetic information encoding these functions are the 'microbiome', which contain 100 times more genetic information than the human genome (5) and microbiota are typically inferred by use of taxa specific genetic markers within the microbiome. Microbiota studies rely almost exclusively on genetic inference of the microbial community structure. Commonly these studies discuss the "microbiome" as the topic of research. However, I shall refer to "microbiota" throughout this thesis because the studies discussed and detailed focus on inferring and exploring the structure and functional capacity of the microbial community, not the structure of the genomes of these organisms.

Large-scale microbiota studies only began in 2007 with the Human Microbiome Project (HMP) (6,7), which demonstrated that microbiota interactions with the human body are essential to health and that environmental pressures can alter microbiota and create a dysbiosis that can lead to severe disease (8). Despite the increasing realization of microbiota importance, there is still limited understanding of how the microbial communities are defined and altered across a human population. This thesis addresses this issue, and investigates the factors that define the microbiota at a human population level and have caused alteration to the microbial community over time.

History offers a large-scale, natural experiment to elucidate how microbiota are impacted by evolution, migration, culture, living environment, diet, and disease. Studies of historical microbiota can observe how these factors impact individuals across multiple generations by associating

microbiota with additional archaeological, anthropological, and historical data (9). Such analyses have the potential to inform medical studies of the key factors that could be used as medical manipulations of the microbiota in future treatments of microbiota-associated disease (10). In addition, studies of past microbiota will reveal a novel element of human history and may provide a tool to better understand the history of health and disease. To date, ancient DNA (aDNA) is the only method to allow the recovery of historical microbiota (11). Unfortunately, current ancient DNA studies have lacked the resolution to define the specific cultural or environmental factors that drove microbiota change.

This thesis presents the first high-resolution ancient DNA studies of historical microbiota and provides the first detailed insight into the factors that influenced the microbiota of a historical population. To achieve this goal, I analysed 258 ancient dental calculus samples from Great Britain. These samples were acquired from individuals with detailed archaeological, anthropological, and historical contexts, allowing association of microbiota patterns with multiple cultural and environmental factors. I correlated microbiota community composition and functional profile with a range of population and individual level metrics, including burial location, urban or rural living, age, sex, religion, social status, key historical events, diet, and paleopathology. In addition, I demonstrate new methods and considerations that must be considered in all future studies of dental calculus. Finally, I provide a framework for research groups to effectively use online communication tools to promote interdisciplinary distribution of research concepts and results to peers and the public. The goal of this framework is to allow, research projects, such as the work presented here, to expand and improve through the effective collaboration of multiple historical and medical research fields.

In this introductory chapter, I discuss the ancient DNA field and the developments in technology that have permitted ancient DNA studies of dental calculus. I go on to describe microbiota and their medical importance, with a specific focus on oral microbiota, before I review the studies that have

indicated how evolutionarily distinct the modern, industrialized human microbiota are, which demonstrates the need for the high-resolution studies presented in later chapters. I briefly provide a history of Great Britain, to demonstrate the suitability of using this country as a model to explore population and individual level impacts on the microbiota and I indicate the key events I hypothesise will have defined or altered the historical microbiota. Finally, I discuss how communication of science into the public forum (public facing science communication) has aided in academia, and has the potential to bring together multiple research fields for collaboration and to have a positive impact on society. This introduction concludes with an outline of the following chapters and identifies the key questions they address.

Working with ancient DNA

Ancient DNA analyses provide a direct measure of the genetic state of an ancient individual by recovering preserved DNA (12). To date, the oldest genome successfully recovered and analysed is 700,000 years old (13). However, the use of multiple individuals through time and across space provides insight into the history of populations, lifestyles, and health (14,15). Ancient DNA has been applied to the ancestors of many living organisms, and to multiple extinct groups, including quagga (16), mammoths (17), horses (13), dogs (18), bovids (19), chickens (20), rats (21), and humans (14). Samples commonly used for these analyses include bones, teeth, mummified tissue, coprolites, soils, sediments, and ice cores. In this thesis, I use ancient DNA protocols to identify multiple unknown taxa within a single sample to reconstruct historical microbiota.

Analysis of ancient DNA is difficult due to the degraded and damaged nature of surviving DNA. DNA repair mechanisms cease working at the point of cell death, and over time DNA strands begin to fragment (22). The rate of this degradation is not constant, and increases in warm and wet conditions that promote the chemical reactions that break down the DNA sugar-phosphate backbone (22). As DNA fragments become shorter, the amount

of original endogenous DNA that can be analysed is reduced, potentially to less than 1% of the original content (23). The estimated maximum DNA preservation time for analysable fragments is 1,000,000 years (24). To date this theoretical limit has not been reached (13). Analyses are further confounded by damage occurring within DNA fragments. Blocking lesions and coding lesions can prevent or limit analysis: (22). Blocking lesions prevent ('block') DNA polymerases from moving along the DNA fragment, limiting or inhibiting several laboratory processes. This kind of damage may be a modification of the nucleotides or the result of cross-linking of the DNA strands. The alternative form of damage, coding lesions, do not block DNA polymerases, but result in alterations to the DNA base sequence. A common miscoding lesion is the deamination of cytosine to a uracil, although other base modifications can be observed (25). During laboratory procedure, this uracil is 'read' by DNA polymerases as a thymine. This results in a cytosine to thymine transition in the final data (with a corresponding guanine to adenine transition on the opposing strand) (25). This can result in the inference of sequence variation between individuals that is an artifact and not biological. Ultimately, degradation and damage result in very small concentrations of DNA available for analysis (26). However, ancient DNA protocols have been developed to recover and handle these small, damaged fragments.

To study ancient DNA, preserved DNA must be extracted from the sample, amplified, and sequenced before analysis can begin. There have been two major technological developments that have allowed the ancient DNA field to advance. Firstly, the Polymerase Chain Reaction (PCR), which allows rapid amplification of DNA fragments, and subsequently High Throughput Sequencing (HTS), which allows millions of unique DNA fragments to be sequenced simultaneously (27,28). PCR is considered one of the most important innovations in genetics and amplifies DNA exponentially, allowing the extremely low concentrations of DNA from ancient samples to be amplified to useable concentrations (27). A key element of PCR is that it requires primers (short DNA fragments) that have

complementary base pairing to DNA flanking the region to be amplified. These primers allow the DNA polymerase to bind and the DNA duplication to occur. Primers can be designed to target different sequences within the genome, allowing researchers to specifically amplify a chosen, diagnostically important genetic region. This was particularly important with pre-HTS studies, as the sequencing machines could not differentiate different DNA molecules, and thus required a single, clonal DNA molecule (29). However, HTS machines are capable of sequencing millions of different DNA fragments simultaneously (producing up to 120 Giga-bases of DNA sequence on the recently released Illumina NextSeq) (28). This allowed two key developments: firstly, different DNA targets from a single sample could be sequenced simultaneously, and second, multiple samples could be sequenced together (samples are differentiated by a unique, artificial DNA fragment (barcode or index) added during laboratory processing (30)). This allows for two different sequencing approaches: targeted and shotgun sequencing. When using targeted sequencing, researchers pre-select specific genetic regions for analysis and can apply a number of strategies, such as metabarcoding, to recover these informative regions. Metabarcoding uses PCR to target and amplify the same diagnostic region (barcode region) from multiple taxa. Requirements for these regions are that they appear in every organism of interest (so that a single primer pair can amplify from all taxa and the data are comparable between organisms) but have sufficient variation that each taxon has a unique, differentiating sequence (31). For bacterial taxa, a common diagnostic region used is the 16S ribosomal RNA (rRNA) gene (30). This is an ideal candidate as it is fundamental to cellular processes (coding for a ribosomal subunit) meaning it is conserved across all taxa, but it has sufficient variation to differentiate genus level taxonomy reliably. The effectiveness of 16S rRNA gene sequencing means that it is the standard for cataloging bacterial taxonomy (30). Targeted sequencing methods, such as metabarcoding, allow identification of specific informative regions, thus reducing the amount of DNA to be sequenced to provide appropriate information for the study.

However, the flexibility of HTS allows researchers to sequence DNA without any pre-selection.

Shotgun sequencing allows recovery of a representative subset of all the DNA fragments preserved within a single sample without any pre-selection. Rather than having primers that identify a specific region of the DNA, adaptor sequences are ligated (bound) to every DNA fragment in the extract. PCR primers that complement these adaptors allow all DNA fragments to be amplified for sequencing (23). However, the adaptor sequences need a blunt end (where both strands of the DNA terminate at the same point) for ligation. As ancient DNA commonly has single stranded overhangs, due to the manner in which DNA degrades, the single stranded sections have to be removed. To achieve this, an endo-nuclease can be used to remove a 5' over hang in line with the 3' end of the opposing strand, and a polymerase can be used to extend from the 5' end of the opposing strand to fill a 3' overhang (32). Depending on the specific protocol used, the damage on the 3' overhang will be retained or removed. Once sequenced, the DNA data need to be analysed, this often includes mapping the reads to known references. When the aim of sequencing is to identify the organism/organisms in the sample, this mapping is done against curated databases of sequences representative of known taxa.

Bioinformatic comparison of sequence data allows the identification of the organisms that contributed DNA to the sample. If this data are from a targeted sequencing approach, it is compared to a specialized reference database that contains only information relevant to the recovered region. For example, for bacterial identifications based on the 16S rRNA gene we can use the GreenGenes database, which is a curated database of only 16S rRNA gene sequences from known bacterial taxa (33). However, for shotgun data, more comprehensive, genomic databases are required as the DNA fragments could come from any region of the genome. Microbial communities also require large, comprehensive databases for identification of all members. Databases for these analyses include the NCBI non-redundant database (34). However, millions of reads being queried against

large databases is computationally expensive because the number of potential matches that must be assessed is large. To combat this, new algorithms have been developed to allow rapid alignments of sequences. For example, MALT and MALTX are up to 10,000X faster than the traditional algorithms available in BLAST or BLASTX (35) and have now been applied in ancient DNA research (36), However, the data of a single Illumina NextSeq run can still take several days to taxonomically identify. Functional information can be gained by the same method, although the sequences are compared to a database of known functions, such as the Kyoto Encyclopedia of Genes and Genomes (KEGG) or SEED databases. Bioinformatic analysis of shotgun data sets can also provide other information, such as damage profiling to ensure ancient authenticity. The single stranded over-hangs at the end of a DNA fragment get damaged more rapidly than the double stranded regions. Identifying the proportion of mismatches between a sequence and the reference at each base position allow the distribution of damage across the DNA fragments to be quantified. If the DNA is ancient and had the single strand overhang, the number of mismatches will increase toward the ends. However, modern DNA will not display this pattern, having equal (low) damage along the full length of the fragment (25,37). Not only does this help confirm a sample is ancient, it can also be used to identify modern contaminating DNA that has been mixed into the ancient DNA pool.

Contamination is a critical issue for ancient DNA studies, as contaminating DNA can overwhelm the low concentrations of endogenous DNA (38). Contaminating DNA can be broadly split into two groups: environmental and laboratory. Environmental contamination is non-endogenous DNA that is present on or in the sample prior to entering an ancient DNA facility. Laboratory contamination is DNA coming from the laboratory environment and reagents (39). Environmental DNA contaminants that coat the surface or have penetrated the sample typically originate from the matrix surrounding the sample during preservation (*e.g.* soil microorganisms). However, environmental contaminant DNA can leach into

ancient samples from other sources (e.g. bones elsewhere in the site or macroorganisms using the site) (40), particularly if there is water flow to leach and carry the DNA through the matrix at any point while the sample is in situ (40). Further contamination can come during excavation and sample handling (11). To combat this, samples are typically decontaminated as the first step in the ancient DNA facility. This process physically or chemically removes or destroys any DNA on the surface of the sample prior to DNA extraction (41,42). With large samples (*e.g.* bone), the outer layer will be removed, while smaller samples may only be treated with UV radiation or chemicals to destroy DNA on the surface. While important, these methods will not completely remove contaminant DNA. Consequently, bioinformatics assessment of sequenced DNA against known contaminant databases or by processing associated environmental samples in parallel to assess the similarity of DNA profiles are important methods to filter contaminant DNA from the data set prior to analysis (40,43).

Laboratory contamination comes from the laboratory environment and reagents. The first step to avoid laboratory contamination is to perform ancient DNA research in a dedicated facility. Ancient DNA laboratories should be isolated from other genetic facilities and have no PCR products or other solutions with high concentrations of DNA. There should be positive air pressure to constantly force contaminated air out of the lab, while avoiding influx of contaminated air via a back draft. The entire laboratory should be regularly cleaned with DNA destructive methods (*e.g.* bleach and UV radiation), and individual workspaces should be similarly cleaned prior to and following any work. These workspaces are typically still air environments to limit cross contamination of DNA between samples being handled simultaneously. Personnel should wear sterile, full body suits, gloves, boots, masks and visors to limit the introduction of contaminant DNA from the person processing the samples. Clean equipment and sterile technique are essential (44,45). However, to account for the inevitable contamination, particularly from microorganisms, negative controls must be included in all sample processing (26). Post sequencing, a bioinformatics assessment of

these negative controls can identify contamination from within the laboratory and exclude matching DNA sequences or taxa assignments from samples. Once these controls are complete, it is important to correlate the data set with records of laboratory procedures (e.g. the batches samples were handled in during extraction) to confirm that the data set is not associated with a systematic bias not accounted for by the previous filtering steps. Each of these steps is critical to maximize the analysis of genuine ancient patterns.

The combined use of modern sequencing technologies and careful contamination control has allowed ancient DNA researchers to study a wider and more ancient range of organisms than ever before. Population level studies are now practically, financially, and computationally feasible, and more detailed studies on human behavior, lifestyle, and diet have emerged. These advances have only recently made studies of historical and ancient human microbiota possible. The development of HTS was critical to allowing whole communities to be studied, and shotgun sequencing is beginning to allow direct identification of community functional capacity. Understanding the extent of human microbiota and their role in human health is now possible, and is providing important insight into a previously unrealized element of human-environment interaction.

Human microbiota

Microbiota are the microbial communities on the body and include bacteria, lower eukaryotes, and archaea (3). The functional capacity of the microbiota has caused it be considered an additional organ (46). The term 'supraorganism' has been applied to describe the mutualistic links between microbiota and the human organism (6). Each body site harbors a unique mixture of microorganisms, for example, the gut hosts species from the *Bacteroidetes, Dorea/Eubacterium/Ruminococcus, Bifidobacteria, Proteobacteria,* and *Streptococci/Lactobacilli* groups (47). In contrast, the skin has microbial species that are dominated by *Corynebacteria*, *Propionibacteria*, and *Staphylococci* (48). The oral cavity, as a third example,

contains *Streptococcus*, *Haemophilus*, *Actinomyces*, and *Prevotella* species (49). However, the extent and diversity of these microbiota within the human body is still being realized, primarily through the use HTS DNA sequencing (50). Analyses of microbiota from living people have demonstrated that these communities are seeded at birth and are molded throughout life by numerous genetic and environmental factors (8).

The different sites of the human body have different microbiota that perform specialized functions (49). During a natural, vaginal birth, the child is inoculated with the vaginal and faecal microbiota of the mother (8). During the first years of life, the specific environmental pressures at each body site alter and define the community at that site (51). This differentiation results in unique functional profiles that can complete roles the human body has not evolved to do, resulting in an ecological balance that benefits the human host and the microbial communities. The most highly studied microbiota functions are those in the gut. Gut microbiota are associated with a series of functions including the digestion of compounds the human body cannot metabolize (releasing otherwise inaccessible nutrients) and the production of vitamin K (52). The processing of substances in the digestive system also impacts how the body perceives drugs, as xenobiotic breakdown by microbes determines the chemical structure of the drug as it enters the blood stream (4). In addition, microbial action can directly impact bodily functions. For example, uptake and storage of nutrients in the form of fat is altered depending on the microbial community present, and transplants of gut microbiota have resulted in obesity in mice without a dietary change (53). Further, microbial functions also impact the human immune system (54). The gut microbiota are critical in immune development, and microbial imbalance in the gut is associated with inappropriate immune response elsewhere in the body (*e.g.* allergic reactions in the respiratory system) (55). In addition, the impacts on the immune system can alter the brain environment (*i.e.* metabolites produced by gut microbes permeate into the blood and impact brain function) (56). As a result, investigations into gut microbiota and their

functions are a major focus of medical research to improve and maintain health.

In comparison to the gut, microbial functions at other body sites have not been examined in great detail. Consequently, the roles that non-gut communities play in human health are not well understood. However, initial studies indicate that these communities similarly impact human health. For example, skin microbiota help maintain skin health by digesting dead epithelial cells and produce moisturizing oils (57). They also impact immunity, as native bacteria can directly trigger an immune response to a pathogen (58). In the mouth, the microbial communities aid in enamel restoration (59). However, for other sites, the direct beneficial functions of the microbiota are unknown. It is commonly noted that the microbiota are altered during disease. In the lung microbiota, shift from the Bacteroidetes phylum to the Proteobacteria phylum is commonly seen in chronic disease states. The effectiveness of corticosteroids and antibiotics in asthma and death rates related to idiopathic pulmonary fibrosis are associated with microbiota composition (60). In the vaginal microbiota, *Lactobacillus* is dominant in healthy individuals. However, imbalances of this taxa have been associated with disease (vaginosis) (61). In many cases it is still unclear if microbiota alteration is a result of disease or a causal element (60).

The reasons why microbiota change are not well understood. Change in diet, environmental alterations, or exposure to disease can alter the microbiota and contribute to the introduction of novel microbes (62–65). The microbiota also naturally alter over time with human development and aging (51). Of these factors, dietary alterations appear to be the major driver of microbiota change, especially in the gut microbiota. Switches in diet, from high protein and fat to high fibre, have been shown to trigger shifts in the abundance of microbes within 24 hours (62). Longer-term changes in diet, such as those in individuals on diets to lose weight, see a shift in gut microbiota from one associated with obesity to a structure seen in healthy weight individuals (53). Dietary shifts alter the nutrients available to the microbiota, thus impacting the specific community structure that can be

supported. Dietary items are also a source of novel microbes being introduced to the body. For example, bacteria present on food items could be identified and were metabolically active in the faecal microbiota, suggesting that these microbes could have impacted or being integrated into the oral or gut communities (62). The impact on the overall ecosystem of these introduced microbes is unknown but potentially significant.

In addition to the diet, novel microbes can be introduced from multiple environmental sources, such as other individuals. Enclosed living spaces trap human-associated microorganisms and exclude environmental microorganisms, which can result in an increased number of microbes shared between individuals (65). While infants typically inherit microorganisms from their mother, they can also receive bacteria from their primary caregivers and siblings (66). Non-human animals (*e.g.* pets) are also a source of unique microorganisms in the home. In comparison to children without pets, children living with pets (cats and dogs), had increased levels of *Clostridiaceae*, *Veillonella*, *Peptostreptococcaceae* and *Coprococcus*, and decreased levels of *Bifidobacteria* (64). The overall increased diversity has been associated with the healthy development of the immune system and negatively correlated with immune related diseases, such as allergies and asthma (67). Interestingly, children with older siblings are noted to have a decreased diversity of gut species (including *Peptostreptococcaceae,* which was enriched when pets were present) (64). Consequently, identifying the key drivers of microbiota structure and change, even within a single home, is complex.

Chronic and infectious diseases are associated with the microbiota within the body. Imbalances in microbiota structure as a result of environmental factors or disease can trigger or exacerbate disease (68). However, microbiota also aid directly and indirectly in the defence against disease. For example, infection of the skin by *Staphylococcus aureus* can be inhibited by skin microbiota (69). In addition, immune response may not be triggered without signaling from the microbiota (58,70,71). Microbiota community members also cause disease if they are able to grow in excess.

Examples include Helicobacter pylori, which is a native bacteria of the stomach but can cause stomach ulcers if present in excessive quantities (72). Microbiota are also associated with systemic diseases such as arthritis, cancer, and mental disorders. In arthritis patients, the gut microbiota member *Prevotella copri* has been associated with arthritis (73), and multiple other joint related diseases are associated with members of the gut microbiota, including Whipple disease, which is triggered by *Tropheryma whipplei* (74). Alterations in the gut microbiota can result in chronic inflammation, inappropriate immune responses, bacterial translocation, and epithelial toxicity, which drive the formation of cancerous tumors. Notably, treatment of cancer is also impacted by microbiota. Common treatments such as immunotherapy, chemotherapy, and radiotherapy perturb the microbial community, which can result in further cancer supporting actions from the altered microbiota structure (75). Microbiota impacts of immune, hormonal, and neural interactions can also lead to alteration of brain function (76). Studies have indicated the role of multiple microbial species in aspects of metal health. In humans, for example, *Lactobacillus helveticus* and *Bifidobacterium longum* are associated with reducing anxiety and depression (77), and *Lactobacillus casei* was linked to an improvement in mood (78). However, *Bifidobacterium animalis, Streptococcus thermophiles, Lactobacillus bulgaricus,* and *Lactococcus lacti* have been associated with negative impacts on the control of emotion and sensation (79). These examples indicate the importance of appreciating microbiota interactions that control or trigger disease, but also microbiota as an element of the treatments, which must be returned to a healthy state to support the body's recovery.

Oral microbiota

Oral microbiota contains ~ 700 microbial species that coat the three major tissues in the mouth: epithelial cells (e.g. cheek cells), the tongue, and the teeth (80). The oral environment is a key selector for microbial diversity. The oral cavity is warm (~35°C), moist, and has a relatively neutral if

fluctuating pH-value. Saliva is a source of ionic and nitrogenous compounds, enzymes, glycoproteins, and proteins, which can be used as resources by microbes (63,81–83). *Streptococcus* species dominate the mouth almost universally. Other common community constituents are *Veillonella, Corynebacterium, Actinomyces, Prevotella, Fusobacterium,* and *Treponema* species (80). These microbes must adhere to the oral surfaces or to each other to colonize the mouth and avoid being swallowed. *Streptococcus* species and *Lactobacillus* species predominantly adhere to surfaces (e.g. tooth enamel) (84). Following this initial binding, further bacteria, such as *Actinomyces, Capnocytophagae, Haemophili, Prevotellae, Propionibacteria,* and *Veillonellae* can then bind, conglomerate, and begin to form a biofilm. The biofilm is most obvious on the tooth surface and is called dental plaque. This biofilm is associated with healthy oral functions and aids in restoration of tooth enamel (59). However, the high presence of simple sugars in the modern diet is related to excessive plaque formation, which can result in an outgrowth of pathogenic bacteria (83).

Oral disease is often triggered by an overgrowth of specific pathogenic microbes (*i.e.* a dysbiosis within the oral microbiota). The two major oral diseases linked to microbiota alteration are dental caries (cavities) and periodontal disease (gum disease). Dental caries were classically caused by *Streptococcus mutans*, which produce acids that dissolve the enamel surface and erode the tooth (85). However, the disease has been recently found to be much more polymicrobial (86), and requires multiple species within the oral microbiota. Dental caries impact 60-90% of children and almost 100% of adults worldwide (87). Additionally, inflammation of the gums occurs with the presence of the Red Three complex (*Porphymonas gingivalis, Tannerella forsythia,* and *Treponema denticola*) (88) and afflicts 15-20% of middle aged adults globally (87). Expenditure on oral health accounts for 5-10% of all public health funding (87). Consequently, treatments to manage the oral microbiota and prevent the excessive growth of pathogens are of major importance.

Oral microbes are also related to systemic disease outside of the mouth. The oral cavity is a reservoir of pathogens that can spread to physically nearby sites, such as the larynx (89). Poor oral health has been linked to lung diseases such as pneumonia due to the transfer of oral microbes to the lungs, particularly in immune-compromised individuals (90). Transmission of microbes can also occur as a result of the ulceration and bleeding of the gums during periodontal disease allowing microbes to enter the blood stream and to be distributed throughout the body (91–93), including the heart, lungs, and peripheral capillaries, within one minute (94). Oral microbes have been associated with heart disease, particularly linked to periodontitis, which is commonly the result of the Red Three complex (*Porphymonas gingivalis, Tannerella forsythia,* and *Treponema denticola* (88)). However, dental caries do not have an association with heart disease (90), indicating a specific relationship between arthritis and periodontal disease. Arthritis is also associated with oral microbiota, with *Porphymonas gingivalis* and *Prevotella* species being potential trigger taxa. However, microbial products (*e.g.* heat shock proteins) may be the factors triggering immune response within the joints (95). As mentioned above, links between cancer and oral health have been observed, particularly noting that oral and esophageal cancers are closely related. There is also evidence that pancreatic and gastric cancer are associated with oral health, distinct oral microbiota have been identified in cancer patients. A key genus associated with cancer is *Streptococcus,* which has a reduced relative abundance in individuals with oral cancer, compared to healthy individuals (96). However, the causal relationship between oral bacteria and disease is not yet fully understood. Further studies are required to specify causal relations between oral microbiota and disease states, where present, to develop future medical treatments for these diseases.

The importance of the oral microbiota in both oral and systemic health and disease make it a key community to examine in greater detail. While this community is among the most stable on the body (49), environmental factors, such as diet, can significantly alter oral microbiota and

can have systemic impacts. Further, it is the point of entry to both the intestinal and respiratory systems and can also serve as a reservoir or site of interaction for the transfer of novel microbial species into different body sites (60). The range of potential information and direct medical applicability indicate the need to study the oral microbiota in detail.

Modern, industrialized human microbiota are distinct from the ancestral state

Humans have unique microbiota in comparison to other mammals, and human microbiota vary between human groups. Studies have used comparisons of different mammalian and human groups to explore the drivers of this variation. Throughout great apes and other mammals, microbiota differentiation largely follows the evolutionary history of the host (97,98). Studies focused on chimpanzees indicate that the microbiota could differentiate chimpanzee species, but not necessarily indicate family relations within a species (99,100). However, geographic location or resources availability appears to play a role because the gut microbiota of gorillas and chimpanzees at the same location share more similarity with each other than with members of the same species in different locations (101). However, comparisons of human microbiota to that in African great apes indicate that humans have diverged dramatically from the great ape ancestral community, even more than is expected from species to species (102). This human distance is not explained by geography. When humans, chimpanzees, and bonobos that were living in the same geographical area (wildlife reserves in Africa) had their oral microbiota examined, the chimpanzees and bonobos were more similar to chimpanzees and bonobos living in other reserves than to the local human group (103). This suggests that humans have been subject to another driver of microbiota change that has made them distinct from other Great Apes. Studies indicate that this driver may be diet. Emphasis has been placed on an increased intake of meat being associated with human gut microbiota composition (102). The role of diet in the

development of the human microbiota was further inferred when non-human primates were demonstrated to develop a more human-like microbiota as their diet was humanized. Specifically, the change in diet was a reduction in fibre and diversity of plants eaten (104). Further studies of great apes have indicated the importance of social roles in seeding and defining the microbiota (105), suggesting that human cultural development and the changing social roles between individuals will also have resulted in microbiota alteration. Despite these differences, human gut microbiota still harbor ancestral bacterial lineages that can be found in chimpanzees (106), indicating that humans have not undergone a complete microbial replacement. However, these studies lack the temporal resolution to identify the specific factors or the timing of the shifts that caused the human microbiota to differentiate from other primates. Many of the shifts may have occurred during hominin development, while others will be specific to anatomically or behaviorally modern humans and thus cannot be resolved by studying beyond the human lineage.

Comparisons of modern human groups living different lifestyles can provide increased resolution on microbiota change within behaviourally modern humans, indicating changes since the last common ancestral population. Studies have compared the microbiota within modern hunter-gatherer groups to that observed within industrialized individuals (*i.e.* people living a traditional, ancestral-like hunter-gatherer lifestyle versus individuals from Western countries with agriculture and industrialized technologies). Peoples living hunter-gatherer lifestyles had microbiota distinct from that of Western individuals (107–109). Hadza hunter-gatherers from Africa maintained unique gut microbiota that is enriched in *Prevotella* and *Treponema* but lacking in *Bacteroides* due to the heavily plant based diet, which differs from the diet of modern Westerners. In addition, a decrease in *Bifidobacterium* was linked to the lack of dairy and cattle exposure (Hadza do not have domestic cattle) (107). Shifts in *Prevotella* and *Bacteroides* are repeated in multiple non-industrialized groups (108). In a previously uncontacted South American group, oral microbiota were also shown to be

significantly different in comparison to Westerners. Both the South Americans and the USA comparison groups were dominated by *Streptococcus*. However, in comparison to the USA group, the South American group had higher proportions of several oral taxa, including *Prevotella*, *Fusobacterium*, and *Gemella*, and lower proportions of *Rothia*, *Stenotrophomonas*, *Acinetobacter*, or *Pseudomonas.* Interestingly, while the gut and skin microbiota of the South Americans were more diverse than the USA group (*i.e.* had a greater number of unique taxa present), the oral microbiota had the same level of diversity. This was suggested to be the result of tobacco chewing (108), which is known to lower oral diversity (110). The distinct cultural histories of the South Americans and the USA group were also suggested to be a factor causing the microbiota differences (108). The different hunter-gatherer groups around the world have distinct microbiota from one another, supporting the hypothesis that cultural factors and resources impact the microbiota (107). Notably, this trend is not observed between different Western groups with similar geographical separation (107), suggesting that shared culture and diet define the Western microbiota (111,112). Studies of microbiota from groups living hunter-gatherer lifestyles do not provide information on when the microbiota became different between groups, nor do they explain why certain microbiota shifts occur in specific populations. Inferences of causality (*e.g.* cultural practices and diet) may be accurate but it is likely that non-industrialized ancestors of now industrialized populations were distinct from the modern non-industrialized groups, as the modern non-industrial are from each other. Consequently, modern non-industrial groups may not represent an ancestral state for individuals in industrialized countries, such as those in Europe and the United States of America. Consequently, direct study of microbiota from historical individuals is necessary to identify ancestral states and to date the point of microbiota alteration.

Ancient DNA analysis of human microbiota

In recent years, ancient DNA analyses have reconstructed whole, historical microbial communities, allowing researchers to study microbiota composition and function over multiple generations (11). This allows identification and dating of microbiota, and the observation of historical periods when the microbiota was altered. Information on microbiota structure can then be compared to the wealth of existing archaeological, anthropological, and historical data to identify the specific events and factors that defined and altered the microbiota in the past. Ancient DNA studies of microbiota can offer new insights into health, disease, and lifestyle in ancient populations, demonstrating the degree and tempo of change leading to the modern microbiota structure. Unfortunately, there are limited sample types available within the archaeological record to examine microbiota, as microbial cells primarily inhabit soft tissue that degrades after death (except in exceptional circumstances, *e.g.* mummification). However, coprolites (preserved faecal matter) and dental calculus (calcified dental plaque) preserve ancient microbiota for routine ancient DNA analysis.

Coprolites are sub-fossil faecal matter, and contain bacteria that are representative of the gut microbiota (113). However, the preservation of the gut community is often biased, limiting routine use (114). For example, initial studies examined gut microorganisms in 1,400-year-old coprolites from Northern-Chile and Mexico (115). The Mexican sample grouped more closely with modern African children and primates than with modern USA individuals, as expected. However, the Chilean analyses revealed that the bacteria within the coprolite matched mostly to compost (115,116). While modern-like gut functions were identified in some of the ancient samples (116), the data suggests alteration of the community due to continued metabolism of the organic matter by gut microorganisms and/or soil contamination prior to preservation (11). Rapid preservation is needed to prevent compost like bacterial signatures, and it is difficult to identify coprolites that have been rapidly desiccated. Alongside concerns that coprolites do not accurately represent the faecal bacterial community, it has

also been noted that faeces may not be representative of the gut microbiota. Faecal matter contains primarily the planktonic bacteria, that do not adhere to the gut wall, and are functionally distinct (117). Consequently, coprolites have not yet been used extensively for studying ancient microbiota.

Dental calculus is an alternative source of ancient microbiota. Dental calculus is formed from dental plaque (the microbial biofilm that grows on the surface of the teeth). This biofilm is then calcified, primarily at night when the pH level of the oral cavity is reduced. Hydroxy-apatite, whitlockite, octacalcium phosphate, and brushite precipitate from the saliva among the microbial cells, hardening the plaque biofilm into a solid structure (calculus) (118). This process occurs repeatedly and layers of calculus are laid down trapping, and preserving the microbiota (119). The advantage of dental calculus as an archaeological record is that the preserved microbes were not metabolically active outside of a living oral environment. Dental calculus is also very dense and resistant to degradation and microbial entry, making it a well-preserved sample with limited contamination (11). In contrast to coprolites, dental calculus is also found associated with human remains, and additional archaeological, morphometric, paleopathological, human DNA, or isotope analysis can be conducted using this human material.

Preserved microorganisms within the calculus were not identified until the 1980s. At that time, scanning electron microscopy (SEM) was used to identify dietary debris and phytolithis (plant microfossils) trapped within dental calculus, providing information on dietary patterns (119–121). However, SEM studies also revealed the preserved bacterial cells within dental calculus. The bacterial cell morphology visible was not useful for fine scale taxonomic identification (119). Later, preserved DNA was identified with the use of transmission electron microscopy (122). While the transmission electron microscopy technique could not provide any sequence information, or indicate if the preservation was suitable for genetic analysis, this was positive evidence that DNA was preserved within ancient dental calculus and that ancient DNA analysis was a potential tool to study ancient microbiota.

The first ancient DNA study used PCR amplification of key marker regions for five known oral bacteria (123). This proved that DNA was preserved in sufficient quality to provide positive identification of oral bacteria (123). Subsequently, whole community analysis was completed using HTS sequencing of 16S rRNA gene sequences to explore European oral microbiota from pre-agricultural revolution through to the modern day (124). The findings from this study indicated that microbiota had undergone significant shifts in diversity during the agricultural and industrial revolutions (~8,000 BCE and 1750 CE, respectively). This finding was associated with the dramatic increase in dietary carbohydrates during these two revolutions (124). However, the study lacked the microbial species and metadata resolution to determine the specific drivers of this change and the impacts of other cultural alterations, such as other dietary changes (*e.g.* the introduction of dairy) urbanization, increasing trade, and disease events. Following this study, it became apparent that using the 16S rRNA gene as a diagnostic region is not appropriate for examining ancient microbiota, as this method can introduce taphonomic (preservation) bias into the data (125). To overcome this bias and gain species level resolution and functional information, following studies have used shotgun sequencing to reconstruct ancient microbiota from dental calculus.

The first study to employ shotgun sequencing was completed on two medieval German individuals (~45 quality filtered reads per sample) (126). This study explored microbiota composition and demonstrated the presence of characteristic oral pathogens, including *Streptococcus mutans* (causative of dental caries (85)) and the Red Three complex (*Tannerella forsythia, Porphyromonas gingivalis* and *Treponema denticola*, which trigger severe periodontal disease (88)). Functional analysis revealed antibiotic resistance and virulence factors. Importantly, initial information on lifestyle was revealed due to the recovery of very low levels of eukaryotic DNA. This eukaryotic DNA allowed direct inference of organisms that were likely to have been part of the diet of the two individuals (126). Similarly, a recent study has also utilized shotgun sequencing to explore microbiota and the dietary elements

in ancient hominins. In this study, dental calculus from five Neandertals, a wild chimpanzee, and a modern human were analysed (~17-50 million reads per sample) (36). Combining the Neandertals, chimpanzee, and modern human data with the two medieval samples from the initial shotgun-based study revealed a correlation between the dietary items identified and a split in microbiota diversity that suggested the microbiota was defined by the presence of meat in the diet. Recovery of ancient microbiota from ~48,000 year old Neandertals demonstrates the potential for studying ancient microbiota in detail through time. However, minimal numbers of well preserved samples limited the resolution of the study to identify factors that altered modern human microbiota through time beyond the broad dietary association (36).

Great Britain allows fine scale analysis of microbiota

High-resolution studies of single populations through time are required to gain insight into the specific factors that drive microbiota variation. Initially, such studies will need to investigate populations with existing detailed archaeological, anthropological, and historical information to allow changes in the microbiota to be associated with the events of the time. Such studies will be able to explore the factors defining and altering the microbiota on an individual and population level. In addition, studying populations with large amounts of metadata may also provide the understanding of microbiota response to culture and environment to allow inference of the cultural and environmental background for individuals lacking contextual information.

Great Britain is an excellent region to conduct high-resolution studies of historical microbiota. Great Britain is the major geographical area of the United Kingdom, and comprises England, Scotland and Wales. This region has detailed, extensive historical and archaeological records (127). In addition, there are large, curated collections of human remains associated with these data (*e.g.* (128)). This allows detailed associations of dental

calculus samples (and the resulting microbiota data) with an individual's life history and the broader cultural and environmental landscape. A detailed study of microbiota from Britain offers a new element to the history of the country, providing the first insights into previously unappreciated aspects of human health and disease. In a broader context, the history of Great Britain is also a key part of the history of Western civilization. Thus, ancient microbiota studies in the UK offer direct insight into the development of the Western microbiota and provide baseline data for detailed studies of European and world history, including defining the microbiota introduced by British colonialists to native groups around the world.

In this thesis, I utilized 258 dental calculus samples for microbiota analysis, which range from the Bronze Age to the Early Victorian period (~6,000 BCE – 1853 CE). Due to sample availability, I concentrated my efforts on the post-Roman period (410 CE onwards) with particular focus on the post-Norman Invasion period (1066 CE onwards). Using this large data set, I explore how the microbiota were altered through time, examine the structure of microbiota communities within a single population, and interrogate the drivers of these patterns. However, British history is complex and contains many events that may have impacted the microbiota. Therefore, I targeted specific events and cultural or environmental shifts to determine their impact on the microbiota of historical Britain. I also compared this information to previously published modern studies to investigate the importance of these changes to the development of the modern, industrialized oral microbiota. The following is a brief summary of British history, with attention to how the microbiota may have been affected in each period.

Pre-Roman Britain (Colonization – 43 CE)

Britain has been continually inhabited since ~10,000 BCE (129), when the glaciers of the Younger Dryas retreated. Agriculture arrived in ~4,000 BCE (130) and is expected to have drastically altered the microbiota due to the increase in dietary carbohydrates, as previously observed (124).

However, permanent settlements were not utilized until ~2,200 BCE. This sedentary lifestyle would have impacted the microbiota, through impacts of walled homes increasing human-to-human transfer of microbes (65). In addition, domestication of animals ~3,500 BCE (131) would have increased transfer of microbes from animals. Early settlements also lacked waste removal, increasing transfer of infectious diseases. A change in endemic disease loads would likely have impacted the microbiota. Notably, throughout this period, trade occurred with continental Europe, which, while minimal compared to later periods, may have impacted the resources and their associated microbiota available to trading communities.

Roman Britain (43 CE – 410 CE)

Roman rule resulted in major changes to the country's infrastructure by building roads and urban centres (*e.g.* York, London, Lincoln, and Chester). Towns increasingly became a focal point for society, and the living conditions were improved with imported foods, which likely impacted the microbiota. However, the lifestyle in these urban centres is unlikely to have been representative of the population, who continued working small farms (132,133). This divide between the rich and poor, urban and rural may be evident in the microbiota, as compositional variation may exist from the resources available to different classes. London became the capital during this period and a major hub for trade and movement of people. In contrast, unconquered regions, such as Scotland, developed local kingdoms based on farming and trade. Consequently, geography and its links to resource availability may be a key defining factor for microbiota structure. The Romans introduced many foods (including lentils) and cooking practices (such as use of eggs in cooking) to Britain (131). However, the degree to which the culinary imports reached the majority of the population will impact the degree of change within the microbiota. Differences in diet between population sub-groups may drive microbiota differences that can be used to identify these groups. Ultimately, the Roman Empire retreated from Britain in ~410 CE. This was accompanied with a breakdown of state, and much of the trade

and use of towns were reduced until the 1100's (132). Microbiota around this time period may offer insight into how the Roman retreat impacted the lifestyle and health of the population and the change in population sub-structure, such as the loss of a small number of social groups (*e.g.* individuals of military or merchant status).

Anglo-Saxon Britain (410 CE – 1066 CE)

The first invasion of this period was by the Angles, Saxons, and Jutes, who arrived from Eastern Europe and settled throughout England. These different groups settled in separate regions. The Angles took central and northern territories (Northumberland and Mercia), while the Saxons were in the south (Wessex). The Jutes settled in modern day Kent (133). The arrival of these different groups may be indicated by unique microbiota they brought to Britain. In addition, change in composition over time in these microbiota may reveal replacement of, co-existence with, or integration between the local peoples. During this period, Christianity was reintroduced to Britain, after it was lost following the Roman retreat (132). A shift in religion is often accompanied by a change in lifestyle due to religious rules. For example, repetitive and non-diverse food use, particularly in monastic individuals, is likely to have altered the microbiota. Marine resources became a larger part of the diet as fishing technologies advanced, and by the end of this period, herring were a major economy and dietary component (131). Lastly, trade was less important than in Roman Britain, but slowly increased during the later 600s CE. Britain exported wool in the 800's (132). The large sheep industry and culture that worked closely with animals to support this trade may have resulted in microbial transfer from domesticated animals.

The Viking invasion occurred during the 800-900s, as Scandinavian people moved into the north of Britain. Despite the influx of Scandinavians, the cultural changes appear to have been minimal, and Anglo-Saxon and Viking settlements are difficult to distinguish archaeologically (133). Coinciding with this period is the re-emergence of urban settlements, suggesting the rise of dedicated craft workers. Profession has been shown

to impact microbiota (70), and this increase in trade specialization may have impacted the microbiota by introducing novel microbial species into the body or exposing the body to chemicals or other factors that alter bacterial community structure. While relations between Anglo-Saxons and Vikings were more peaceful than commonly believed, there were also disputes, raids, and wars between the two. While Britain was "united" in the 900s CE, Viking invasion continued until 1066 CE. The influx of soldiers during invasions may have resulted in a complex distribution of microbiota across the country, as integration varied across the landscape and through time.

Norman Britain and the Middle Ages (1066 CE – 1547 CE)

As with previous invasions, the Norman invasion replaced England's elite but likely left the majority of the population in place and unchanged (133). However, social differentiation between individuals based on their Norman or Saxon heritage was no longer in use by 1200 CE (132). Towns with local markets became a major part of society again (133), and this local and broader trade would have made dietary items, such as herring and other marine fish, available inland. The expansion of herring included smoking and salting to preserve the fish (131). These practices would have modified the nutrient availability in the food, likely altering microbiota composition. In 1347 CE, the Black Death reached England, and was a continual presence until the mid-1600s CE. This major disease epidemic wiped out ~30% of the population and up to 50% of the inhabitants of London in the first wave (15). The loss of large sections of the population, disrupting population structure, is likely to have altered many interpersonal connections and trade patterns, resulting in a microbiota shift. Of particular interest is if host microbiota offered any protection or immunity to the plague bacteria (*Yersinia pestis*).

Reformation (1547 CE – 1750 CE)

In contrast to previous periods, this is not marked by invasion, but by a cultural shift. Religion was a key factor in society structure and the reformation in 1547 CE (the breakaway of the Church of England from the

Roman Catholic Church) was disruptive to social cohesion (132). This alteration in social interactions throughout the population may have changed the structures of microbiota within society. This transition was also associated with Saturday become a day to eat fish (131), a dietary pattern alteration that potentially trigged microbiota shifts. This period was also demarked by periodic food shortages during late 1500's and 1600's. Poor nutrient intake during this time may have put stress on certain species within the microbiota. This period also saw improvements in hygiene. In London, movements were made to prevent sewage flowing in the streets (132). This sewer system would have improved health and altered environmental stress on the microbiota. The final instance of the plague occurred in 1666 CE, and again had significant impacts on population size. The absence of *Y. pestis* following this may have reduced a selection pressure on specific microbiota that can compete with the pathogen. In addition, this year saw the Great Fire of London, which damaged large areas of the city. The resulting population movements within the city, the environment of rebuilding, and the temporary infrastructure (*i.e.* camping in fields following the destruction of the city) may have each placed unique pressures on the microbiota. This period also saw the introduction of new food products such as coffee, tea, and chocolate. These products have grown in popularity and as major modern dietary elements are likely to have altered the microbiota.

Industrial Revolution (1750 CE – 1900 CE)

The industrial revolution is currently seen as a period of major change in the oral microbiota that has been associated with refined carbohydrates and sugar (124). However, multiple other dietary items increased during this period. The development of pasteurisation allowed for the large-scale production of milk, and importation introduced additional new food items. Large-scale production of preserved foods required the sterilisation of food for canning, which also dramatically altered the microbial content of the diet and the possibility for the introduction of novel microorganisms. Despite these changes, access to these resources was limited, and did not become

widely available until the end of this period. There was also a dramatic intensification of urban environments with little open, green space (132). The intense use of fossil fuels resulted in severe air pollution. Heavy metals, which would be present in the atmosphere as a result of coal burning have been shown to impact the gut microbiota (134), and would be expected to similarly impact oral microbiota.

Communication benefits science and society

Science communication informs and engages non-specialists on the state and relevance of research (135). There have been many calls for academics to increase their engagement with science communication, both via interaction with the press (136) and by actively participating in communication activities such as science outreach events (137) and social media (138). The literature surrounding the need for science communication focuses on the importance of science communication to help create a science-literate society (136,139). However, science communication is not only critical to society but directly beneficial to academics (136,140–142). In addition, academia is increasingly bringing together disciplines to address single questions and goals (143), which requires communication between academics of different disciplines (144). Differing fields of research have distinct methods, formats, and standards of work, and the incompatibility of these elements can limit collaborations between disciplines (145). However, science communication, with its focus on communicating with the layperson, may offer a mechanism to break down disciplinary barriers by placing understandable descriptions of research methods, ideas, and results into a forum accessible to the public, which includes academics in different disciplines.

Science communication has direct benefits to academics. Yearly citation counts of research articles also disseminated in a traditional media publication, (which has a public audience) received higher yearly citation counts over ten years compared to research articles without additional

coverage (140). Notably, this shift in citations was not a result of journalists and scientists independently recognizing the research but was a direct consequence of public exposure. This direct link between public exposure and citation counts could be demonstrated because of a printing strike. Research articles selected by journalists for publication but never published did not receive a higher citation count. This demonstrated that it was specifically the dissemination of the media article that resulted in the increased citations (140). This indicates the importance of public exposure and the press in disseminating academic research to peers within the field. Similar correlations have been made in analyses of new media (*i.e.* online media and social media). The number of times a scientific article was mentioned on the social networking platform Twitter was correlated with increased citation counts (141). As the importance of science communication is increasingly recognized by research organizations (146), the coverage of scientific literature in new media is a new metric by which academics can be ranked. New "altmetrics" (alternative metrics) track mentions on social media, blogs, online news outlets, Wikipedia, etc., has grown in recent times (142). Altmetrics provide a more rapid assessment of research impact and, compared to metrics that track longer-term impact *(e.g*. citation counts) are likely to be incorporated in many assessments of academic performance in the coming years.

The importance of science communication is not limited to direct research returns. In a climate where science literacy is a concern, science communication offers an opportunity to make scientific information publicly available and accessible. Scientific theories can be complex and counterintuitive (139). Consequently, it is critical that science is made publicly available in a format accessible to non-specialists. This would include explanations of the concepts, logic of method, and (where possible) links to broader social aspects and issues the research is investigating. To achieve this, science communication needs to place factual info into a personal, emotional, and engaging context (147,148). However, this is a difficult task, particularly for scientists trained to write for a specialized audience. Writing

for a non-specialized public requires adjustment and training in how to prepare scientific information for the public forum (149).

Multiple papers have been publish outlining social media and science communication to individual researchers (150,151), as have discussions of large scale science communication projects that require dedicated staff (137). While critical, these discussions do not take advantage of the existing support structures of a research group. In this thesis, I explore methods for a research group to use science communication to engage the public with their science and, in doing so, to communicate with their peers. Interdisciplinary research is limited by specialist language that is not coherent outside of individual academic disciplines (152). Consequently, the skills and time devoted to science communication are also applicable to communicating between disciplines. A public forum is likely accessed by many members of multiple academic disciplines and should be used by researchers to demonstrate ideas, methods, and results and to observe these elements in other disciplines. Active production and consumption of science communication material offers an opportunity to explore how interdisciplinary collaboration could expand upon the current methods and scope of research questions.

Thesis overview

This thesis compiles five manuscripts to address the issues raised in this introductory chapter. Specifically, an exploration of microbiota across a single population through time is conducted to identify the cultural and environmental factors that define and alter the oral microbial community. Several factors predicted to have impacted the oral microbiota are assessed, including diet, disease, living environment, human movement, religion, and war. In an emerging research field, this research investigates best practice to retrieve accurate representations of historical microbiota, applies high-resolution HTS studies of microbiota from a single geographical population,

and explores methods by which researchers can effectively engage public and peers with research ideas and results.

Chapter I

Ancient bacteria from dental calculus as an archaeological and anthropological tool

In this chapter, I review and explore the potential of dental calculus to address archaeological and anthropological questions. The current studies of modern microbiota link microbial community structure to a range of factors experienced in today's world. However, these studies indicate how studies of dental calculus could infer the potential impacts of similar factors in ancient individuals and populations. This chapter also demonstrates how ancient dental calculus research can complement existing archaeological analyses and offer considerable insight into the lifestyles and experiences of past individuals with little or no other archaeological context.

Chapter II

Effectiveness of decontamination protocols when analysing ancient DNA preserved in dental calculus.

Contamination is a serious risk for ancient DNA studies, and particularly for paleomicrobiological studies. Therefore, successful removal of contaminating DNA during laboratory processing maximizes sequencing efficiency and improves data quality prior to bioinformatics assessment. In this chapter, I assess four different protocols for the removal of environmental contamination from the surface of dental calculus prior to DNA extraction. Ultimately, I recommend a UV irradiation and bleach submersion as the optimal decontamination method for ancient DNA samples.

Chapter III

Diet driven differentiation of oral microbiota in ancient Britain

How microbiota were structured within historical populations and how this relates to the modern microbiota remains unknown. To provide the first population level insight into historical microbiota, Chapter III applies the concepts and protocols from Chapters I and II to investigate the microbiota in Great Britain from the Pre-Roman (pre-43 CE) to the Medieval (~1700 CE) period. Two distinct groups were identified based on microbiota composition. Functional analyses of these groups indicate that diet was the driving factor and indicative of socio-economic rank. This study is the first to examine function in historical microbiota and provides a broad insight into the diet of historical individuals. Further, comparison with previously published oral microbiota data indicated that modern, industrialized populations have microbiota that are derived from one of these groups.

Chapter IV

Biological and cultural drivers of oral microbiota in Medieval and Post-Medieval London

In a collaboration I led with the Museum of London, Chapter IV recovers microbiota from 128 dental calculus samples to build upon the observations within Chapter III and explore associations beyond diet, including disease, religion, status, location, key historical events, and change through time. In addition, this large sample size allowed the first exploration of potential technical biases, such as oral geography, in ancient samples. I identified that there is a significant association between oral geography and microbiota composition. Once I controlled for oral geography, this study provided the first direct association of microbiota structure with past oral and systemic disease.

Chapter V

Incorporating science communication into standard research group practice is described in Chapter V. This chapter describes a framework that I co-developed at the Australian Centre for Ancient DNA to provide the centre with an online science communication network, focused on promoting both the centre and the individual researchers, while simultaneously minimizing the time spent per individual. From this framework, the metrics from a year's worth of communication efforts are analysed and discussed. These results indicate the success of engaging both the public and the centre's peers with research, including direct benefits to research.

References

1. Komaroff AL. Increasing Evidence That the Human Microbiome Plays a Major Role in Health. N Engl Med J. 2013 Dec 31;

2. Sender R, Fuchs S, Milo R. Are We Really Vastly Outnumbered? Revisiting the Ratio of Bacterial to Host Cells in Humans. Cell. 2016 Jan 28;164(3):337–40.

3. Marchesi JR, Ravel J. The vocabulary of microbiome research: a proposal. Microbiome. 2015 Jul 30;3(1):31.

4. Jandhyala SM, Talukdar R, Subramanyam C, Vuyyuru H, Sasikala M, Reddy DN. Role of the normal gut microbiota. World J Gastroenterol WJG. 2015 Aug 7;21(29):8787–803.

5. Redford KH, Segre JA, Salafsky N, del Rio CM, McAloose D. Conservation and the Microbiome. Conserv Biol J Soc Conserv Biol. 2012 Apr;26(2):195–7.

6. Turnbaugh PJ, Ley RE, Hamady M, Fraser-Liggett CM, Knight R, Gordon JI. The human microbiome project. Nature. 2007 Oct 18;449(7164):804–10.

7. Peterson J, Garges S, Giovanni M, McInnes P, Wang L, Schloss JA, et al. The NIH Human Microbiome Project. Genome Res. 2009 Dec;19(12):2317–23.

8. Xu X, Wang Z, Zhang X. The human microbiota associated with overall health. Crit Rev Biotechnol. 2015 Jan 2;35(1):129–40.

9. Weyrich LS. Evolution of the Human Microbiome and Impacts on Human Health, Infectious Disease, and Hominid Evolution. In: Gontier N, editor. Reticulate Evolution [Internet]. Springer International Publishing; 2015 [cited 2015 Oct 27]. p. 231–53. (Interdisciplinary Evolution Research). Available from: http://link.springer.com/chapter/10.1007/978-3-319-16345-1_9

10. Warinner C, Speller C, Collins MJ. A new era in palaeomicrobiology: prospects for ancient dental calculus as a long-term record of the human oral microbiome. Phil Trans R Soc B. 2015 Jan 19;370(1660):20130376.

11. Weyrich LS, Dobney K, Cooper A. Ancient DNA analysis of dental calculus. J Hum Evol. 2015 Feb;79:119–24.

12. Orlando L, Cooper A. Using Ancient DNA to Understand Evolutionary and Ecological Processes. Annu Rev Ecol Evol Syst. 2014;45(1):573–98.

13. Orlando L, Ginolhac A, Zhang G, Froese D, Albrechtsen A, Stiller M, et al. Recalibrating Equus evolution using the genome sequence of an early Middle Pleistocene horse. Nature. 2013 Jul 4;499(7456):74–8.

14. Haak W, Lazaridis I, Patterson N, Rohland N, Mallick S, Llamas B, et al. Massive migration from the steppe was a source for Indo-European languages in Europe. Nature [Internet]. 2015 Mar 2 [cited 2015 Mar 5];advance online publication. Available from: http://www.nature.com/nature/journal/vaop/ncurrent/full/nature14317.html

15. Bos KI, Schuenemann VJ, Golding GB, Burbano HA, Waglechner N, Coombes BK, et al. A draft genome of Yersinia pestis from victims of the Black Death. Nature. 2011 Oct 27;478(7370):506–10.

16. Higuchi R, Bowman B, Freiberger M, Ryder OA, Wilson AC. DNA sequences from the quagga, an extinct member of the horse family. Nature. 1984 Nov 15;312(5991):282–4.

17. Haile J, Froese DG, MacPhee RDE, Roberts RG, Arnold LJ, Reyes AV, et al. Ancient DNA reveals late survival of mammoth and horse in interior Alaska. Proc Natl Acad Sci. 2009 Dec 29;106(52):22352–7.

18. Thalmann O, Shapiro B, Cui P, Schuenemann VJ, Sawyer SK, Greenfield DL, et al. Complete Mitochondrial Genomes of Ancient Canids Suggest a European Origin of Domestic Dogs. Science. 2013 Nov 15;342(6160):871–4.

19. Shapiro B, Drummond AJ, Rambaut A, Wilson MC, Matheus PE, Sher AV, et al. Rise and Fall of the Beringian Steppe Bison. Science. 2004 Nov 26;306(5701):1561–5.

20. Thomson VA, Lebrasseur O, Austin JJ, Hunt TL, Burney DA, Denham T, et al. Using ancient DNA to study the origins and dispersal of ancestral Polynesian chickens across the Pacific. Proc Natl Acad Sci. 2014 Apr 1;111(13):4826–31.

21. Larson G, Cucchi T, Fujita M, Matisoo-Smith E, Robins J, Anderson A, et al. Phylogeny and ancient DNA of Sus provides insights into neolithic expansion in Island Southeast Asia and Oceania. Proc Natl Acad Sci. 2007 Mar 20;104(12):4834–9.

22. Dabney J, Meyer M, Paabo S. Ancient DNA damage. Cold Spring Harb Perspect Biol [Internet]. 2013 Jul;5(7). Available from: http://www.ncbi.nlm.nih.gov/pubmed/23729639

23. Carpenter ML, Buenrostro JD, Valdiosera C, Schroeder H, Allentoft ME, Sikora M, et al. Pulling out the 1%: Whole-Genome Capture for the

Targeted Enrichment of Ancient DNA Sequencing Libraries. Am J Hum Genet. 2013 Nov 7;93(5):852–64.

24. Allentoft ME, Collins M, Harker D, Haile J, Oskam CL, Hale ML, et al. The half-life of DNA in bone: measuring decay kinetics in 158 dated fossils. Proc R Soc B. 2012 Dec 7;279(1748):4724–33.

25. Ginolhac A, Rasmussen M, Gilbert MTP, Willerslev E, Orlando L. mapDamage: testing for damage patterns in ancient DNA sequences. Bioinformatics. 2011 Aug 1;27(15):2153–5.

26. Cooper A, Poinar HN. Ancient DNA: Do It Right or Not at All. Science. 2000 Aug 18;289(5482):1139–1139.

27. Bartlett JMS, Stirling D. A short history of the polymerase chain reaction. Methods Mol Biol Clifton NJ. 2003;226:3–6.

28. Reuter JA, Spacek DV, Snyder MP. High-Throughput Sequencing Technologies. Mol Cell. 2015 May 21;58(4):586–97.

29. Sanger F, Nicklen S, Coulson AR. DNA sequencing with chain-terminating inhibitors. Proc Natl Acad Sci U S A. 1977 Dec;74(12):5463–7.

30. Caporaso JG, Lauber CL, Walters WA, Berg-Lyons D, Huntley J, Fierer N, et al. Ultra-high-throughput microbial community analysis on the Illumina HiSeq and MiSeq platforms. ISME J. 2012 Aug;6(8):1621–4.

31. Taberlet P, Coissac E, Pompanon F, Brochmann C, Willerslev E. Towards next-generation biodiversity assessment using DNA metabarcoding. Mol Ecol. 2012 Apr 1;21(8):2045–50.

32. Meyer M, Kircher M. Illumina Sequencing Library Preparation for Highly Multiplexed Target Capture and Sequencing. Cold Spring Harb Protoc. 2010 Jun 1;2010(6):pdb.prot5448.

33. DeSantis TZ, Hugenholtz P, Larsen N, Rojas M, Brodie EL, Keller K, et al. Greengenes, a Chimera-Checked 16S rRNA Gene Database and Workbench Compatible with ARB. Appl Environ Microbiol. 2006 Jul 1;72(7):5069–72.

34. O'Leary NA, Wright MW, Brister JR, Ciufo S, Haddad D, McVeigh R, et al. Reference sequence (RefSeq) database at NCBI: current status, taxonomic expansion, and functional annotation. Nucleic Acids Res. 2016 Jan 4;44(Database issue):D733–45.

35. Herbig A, Maixner F, Bos KI, Zink A, Krause J, Huson DH. MALT: Fast alignment and analysis of metagenomic DNA sequence data applied to the Tyrolean Iceman. bioRxiv. 2016 Apr 27;50559.

36. Weyrich LS, Duchane S, Soubrier J, Arriola L, Llamas B, Breen J, et al. Reconstructing Neandertal behavior, diet, and disease using ancient DNA from dental calculus. Nature. 2016;

37. Mouttham N, Klunk J, Kuch M, Fourney R, Poinar H. Surveying the repair of ancient DNA from bones via high-throughput sequencing. BioTechniques. 2015;59(1):19–25.

38. Pääbo S, Poinar H, Serre D, Jaenicke-Després V, Hebler J, Rohland N, et al. Genetic Analyses from Ancient DNA. Annu Rev Genet. 2004;38(1):645–79.

39. Weiss S, Amir A, Hyde ER, Metcalf JL, Song SJ, Knight R. Tracking down the sources of experimental contamination in microbiome studies. Genome Biol. 2014 Dec 17;15(12):1–3.

40. Rawlence NJ, Lowe DJ, Wood JR, Young JM, Churchman GJ, Huang Y-T, et al. Using palaeoenvironmental DNA to reconstruct past environments: progress and prospects. J Quat Sci. 2014 Oct 1;29(7):610–26.

41. Kemp BM, Smith DG. Use of bleach to eliminate contaminating DNA from the surface of bones and teeth. Forensic Sci Int. 2005 Nov 10;154(1):53–61.

42. Cone RW, Fairfax MR. Protocol for ultraviolet irradiation of surfaces to reduce PCR contamination. Genome Res. 1993 Dec 1;3(3):S15–7.

43. Salter SJ, Cox MJ, Turek EM, Calus ST, Cookson WO, Moffatt MF, et al. Reagent and laboratory contamination can critically impact sequence-based microbiome analyses. BMC Biol. 2014 Nov 12;12(1):87.

44. Fulton T. Setting Up an Ancient DNA Laboratory. In: Shapiro B, Hofreiter M, editors. Ancient DNA [Internet]. Humana Press; 2012 [cited 2016 Sep 2]. p. 1–11. (Methods in Molecular Biology). Available from: http://dx.doi.org/10.1007/978-1-61779-516-9_1

45. Knapp M, Clarke AC, Horsburgh KA, Matisoo-Smith EA. Setting the stage – Building and working in an ancient DNA laboratory. Ann Anat - Anat Anz. 2012 Jan 20;194(1):3–6.

46. O'Hara AM, Shanahan F. The gut flora as a forgotten organ. EMBO Rep. 2006 Jul;7(7):688–93.

47. Qin J, Li R, Raes J, Arumugam M, Burgdorf KS, Manichanh C, et al. A human gut microbial gene catalogue established by metagenomic sequencing. Nature. 2010 Mar 4;464(7285):59–65.

48. Grice EA, Kong HH, Conlan S, Deming CB, Davis J, Young AC, et al. Topographical and temporal diversity of the human skin microbiome. Science. 2009 May 29;324(5931):1190–2.

49. Consortium THMP. Structure, function and diversity of the healthy human microbiome. Nature. 2012 Jun 14;486(7402):207–14.

50. Wade WG. Has the use of molecular methods for the characterization of the human oral microbiome changed our understanding of the role of bacteria in the pathogenesis of periodontal disease? J Clin Periodontol. 2011 Mar 1;38:7–16.

51. Koenig JE, Spor A, Scalfone N, Fricker AD, Stombaugh J, Knight R, et al. Succession of microbial consortia in the developing infant gut microbiome. PNAS. 2011;108(1):8.

52. LeBlanc JG, Milani C, de Giori GS, Sesma F, van Sinderen D, Ventura M. Bacteria as vitamin suppliers to their host: a gut microbiota perspective. Curr Opin Biotechnol. 2013 Apr;24(2):160–8.

53. Turnbaugh PJ, Backhed F, Fulton L, Gordon JI. Diet-induced obesity is linked to marked but reversible alterations in the mouse distal gut microbiome. Cell Host Microbe. 2008 Apr 17;3(4):213–23.

54. Neuman H, Debelius JW, Knight R, Koren O. Microbial endocrinology: the interplay between the microbiota and the endocrine system. FEMS Microbiol Rev. 2015 Jul 1;39(4):509–21.

55. Noverr MC, Huffnagle GB. Does the microbiota regulate immune responses outside the gut? Trends Microbiol. 2004 Dec;12(12):562–8.

56. Hsiao EY, McBride SW, Hsien S, Sharon G, Hyde ER, McCue T, et al. Microbiota Modulate Behavioral and Physiological Abnormalities Associated with Neurodevelopmental Disorders. Cell. 2013 Dec 19;155(7):1451–63.

57. Trivedi B. Microbiome: The surface brigade. Nature. 2012;492(7429):2.

58. Naik S, Bouladoux N, Wilhelm C, Molloy MJ, Salcedo R, Kastenmuller W, et al. Compartmentalized Control of Skin Immunity by Resident Commensals. Science. 2012 Aug 31;337(6098):1115–9.

59. Li X, Wang J, Joiner A, Chang J. The remineralisation of enamel: a review of the literature. J Dent. 2014 Jun;42, Supplement 1:S12–20.

60. Dickson RP, Huffnagle GB. The Lung Microbiome: New Principles for Respiratory Bacteriology in Health and Disease. PLOS Pathog. 2015 Jul 9;11(7):e1004923.

61. Nunn KL, Forney LJ. Unraveling the Dynamics of the Human Vaginal Microbiome. Yale J Biol Med. 2016 Sep 30;89(3):331–7.

62. David LA, Maurice CF, Carmody RN, Gootenberg DB, Button JE, Wolfe BE, et al. Diet rapidly and reproducibly alters the human gut microbiome. Nature. 2014 Jan 23;505(7484):559–63.

63. Zarco MF, Vess TJ, Ginsburg GS. The oral microbiome in health and disease and the potential impact on personalized dental medicine. Oral Diseses. 2012 Mar;18(2):109–20.

64. Azad MB, Konya T, Maughan H, Guttman DS, Field CJ, Sears MR, et al. Infant gut microbiota and the hygiene hypothesis of allergic disease: impact of household pets and siblings on microbiota composition and diversity. Allergy Asthma Clin Immunol. 2013;9(15):9.

65. Ruiz-Calderon JF, Cavallin H, Song SJ, Novoselac A, Pericchi LR, Hernandez JN, et al. Walls talk: Microbial biogeography of homes spanning urbanization. Sci Adv. 2016 Feb 1;2(2):e1501061.

66. Grice EA, Segre JA. The skin microbiome. Nat Rev Microbiol. 2011 Apr;9(4):244–53.

67. Riiser A. The human microbiome, asthma, and allergy. Allergy Asthma Clin Immunol Off J Can Soc Allergy Clin Immunol [Internet]. 2015 Dec 10 [cited 2016 Oct 14];11. Available from: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC4674907/

68. Ojetti V, Gigante G, Ainora ME, Fiore F, Barbaro F, Gasbarrini A. Microflora imbalance and gastrointestinal diseases. Dig Liver Dis Suppl. 2009 Jul 1;3(2):35–9.

69. Iwase T, Uehara Y, Shinji H, Tajima A, Seo H, Takada K, et al. Staphylococcus epidermidis Esp inhibits Staphylococcus aureus biofilm formation and nasal colonization. Nature. 2010 May 20;465(7296):346–9.

70. Chen YE, Tsao H. The skin microbiome: Current perspectives and future challenges. J Am Acad Dermatol. 2013 Jul;69(1):143–155.e3.

71. Shu M, Wang Y, Yu J, Kuo S, Coda A, Jiang Y, et al. Fermentation of Propionibacterium acnes, a Commensal Bacterium in the Human Skin Microbiome, as Skin Probiotics against Methicillin-Resistant Staphylococcus aureus. PLoS One. 2013;8(2).

72. Engstrand L, Lindberg M. Helicobacter pylori and the gastric microbiota. Best Pr Res Clin Gastroenterol. 2013 Feb;27(1):39–45.

73. Scher JU, Sczesnak A, Longman RS, Segata N, Ubeda C, Bielski C, et al. Expansion of intestinal Prevotella copri correlates with enhanced susceptibility to arthritis. eLife. 2013 Nov 5;2.

74. Scher JU, Abramson SB. The microbiome and rheumatoid arthritis. Nat Rev Rheumatol. 2011 Aug 23;7(10):569–78.

75. Zitvogel L, Galluzzi L, Viaud S, Vétizou M, Daillère R, Merad M, et al. Cancer and the gut microbiota: An unexpected link. Sci Transl Med. 2015 Jan 21;7(271):271ps1-271ps1.

76. Flowers SA, Ellingrod VL. The Microbiome in Mental Health: Potential Contribution of Gut Microbiota in Disease and Pharmacotherapy Management. Pharmacother J Hum Pharmacol Drug Ther. 2015 Oct 1;35(10):910–6.

77. Benton D, Williams C, Brown A. Impact of consuming a milk drink containing a probiotic on mood and cognition. Eur J Clin Nutr. 2006 Dec 6;61(3):355–61.

78. Messaoudi M, Lalonde R, Violle N, Javelot H, Desor D, Nejdi A, et al. Assessment of psychotropic-like properties of a probiotic formulation (Lactobacillus helveticus R0052 and Bifidobacterium longum R0175) in rats and human subjects. Br J Nutr. 2011 Mar;105(5):755–64.

79. Tillisch K, Labus J, Kilpatrick L, Jiang Z, Stains J, Ebrat B, et al. Consumption of Fermented Milk Product With Probiotic Modulates Brain Activity. Gastroenterology. 2013 Jun;144(7):1394–1401.e4.

80. Palmer Jr. RJ. Composition and development of oral bacterial communities. Periodontol 2000. 2014;64:20.

81. Avila M, Ojcius DM, Yilmaz O. The Oral Microbiota: Living with a Permanent Guest. DNA Cell Biol. 2009 Nov 8;28(8):7.

82. Dewhirst FE, Chen T, Izard J, Paster BJ, Tanner AC, Yu WH, et al. The human oral microbiome. J Bacteriol. 2010 Oct;192(19):5002–17.

83. Marsh PD. Are dental diseases examples of ecological catastrophes? Microbiology. 2003;149(2):279–94.

84. Wessel SW, Chen Y, Maitra A, van den Heuvel ER, Slomp AM, Busscher HJ, et al. Adhesion forces and composition of planktonic and adhering oral microbiomes. J Dent Res. 2014 Jan;93(1):84–8.

85. Ajdić D, McShan WM, McLaughlin RE, Savić G, Chang J, Carson MB, et al. Genome sequence of Streptococcus mutans UA159, a cariogenic dental pathogen. Proc Natl Acad Sci. 2002 Oct 29;99(22):14434–9.

86. Belda-Ferre P, Alcaraz LD, Cabrera-Rubio R, Romero H, Simón-Soro A, Pignatelli M, et al. The oral metagenome in health and disease. ISME J. 2012 Jan;6(1):46–56.

87. Petersen PE, Bourgeois D, Ogawa H, Estupinan-Day S, Ndiaye C. The global burden of oral diseases and risks to oral health. Bull World Health Organ. 2005 Sep;83(9):661–9.

88. Rocas IN, Siqueira JF Jr, Santos KR, Coelho AM. "Red complex" (Bacteroides forsythus, Porphyromonas gingivalis, and Treponema denticola) in endodontic infections: a molecular approach. Oral Surg Oral Med Oral Pathol Oral Radiol Endod. 2001 Apr;91(4):468–71.

89. Rautemaa R, Lauhio A, Cullinan MP, Seymour GJ. Oral infections and systemic disease—an emerging problem in medicine. Clin Microbiol Infect. 2007 Nov 1;13(11):1041–7.

90. Kumar PS. Oral microbiota and systemic disease. Anaerobe. 2013;24.

91. Billing F. Focal Infection: Its broader application in the etiology of general disease. J Am Med Assoc. 1914;63(11):5.

92. Gendron R, Grenier D, Maheu-Robert L-F. The oral cavity as a reservoir of bacterial pathogens for focal infections. Microbes Infect. 2000 Jul;2(8):897–906.

93. Pizzo G, Guiglia R, Lo Russo L, Campisi G. Dentistry and internal medicine: from the focal infection theory to the periodontal medicine concept. Eur J Intern Med. 2010 Dec;21(6):496–502.

94. Souto R, Andrade AFB, Uzeda M, Colombo APV. Prevalence of "non-oral" pathogenic bacteria in subgingival biofilm of subjects with chronic periodontitis. Braz J Microbiol. 2006;(37):8.

95. Loyola-Rodriguez JP, Martinez-Martinez RE, Abud-Mendoza C, Patiño-Marin N, Seymour GJ. Rheumatoid arthritis and the role of oral bacteria. J Oral Microbiol [Internet]. 2010 Dec 21 [cited 2016 Sep 12];2. Available from: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3084578/

96. Ahn J, Chen CY, Hayes RB. Oral microbiome and oral and gastrointestinal cancer risk. Cancer Causes Control. 2012 Jan 22;23(3):399–404.

97. Ley RE, Hamady M, Lozupone C, Turnbaugh PJ, Ramey RR, Bircher JS, et al. Evolution of Mammals and Their Gut Microbes. Science. 2008 Jun 20;320(5883):1647–51.

98. Ochman H, Worobey M, Kuo C, Ndjango JN, Peeters M, Hahn BH, et al. Evolutionary Relationships of Wild Hominids Recapitulated by Gut Microbial Communities. PLoS One. 2010;8(11):8.

99. Degnan PH, Pusey AE, Lonsdorf EV, Goodall J, Wroblewski EE, Wilson ML, et al. Factors associated with the diversification of the gut microbial communities within chimpanzees from Gombe National Park. Proc Natl Acad Sci. 2012 Aug 7;109(32):13034–9.

100. Moeller AH, Ochman H. Factors that drive variation among gut microbial communities. Gut Microbes. 2013 Sep 28;4(5):403–8.

101. Moeller AH, Peeters M, Ndjango J-B, Li Y, Hahn BH, Ochman H. Sympatric chimpanzees and gorillas harbor convergent gut microbial communities. Genome Res. 2013 Oct 1;23(10):1715–20.

102. Moeller AH, Li Y, Ngole EM, Ahuka-Mundeke S, Lonsdorf EV, Pusey AE, et al. Rapid changes in the gut microbiome during human evolution. Proc Natl Acad Sci. 2014 Nov 18;111(46):16431–5.

103. Li J, Nasidze I, Quinque D, Li M, Horz H-P, André C, et al. The saliva microbiome of Pan and Homo. BMC Microbiol. 2013;13:204.

104. Clayton JB, Vangay P, Huang H, Ward T, Hillmann BM, Al-Ghalith GA, et al. Captivity humanizes the primate microbiome. Proc Natl Acad Sci. 2016 Aug 29;201521835.

105. Moeller AH, Foerster S, Wilson ML, Pusey AE, Hahn BH, Ochman H. Social behavior shapes the chimpanzee pan-microbiome. Sci Adv [Internet]. 2016 Jan 15 [cited 2016 Sep 12];2(1). Available from: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC4730854/

106. Moeller AH, Caro-Quintero A, Mjungu D, Georgiev AV, Lonsdorf EV, Muller MN, et al. Cospeciation of gut microbiota with hominids. Science. 2016 Jul 22;353(6297):380–2.

107. Schnorr SL, Candela M, Rampelli S, Centanni M, Consolandi C, Basaglia G, et al. Gut microbiome of the Hadza hunter-gatherers. Nat Commun [Internet]. 2014 Apr 15 [cited 2015 Feb 21];5. Available from: http://www.nature.com/ncomms/2014/140415/ncomms4654/full/ncomms4654.html?message-global=remove

108. Clemente JC, Pehrsson EC, Blaser MJ, Sandhu K, Gao Z, Wang B, et al. The microbiome of uncontacted Amerindians. Sci Adv. 2015 Apr 1;1(3):e1500183.

109. Martínez I, Stegen JC, Maldonado-Gómez MX, Eren AM, Siba PM, Greenhill AR, et al. The Gut Microbiota of Rural Papua New Guineans: Composition, Diversity Patterns, and Ecological Processes. Cell Rep. 2015 Apr 28;11(4):527–38.

110. Lie MA, van der Weijden GA, Timmerman MF, Loos BG, van Steenbergen TJM, van der Velden U. Oral microbiota in smokers and

non-smokers in natural and experimentally-induced gingivitis. J Clin Periodontol. 1998;25(8):10.

111. Filippo CD, Cavalieri D, Paola MD, Ramazzotti M, Poullet JB, Massart S, et al. Impact of diet in shaping gut microbiota revealed by a comparative study in children from Europe and rural Africa. Proc Natl Acad Sci. 2010 Aug 17;107(33):14691–6.

112. Agus A, Denizot J, Thévenot J, Martinez-Medina M, Massier S, Sauvanet P, et al. Western diet induces a shift in microbiota composition enhancing susceptibility to Adherent-Invasive E. coli infection and intestinal inflammation. Sci Rep. 2016 Jan 8;6:19032.

113. Gilbert MTP, Jenkins DL, Gotherstrom A, Naveran N, Sanchez JJ, Hofreiter M, et al. DNA from pre-Clovis human coprolites in Oregon, North America. Science. 2008 May 9;320(5877):786–9.

114. Santiago-Rodriguez TM, Narganes-Storde YM, Chanlatte L, Crespo-Torres E, Toranzos GA, Jimenez-Flores R, et al. Microbial communities in pre-columbian coprolites. PLoS One. 2013;8(6):e65191.

115. Tito RY, Macmill S, Wiley G, Najar F, Cleeland L, Qu C, et al. Phylotyping and Functional Analysis of Two Ancient Human Microbiomes. PLoS One. 2008 Nov;3(11).

116. Tito RY, Knights D, Metcalf J, Obregon-Tito AJ, Cleeland L, Najar F, et al. Insights from characterizing extinct human gut microbiomes. PLoS One. 2012;7(12):e51146.

117. Dave M, Higgins PD, Middha S, Rioux KP. The human gut microbiome: current knowledge, challenges, and future directions. Transl Res. 2012 Oct;160(4):246–57.

118. White DJ. Processes contributing to the formation of dental calculus. Biofouling. 1991 Aug 1;4(1–3):209–18.

119. Dobney K, Brothwell D. Dental Calculus: Its Relevance to Ancient Diet and Oral Ecology. In: Cruwys E, Foley RA, editors. Teeth and Anthropology. 1986. p. 55–81. (British Archaeological Reports).

120. Fox CL, Juan J, Albert RM. Phytolith analysis on dental calculus, enamel surface, and burial soil: information about diet and paleoenvironment. Am J Phys Anthropol. 1996 Sep;101(1):101–13.

121. Henry AG, Piperno DR. Using plant microfossils from dental calculus to recover human diet: a case study from Tell al-Raqā'i, Syria. J Archaeol Sci. 2008 Jul;35(7):1943–50.

122. Preus HR, Marvik OJ, Selvig KA, Bennike P. Ancient bacterial DNA (aDNA) in dental calculus from archaeological human remains. J Archaeol Sci. 2011;38(8):1827–31.

123. De La Fuente C, Flores S, Moraga M. DNA from Human Ancient Bacteria: A Novel Source of Genetic Evidence from Archaeological Dental Calculus. Archaeometry. 2013;55(4):767–78.

124. Adler CJ, Dobney K, Weyrich LS, Kaidonis J, Walker AW, Haak W, et al. Sequencing ancient calcified dental plaque shows changes in oral microbiota with dietary shifts of the Neolithic and Industrial revolutions. Nat Genet. 2013 Apr;45(4):450–5.

125. Ziesemer KA, Mann AE, Sankaranarayanan K, Schroeder H, Ozga AT, Brandt BW, et al. Intrinsic challenges in ancient microbiome reconstruction using 16S rRNA gene amplification. Sci Rep. 2015 Nov 13;5:16498.

126. Warinner C, Rodrigues JFM, Vyas R, Trachsel C, Shved N, Grossmann J, et al. Pathogens and host immunity in the ancient human oral cavity. Nat Genet. 2014;

127. Kansa EC, Kansa SW, Watrall E. Archaeology 2.0: New Approaches to Communication and Collaboration. Cotsen Digit Archaeol Ser [Internet]. 2011 Jan 1 [cited 2016 Oct 14]; Available from: http://escholarship.org/uc/item/1r6137tb

128. Connell B, Rauxloh P. A rapid method for recording human skeletal data. Unpublished Museum of London report; 2003.

129. Jacobi RM, Higham TFG. The early Lateglacial re-colonization of Britain: new radiocarbon evidence from Gough's Cave, southwest England. Quat Sci Rev. 2009 Sep;28(19–20):1895–913.

130. Dinnis R. The archaeology of Britain's first modern humans. Antiquity. 2012 Sep;86(333):627–41.

131. Spencer C. The British Isles. In: Kiple KF, Ornelas KC, editors. The Cambridge World History of Food [Internet]. Cambridge: Cambridge University Press; 2012 [cited 2016 Oct 14]. p. 1217–26. Available from: http://universitypublishingonline.org/ref/id/histories/CBO9781139058643A017

132. Black J. London: a history. Carnegie; 2009.

133. Hinton DA, Crawford S, Hamerow H, editors. The Oxford Handbook of Anglo-Saxon Archaeology [Internet]. 1st ed. Oxford University Press; 2011 [cited 2016 Aug 26]. Available from: http://www.oxfordhandbooks.com/view/10.1093/oxfordhb/9780199212149.001.0001/oxfordhb-9780199212149

134. Breton J, Massart S, Vandamme P, De Brandt E, Pot B, Foligné B. Ecotoxicology inside the gut: impact of heavy metals on the mouse microbiome. BMC Pharmacol Toxicol. 2013;14:62.

135. Burns TW, O'Connor DJ, Stocklmayer SM. Science Communication: A Contemporary Definition. Public Underst Sci. 2003 Apr 1;12(2):183–202.

136. Russo G. Outreach: Meet the press. Nature. 2010 Nov 1;468(7322):465–7.

137. Balkwill F, Chambers K. Centre of the Cell: Science Comes to Life. PLoS Biol. 2015 Sep 4;13(9):e1002240.

138. Osterrieder A. The value and use of social media as communication tool in the plant sciences. Plant Methods. 2013;9:26.

139. Merton RK. The sociology of science : theoretical and empirical investigations. 1973.

140. Phillips DP, Kanter EJ, Bednarczyk B, Tastad PL. Importance of the Lay Press in the Transmission of Medical Knowledge to the Scientific Community. N Engl J Med. 1991 Oct 17;325(16):1180–3.

141. Eysenbach G. Can Tweets Predict Citations? Metrics of Social Impact Based on Twitter and Correlation with Traditional Metrics of Scientific Impact. J Med Internet Res. 2011 Dec 16;13(4):e123.

142. Priem J, Piwowar HA, Hemminger BM. Altmetrics in the wild: Using social media to explore scholarly impact. ArXiv12034745 Cs [Internet]. 2012 Mar 20 [cited 2014 Sep 30]; Available from: http://arxiv.org/abs/1203.4745

143. Wagner CS, Roessner JD, Bobb K, Klein JT, Boyack KW, Keyton J, et al. Approaches to understanding and measuring interdisciplinary scientific research (IDR): A review of the literature. J Informetr. 2011 Jan;5(1):14–26.

144. Körner M. Interprofessional teamwork in medical rehabilitation: a comparison of multidisciplinary and interdisciplinary team approach. Clin Rehabil. 2010 Aug 1;24(8):745–55.

145. O'Cathain A, Murphy E, Nicholl J. Multidisciplinary, Interdisciplinary, or Dysfunctional? Team Working in Mixed-Methods Research. Qual Health Res. 2008 Nov 1;18(11):1574–85.

146. Ecklund EH, James SA, Lincoln AE. How Academic Biologists and Physicists View Science Outreach. PLOS ONE. 2012 May 9;7(5):e36240.

147. Roeser S. Risk Communication, Public Engagement, and Climate Change: A Role for Emotions. Risk Anal. 2012 Jun 1;32(6):1033–40.

148. Fiske ST, Dupree C. Gaining trust as well as respect in communicating to motivated audiences about science topics. Proc Natl Acad Sci. 2014 Sep 16;111(Supplement 4):13593–7.

149. Brownell SE, Price JV, Steinman L. Science Communication to the General Public: Why We Need to Teach Undergraduate and Graduate Students this Skill as Part of Their Formal Scientific Training. J Undergrad Neurosci Educ. 2013 Oct 15;12(1):E6–10.

150. Bik HM, Goldstein MC. An Introduction to Social Media for Scientists. PLoS Biol. 2013 Apr 23;11(4):e1001535.

151. Kietzmann JH, Hermkens K, McCarthy IP, Silvestre BS. Social media? Get serious! Understanding the functional building blocks of social media. Bus Horiz. 2011 May;54(3):241–51.

152. Bracken LJ, Oughton EA. “What do you mean?” The importance of language in developing interdisciplinary research. Trans Inst Br Geogr. 2006 Sep 1;31(3):371–82.

Chapter I

Ancient bacteria from dental calculus as an archaeological and anthropological tool

Statement of Authorship

Title of Paper	Ancient bacteria from dental calculus as an archaeological and anthropological tool
Publication Status	Unpublished and Unsubmitted work written in manuscript style
Publication Details	In preparation for: Science and Technology of Archaeological Research (STAR)

Principal Author

Name of Principal Author (Candidate)	Andrew G. Farrer		
Contribution to the Paper	Reviewed literature. Wrote the manuscript		
Overall percentage (%)	80%		
Certification:	This paper reports on original research I conducted during the period of my Higher Degree by Research candidature and is not subject to any obligations or contractual agreements with a third party that would constrain its inclusion in this thesis. I am the primary author of this paper.		
Signature		Date	6th October 2016

Co-Author Contributions

By signing the Statement of Authorship, each author certifies that:

i. the candidate's stated contribution to the publication is accurate (as detailed above);

ii. permission is granted for the candidate in include the publication in the thesis; and

iii. the sum of all co-author contributions is equal to 100% less the candidate's stated contribution.

Name of Co-Author	Laura S. Weyrich		
Contribution to the Paper	Advised on content, and discussed and edited the manuscript		
Signature		Date	7th October 2016

Name of Co-Author	Keith Dobney		
Contribution to the Paper	Discussed and edited the manuscript		
Signature		Date	10th October 2016

Name of Co-Author	Alan Cooper		
Contribution to the Paper	Edited the manuscript		
Signature		Date	11th October 2016

Ancient bacteria from dental calculus as an archaeological and anthropological tool

Authors:

A. G. Farrer[1*], K. Dobney[2], A. Cooper[1+], and L. S. Weyrich[1+]

Affiliations:

[1]Australian Centre for Ancient DNA, School of Biological Sciences, University of Adelaide, Adelaide, South Australia, Australia

[2]Department of Archaeology, Classics and Egyptology, School of Histories, Languages and Cultures, University of Liverpool, Liverpool, United Kingdom

[+]These authors contributed equally to this work

Corresponding author:

[*]andrew.farrer@adelaide.edu.au

Abstract:

A critical aspect of archaeology and anthropology is the identification of the culture and life experiences of historical individuals. Identification of dietary patterns, disease exposure, migrations, trade, and socio-economic group membership is currently difficult too establish unless a substantial array of contextual evidence is available. However, many life events alter the structure of the commensal bacterial communities (microbiota) that live on and in the human body. Recent studies demonstrate that elements of the microbiota can be recovered from archaeological samples, especially calcified dental plaque (calculus). Here, we discuss the potential for using ancient microbiota from dental calculus present on archaeological human remains to identify an array of life events and histories, and indicate some aspects of cultural affinity.

Introduction

Understanding human evolutionary and cultural development is a multi-disciplinary endeavour. These disciplines involve direct study of human remains and associated evidence of human behaviour, placing this information in a broader context in order to first describe and reconstruct the historical context, and then attempt to explain the diverse and complex bio-cultural history. However, it remains an immense challenge to go from initial analyses of an archaeological site to a confident description of the culture of the people who created it, even with many lines of evidence used, including archaeological artifacts and features, osteology, dating techniques, zooarchaeology, and molecular archaeology. While artifacts and features such as pottery and building remains are powerful and traditional cultural indicators, movement of people and goods, and the use of similar architecture, manufacturing, and artistic styles can obscure more complex population structures and individual or group dynamics. The direct study of human remains (and associated grave goods) may be the only reliable source of information about an individual's cultural affinity/identity.

Excavation of human remains includes extensive analyses. Contextual information such as grave goods, or use of a grave marker or coffin, can indicate date, an individual's religion and their socio-economic status (Renfrew and Bahn 2004). If the burial is undisturbed, burial position and orientation can also provide further insights into these factors (Lilley et al. 1994). Osteological analysis can infer sex, age, dietary information, and health and disease (Wood et al. 1992; Cox and Mays 2000). These analyses can be conducted even if the remains have been disturbed, although destruction or loss of the skeleton limits interpretation. Standardized assessment of morphology helps normalize across different analyses, but morphology is a difficult trait to assess, with identification of sex and age not always possible (Lovejoy et al. 1985; Bruzek 2002). Patterns of bone wear and remodeling can indicate the stresses placed upon the body, thus providing some basic but limited information about an individual's lifestyle and (perhaps) societal rank (Macintosh, Pinhasi, and Stock 2014). Tooth

wear is an important element of analysing diet (Molnar 1971), and can be further coupled with isotopic analysis to determine general dietary trends (e.g. Frei et al. 2015; Richards, et al. 2006). It has also been used to explore specific cultural behavior such as including tool use (Molnar 1971).

The recent development of molecular archaeology has added a further powerful tool with which to explore the human journey. It involves the use of preserved biological molecules, such as ancient DNA (aDNA) and proteins, and has been successful in elucidating population dynamics (Haak et al. 2015), demography (Harpending et al. 1998; Orlando and Cooper 2014), and kinship in humans (Bouwman et al. 2008; Haak et al. 2008). While whole genome or proteome evolution in an organism is relatively slow (1.1×10^{-8} per base pair per generation (Roach et al. 2010)), targeted analyses of specific regions with proportionally higher mutational rates (e.g. microsatellites) can provide higher resolution and begin to distinguish between and within populations (Cavalli-Sforza and Feldman 2003). Additionally, human DNA analysis can reveal definitive information about sex and genetic based disease. However, identification of specific events that occurred during life is limited, as DNA is not altered by factors such as disease or diet. Some infectious diseases, notable examples including plague (*Y. pestis*) and Tuberculosis (Bos et al. 2011; Donoghue et al. 2015), have recently been detected by extracting the actual pathogen DNA preserved in archaeological human skeletal material.

A recently developed approach in biomolecular archaeology studies the ancient bacterial DNA preserved in dental calculus (calcified dental plaque) (De La Fuente, Flores, and Moraga 2013; Adler et al. 2013; Warinner et al. 2014). The well-preserved DNA of oral bacteria trapped within dental calculus are currently the only known record of an ancient microbiota (Weyrich, Dobney, and Cooper 2015). Microbiota are the microbial communities across the human body, including the oral cavity, are shaped by the diet, disease, environment and bio-cultural dynamics of the host individual (Consortium 2010). As these same factors help define different cultural and socio-economic groups, this implies that analysis of the oral

bacterial communities preserved in dental calculus could potentially be used to infer life and cultural histories of individuals and population groups (Weyrich, Dobney, and Cooper 2015; Warinner, Speller, and Collins 2015). In this paper, we detail the relationships between microbiota, the human body, culture, health, and environment in the modern world, and consider these a proxy to understand how human microbiota may have been impacted by these factors in the past. We discuss how ancient microbiota may be examined using aDNA, including recent discoveries, technologies, and technical considerations. Lastly, we examine the archaeological insights and applications that aDNA analysis of dental calculus can provide to help unravel the diverse and often complex bio-cultural human histories.

Human microbiota

Human microbiota are communities of microorganisms, primarily bacteria, living on and in the human body (Marchesi and Ravel 2015), and comprises as many microbial cells as human cells (totaling ~ 7 x 10^{13} cells) (Sender, Fuchs, and Milo 2016). Different body regions harbor unique bacterial profiles, which form biofilms across the various ecological niches of the body (*e.g.* skin, mouth, and gut) (Consortium 2010). These diverse biofilms, play key roles in human health (Flowers and Ellingrod 2015; Xu, Wang, and Zhang 2015), and perform a multitude of daily functions that the human body is incapable of completing alone, including functions as diverse as food digestion and vitamin production in the gut (Rosenbaum, Knight, and Leibel 2015), repair of tooth enamel in the mouth (Li et al. 2014), and digestion of oils and dead cells on the skin (Trivedi 2012). Microbiota also play critical roles in disease prevention by directly and indirectly excluding pathogenic microbes (Iwase et al. 2010; Weyrich et al. 2013; Chen and Tsao 2013; Shu et al. 2013), and by promoting immune tolerance and development (Noverr and Huffnagle 2004). Given the crucial roles microbiota play, and the complexity of the interactions with the human host, it is unsurprising that microbiota and their functions are inherited, and that disruptions to these communities can result in disease.

Microbiota are passed down from primary caregivers (predominantly the mother) to initiate development of the bacterial community in the child (Ebersole, Holt, and Delaney 2014). The initial microbial exposure occurs during birth, when the infant is inoculated with the mother's faecal and vaginal microbiota (Morgan and Huttenhower 2012). Differentiation of bacterial communities then occurs throughout the first few years of life (Funkhouser and Bordenstein 2013; Koenig et al. 2011). Body environment defines this differentiation, although the influx of from the surrounding environment is a source of further diversity. Generally, the composition of adult microbiota are reached within three years (for the gut community (Koenig et al. 2011)), although microbiota continue to be shaped throughout life.

Change in factors such as diet, disease, and environment trigger shifts in bacterial community structure during life. Diet is one of the most highly studied drivers of microbiota alteration. Nutrients and metabolically active bacteria introduced by the food, can alter the composition and functional capacity of gut microbiota (David et al. 2013). Altering an entirely animal- or plant-based diet can cause changes to the gut microbiota in as little as 24 hours (David et al. 2013). This difference in dietary-associated microbiota has also been observed more generally in the different bacteria states of obese and lean individuals (Turnbaugh et al. 2006, 2008; Turnbaugh, Hamady, et al. 2009), and rural and urban societies (Obregon-Tito et al. 2015). Comparison of Western gut microbiota with traditional hunter-gatherer groups has shown dramatic variation between the two, with dietary differences assumed to be the main cause (Schnorr et al. 2014; Clemente et al. 2015). The microbiota of a diet high in carbohydrates and refined sugars, as commonly seen in Western societies today (Cordain et al. 2005), results in increased weight gain (Turnbaugh et al. 2006, 2008; Turnbaugh, Ridaura, et al. 2009). Additional impacts of the Western diet have been an increase in the pathogens responsible for dental caries and periodontitis in the mouth (Selwitz, Ismail, and Pitts 2007), and poorer oral

health is generally more prevalent in post agricultural and industrial societies than in hunter-gatherers (Adler et al. 2013).

Microbiota have been linked to a wide range of local and systemic disease, including cardiovascular and diabetes (He et al. 2014), kidney disease (Sabatino et al. 2015), cancer (Zitvogel et al. 2015), and mental disorders (Flowers and Ellingrod 2015). However, microbiota play key roles in disease prevention by aiding the host in controlling pathogenic bacteria. For example, on the skin, commensal bacteria inhibit growth of the pathogen responsible for atopic dermatitis (Iwase et al. 2010). However, if a pathogenic bacterium does take hold, infections can alter microbiota composition, impacting immune function, and supporting additional bacterial infections (Weyrich et al. 2013), which can result in a wide range of diseases. Non-pathogen triggered alterations of microbiota can also cause disease. For example, kwashiorkor patients suffer from anorexia like symptoms and these individuals have microbiota associated with malnutrition, even when dietary supplements were provided. Transplantation of microbiota from healthy individuals provided a larger benefit than the improvement in diet alone (Smith et al. 2013). Further relating to the balance between disease prevention and causation, the microbiota interacts with the human immune system, and can play a regulatory role in inflammatory response (Stefka et al. 2014). The development of a mature immune system in infants is also closely linked to the developing microbiota, and the disruption of microbiota development following birth can have long-term consequences for disease susceptibility in an individual (Koenig et al. 2011; Azad et al. 2013). Thus, coupled with disease exposure, microbiota are a major factor defining the health of an individual or population.

Environmental factors, other than diet and disease, have also been shown to impact microbiota, though the range of influencing factors are not fully realized or understood. For example, if a person has older siblings, they tend to have a less diverse gut microbiota than those without. Conversely, having animals (pets) increases microbiota diversity (Azad et al. 2013), while urban living and industrialization has reduced our exposure to environmental

bacteria (Ruiz-Calderon et al. 2016). Urban housing (i.e. with more walls segregating living spaces) contain more human related bacteria, and individual rooms have a bacterial profile that appear correlated with room use (Ruiz-Calderon et al. 2016; Chase et al. 2016). It has been suggested that, for developing children, the bacterial exposure of an urban environment will be a defining factor in adult microbiota structure and disease (Ruiz-Calderon et al. 2016). Locations within the same city are likely to resemble each other more than between cities (Chase et al. 2016), indicating the potential for the oral microbiota to differentiate populations/groups at the city level.

Ancient DNA analysis of dental calculus

Ancient DNA studies have provided insight into the evolution and adaptation of human microbiota, and are a powerful new archaeological tool to understand past societies. Microbiota have been studied using dental calculus, soft tissues, and coprolites. However, while studies of 1,300 year old human coprolites from Mexico indicated that the microbiota more closely resembled samples from modern African than USA (Tito et al. 2008), as expected, a later study using coprolites and mummified remains revealed bacterial compositions that matched compost (Tito et al. 2012). Soft tissues are subject to post-deposit taphonomic changes and decay (Pechal et al. 2014), which alter the bacterial composition, and this has resulted in minimal use of coprolites or mummified remains to study human microbiota. In contrast, the well-preserved remains of bacterial communities in dental calculus are representative of microbiota present during the life of the individual (Weyrich 2015), and are not greatly impacted by post-mortem decay or alterations (Warinner et al. 2015). Consequently, dental calculus represents the only known reliable deep-time record of human microbiota (Weyrich, Dobney, and Cooper 2015). Bacteria within calculus are trapped and preserved in a calcified matrix formed from the congregation of calcium ions throughout the life of the individual (White 1991, 1997). This matrix is securely attached to the teeth and, once formed, preserves the bacteria

prior to and after the individual's death. From an archaeological perspective, dental calculus is regularly present on ancient human remains (Dobney and Brothwell 1986) and can be directly associated with a specific individual. Although bacteria were first observed and identified within ancient dental calculus in 1986 using Scanning Electron Microscopy (SEM) (Dobney and Brothwell 1986), modern genetic techniques applied to identify these microorganisms have drastically improved our understanding of ancient oral microbiota.

Preservation of recoverable genetic material was first shown by targeting key oral bacteria (De La Fuente, Flores, and Moraga 2013). Five bacterial species were tested using DNA barcodes (DNA fragments known to be unique to a species) in dental calculus samples from 38 individuals over 4,000 years old. At least one barcode was successfully recovered from 16 dental calculus samples, demonstrating that DNA recovery from dental calculus was possible. However, as oral microbiota comprise hundreds of species, individually targeting single species provides a limited view of the overall community structure. Consequently, high throughput sequencing (HTS), which allows researchers to obtain DNA sequences from multiple organisms simultaneously, has been applied in further published studies. To date, two HTS methods have been applied to dental calculus: amplicon and shotgun sequencing.

Initial whole community analyses of dental calculus (undertaken on 34 ancient European specimens) revealed major changes in the oral microbiota. These changes occured specifically during the prehistoric agricultural and much later industrial revolutions (~ 8,000 BCE and ~ 1750 CE, respectively), with an increasing pathogen load also seen through time (Adler et al. 2013). Increase in the production and consumption of carbohydrates and sugars during these two key bio-cultural transitions in the human diet is thought to be the key driving-factor in the observed shifts in bacterial communities. Along with these major changes, smaller differences between the distinct period and cultural groups represented by these

samples was also apparent, indicating that calculus might also provide a signal indicative of cultural affinity/identity (Adler et al. 2013).

Amplicon metabarcoding made it possible to recover whole community data in this study by targeting a specific gene or DNA region (amplicon) present in all of the organisms of interest. For bacteria, this amplicon target is a region within the 16S ribosomal RNA (rRNA) gene (Caporaso et al. 2012). This gene has sufficient variation to differentiate bacterial taxa while being conserved enough to be present and identifiable in all taxa. Once sequenced, the variety of 16S rRNA gene sequences recovered can be compared to bacterial databases (such as GreenGenes (DeSantis et al. 2006)) in order to identify the taxa present. 16S rRNA gene sequencing is the standard for bacterial identification in modern samples, but was not used in later ancient DNA.

A following study identifed a broad range of information from two medieval German human dental calculus samples dating from 950-1200 CE (Warinner et al. 2014). Analysis revealed oral bacterial pathogens similar to those present in modern humans, which were present despite the major cultural shifts occurring since the medieval period. Eukaryotic DNA was recovered in addition to bacterial DNA, which allowed direct identification of plant and animal DNA from dietary items that had also been preserved within the dental calculus. With this method, the medieval individuals were shown to have eaten pork, crucifer, and bread wheat. Additionally host (i.e. human) DNA was recovered. Moving beyond the presence/absence of specific taxa, functional analysis indicated that medieval German oral microbiota also showed low-level resistance to antibiotics (Warinner et al. 2014). Shotgun sequencing was required to recover the diversity of DNA to identify bacteria, eukaryotes, and functional information. Rather than targeting a specific gene, shotgun sequencing recovers a random sub-selection of the total DNA present in a sample. These fragments are mapped (matched) to known reference genomes from a large database, such as the NCBI taxonomy database (Benson et al. 2009; Sayers et al. 2009), to identify specific species. This is a financially and computationally expensive procedure in

relation to amplicon sequencing, but offers more detailed resolution and the ability to identify function. Recovery of human and dietary DNA was limited as 99.3 % of the DNA recovered was of bacterial origin (Warinner et al. 2014). Despite this, recent studies have been able to obtain human mitochondrial genomes from dental calculus by using methods to enrich for the low levels of human mitochondrial DNA (Ozga et al. 2016). While this demonstrates potential to use dental calculus as a reservoir for human DNA, recovery rates were lower than from teeth or bone.

Ethical considerations must be considered when isolating ancient human DNA. Cultural descendants of ancient individuals may not wish for genetic material from their ancestors to be analysed. Consequently, ethical approval should be sought prior to collecting samples or, if samples have been stored previously, before DNA extraction. In some cases, this ethical concern may extend to the microbiome, depending on how different groups recognize the microbiota as part of the human body. The method of DNA analysis will impact the ethical considerations. 16S rRNA gene sequencing will not result in human DNA been sequenced, thus is suitable when groups are interested in the microbiota but not the human genome. However, shotgun sequencing will sequence human DNA (if present) and thus removal of certain sequences may be required (*e.g.* human sequences) for ethical analysis.

Shotgun sequencing is now the standard for analysing ancient microbiota. Recent studies have shown that, despite the 16S rRNA gene being the international standard for characterizing modern microbiota (Caporaso et al. 2012), there are biases when applied to aDNA studies (Ziesemer et al. 2015). However, agreement of ancient bacterial community recovery between methods indicates that the 16S rRNA gene based technique can identify genuine patterns, likely in better-preserved samples (Warinner et al. 2014). Nevertheless, shotgun sequencing provides the ability to reproducibly recover ancient bacterial communities with high resolution, allowing for links between finer-scale human population groupings to be determined. Beyond the standardization of DNA extraction, shotgun

sequencing preparation, and taxonomic identification protocols, other key procedures developed during these initial studies are required to obtain accurate community identifications from ancient dental calculus samples.

Considerations when using ancient DNA

Substantial hurdles have to be overcome to successfully analyse microbial communities preserved within ancient dental calculus. Ancient samples contain very little endogenous (true) DNA, making inclusion of any contaminant DNA a major proportion of the total sequenced DNA (Cooper and Poinar 2000; Weiss et al. 2014). All ancient samples, including dental calculus, have this low DNA concentration in comparison to modern samples because cellular DNA repair mechanisms cease to function following cell death and DNA molecules begin to degrade (Dabney, Meyer, and Paabo 2013). Degradation causes the DNA molecules to break, creating progressively shorter fragments over time. Additional oxidative damage also occurs and modifies the DNA sequence (Pääbo et al. 2004). It is notable that dental calculus has a high DNA concentration for an ancient DNA source, up to 1,000 times the concentration found in bone (Weyrich, Dobney, and Cooper 2015). However, the DNA fragments surviving have undergone both of the above processes, which result in substantial biases, and must be considered when analysing ancient microbiota.

Alongside age, preservation conditions impact DNA breakage and damage. DNA degradation is slower in dry and cold environments, and quicker in warmer and wetter ones (Dabney, Meyer, and Paabo 2013). However, even under perfect preservation conditions, the theoretical limit of aDNA recovery is around 1 million years (Allentoft et al. 2012). To date, the oldest recovered and analysed aDNA comes from a ~700,000-year-old horse preserved in permafrost (Orlando et al. 2013). Reconstructing a community of multiple bacteria requires DNA fragments distinguishable between closely related taxa. As DNA degradation proceeds, it will become impossible to confidently assign a DNA sequence to one taxon over another. It has not been determined if the degradation limit is lower than for studies of

eukaryotic remains. However, in contrast to the oldest eukaryotic genome, the oldest published oral microbiota is only 8,000 years old (Adler et al. 2013).

In addition to DNA damage, contamination of the sample must be assessed. Bacteria are ubiquitous in the environment (Whitman, Coleman, and Wiebe 1998) and can be easily introduced into dental calculus samples prior to sequencing. This is a significant issue for all studies exploring bacterial communities (Salter et al. 2014), but is an even bigger risk for researchers exploring ancient bacterial DNA. Failure to account for contamination has resulted in dramatically incorrect conclusions. In 2014, bacterial contamination from laboratory equipment and reagents was identified in sequencing data, particularly in samples that had a low starting DNA concentration (analogous to aDNA samples) (Salter et al. 2014). Individual laboratories and reagent batches had their own unique signal, indicating that standardized taxa removal following DNA sequencing cannot account for the contamination, and that contamination in the laboratory must be continually monitored. Removal of contaminant sequences from a published amplicon dataset (Turner et al. 2012) demonstrated how the authors had incorrectly identified contamination from two different DNA extraction kits as an alteration in microbiota as children aged (Salter et al. 2014). Consequently, it is critical to limit, monitor, and remove contamination when processing ancient microbiota samples. Contamination in ancient samples typically comes from the two major sources: environment and laboratory (Warinner, Speller, and Collins 2015). Environmental contamination occurs after the death of the human as the calculus is exposed to the altering microbial community of the body as it decomposes and the microorganisms from the surrounding substrate. Additional contamination can occur during excavation, analysis, and storage (Yang and Watt 2005). Laboratory contamination comes from bacteria present in the laboratory facility and in the chemicals and equipment used to handle the samples. Each of these sources requires different methods and approaches to monitor and control for incoming bacterial DNA.

To limit environmental contamination from researchers, samples should be collected and handled by personnel with gloves and facemasks (Weyrich, Dobney, and Cooper 2015). Dental calculus samples should be collected in a clean environment, and removed from the tooth surface using a sterile pick. Samples should be stored in sterile, sealed bags to prevent contamination, kept at 4°C or cooler to minimize DNA degredation, and analysed in a dedicated aDNA laboratory to prevent contamination with modern DNA. Once samples are in an aDNA laboratory, environmental contamination (microbial cells or DNA) on the outside of the calculus sample can be removed and destroyed. Published methods include UV irradiation followed by a bleach submersion (Adler et al. 2013), or a predigestion with Ethylenediaminetetraacetic acid (EDTA) to break down cells on the outer surface (Warinner et al. 2014). Care must be taken when applying these methods as endogenous DNA may be damaged (reducing the true signal) (Malmström et al. 2007). Further, contaminant DNA may be damaged but not destroyed, potentially making the DNA appear ancient when assessed with bioinformatic tools such as MapDamage. However, these methods improve oral microbiota recovery (Adler et al. 2013; Warinner et al. 2014), they do not account for contamination introduced during laboratory analysis.

Once the extraction of DNA has begun, it is no longer possible to decontaminate the sample. However, contamination can continue to enter even if reagents have been sterilized before use (Salter et al. 2014). To minimize laboratory contamination, work should be conducted in a specialized aDNA facility that follows strict operational protocols to reduce contamination. The facility should have positive air pressure, to prevent entry of airborne contaminants; dedicated still-air workrooms, to minimize cross contamination of samples; regular procedures to decontaminate the facility (e.g. UV irradiation during no-work periods); and access only for trained personnel (Cooper and Poinar 2000; Knapp et al. 2012). Reagents should be certified DNA-free (although absence of DNA is not always certain) and equipment should be sterilized with bleach to remove trace DNA between uses (Willerslev and Cooper 2005).

These methods will limit, but not eliminate, contamination from the laboratory environment. Therefore, multiple negative controls should be included throughout the laboratory process (Cooper and Poinar 2000), including both extraction and amplification (or library construction) steps. Negative extraction controls (often referred to as extraction blank controls) should be included for each extraction, be exposed to the lab environment prior to samples being handled, and undergo the exact same treatment as a sample. Consequently, any microbial DNA detected in the extraction controls is indicative of laboratory contaminants during DNA extraction. Negative amplification controls should also be included during library preparation, to examine any additional contamination that could be introduced through amplification reagents. When amplicon sequencing is used, the microbial DNA identified in negative controls can be directly removed from the reconstructed microbiota (Adler et al. 2013). However, removing contaminant microbial DNA from shotgun sequencing can be more difficult. DNA fragments from different parts of the genome can be assigned to the same taxon; meaning DNA sequences cannot be directly compared between controls and samples. Despite this, DNA identified as corresponding to a contaminant species can still be filtered using bioinformatics tools. Another method for removing modern microbial contaminants is to analyse DNA fragments for characteristic aDNA damage patterns (e.g. C to T base transitions) (Ginolhac et al. 2011). Consequently, careful consideration must be given to which taxa are filtered, and diligence must be enforced when comparing data across studies (Kennedy et al. 2014).

Mapping-to-source analyses provide an indication of how contaminated a sample may be. This analysis was developed to identify the proportion of a microbial community that derived from a source population, for example where the bacterial community on a hospital surface originated (Knights et al. 2011). However, by comparing dental calculus samples to microbiota profiles that represent possible contamination sources the proportion of the data likely to have originated from these contamination

sources can be identified. Ideally the microbiota data sets compared to will come from samples of the environment surrounding the sample, such as the matrix surrounding the sample during preservation (*e.g.* soil), the sample storage containers and handling equipment, and the humans that handled the sample. However, representations of potential contamination sources (*e.g.* human skin and different soil types, etc.) are available in online data repositories (*e.g.* Qiita). Wherever the comparison data is from, it is critical that is of high quality. If the comparison data is contaminated, then the assignments of sequences in the samples to sources will be non-representative. While SourceTracker (a mapping-to-source analysis program (Knights et al. 2011)) has been used to assess the effectiveness of the above described filtering methods (Weyrich et al. 2016), to date this method has not been used to identify specific sequences to be removed from the sample data set. Identifying specific sequences that assign a proportion of a sample to a source population is not the aim of these analyses, rather these tests are perhaps most useful to indicate overall sample quality. Further, such analyses have only so far been applied to 16S rRNA gene data, and the results of application to Shotgun data should be assessed prior to any conclusions been drawn. While shotgun data offers a higher taxonomic resolution (which should increase the accuracy of these methods), current taxonomic assignment methods do not differentiate sub-groups of sequences within each taxon, thus reducing potential resolution. Whether the increased taxonomic resolution renders the loss of within taxon resolution redundant has not yet been assessed. However, the indicated effectiveness of the above described filtering methods on removing contamination in 16S rRNA gene datasets (Weyrich et al. 2016) is an indication of the importance of contamination control throughout the laboratory and analytical pipeline.

Using oral microbiota to unravel past human behaviour

Ancient microbiota offer an exciting new method to explore the life histories of individuals and populations. Modern studies have demonstrated

that microbiota respond to, and thus record, features of an individual's life, including dietary patterns, disease exposure, and environmental experience such as living conditions, animals, and people. Dental calculus allows analysis of microbiota post mortem, making it an ideal target for biomolecular archaeology studies. Here, we discuss several key archaeological questions for which dental calculus studies can provide insight, and examine how these findings can be used to link ancient individuals and cultures.

Diet

Access to food items is dependent on local availability, technological level of society, degree of trade, and status of an individual (Renfrew and Bahn 2004). Dental calculus allows direct identification of foodstuffs, e.g. pork and cabbage DNA was identified in ancient dental calculus from Medieval German populations (Warinner et al. 2014). This allows specific dietary components to be characterized in each individual, providing a deeper insight into dietary habits, available nutrition, and potential health of a specific individual. These dietary signals can also be used to identify food items that were not locally available, and thus can indicate the presence of product trade or population mobility. Furthermore, ancient bacterial community structures have been shown to respond to alterations in diet (Adler et al. 2013). In ancient calculus samples, shifts in the microbial community structure were linked to major dietary shifts, including the transition from foraging to farming, and later the impact of processed carbohydrates and food preservation (Adler et al. 2013). Food preparation also alters the nutrient and bacterial content of the diet (Monteiro 2009), and may alter the selection pressures on specific microorganisms. Consequently, many aspects of diet are reflected in microbiota structure of dental calculus, allowing researchers to identify both dietary components and health consequences without the need to locate or identify preserved, archaeological food items or utilities.

Disease

Infectious and chronic diseases can be detected in dental calculus through either direct identification of pathogenic microorganisms or by observation of an altered microbiota structure associated with disease. Direct identifications from the oral cavity include the 'red three' (*Porphymonas gingivalis*, *Tannerella forsythia*, and *Treponema denticola*), which are the pathogenic bacteria linked to severe periodontal disease (Rocas et al. 2001), and *Streptococcus mutans*, a causative agent of dental caries (Ajdić et al. 2002). Ancient DNA has revealed multiple examples of pathogenic bacteria from dental calculus (Adler et al. 2013; Warinner et al. 2014). Many infectious diseases also alter the microbiota directly, by manipulating the immune system and its response to commensal bacterial species (Weyrich et al. 2013). This responsive alteration of microbiota structure can be indicative of ancient diseases that are not identifiable to a specific pathogen, providing insight into endemic health. In addition, pathogen identification using aDNA from bone has already been noted as a potential method for circumventing the osteological paradox (Wright and Yoder 2003). The paradox notes that not all diseases are evident from the skeletal remains, for multiple reasons, which can result in misclassification of individual, and therefore population, health status (Wood et al. 1992). However, pathogen identification from bone allows characterization of a disease state that had no physiological impact on the skeleton. Microbiota studies may expand on this by identifying specific pathogens not transferred to bone during infection, diseases that killed the infected individual before bone remodeling occurred, and microbiota that have community structures corresponding to known disease states (as identified in modern studies and by characterizing microbiota of historical individuals known to be diseased from other contextual information). Comparison of disease presence across individuals can also provide an insight into a population's morbidity, and how this changes through time. Analysis across key historical events, such as wars, disease epidemics, introduction of sanitation, or societal collapse, could reveal how the population was impacted in different regions, between

socio-economic groups, and whether there were permanent or short-term impacts on health.

Human Environment

The human environment has varied markedly through history, and corresponding alterations in modern populations have been linked to changes in the microbiota. These environmental alterations include changes in housing, animal handling, occupation, medical practices and availability, cultural drug use, and historical events such as war. For example, built environments have bacterial communities that differ between cities, and, to a lesser degree, between buildings (Chase et al. 2016). Further, urban living can increase the potential for bacterial transmission between people (Ruiz-Calderon et al. 2016). This signal can provide insight into urbanization and mark individuals who were present in specific locations. Living with animals has also been shown to alter microbiota, as microorganisms are shared between co-habitating pets and humans (Lax et al. 2014). Thus the microbiota may be able to provide insight into the presence of animal cohabitation, husbandry, domestication and zoonoses. Similar environmental exposure from places of work, rather than the home, have been shown to impact modern skin microbiota (Chen & Tsao, 2013), suggesting that an individual's profession could also influence their bacterial community.

Habitual behaviours such as smoking tobacco also affect an individual's microbiota (Leeuwenhoek, 1683; Lie et al., 1998). Signatures of tobacco or coffee consumption (caffeine (Cowan et al. 2014)) could be used to identify status within a population. Other drugs, *e.g.* the use of betel nut in Pacific Islander populations (Dahlén et al. 2010), should also have specific microbial signatures that can likely be traced through time and space. The signatures of diet, disease, and environmental exposure, including personal behaviours, will vary within and between populations, allowing appreciation of the complexities of social structure and population movement across a region or country.

Trade, travel, and migration

Movement of people into new areas and between existing settlements, and the transport of goods, are major aspects of human culture. Migration and trade alter people's exposure to goods and microbes. The identification of dietary components, disease, or any other factors that are not in the local environment could be a signal of population and culture contact, admixture, or trade. Alterations in microbiota linked to the use of a specific product (e.g. tobacco) could be used to identify when it was introduced and when different groups began using it. Additionally, travel exposes individuals to different environments. People traveling to new regions may be exposed to and bring back unique commensal microbial species, similar to the transmission and spread of bubonic plague and cholera (Tatem, Rogers, and Hay 2006). Novel microbiota may also indicate recent migrants to a population, and the maintenance of these unique microbiota over time would indicate the level of integration. Similarly, differences in microbiota structure within a population, defined by resource availability, social movement, and exposure to alterations through trade and travel, also provides information about sub-populations, that may allow the identification of different socio-economic groups.

Developing dental calculus as an archaeological tool

The combined impacts of these factors will result in various different configurations of microbiota that may be indicative of a cultural group and the socio-economic sub-groups within populations. However, the number of samples required to discern many of these levels is yet to be determined. Using no more than 8 skeletons per culture, it has been observed that inter-country cultural differences could be determined, and these may also have been related to sub-groupings such as rural vs. urban living or differing religious practices (Adler et al. 2013). However, there is currently little baseline data and insufficient resolution to determine how populations and sub-groups are defined. Future studies should look to sample single

populations in greater depth to be able to observe the microbiota trends linked to sub-cultures. Furthermore, initial applications of these studies will require historical populations and individuals with detailed records and inferences of lifestyle, health, trade, diet, etc., allowing patterns in microbiota to be linked with specific behavioural and cultural traits. Such studies will elucidate patterns of microbiota structure that will be informative when examining dental calculus from human remains with little or no archaeological context.

Conclusions

Archaeology and anthropology look to understand past peoples and their cultures, and a large array of disciplines contribute to this goal. Yet, it is still difficult to identify the socio-economic and cultural structures within many past populations. In many situations, evidence of social rank or the dynamics of cultural affinity/identity is limited or missing. Here, we propose that dental calculus, a calcified and (technically) non-human tissue, found commonly in many archaeological human skeletal assemblages) can now be used as a line of evidence to resolve fine-scale population divisions. Dental calculus is the only accurate fossil record of human microbiota, which record an individual's diet, culture, and environment. Modern studies continue to identify factors that impact microbiota, and current aDNA studies are retrieving signals preserved in ancient and historical dental calculus. Using established aDNA protocols, future studies should collect, process, and analyse samples from well-described populations to identify the preserved patterns. These studies will define and calibrate microbiota analysis as a powerful new tool with which to reconstruct key aspects of human history.

References

Adler, Christina J., Keith Dobney, Laura S. Weyrich, John Kaidonis, Alan W. Walker, Wolfgang Haak, Corey J. A. Bradshaw, et al. 2013. "Sequencing Ancient Calcified Dental Plaque Shows Changes in Oral Microbiota with Dietary Shifts of the Neolithic and Industrial Revolutions." Nature Genetics 45 (4): 450–55. doi:10.1038/ng.2536.

Ajdić, Dragana, William M. McShan, Robert E. McLaughlin, Gorana Savić, Jin Chang, Matthew B. Carson, Charles Primeaux, et al. 2002. "Genome Sequence of Streptococcus Mutans UA159, a Cariogenic Dental Pathogen." Proceedings of the National Academy of Sciences 99 (22): 14434–39. doi:10.1073/pnas.172501299.

Allentoft, Morten E., Matthew Collins, David Harker, James Haile, Charlotte L. Oskam, Marie L. Hale, Paula F. Campos, et al. 2012. "The Half-Life of DNA in Bone: Measuring Decay Kinetics in 158 Dated Fossils." Proc. R. Soc. B 279 (1748): 4724–33. doi:10.1098/rspb.2012.1745.

Azad, M. B., T. Konya, H. Maughan, D. S. Guttman, C. J. Field, M. R. Sears, A. B. Becker, J. A. Scott, A. L. Kozyrskyj, and CHILD Study Investigators. 2013. "Infant Gut Microbiota and the Hygiene Hypothesis of Allergic Disease: Impact of Household Pets and Siblings on Microbiota Composition and Diversity." Allergy, Asthma & Clinical Immunology 9 (15): 9. doi:10.1186/1710-1492-9-15.

Benson, Dennis A., Ilene Karsch-Mizrachi, David J. Lipman, James Ostell, and Eric W. Sayers. 2009. "GenBank." Nucleic Acids Research 37 (suppl 1): D26–31. doi:10.1093/nar/gkn723.

Bos, K. I., V. J. Schuenemann, G. B. Golding, H. A. Burbano, N. Waglechner, B. K. Coombes, J. B. McPhee, et al. 2011. "A Draft Genome of Yersinia Pestis from Victims of the Black Death." Nature 478 (7370): 506–10. doi:10.1038/nature10549.

Bouwman, Abigail S., Keri A. Brown, A. John N. W. Prag, and Terence A. Brown. 2008. "Kinship between Burials from Grave Circle B at Mycenae Revealed by Ancient DNA Typing." Journal of

Archaeological Science 35 (9): 2580–84. doi:10.1016/j.jas.2008.04.010.

Bruzek, Jaroslav. 2002. “A Method for Visual Determination of Sex, Using the Human Hip Bone.” American Journal of Physical Anthropology 117 (2): 157–68. doi:10.1002/ajpa.10012.

Caporaso, J. Gregory, Christian L. Lauber, William A. Walters, Donna Berg-Lyons, James Huntley, Noah Fierer, Sarah M. Owens, et al. 2012. “Ultra-High-Throughput Microbial Community Analysis on the Illumina HiSeq and MiSeq Platforms.” The ISME Journal 6 (8): 1621–24. doi:10.1038/ismej.2012.8.

Cavalli-Sforza, L. Luca, and Marcus W. Feldman. 2003. “The Application of Molecular Genetic Approaches to the Study of Human Evolution.” Nature Genetics 33 (March): 266–75. doi:10.1038/ng1113.

Chase, John, Jennifer Fouquier, Mahnaz Zare, Derek L. Sonderegger, Rob Knight, Scott T. Kelley, Jeffrey Siegel, and J. Gregory Caporaso. 2016. “Geography and Location Are the Primary Drivers of Office Microbiome Composition.” mSystems 1 (2): e00022-16. doi:10.1128/mSystems.00022-16.

Chen, Yiyin Erin, and Hensin Tsao. 2013. “The Skin Microbiome: Current Perspectives and Future Challenges.” Journal of the American Academy of Dermatology 69 (1): 143–155.e3. doi:10.1016/j.jaad.2013.01.016.

Clemente, Jose C., Erica C. Pehrsson, Martin J. Blaser, Kuldip Sandhu, Zhan Gao, Bin Wang, Magda Magris, et al. 2015. “The Microbiome of Uncontacted Amerindians.” Science Advances 1 (3): e1500183. doi:10.1126/sciadv.1500183.

Consortium, The 1000 Genomes Project. 2010. “A Map of Human Genome Variation from Population-Scale Sequencing.” Nature 467 (7319): 1061–73. doi:10.1038/nature09534.

Cooper, Alan, and Hendrik N. Poinar. 2000. “Ancient DNA: Do It Right or Not at All.” Science 289 (5482): 1139–1139. doi:10.1126/science.289.5482.1139b.

Cordain, Loren, S. Boyd Eaton, Anthony Sebastian, Neil Mann, Staffan Lindeberg, Bruce A. Watkins, James H. O'Keefe, and Janette Brand-Miller. 2005. "Origins and Evolution of the Western Diet: Health Implications for the 21st Century." The American Journal of Clinical Nutrition 81 (2): 341–54.

Cowan, Theresa E., Marie S. A. Palmnäs, Jaeun Yang, Marc R. Bomhof, Kendra L. Ardell, Raylene A. Reimer, Hans J. Vogel, and Jane Shearer. 2014. "Chronic Coffee Consumption in the Diet-Induced Obese Rat: Impact on Gut Microbiota and Serum Metabolomics." The Journal of Nutritional Biochemistry 25 (4): 489–95. doi:10.1016/j.jnutbio.2013.12.009.

Cox, Margaret, and Simon Mays. 2000. Human Osteology: In Archaeology and Forensic Science. Cambridge University Press.

Dabney, J., M. Meyer, and S. Paabo. 2013. "Ancient DNA Damage." Cold Spring Harb Perspect Biol 5 (7). doi:10.1101/cshperspect.a012567.

Dahlén, Gunnar, Camilla Nauclér, Sofia Nordwall, and Narong Suksu-art. 2010. "Oral Microflora in Betel-Chewing Adults of the Karen Tribe in Thailand." Anaerobe 16 (4): 331–36. doi:10.1016/j.anaerobe.2010.04.003.

David, L. A., C. F. Maurice, R. N. Carmody, D. B. Gootenberg, J. E. Button, B. E. Wolfe, A. V. Ling, et al. 2013. "Diet Rapidly and Reproducibly Alters the Human Gut Microbiome." Nature, December. doi:10.1038/nature12820.

De La Fuente, C., S. Flores, and M. Moraga. 2013. "DNA from Human Ancient Bacteria: A Novel Source of Genetic Evidence from Archaeological Dental Calculus." Archaeometry 55 (4): 767–78. doi:10.1111/j.1475-4754.2012.00707.x.

DeSantis, T. Z., P. Hugenholtz, N. Larsen, M. Rojas, E. L. Brodie, K. Keller, T. Huber, D. Dalevi, P. Hu, and G. L. Andersen. 2006. "Greengenes, a Chimera-Checked 16S rRNA Gene Database and Workbench Compatible with ARB." Applied and Environmental Microbiology 72 (7): 5069–72. doi:10.1128/AEM.03006-05.

Dobney, Keith, and Don Brothwell. 1986. “Dental Calculus: Its Relevance to Ancient Diet and Oral Ecology.” In Teeth and Anthropology, edited by Elizabeth Cruwys and Robert. A. Foley, 55–81. British Archaeological Reports 291.

Donoghue, Helen D., Mark Spigelman, Justin O’Grady, Ildikó Szikossy, Ildikó Pap, Oona Y. -C. Lee, Houdini H. T. Wu, Gurdyal S. Besra, and David E. Minnikin. 2015. “Ancient DNA Analysis – An Established Technique in Charting the Evolution of Tuberculosis and Leprosy.” Tuberculosis, Supplement issue: Tuberculosis in Evolution, 95, Supplement 1 (June): S140–44. doi:10.1016/j.tube.2015.02.020.

Ebersole, J. L., S. C. Holt, and J. E. Delaney. 2014. “Acquisition of Oral Microbes and Associated Systemic Responses of Newborn Nonhuman Primates.” Clin Vaccine Immunol 21 (1): 21–28. doi:10.1128/CVI.00291-13.

Flowers, Stephanie A., and Vicki L. Ellingrod. 2015. “The Microbiome in Mental Health: Potential Contribution of Gut Microbiota in Disease and Pharmacotherapy Management.” Pharmacotherapy: The Journal of Human Pharmacology and Drug Therapy 35 (10): 910–16. doi:10.1002/phar.1640.

Funkhouser, L. J., and S. R. Bordenstein. 2013. “Mom Knows Best: The Universality of Maternal Microbial Transmission.” PLoS Biol 11 (8): e1001631. doi:10.1371/journal.pbio.1001631.

Ginolhac, Aurelien, Morten Rasmussen, M. Thomas P. Gilbert, Eske Willerslev, and Ludovic Orlando. 2011. “mapDamage: Testing for Damage Patterns in Ancient DNA Sequences.” Bioinformatics 27 (15): 2153–55. doi:10.1093/bioinformatics/btr347.

Haak, Wolfgang, Guido Brandt, Hylke N. de Jong, Christian Meyer, Robert Ganslmeier, Volker Heyd, Chris Hawkesworth, Alistair W. G. Pike, Harald Meller, and Kurt W. Alt. 2008. “Ancient DNA, Strontium Isotopes, and Osteological Analyses Shed Light on Social and Kinship Organization of the Later Stone Age.” Proceedings of the

National Academy of Sciences 105 (47): 18226–31. doi:10.1073/pnas.0807592105.

Haak, Wolfgang, Iosif Lazaridis, Nick Patterson, Nadin Rohland, Swapan Mallick, Bastien Llamas, Guido Brandt, et al. 2015. “Massive Migration from the Steppe Was a Source for Indo-European Languages in Europe.” Nature advance online publication (March). doi:10.1038/nature14317.

Harpending, Henry C., Mark A. Batzer, Michael Gurven, Lynn B. Jorde, Alan R. Rogers, and Stephen T. Sherry. 1998. “Genetic Traces of Ancient Demography.” Proceedings of the National Academy of Sciences 95 (4): 1961–67.

He, Jinzhi, Yan Li, Yangpei Cao, Jin Xue, and Xuedong Zhou. 2014. “The Oral Microbiome Diversity and Its Relation to Human Diseases.” Folia Microbiologica 60 (1): 69–80. doi:10.1007/s12223-014-0342-2.

Iwase, T., Y. Uehara, H. Shinji, A. Tajima, H. Seo, K. Takada, T. Agata, and Y. Mizunoe. 2010. “Staphylococcus Epidermidis Esp Inhibits Staphylococcus Aureus Biofilm Formation and Nasal Colonization.” Nature 465 (7296): 346–49. doi:10.1038/nature09074.

Kennedy, Nicholas A., Alan W. Walker, Susan H. Berry, Sylvia H. Duncan, Freda M. Farquarson, Petra Louis, and John M. Thomson. 2014. “The Impact of Different DNA Extraction Kits and Laboratories upon the Assessment of Human Gut Microbiota Composition by 16S rRNA Gene Sequencing.” PLOS ONE 9 (2): e88982. doi:10.1371/journal.pone.0088982.

Knapp, Michael, Andrew C. Clarke, K. Ann Horsburgh, and Elizabeth A. Matisoo-Smith. 2012. “Setting the Stage – Building and Working in an Ancient DNA Laboratory.” Annals of Anatomy - Anatomischer Anzeiger, Special Issue: Ancient DNA, 194 (1): 3–6. doi:10.1016/j.aanat.2011.03.008.

Knights, Dan, Justin Kuczynski, Emily S. Charlson, Jesse Zaneveld, Michael C. Mozer, Ronald G. Collman, Frederic D. Bushman, Rob Knight, and Scott T. Kelley. 2011. “Bayesian Community-Wide Culture-

Independent Microbial Source Tracking." Nature Methods 8 (9): 761–63. doi:10.1038/nmeth.1650.

Koenig, Jeremy E., A. Spor, N. Scalfone, A. D. Fricker, Jesse Stombaugh, R. Knight, L. T. Angenent, and R. E. Ley. 2011. "Succession of Microbial Consortia in the Developing Infant Gut Microbiome." PNAS 108 (1): 8. doi:10.1073/pnas.1000081107.

Lax, Simon, Daniel P. Smith, Jarrad Hampton-Marcell, Sarah M. Owens, Kim M. Handley, Nicole M. Scott, Sean M. Gibbons, et al. 2014. "Longitudinal Analysis of Microbial Interaction between Humans and the Indoor Environment." Science 345 (6200): 1048–52. doi:10.1126/science.1254529.

Li, Xiaoke, Jinfang Wang, Andrew Joiner, and Jiang Chang. 2014. "The Remineralisation of Enamel: A Review of the Literature." Journal of Dentistry, Enamel erosion – Advanced mineralisation technology, 42, Supplement 1 (June): S12–20. doi:10.1016/S0300-5712(14)50003-6.

Lilley, J. M., G. Stroud, Don Brothwell, and M. H. Williamson. 1994. The Jewish Burial Ground at Jewbury. Vol. 12, Fascicule 3. 19 vols. The Archaeology of York: The Medieval Cemeteries. Council for British Archaeology.

Lovejoy, C. Owen, Richard S. Meindl, Robert P. Mensforth, and Thomas J. Barton. 1985. "Multifactorial Determination of Skeletal Age at Death: A Method and Blind Tests of Its Accuracy." American Journal of Physical Anthropology 68 (1): 1–14. doi:10.1002/ajpa.1330680102.

Macintosh, Alison A., Ron Pinhasi, and Jay T. Stock. 2014. "Divergence in Male and Female Manipulative Behaviors with the Intensification of Metallurgy in Central Europe." PLOS ONE 9 (11): e112116. doi:10.1371/journal.pone.0112116.

Malmström, Helena, Emma M. Svensson, M. Thomas P. Gilbert, Eske Willerslev, Anders Götherström, and Gunilla Holmlund. 2007. "More on Contamination: The Use of Asymmetric Molecular Behavior to Identify Authentic Ancient Human DNA." Molecular Biology and Evolution 24 (4): 998–1004. doi:10.1093/molbev/msm015.

Marchesi, Julian R., and Jacques Ravel. 2015. “The Vocabulary of Microbiome Research: A Proposal.” Microbiome 3 (1): 31. doi:10.1186/s40168-015-0094-5.

Margarita Frei, Karin, Ulla Mannering, Kristian Kristiansen, Morten E. Allentoft, Andrew S. Wilson, Irene Skals, Silvana Tridico, et al. 2015. “Tracing the Dynamic Life Story of a Bronze Age Female.” Scientific Reports 5 (May). doi:10.1038/srep10431.

Molnar, Stephen. 1971. “Human Tooth Wear, Tooth Function and Cultural Variability.” American Journal of Physical Anthropology 34 (2): 175–89. doi:10.1002/ajpa.1330340204.

Monteiro, Carlos A. 2009. “Nutrition and Health. The Issue Is Not Food, nor Nutrients, so Much as Processing.” Public Health Nutrition 12 (5): 729–31. doi:10.1017/S1368980009005291.

Morgan, X. C., and C. Huttenhower. 2012. “Chapter 12: Human Microbiome Analysis.” PLoS Comput Biol 8 (12): e1002808. doi:10.1371/journal.pcbi.1002808.

Noverr, Mairi C., and Gary B. Huffnagle. 2004. “Does the Microbiota Regulate Immune Responses Outside the Gut?” Trends in Microbiology 12 (12): 562–68. doi:10.1016/j.tim.2004.10.008.

Obregon-Tito, Alexandra J., Raul Y. Tito, Jessica Metcalf, Krithivasan Sankaranarayanan, Jose C. Clemente, Luke K. Ursell, Zhenjiang Zech Xu, et al. 2015. “Subsistence Strategies in Traditional Societies Distinguish Gut Microbiomes.” Nature Communications 6 (March): 6505. doi:10.1038/ncomms7505.

Orlando, Ludovic, and Alan Cooper. 2014. “Using Ancient DNA to Understand Evolutionary and Ecological Processes.” Annual Review of Ecology, Evolution, and Systematics 45 (1): 573–98. doi:10.1146/annurev-ecolsys-120213-091712.

Orlando, Ludovic, Aurélien Ginolhac, Guojie Zhang, Duane Froese, Anders Albrechtsen, Mathias Stiller, Mikkel Schubert, et al. 2013. “Recalibrating Equus Evolution Using the Genome Sequence of an

Early Middle Pleistocene Horse." Nature 499 (7456): 74–78. doi:10.1038/nature12323.

Ozga, Andrew T., Maria A. Nieves-Colón, Tanvi P. Honap, Krithivasan Sankaranarayanan, Courtney A. Hofman, George R. Milner, Cecil M. Lewis, Anne C. Stone, and Christina Warinner. 2016. "Successful Enrichment and Recovery of Whole Mitochondrial Genomes from Ancient Human Dental Calculus." American Journal of Physical Anthropology, March, n/a-n/a. doi:10.1002/ajpa.22960.

Pääbo, Svante, Hendrik Poinar, David Serre, Viviane Jaenicke-Després, Juliane Hebler, Nadin Rohland, Melanie Kuch, Johannes Krause, Linda Vigilant, and Michael Hofreiter. 2004. "Genetic Analyses from Ancient DNA." Annual Review of Genetics 38 (1): 645–79. doi:10.1146/annurev.genet.37.110801.143214.

Pechal, J. L., T. L. Crippen, M. E. Benbow, A. M. Tarone, S. Dowd, and J. K. Tomberlin. 2014. "The Potential Use of Bacterial Community Succession in Forensics as Described by High Throughput Metagenomic Sequencing." Int J Legal Med 128 (1): 193–205. doi:10.1007/s00414-013-0872-1.

Renfrew, Colin, and Paul Bahn. 2004. Archaeology: Theories, Methods, and Practice. Fourth. Thames and Hudson.

Richards, M. P., B. T. Fuller, and T. I. Molleson. 2006. "Stable Isotope Palaeodietary Study of Humans and Fauna from the Multi-Period (Iron Age, Viking and Late Medieval) Site of Newark Bay, Orkney." Journal of Archaeological Science 33 (1): 122–31. doi:10.1016/j.jas.2005.07.003.

Roach, Jared C., Gustavo Glusman, Arian F. A. Smit, Chad D. Huff, Robert Hubley, Paul T. Shannon, Lee Rowen, et al. 2010. "Analysis of Genetic Inheritance in a Family Quartet by Whole-Genome Sequencing." Science 328 (5978): 636–39. doi:10.1126/science.1186802.

Rocas, I. N., J. F. Siqueira Jr., K. R. Santos, and A. M. Coelho. 2001. "'Red complex' (Bacteroides Forsythus, Porphyromonas Gingivalis, and

Treponema Denticola) in Endodontic Infections: A Molecular Approach." Oral Surg Oral Med Oral Pathol Oral Radiol Endod 91 (4): 468–71. doi:10.1067/moe.2001.114379.

Rosenbaum, Michael, Rob Knight, and Rudolph L. Leibel. 2015. "The Gut Microbiota in Human Energy Homeostasis and Obesity." Trends in Endocrinology & Metabolism 26 (9): 493–501. doi:10.1016/j.tem.2015.07.002.

Ruiz-Calderon, Jean F., Humberto Cavallin, Se Jin Song, Atila Novoselac, Luis R. Pericchi, Jean N. Hernandez, Rafael Rios, et al. 2016. "Walls Talk: Microbial Biogeography of Homes Spanning Urbanization." Science Advances 2 (2): e1501061. doi:10.1126/sciadv.1501061.

Sabatino, Alice, Giuseppe Regolisti, Irene Brusasco, Aderville Cabassi, Santo Morabito, and Enrico Fiaccadori. 2015. "Alterations of Intestinal Barrier and Microbiota in Chronic Kidney Disease." Nephrology Dialysis Transplantation 30 (6): 924–33. doi:10.1093/ndt/gfu287.

Salter, Susannah J., Michael J. Cox, Elena M. Turek, Szymon T. Calus, William O. Cookson, Miriam F. Moffatt, Paul Turner, Julian Parkhill, Nicholas J. Loman, and Alan W. Walker. 2014. "Reagent and Laboratory Contamination Can Critically Impact Sequence-Based Microbiome Analyses." BMC Biology 12 (1): 87. doi:10.1186/s12915-014-0087-z.

Sayers, Eric W., Tanya Barrett, Dennis A. Benson, Stephen H. Bryant, Kathi Canese, Vyacheslav Chetvernin, Deanna M. Church, et al. 2009. "Database Resources of the National Center for Biotechnology Information." Nucleic Acids Research 37 (suppl 1): D5–15. doi:10.1093/nar/gkn741.

Schnorr, Stephanie L., Marco Candela, Simone Rampelli, Manuela Centanni, Clarissa Consolandi, Giulia Basaglia, Silvia Turroni, et al. 2014. "Gut Microbiome of the Hadza Hunter-Gatherers." Nature Communications 5 (April). doi:10.1038/ncomms4654.

Selwitz, Robert H, Amid I Ismail, and Nigel B Pitts. 2007. “Dental Caries.” The Lancet 369 (9555): 51–59. doi:10.1016/S0140-6736(07)60031-2.

Sender, Ron, Shai Fuchs, and Ron Milo. 2016. “Are We Really Vastly Outnumbered? Revisiting the Ratio of Bacterial to Host Cells in Humans.” Cell 164 (3): 337–40. doi:10.1016/j.cell.2016.01.013.

Shu, M., Y. Wang, J. Yu, S. Kuo, A. Coda, Y. Jiang, R. L. Gallo, and C. Huang. 2013. “Fermentation of Propionibacterium Acnes, a Commensal Bacterium in the Human Skin Microbiome, as Skin Probiotics against Methicillin-Resistant Staphylococcus Aureus.” PLoS One 8 (2). doi:10.1371/journal.pone.0055380.g001.

Smith, M. I., T. Yatsunenko, M. J. Manary, I. Trehan, R. Mkakosya, J. Cheng, A. L. Kau, et al. 2013. “Gut Microbiomes of Malawian Twin Pairs Discordant for Kwashiorkor.” Science 339 (6119): 548–54. doi:10.1126/science.1229000.

Stefka, Andrew T., Taylor Feehley, Prabhanshu Tripathi, Ju Qiu, Kathy McCoy, Sarkis K. Mazmanian, Melissa Y. Tjota, et al. 2014. “Commensal Bacteria Protect against Food Allergen Sensitization.” Proceedings of the National Academy of Sciences 111 (36): 13145–50. doi:10.1073/pnas.1412008111.

Tatem, A. J., D. J. Rogers, and S. I. Hay. 2006. “Global Transport Networks and Infectious Disease Spread.” In Advances in Parasitology, edited by Alastair Graham and David J. Rogers Simon I. Hay, 62:293–343. Global Mapping of Infectious Diseases: Methods, Examples and Emerging Applications. Academic Press. http://www.sciencedirect.com/science/article/pii/S0065308X0562009X.

Tito, R. Y., D. Knights, J. Metcalf, A. J. Obregon-Tito, L. Cleeland, F. Najar, B. Roe, et al. 2012. “Insights from Characterizing Extinct Human Gut Microbiomes.” PLoS One 7 (12): e51146. doi:10.1371/journal.pone.0051146.

Tito, R. Y., S. Macmill, G. Wiley, F. Najar, L. Cleeland, C. Qu, P. Wang, et al. 2008. “Phylotyping and Functional Analysis of Two Ancient Human Microbiomes.” PLoS One 3 (11). doi:10.1371/journal.pone.0003703.g001.

Trivedi, B. 2012. “Microbiome: The Surface Brigade.” Nature 492 (7429): 2. doi:10.1038/492S60a.

Turnbaugh, P. J., F. Backhed, L. Fulton, and J. I. Gordon. 2008. “Diet-Induced Obesity Is Linked to Marked but Reversible Alterations in the Mouse Distal Gut Microbiome.” Cell Host Microbe 3 (4): 213–23. doi:10.1016/j.chom.2008.02.015.

Turnbaugh, P. J., M. Hamady, T. Yatsunenko, B. L. Cantarel, A. Duncan, R. E. Ley, M. L. Sogin, et al. 2009. “A Core Gut Microbiome in Obese and Lean Twins.” Nature 457 (7228): 480–84. doi:10.1038/nature07540.

Turnbaugh, P. J., R. E. Ley, M. A. Mahowald, V. Magrini, E. R. Mardis, and J. I. Gordon. 2006. “An Obesity-Associated Gut Microbiome with Increased Capacity for Energy Harvest.” Nature 444 (7122): 1027–31. doi:10.1038/nature05414.

Turnbaugh, P. J., Vanessa K. Ridaura, Jeremiah J. Faith, Federico E. Rey, Rob Knight, and Jeffrey I. Gordon. 2009. “The Effect of Diet on the Human Gut Microbiome: A Metagenomic Analysis in Humanized Gnotobiotic Mice.” Science Translational Medicine 1 (6): 6ra14-6ra14. doi:10.1126/scitranslmed.3000322.

Turner, Paul, Claudia Turner, Auscharee Jankhot, Naw Helen, Sue J. Lee, Nicholas P. Day, Nicholas J. White, Francois Nosten, and David Goldblatt. 2012. “A Longitudinal Study of Streptococcus Pneumoniae Carriage in a Cohort of Infants and Their Mothers on the Thailand-Myanmar Border.” PLOS ONE 7 (5): e38271. doi:10.1371/journal.pone.0038271.

Warinner, Christina, João F. Matias Rodrigues, Rounak Vyas, Christian Trachsel, Natallia Shved, Jonas Grossmann, Anita Radini, et al. 2014.

"Pathogens and Host Immunity in the Ancient Human Oral Cavity." Nature Genetics. doi:10.1038/ng.2906.

Warinner, Christina, Camilla Speller, and Matthew J. Collins. 2015. "A New Era in Palaeomicrobiology: Prospects for Ancient Dental Calculus as a Long-Term Record of the Human Oral Microbiome." Philosophical Transactions of the Royal Society of London B: Biological Sciences 370 (1660): 20130376. doi:10.1098/rstb.2013.0376.

Warinner, Christina, Camilla Speller, Matthew J. Collins, and Cecil M. Lewis Jr. 2015. "Ancient Human Microbiomes." Journal of Human Evolution, Special Issue: Ancient DNA and Human Evolution, 79 (February): 125–36. doi:10.1016/j.jhevol.2014.10.016.

Weiss, Sophie, Amnon Amir, Embriette R. Hyde, Jessica L. Metcalf, Se Jin Song, and Rob Knight. 2014. "Tracking down the Sources of Experimental Contamination in Microbiome Studies." Genome Biology 15 (12): 1–3. doi:10.1186/s13059-014-0564-2.

Weyrich, Laura S. 2015. "Evolution of the Human Microbiome and Impacts on Human Health, Infectious Disease, and Hominid Evolution." In Reticulate Evolution, edited by Nathalie Gontier, 231–53. Interdisciplinary Evolution Research 3. Springer International Publishing. doi:10.1007/978-3-319-16345-1_9.

Weyrich, Laura S., Keith Dobney, and Alan Cooper. 2015. "Ancient DNA Analysis of Dental Calculus." Journal of Human Evolution, Special Issue: Ancient DNA and Human Evolution, 79 (February): 119–24. doi:10.1016/j.jhevol.2014.06.018.

Weyrich, Laura S., Sebastian Duchane, Julien Soubrier, Luis Arriola, Bastien Llamas, James Breen, Alan G. Morris, et al. 2016. "Reconstructing Neandertal Behavior, Diet, and Disease Using Ancient DNA from Dental Calculus." Nature.

Weyrich, Laura S., H. A. Feaga, J. Park, S. J. Muse, C. Y. Safi, O. Y. Rolin, S. E. Young, and E. T. Harvill. 2013. "Resident Microbiota Affect Bordetella Pertussis Infectious Dose and Host Specificity." J Infect Dis, December. doi:10.1093/infdis/jit597.

White, Donald J. 1991. “Processes Contributing to the Formation of Dental Calculus.” Biofouling 4 (1–3): 209–18. doi:10.1080/08927019109378211.

———. 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 (5): 508–22. doi:10.1111/j.1600-0722.1997.tb00238.x.

Whitman, William B., David C. Coleman, and William J. Wiebe. 1998. “Prokaryotes: The Unseen Majority.” Proceedings of the National Academy of Sciences of the United States of America 95 (12): 6578–83.

Willerslev, Eske, and Alan Cooper. 2005. “Review Paper. Ancient DNA.” Proceedings of the Royal Society of London B: Biological Sciences 272 (1558): 3–16. doi:10.1098/rspb.2004.2813.

Wood, James W., George R. Milner, Henry C. Harpending, Kenneth M. Weiss, Mark N. Cohen, Leslie E. Eisenberg, Dale L. Hutchinson, et al. 1992. “The Osteological Paradox: Problems of Inferring Prehistoric Health from Skeletal Samples [and Comments and Reply].” Current Anthropology 33 (4): 343–70.

Wright, Lori E., and Cassady J. Yoder. 2003. “Recent Progress in Bioarchaeology: Approaches to the Osteological Paradox.” Journal of Archaeological Research 11 (1): 43–70. doi:10.1023/A:1021200925063.

Xu, Xiaofei, Zhujun Wang, and Xuewu Zhang. 2015. “The Human Microbiota Associated with Overall Health.” Critical Reviews in Biotechnology 35 (1): 129–40. doi:10.3109/07388551.2013.819485.

Yang, Dongya Y., and Kathy Watt. 2005. “Contamination Controls When Preparing Archaeological Remains for Ancient DNA Analysis.” Journal of Archaeological Science 32 (3): 331–36. doi:10.1016/j.jas.2004.09.008.

Ziesemer, Kirsten A., Allison E. Mann, Krithivasan Sankaranarayanan, Hannes Schroeder, Andrew T. Ozga, Bernd W. Brandt, Egija Zaura,

et al. 2015. “Intrinsic Challenges in Ancient Microbiome Reconstruction Using 16S rRNA Gene Amplification.” Scientific Reports 5 (November): 16498. doi:10.1038/srep16498.

Zitvogel, Laurence, Lorenzo Galluzzi, Sophie Viaud, Marie Vétizou, Romain Daillère, Miriam Merad, and Guido Kroemer. 2015. “Cancer and the Gut Microbiota: An Unexpected Link.” Science Translational Medicine 7 (271): 271ps1-271ps1. doi:10.1126/scitranslmed.3010473.

Chapter II

Effectiveness of decontamination protocols when analyzing ancient DNA preserved in dental calculus

Statement of Authorship

Title of Paper	Effectiveness of decontamination protocols when analyzing ancient DNA preserved in dental calculus.
Publication Status	Unpublished and Unsubmitted work written in manuscript style
Publication Details	In preparation for: PLoS One

Principal Author

Name of Principal Author (Candidate)	Andrew G. Farrer		
Contribution to the Paper	Co-designed the experiments. Performed decontamination protocols, DNA extraction, PCR amplification, and sequencing library preparation. Analysed the data and interpreted the results. Wrote the manuscript.		
Overall percentage (%)	70%		
Certification:	This paper reports on original research I conducted during the period of my Higher Degree by Research candidature and is not subject to any obligations or contractual agreements with a third party that would constrain its inclusion in this thesis. I am the primary author of this paper.		
Signature		Date	6th October 2016

Co-Author Contributions

By signing the Statement of Authorship, each author certifies that:

i. the candidate's stated contribution to the publication is accurate (as detailed above);

ii. permission is granted for the candidate in include the publication in the thesis; and

iii. the sum of all co-author contributions is equal to 100% less the candidate's stated contribution.

Name of Co-Author	Laura S. Weyrich		
Contribution to the Paper	Co-designed the experiments. Assisted in data analysis. Discussed and edited the manuscript		
Signature		Date	7th October 2016

Name of Co-Author	Alan Cooper		
Contribution to the Paper	Edited the manuscript		
Signature		Date	11th October 2016

Effectiveness of decontamination protocols when analysing ancient DNA preserved in dental calculus.

Andrew G. Farrer[1*], Alan Cooper[1+], and Laura S. Weyrich[1+]

Affiliations:

[1]Australian Centre for Ancient DNA, School of Biological Sciences, University of Adelaide, Adelaide, South Australia, Australia

[+]These authors contributed equally to this work

Corresponding author:

[*]andrew.farrer@adelaide.edu.au

Abstract

Ancient DNA analysis of human oral bacterial communities preserved within calcified dental plaque (calculus) allows reconstruction of ancient microbiota (the microbial communities living on and in the human body), and has revealed key insights into the history of human health. However, the risks of contamination are particularly high when analysing ancient microbiota due to the low concentration of endogenous DNA present in a sample and exposure to environmental DNA sources. The outer surface of an ancient calculus sample can be decontaminated prior to DNA extraction to remove environmental DNA sources and so improve the proportion of endogenous DNA in sequencing data. Several different surface decontamination protocols have been published, including Ethylenediaminetetraacetic acid (EDTA) pre-digestion or ultraviolet radiation (UV) and bleach immersion treatments. Here, we examine the efficiency of currently available protocols to remove environmental contaminant DNA from a range of ancient calculus samples from the archaeological site of Jewbury in York, UK. 16S rRNA gene amplicon data was examined for the presence of contaminants using the SourceTracker program and diversity analyses used to explore how the sample diversity is impacted by decontamination protocols. Our results indicate that both published methods are effective at reducing environmental taxa and increasing oral taxa in relation to untreated samples.

Introduction

The complex bacterial communities that make up the human microbiota vary between human individuals and across different body sites (e.g. the gut, skin, and oral cavity). Many bodily functions are dependent on these resident bacterial communities, which release otherwise inaccessible nutrients from food (1), remove dead epithelial cells from the skin (2), and repair tooth enamel (3). These diverse communities are also intricately linked with the human immune and endocrine systems (4), and microbiota alterations have now been linked to a wide range of diseases, including kidney and respiratory conditions (5,6), oral pathologies (7), allergies (8), obesity (9) and mental disorders (10). However, microbiota are also shaped and influenced by the external environment, including diet, disease exposure, and drug use (11–13). This relationship with the environment suggests that the development of future treatments for microbiota-linked diseases will be dependent on understanding how changes in lifestyle impact microbiota. Therefore, an examination of ancient microbiota in populations through time may provide insight into the mechanisms of microbiota change.

Ancient DNA (aDNA) analyses can offer valuable insights into the structure of these human microbial communities and their response to variation in cultural and environmental factors over multiple generations. Calcified dental plaque (calculus) remains the most reliable sample to reconstruct ancient human microbiota (14), as dental calculus is formed by the calcification of the diverse bacterial biofilm that forms on the tooth surface (15). This calcium matrix preserves and protects the bacterial cells from many of the abiotic and biotic factors that degrade soft tissues post-mortem. Recent studies of ancient dental calculus have revealed changes in microbiota that are linked to alteration in diet and lifestyle, including the implementation of agricultural practices (16). These studies have also identified an increase in oral pathogens through time (16). In addition, dental calculus preserves DNA of both the human host and the host's diet, although 99% of the preserved DNA is microbial in origin (17). However, the

preserved DNA in ancient samples is present low concentrations, due to the impacts of DNA degradation in ancient samples. Consequently, such samples can easily become contaminated with DNA from the environment or laboratory during processing. This contaminant DNA can significantly alter the microbial signal within an ancient sample, and therefore poses a significant risk for ancient microbiota analysis (18,19). To combat this, strict aDNA protocols must be followed, including methods to reduce and monitor contaminant DNA from the environment and laboratory.

Failure to account for contaminant DNA can led to severe misinterpretation of results, even in relatively high biomass samples (20). For example, microbial contaminant DNA from different manufacturing batches of DNA extraction kits can create signals within data that appear to be biological (20). Controlling and accounting for laboratory contaminant DNA should be standard practice in all microbiota and aDNA work (21). In aDNA facilities, levels of contaminant DNA are minimized by the use of sterile laboratories, equipment, reagents, and protocols, and is monitoring through the use of frequent extraction and no-template amplification (i.e. PCR negative) controls (22). However, such controls cannot detect contaminant DNA present on the sample prior to entry into the facility (14). Microbial DNA is present on the surface of ancient samples, including DNA from microbes from sediment, storage materials, and handling during and after excavation. Therefore, environmental contaminant DNA must be removed prior to DNA extraction and be further examined using bioinformatic tools to limit the inclusion of false signals and misinterpretation of the data.

To limit the contribution of environmental contaminant DNA to archaeological calculus samples, a decontamination protocol is performed prior to DNA extraction (16,17) to reduce the foreign contaminant DNA on the sample before the endogenous DNA is extracted. This results in a DNA extract that better represents the ancient oral microbiota of the individual, maximizes sequencing capacity (by reducing the proportion of non-endogenous DNA), and minimizes the need for bioinformatic filtering of contaminant DNA downstream. However, published protocols for the

decontamination of calculus differ between research groups. It is important that decontamination is standardized so that data is accurate and comparable between groups. Recovery of endogenous DNA from bone has shown to be improved by pre-extraction treatment with bleach and Ethylenediaminetetraacetic acid (EDTA). However, the microbial community structure was not altered (23). To assess the impacts of a decontamination treatment on a sample of microbial origin (calculus) we test two previously published protocols to identify which is the most effective at removing environmental DNA from ancient dental calculus samples: 1. pre-digestion in EDTA (17) and 2. UV irradiation and 5% sodium hypochlorite (bleach) immersion (16). We also test both the UV irradiation and bleach immersion treatments independently to assess their relative roles, and go on to explore the potential of retrieving environmental information by sequencing the DNA removed from the calculus surface. To test these decontamination protocols in a real-world scenario, we applied each to a 26 samples from a well-preserved, medieval archaeological site (Jewbury, York, UK) (24,25), which has previously been shown to have a robust oral microbial signal (16). These results provide the first examination of aDNA decontamination protocols for ancient dental calculus samples and serve as a resource for future analysis of ancient oral microbiota.

Methods

Archaeological context and site information

Twenty-six dental calculus samples from the Jewbury archaeological site were divided into five groups for analysis. Jewbury is a Medieval, Jewish cemetery in northern England, UK that was excavated in 1983 (24). Archaeological examination of the human remains ended following reburial requests from the Jewish community. However, considerable detail of the site and the skeletons was recorded (25) and dental calculus was collected for later analysis. Analyses were completed under ethical approval to study ancient human dental calculus (University of Adelaide: H-2012-108). The population contained individuals with middle to poor socio-economic

standing, and 98.1% of individuals were buried in single graves. Dental caries were observed in 59.5% of individuals, and periodontal disease was present in >80% (24). Calculus was sampled on-site at the time of excavation, stored in glass vials, and transported to the dedicated aDNA facility based at the Australian Centre for Ancient DNA, Adelaide, Australia for processing. Previous aDNA analysis of calculus from this site has revealed that oral microbial communities were not statistically different between samples, and were distinct from other cultures based on microbial composition alone (16).

Decontamination protocols

Each of the five groups underwent a different decontamination protocol prior to DNA extraction. The protocols are summarized in Figure 1, and were as follows: untreated controls (n = 5); EDTA treatment (n = 5) (17); combined UV and bleach treatment (n = 5) (16); UV treatment (n = 5); and bleach treatment (n = 6). For the EDTA treatment, calculus fragments were submerged in 1 mL 0.5M EDTA for 1 hour (17). The combined UV and bleach treatment exposed individual dental calculus fragments to UV radiation for 30 minutes on each side, followed by submerging the calculus sample in 3 mL of 5% bleach in a sterile petri dish for 3 minutes (16). The individual UV and bleach treatments used the respective element of the combined UV and bleach treatment protocol. Following decontamination, all samples were washed in 1 mL of sterile 80% ethanol for one minute to remove residual chemicals (e.g. EDTA or bleach) prior to extraction. To examine whether DNA was released from the calculus during the ethanol wash, the ethanol washes from each sample (n = 26), as well as control ethanol samples (n=3), were evaporated, and the resulting DNA was suspended in TLE buffer (500 µl Tris(hydroxymethyl)aminomethane hydrochloride (Tris HCL) (1M), 10 µl EDTA (0.5M), and 50 ml dH_2O) (26).

DNA extraction, amplification, and sequencing

All samples, except for the ethanol washes, underwent an in-house, silica based DNA extraction, as previously described (27). To account for small sample sizes, total volumes of lysis and guanidinium DNA binding buffer were reduced as follows: 1.8 mL lysis buffer (1.6 mL 0.5 M EDTA (0.5M), 200 µL SDS (10%), and 20 µL proteinase K (20 mg/ml)) and 3 mL guanidinium DNA binding buffer. Two extraction blank controls were included for every seven dental calculus sample extractions. A 289 base pair stretch from the V4 region of the 16S ribosomal RNA (rRNA) encoding gene (position 515-806 of the E. coli reference genome) was amplified in triplicate from all samples (dental calculus, resuspended ethanol washes, and extraction blank controls) alongside an additional PCR negative control using barcoded primers (28). Each PCR reaction contained: 17.5 µL sterile H_20, 1 µL of DNA extract, 0.25 µL of Hi-Fi taq (Life Technologies), 2.5 µL of 10X Hi-Fi reaction buffer, 1 µL $MgCl_2$ (25mM), 0.2 µL dNTPs (10 mM), and 1 µL each of the forward and reverse primers. Samples were amplified using the following conditions: initial denaturing (95°C, 6 minutes), followed by 37 cycles of denaturing (95°C, 30 seconds), annealing (50°C, 30 seconds), and elongation (72°C, 30 seconds), and finally adenylisation (60°C, 10 minutes). Following amplification, the triplicate reactions were pooled, and PCR products were visualized by electrophoresis on a 2.5% agarose gel. Samples were quantified (Qubit 2.0, Life Technologies) before being pooled at equimolar concentrations and purified (Ampure, Agencourt Bioscience). The pooled sample (e.g. DNA library) was quantified using the Tapestation and the KAPA SYBR Fast Universal master mix qPCR assay (Geneworks), and DNA sequencing was completed using the Illumina MiSeq 150bp paired end chemistry (Illumina, San Diego, CA, USA).

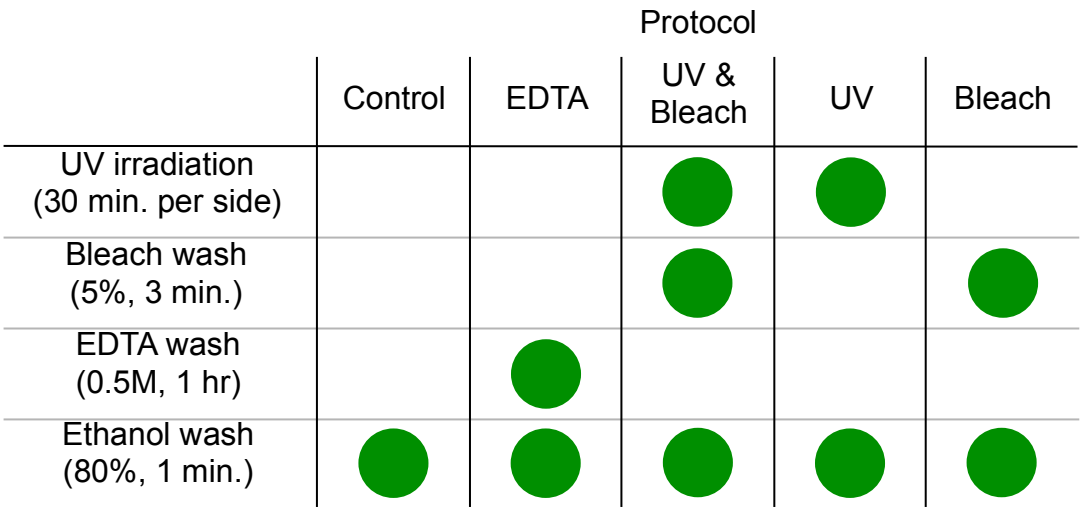


Figure 1: *The five decontamination protocols applied. A workflow, from top to bottom, for each of the decontamination protocols applied to the groups of ancient dental calculus samples in this study. Green dots represent use of that treatment on each sample being analysed.*

OTU picking and filtering

Sequences were demultiplexed and quality filtered in QIIME (V1.8) (29) using the split_libraries_fastq.py script with parameters: barcode error = 0 and quality score > 20. Operational taxonomic unit (OTU) picking was completed against GreenGenes (V13.8) (30) with 97% similarity using both closed and open reference methods. The closed reference OTU dataset only includes sequences that match references within the GreenGenes database; the open reference dataset also included OTUs without reference matches. To remove contaminant DNA introduced through laboratory processing, OTUs identified in negative controls and as common laboratory contaminants (20) were removed from dental calculus samples processed in the same batch. Ethanol washes, which were not expected to have a high biomass or necessarily to be representative of human microbiota, were filtered by the OTUs present in the control ethanol samples. Finally, singletons (OTUs present only once) were removed from the data.

Comparison data

To provide comparative data, dental calculus from modern (n = 6) and Industrial Revolution (n = 3) individuals that had been processed in the same lab were added to the dataset. The ancient samples had previously undergone the combined UV and bleach treatment described above, and both sample sets underwent extraction, amplification, sequencing, and filtering as described above. In addition, preprocessed 16S rRNA gene datasets were downloaded from the Qiita database for comparison (qiita.microbio.me). These samples included: human skin samples (n = 11) (31) and environmental samples from agricultural soil (n = 8), temperate soil (n = 4), forest soil (n = 4), tropical soil (n = 5), and park soil (n = 6) (Study IDs: 232, 808, 846, and 1674. qiita.microbio.me).

Bioinformatic and statistical analysis

Following filtering, all bioinformatics analyses were conducted within QIIME (V1.8). First, SourceTracker (V0.9.6) (32) was used to identify the proportions of endogenous and contaminant signal in each sample. To complete the SourceTracker analysis, closed reference data were compared to the following source populations: modern and Industrial Revolution dental calculus samples (representing high quality oral microbiota), human skin, and varied soil types (representing expected sources of contaminant DNA). Specifically, a subset of skin samples was used to minimize bias from unevenly sized reference groups (samples: P15024, P15268, P15733, P16107, P16187, P16199, P16304, P16320, P16393, P16399, and P16562 were used). SourceTracker was run with default parameters (1,000 subsampling, 10 iterations per sink sample) in R (V3.1.0) (33) using the QIIME wrapper. The suitability of the source populations was confirmed using the “take-one-out” method, which demonstrated that the samples within each reference group were more similar to one another than samples in any other group.

To examine differences in diversity between the different decontamination steps, a variety of analyses were performed. Alpha diversity (observed species) was calculated for each treatment group at rarefaction levels from 0 to 2,000 (in intervals of 10) using closed and open reference datasets in QIIME. A Goodness of fit test (G-test) was applied to detect significant differences in genera-level taxa between untreated samples and each of the decontamination protocols. To reduce false positives generated by rare taxa, OTUs below 0.1% of the total taxa present were removed before performing the G-test. To identify the environmental taxa impacted by the decontamination protocol genera level taxa that were significantly different ($p < 0.001$) were classified as environmental or oral based on their presence or absence (respectively) in the Human Oral Microbiome Database (HOMD) (homd.org).

Several statistical assessments were performed to identify taxa that were significantly altered by the different treatments. First, a one-way ANOVA was applied to test if the average frequency of OTUs in each protocol group had altered in relation to the untreated group. Next, for each sample, OTUs identified in the G-test analysis were ranked as increasing or decreasing relative to the untreated proportion. A one-way ANOVA was performed to identify taxa that significantly differed between the four treatment groups. Finally, taxa released into the ethanol washes were classified as environmental or oral using HOMD, and the ratio of oral to environmental taxa assessed.

Results

Reduction in detectable contaminant OTUs

Our first aim was to identify samples with the lowest level of known contaminant OTUs. SourceTracker analysis was applied to identify the proportion of OTUs within each sample that originated from oral, skin, soil, or unknown bacterial communities (Figure 2A). Interestingly, there was no evidence of skin microbiota in any sample. On average, samples in the untreated group were comprised of 9.1% soil and 27.8% oral OTUs. The

remaining proportion came from unknown sources. The previously published EDTA treatment and the combined UV and bleach treatment both decreased the average proportion of soil OTUs to 4.3% and <0.01%, respectively. Specifically, four out of the five EDTA treated samples had no detectable soil component, while a single EDTA treated sample was comprised of 21.5% soil OTUs. Similarly, only one sample from the combined UV and bleach treatment had a detectable soil signal, which represented <0.01% of the sample. An increase in the average oral proportion was detected in both published treatments (33.8% in the EDTA method and 33.6% in the combined method). When UV treatment and bleach treatment were performed independently, UV treated samples had an average soil component of 6.3%, while bleach treated samples had 5.3%. Surprisingly, UV treated samples had the lowest average oral component of all the groups (19.8%), and bleach treated had the highest (46.7%). Importantly, all treatments reduced environmental contaminant OTUs relative to the untreated samples, although the combined UV and bleach treatment was the most effective at removing soil contaminant OTUs across all samples. The EDTA treatment averages are skewed by a single sample, which may not be representative of the effectiveness of the EDTA treatment.

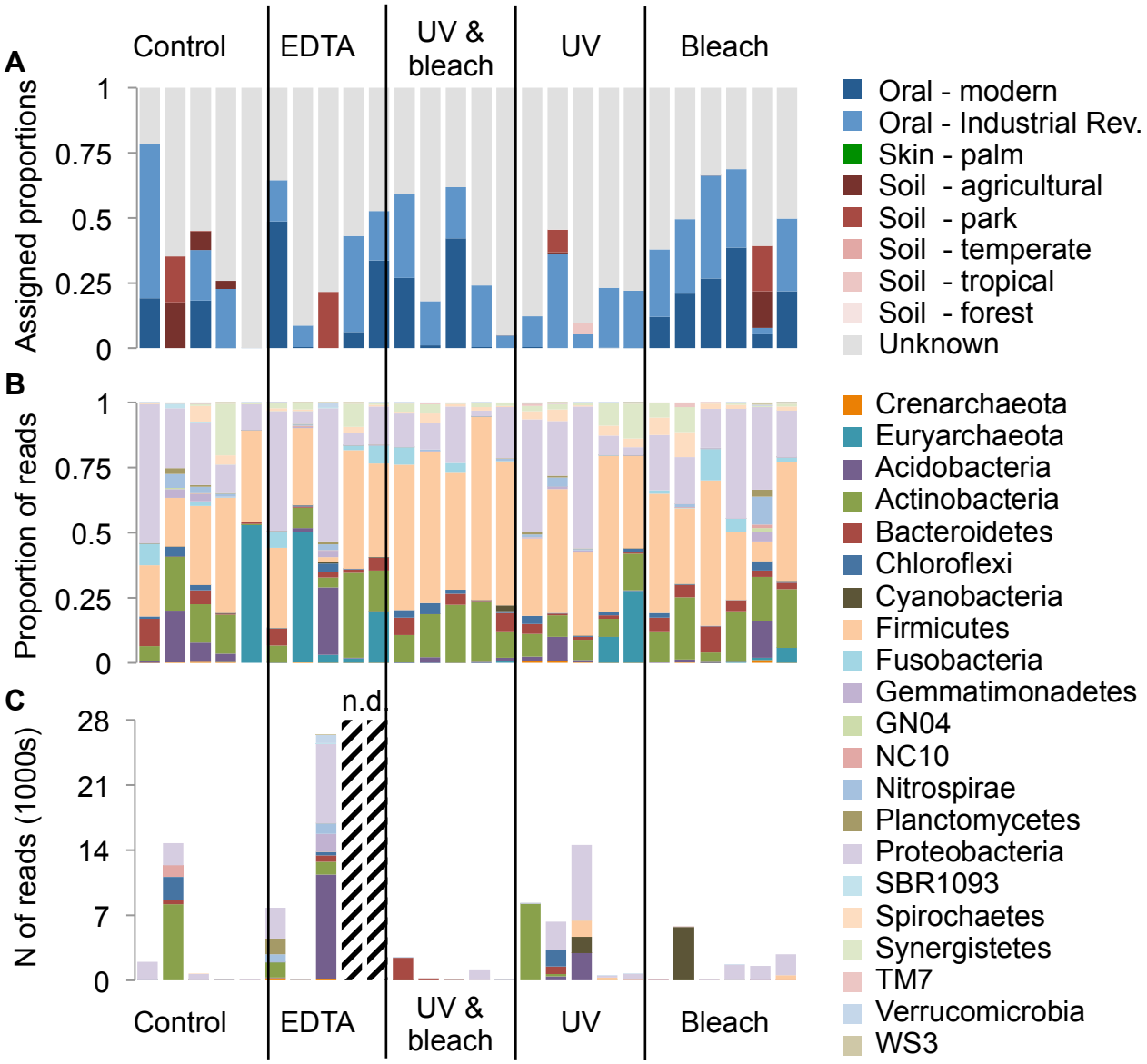


Figure 2: *Contamination and taxonomic profiles of each sample, grouped by decontamination protocol. The proportion of oral, skin, soil, and unknown OTUs in each sample. Proportions were defined by comparison to reference samples of oral, skin and soil microbial communities using the Bayesian modelling program SourceTracker (A). The OTUs identified within each phylum are displayed for each calculus sample (>0.08% of total proportion) (B) and for each ethanol wash (C). These analyses used closed reference OTUs (i.e. OTUs identified in the GreenGenes database).*

Decontamination alters sample diversity

An effective decontamination method should reduce both the number of contaminant taxa and therefore the overall diversity within calculus samples. To test this, alpha diversity was examined in both open and closed reference OTU data sets (Figure 3). The untreated group had the largest variation among samples, resulting in non-significant differences. However, there were interesting trends across the different treatments. Both published protocols reduced the microbial diversity in relation to the untreated samples. For the closed reference data (where only OTUs matching the GreenGenes database were considered), the average diversity in the EDTA treatment was 7.4% lower than the untreated group. In contrast, the UV and bleach treatment had a greater impact, and was 18.5% lower than the untreated group. For open reference OTUs (which also include de novo OTUs), the EDTA treatment was found to reduce diversity by 12.6%, while UV and bleach treatment reduced diversity by 1.1%. UV treatment also reduced diversity in both closed and open reference datasets (26.4% and 0.9%, respectively). Conversely, the bleach treatment increased the diversity in both cases. This is particularly notable in the open reference data, where an average 29.4% increase was observed in the bleach treated group in comparison to the untreated group. This suggests that bleach treatment artificially increases bacterial diversity by causing DNA damage that is recognized as sequence divergence and a unique OTU. Together, this data indicates that the DNA being removed from the calculus surface were from unique taxa in comparison to the taxa with preserved DNA within the sample, supporting the need for effective decontamination protocols.

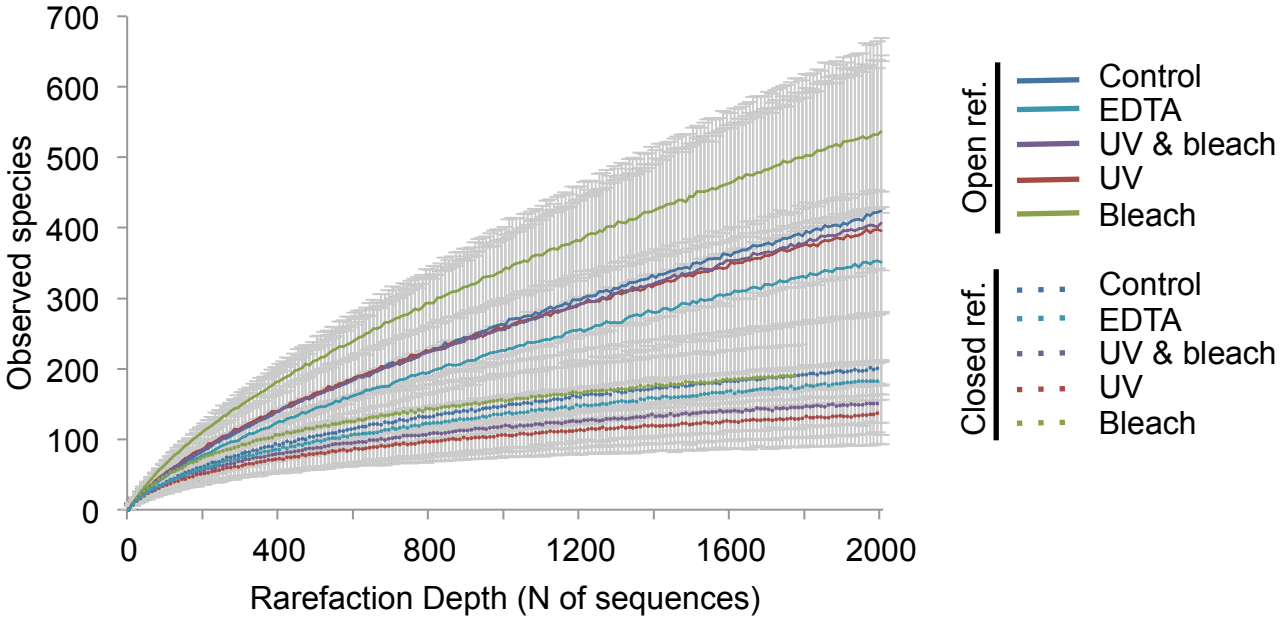


Figure 3: *Average alpha diversity within protocols. Lines represent the average alpha diversity (observed species) of all samples processed with the same decontamination protocol. Dotted lines signify the diversity calculated within the closed reference OTUs (OTUs identified in the GreenGenes database) and solid lines represent diversity detected within the open reference OTUs (Closed reference OTUs plus OTUs without matching sequences in the GreenGenes database). The colours represent the different protocols, and the error bars (one standard deviation) are shown in grey.*

Exclusion of environmental taxa following decontamination

We used a G-test to determine which genus level OTUs significantly changed in frequency during each treatment relative to the untreated group. Methanobrevibacter taxa were excluded from this analysis as abundance measures based on 16S rRNA gene sequencing are heavily biased for this taxon (34). We summarized the percentage of environmental taxa (absent from the Human Oral Microbiome Database (HOMD)) with significantly different frequencies ($p < 0.001$) relative to the no treatment controls (Figure 4A), and the percentage of oral OTUs (present in HOMD) that increased relative to no treatment controls (Figure 4B). Of the published protocols, EDTA treatment reduced a smaller proportion of environmental taxa

compared to the combined UV and bleach protocol (55.6% compared to 63.6%). In both published protocols, the majority of oral taxa were increased relative to no treatment (85.0% for EDTA protocol and 64.8% for UV and bleach protocol). In contrast, UV treatment alone had limited impacts, reducing 53.8% of environmental taxa, and increasing only 28.9% of oral. Bleach treatment also reduced 53.8% of environmental taxa, but promoted the highest proportion of oral taxa (93.3%). However, these differences were non-significant when tested with a one-way ANOVA (Environmental: $p = 0.553$, and Oral: $p = 0.235$). Similarly, when OTUs were ranked as increasing or decreasing in each sample, relative to the untreated samples, the protocols did not have significantly different impacts (one-way ANOVA, Environmental: $p = 0.$ 0.178 and Oral: $p = 0.908$). Despite non-significance, the published protocols perform better than the individual UV or bleach treatments. However, while oral taxa were predominantly increased (except with UV treatment alone), some oral OTUs were reduced following each treatment, indicating that while decontamination is effective, it also impacts endogenous DNA.

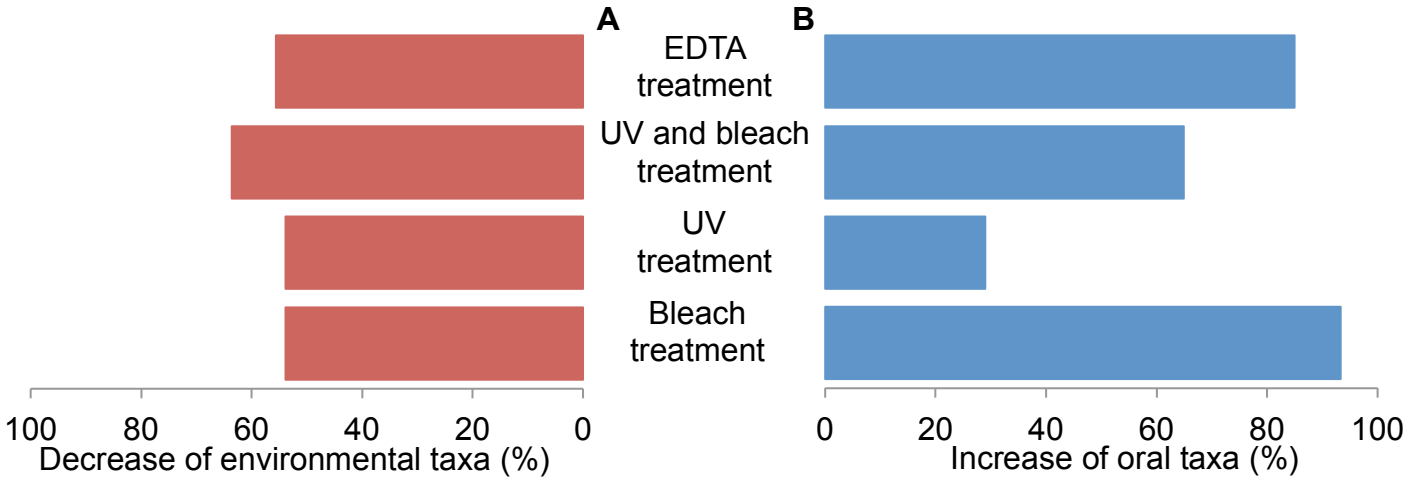


Figure 4: *Percentage change in environmental and oral taxa. Taxa with significantly different abundances from the no treatment controls were identified using the G-test. Of these significant taxa, we summarized the percentage of environmental taxa (absent from HOMD) that decreased relative to the no treatment controls (A), and the percentage of oral OTUs (present in HOMD) that increased relative to no treatment controls (B).*

Contaminant genera are primarily soil taxa

To identify the major sources of environmental contaminant OTUs in calculus samples, the likely habitats of all identified genera not present in the HOMD (12 of 44) were explored. Soil is the habitat of eight of the genera. The Archaeal genus, Candidtus nitrososphaera, is a common soil microorganism (35) and bacteria from soil types expected in the archaeological context (irrigated agricultural soil, landfill and freshwater sediments) were identified, namely Pseudonocardia, Paludibacter, Paenisporosarcina, Pedomicrobium, Propionivibrio, Steroidobacter, and DA101 (36–44). Schwartzia was also identified as an environmental contaminant, and is found in ruminants, particularly cows (45). Cattle markets within York may have resulted in this taxa being present environmentally (46). Several taxa are known to be present in both environmental and human microbiota, and include SHD-231 (ruminants and

human periodontal pockets) (47) and Hydrogenphaga (Daphnia gut and human disease) (48,49) and so may be endogenous or contaminant taxa. The final genera, TG5, which was not recorded in the HOMD has previously been reported in the human mouth (50), indicating disparity between methods of classifying oral taxa. Together, this suggests that soil is the primary source of contaminant DNA within ancient calculus samples, as all non-oral genera were primarily isolated from soils and sediments. Microorganisms from non-oral human body sites were not detected, suggesting minimal contamination from handling. However, other contributing factors, such as storage environment, should be assessed further.

Taxa profile of ethanol washes

DNA within the ethanol washes was sequenced to identify the taxa that were released from the calculus samples during the decontamination process. We assessed the potential of gaining insight into the environmental information preserved on the outer surface of calculus by examining the taxa present within the ethanol washes. Ethanol washes contained low diversity and had a limited numbers of reads, as expected (26). Not all samples could be successfully sequenced, resulting in only three datasets for the samples decontaminated with EDTA (Figure 2C). In total, 77 genera level OTUs were identified within the ethanol washes across samples, and 59 of these were classified as environmental taxa (i.e. not present in HOMD database). Of the 10.2 genera observed on average within the untreated samples, nearly half (47%) were environmental (4.8 genera). The largest proportion of environmental taxa was observed in the EDTA treatment group (n=3); eight of ten genera in ethanol washes following EDTA treatment were environmental. The ethanol washes following the combined UV and bleach treatment had fewer total genera than the untreated samples, and 33% were environmental (2.4 of 7.2 OTUs). Ethanol washes from UV treated samples contained more environmental genera than the untreated (51%; 6.4 of the 12.6 genera present). Ethanol used to wash bleach treated samples had 4.7

taxa, the lowest average of all groups (2.5 environmental genera (53.2%)). While these findings indicate that amplifiable DNA is being recovered from the dental calculus surface following decontamination procedures, only limited, stochastic taxa could be recovered, limiting the effectiveness of ethanol washes to monitor environmental taxa.

Discussion

Using archaeological calculus samples representative of the conditions present during ancient microbiota studies, the results of this study indicate that laboratory based decontamination treatments reduce detectable environmental contaminant taxa and improve oral taxa proportions within ancient samples. The previously published methods using a combined UV radiation and bleach submersion, or a predigestion in EDTA are more effective than single treatment with either UV radiation or bleach.

UV and bleach combined treatment and EDTA treatment are similarly effective decontamination methods

All of the decontamination methods resulted in fewer identifiable contaminant taxa and increased levels of oral taxa in comparison to untreated samples. Both the EDTA treatment and UV and bleach treatment showed a decrease in alpha diversity, and the G-test results show that both methods decrease environmental contaminant OTUs. However, the SourceTracker results indicated that the combined approach of UV and bleach treatment resulted in the lowest levels of known soil contaminant OTUs. Within this study, the average proportion of soil OTUs was higher in the EDTA treatment group. However, this is driven by a heavy load of contaminant OTUs in a single sample within the EDTA was group. Consequently, we note that the two treatments both appear to be effective at removing soil contaminant DNA from calculus samples. However, analysis of more samples, and decontamination analysis on samples from the same individual, should be performed to confirm this finding.

UV treatment and bleach treatment alone are not effective decontamination methods

The individual use of a UV treatment or bleach treatment did not perform comparably to the combined UV and bleach or the EDTA, despite the ability of both single treatments to decrease the level of contaminant OTUs compared to the untreated samples (SourceTracker and the G-test results). There differences between the two methods were likely due to their different coverage of the calculus surface and their actions on DNA. The limited effectiveness of UV treatment may have been due to the inability to irradiate the entire calculus surface. The ridged, three-dimensional surface of dental calculus creates pockets where environmental DNA may be shielded from the UV radiation. In addition, UV radiation is less efficient at destroying short DNA fragments, characteristic of ancient samples, compared to bleach. For example, the 16S rRNA V4 fragment amplified in this study is 289 base pairs (28), and fragments of this size are less susceptible to UV destruction than larger fragments (~700 base pairs) (51). In contrast, bleach treatment demonstrated the ability to decrease environmental taxa. However, despite the reduction in environmental OTUs, the total and oral OTU count increased, particularly in open reference datasets where de novo OTUs are incorporated. It is likely that bleach treatment is creating non-biological and novel DNA fragments by altering the DNA sequence through oxidative action. Low concentration of bleach can cause base modifications and create chlorinated base products, while only higher concentrations of bleach are capable of fully destroying DNA by breaking the phosphate bonds between nucleotides (52). Low concentrations of bleach have been traditionally used in decontamination protocols, because high bleach concentrations may damage the low, endogenous DNA content within the calculus samples. Nevertheless, bleach treatment alone at these concentrations appears to create novel OTUs. Further studies should assess different bleach concentrations and immersion times to explore the varying impacts, particularly as bleach is a recommended treatment for ancient samples (23). UV treatment can also cause base oxidation but this is only

one of several mechanisms that UV radiation can cause DNA destruction (51). Consequently, novel OTUs are likely being produced as a result of UV radiation. However, the proportion is likely lower due to the other impacts on DNA preventing amplification. The results here indicate that use of UV or bleach treatment alone should be avoided.

The combination of UV radiation and bleach immersion did not result in an increased alpha diversity. This is likely because the oxidative effects of bleach treatment are occurring, in part, on DNA fragments that have been crosslinked by the UV irradiation (51). Crosslinking of the DNA strands with surrounding molecules prevents the strand being read by a polymerase enzyme, thus preventing PCR amplification and sequencing. Even if the crosslinked DNA fragments experience sequence modification rather than destruction as a result of the bleach treatment, the novel, non-biological sequences are not available for DNA replication during the PCR. However, the combined use of these treatments does not eliminate the possibility that novel OTUs will be produced as a result of oxidative action. As newly generated DNA sequences are likely to be unique, the OTUs they form will be relatively rare within the dataset. Removal of rare OTUs (i.e. singletons), which is already a common requirement of statistical tests during OTU analysis, may largely reduce the impacts on analysis.

Soil and sediment are the main sources of contaminant OTUs

Environmental contaminant DNA can originate from a wide variety of sources, including soil, water, plant matter, decomposition of the body, archaeologists and museum curators, archaeologist tools, museum dust, etc. Our analysis identified that the matrix that surrounds the archaeological samples (e.g. soil) is the major source of DNA contaminating calculus samples. Of the 12 environmental genera removed by all decontamination protocols, we were able to identify eight of them from known soil or sediment sources. SourceTracker analysis also specifically matched the contaminant DNA within calculus samples to microorganisms within parkland and agriculture soils, which is consistent with the urban gravesites

where these samples were recovered. However, soil from different locations has very different microbial communities (53), and sediment from the actual archaeological site, when and if available, will provide the best comparison of contaminant DNA, rather than relying on soil microbial databases. If soil is not available from the site, then comparisons to these databases serve as the next best option to examine these effects.

Surprisingly, no human skin associated genera were observed in any of the samples. This may reflect particularly minimal handling during this specific excavation. The Jewbury site is a primary example of rescue archaeology, where very limited time was available to complete the survey and very minimal sample handling occurred (24). Further, the samples analysed in this study were not housed in a museum collection, which also likely limited their exposure to human skin. These specific circumstances may have limited exposure to human skin microorganisms, but future studies should continue to monitor for skin contaminant OTUs or unique taxa that could be introduced through additional sample exposure once the archaeological material is unearthed.

As this study was conducted with amplicon data (16S rRNA sequences), only genera level analysis was possible, which may have masked unique environmental contaminant DNA (54). For example, Actinomyces and Streptococcus are genera considered part of the oral bacterial community, and are included in HOMD. However, both of these genera contain multiple species that are commonly found in the environment (55,56) Consequently, 14 of the known 47 Actinomyces species are not included on HOMD, as are 73 of 116 known Streptococcus species. However, with genera level identification, the signal from environmental species cannot be detected and removed from the oral signal. Inference of biologically or culturally relevant patterns may then result from incorrect abundance measures. Shotgun sequencing approaches are preferable as they provide the ability to obtain detailed species and strain information, providing greater resolution of the microbial community within the sample

(34), and allow a more refined identification of contaminants. This approach will also provide data on the functional capacity of the community.

Decontamination of ancient samples is necessary

The combined UV and bleach treatment and EDTA treatment increased the sequencing depth of oral taxa by removing up to 9.1% of environmental contaminants within the sample. This increases data quality and reduces sequencing efforts and costs. Recent studies of dental calculus have applied shotgun sequencing approaches, rather than amplicon based strategies, as several significant biases have been detected in amplicon sequences of ancient microbiota analysis (17,57). However, there are currently few protocols, programs, or databases designed to handle shotgun metagenomic data sets or potential contamination introduced into them. Consequently, basic laboratory practice is still a key element of acquiring high quality data. However, variation still occurs between samples, as indicated by the highly contaminated sample in the EDTA treatment group, and application of a standard protocol does not ensure sample quality.

Conclusion

The susceptibility of ancient microbiota analyses to contaminant DNA is a major concern. While sterile working facilities and effective methods can reduce the influx of contaminating taxa during laboratory work, contaminant DNA gained during preservation, excavation, and storage cannot be accounted for by these methods. To combat this, effective decontamination prior to DNA extraction is critical. The removal of soil taxa and maintenance of oral taxa across samples leads us to suggest the combined UV and bleach treatment or EDTA treatment as suitable protocols to obtain robust, oral microbiota data from ancient dental calculus.

References

1. Rosenbaum M, Knight R, Leibel RL. The gut microbiota in human energy homeostasis and obesity. Trends Endocrinol Metab. 2015 Sep;26(9):493–501.

2. Trivedi B. Microbiome: The surface brigade. Nature. 2012;492(7429):2.

3. Li X, Wang J, Joiner A, Chang J. The remineralisation of enamel: a review of the literature. J Dent. 2014 Jun;42, Supplement 1:S12–20.

4. Noverr MC, Huffnagle GB. Does the microbiota regulate immune responses outside the gut? Trends Microbiol. 2004 Dec;12(12):562–8.

5. Ramezani A, Raj DS. The Gut Microbiome, Kidney Disease, and Targeted Interventions. J Am Soc Nephrol JASN. 2014 Apr;25(4):657–70.

6. Weyrich LS, Feaga HA, Park J, Muse SJ, Safi CY, Rolin OY, et al. Resident Microbiota Affect Bordetella pertussis Infectious Dose and Host Specificity. J Infect Dis [Internet]. 2013 Dec 15; Available from: http://www.ncbi.nlm.nih.gov/pubmed/24227794

7. Wang J, Qi J, Zhao H, He S, Zhang Y, Wei S, et al. Metagenomic sequencing reveals microbiota and its functional potential associated with periodontal disease. Sci Rep [Internet]. 2013 May 15 [cited 2016 Sep 6];3. Available from: http://www.nature.com/articles/srep01843

8. Azad MB, Konya T, Maughan H, Guttman DS, Field CJ, Sears MR, et al. Infant gut microbiota and the hygiene hypothesis of allergic disease: impact of household pets and siblings on microbiota composition and diversity. Allergy Asthma Clin Immunol. 2013;9(15):9.

9. Turnbaugh PJ, Backhed F, Fulton L, Gordon JI. Diet-induced obesity is linked to marked but reversible alterations in the mouse distal gut microbiome. Cell Host Microbe. 2008 Apr 17;3(4):213–23.

10. Foster JA, McVey Neufeld K-A. Gut–brain axis: how the microbiome influences anxiety and depression. Trends Neurosci. 2013 May;36(5):305–12.

11. David LA, Maurice CF, Carmody RN, Gootenberg DB, Button JE, Wolfe BE, et al. Diet rapidly and reproducibly alters the human gut microbiome. Nature [Internet]. 2013 Dec 11; Available from: http://www.ncbi.nlm.nih.gov/pubmed/24336217

12. Benson AK, Kelly SA, Legge R, Ma F, Low SJ, Kim J, et al. Individuality in gut microbiota composition is a complex polygenic trait shaped by multiple environmental and host genetic factors. Proc Natl Acad Sci. 2010 Nov 2;107(44):18933–8.

13. Jakobsson HE, Jernberg C, Andersson AF, Sjölund-Karlsson M, Jansson JK, Engstrand L. Short-Term Antibiotic Treatment Has Differing Long-Term Impacts on the Human Throat and Gut Microbiome. PLOS ONE. 2010 Mar 24;5(3):e9836.

14. Weyrich LS, Dobney K, Cooper A. Ancient DNA analysis of dental calculus. J Hum Evol. 2015 Feb;79:119–24.

15. White DJ. Dental calculus: recent insights into occurrence, formation, prevention, removal and oral health effects of supragingival and subgingival deposits. Eur J Oral Sci. 1997 Oct 1;105(5):508–22.

16. Adler CJ, Dobney K, Weyrich LS, Kaidonis J, Walker AW, Haak W, et al. Sequencing ancient calcified dental plaque shows changes in oral microbiota with dietary shifts of the Neolithic and Industrial revolutions. Nat Genet. 2013 Apr;45(4):450–5.

17. Warinner C, Rodrigues JFM, Vyas R, Trachsel C, Shved N, Grossmann J, et al. Pathogens and host immunity in the ancient human oral cavity. Nat Genet. 2014;

18. Dabney J, Meyer M, Paabo S. Ancient DNA damage. Cold Spring Harb Perspect Biol [Internet]. 2013 Jul;5(7). Available from: http://www.ncbi.nlm.nih.gov/pubmed/23729639

19. Weiss S, Amir A, Hyde ER, Metcalf JL, Song SJ, Knight R. Tracking down the sources of experimental contamination in microbiome studies. Genome Biol. 2014 Dec 17;15(12):1–3.

20. Salter SJ, Cox MJ, Turek EM, Calus ST, Cookson WO, Moffatt MF, et al. Reagent and laboratory contamination can critically impact sequence-based microbiome analyses. BMC Biol. 2014 Nov 12;12(1):87.

21. Cooper A, Poinar HN. Ancient DNA: Do It Right or Not at All. Science. 2000 Aug 18;289(5482):1139–1139.

22. Fulton T. Setting Up an Ancient DNA Laboratory. In: Shapiro B, Hofreiter M, editors. Ancient DNA [Internet]. Humana Press; 2012 [cited 2016 Sep 2]. p. 1–11. (Methods in Molecular Biology). Available from: http://dx.doi.org/10.1007/978-1-61779-516-9_1

23. Boessenkool S, Hanghøj K, Nistelberger HM, Der Sarkissian C, Gondek A, Orlando L, et al. Combining bleach and mild pre-digestion improves ancient DNA recovery from bones. Mol Ecol Resour. :n/a-n/a.

24. Lilley JM, Stroud G, Brothwell D, Williamson MH. The Jewish Burial Ground at Jewbury. Council for British Archaeology; 1994. 578 p. (The Archaeology of York: The Medieval Cemeteries; vol. 12, Fascicule 3).

25. McComish J. The Medieval Jewish Cemetery at Jewbury, York. Jew Cult Hist. 2000 Dec 1;3(2):21–30.

26. Shokralla S, Singer GAC, Hajibabaei M. Direct PCR amplification and sequencing of specimens' DNA from preservative ethanol. Benchmarks. 2010 Mar;48(3):305–6.

27. Brotherton P, Haak W, Templeton J, Brandt G, Soubrier J, Adler CJ, Richards SM, Der Sarkissian C, Ganslmeier R, Friederich S, Dresely V, van Oven M, Kenyon R, Van der Hoek M, Korlach J, Luong K, Ho SYW, Quintana-Murci L, Behar DM, Meller H, Alt KW, Cooper A & The Genographic Project. Neolithic mitochondrial haplogroup H genomes and the genetic origins of Europeans. Nat Commun. Accepted 02/2013.

28. Caporaso JG, Lauber CL, Walters WA, Berg-Lyons D, Huntley J, Fierer N, et al. Ultra-high-throughput microbial community analysis on the Illumina HiSeq and MiSeq platforms. ISME J. 2012 Aug;6(8):1621–4.

29. Caporaso JG, Kuczynski J, Stombaugh J, Bittinger K, Bushman FD, Costello EK, et al. QIIME allows analysis of high-throughput community sequencing data. Nat Methods. 2010 May;7(5):335–6.

30. DeSantis TZ, Hugenholtz P, Larsen N, Rojas M, Brodie EL, Keller K, et al. Greengenes, a Chimera-Checked 16S rRNA Gene Database and Workbench Compatible with ARB. Appl Environ Microbiol. 2006 Jul 1;72(7):5069–72.

31. Fierer N, Lauber CL, Zhou N, McDonald D, Costello EK, Knight R. Forensic identification using skin bacterial communities. Proc Natl Acad Sci. 2010 Apr 6;107(14):6477–81.

32. Knights D, Kuczynski J, Charlson ES, Zaneveld J, Mozer MC, Collman RG, et al. Bayesian community-wide culture-independent microbial source tracking. Nat Methods. 2011 Sep;8(9):761–3.

33. R Core Team. R: A language and environment for statistical computing. R Found Stat Comput Vienna Austria [Internet]. 2014; Available from: http://www.R-project.org/

34. Ziesemer KA, Mann AE, Sankaranarayanan K, Schroeder H, Ozga AT, Brandt BW, et al. Intrinsic challenges in ancient microbiome

reconstruction using 16S rRNA gene amplification. Sci Rep. 2015 Nov 13;5:16498.

35. Swanson CA, Sliwinski MK. Archaeal Assemblages Inhabiting Temperate Mixed Forest Soil Fluctuate in Taxon Composition and Spatial Distribution over Time. Archaea [Internet]. 2013 [cited 2016 Aug 20];2013. Available from: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3747363/

36. GEBERS R. Enrichment, Isolation, and Emended Description of Pedomicrobium ferrugineum Aristovskaya and Pedomicrobium manganicum Aristovskaya. Int J Syst Evol Microbiol. 1981;31(3):302–16.

37. Gebers R, Beese M. Pedomicrobium americanum sp. nov. and Pedomicrobium australicum sp. nov. from Aquatic Habitats, Pedomicrobium gen. emend., and Pedomicrobium ferrugineum sp. emend.†. Int J Syst Evol Microbiol. 1988;38(3):303–15.

38. Ueki A, Akasaka H, Suzuki D, Ueki K. Paludibacter propionicigenes gen. nov., sp. nov., a novel strictly anaerobic, Gram-negative, propionate-producing bacterium isolated from plant residue in irrigated rice-field soil in Japan. Int J Syst Evol Microbiol. 2006;56(1):39–44.

39. Reddy GSN, Manasa BP, Singh SK, Shivaji S. Paenisporosarcinaindica sp. nov., a psychrophilic bacterium from a glacier, and reclassification of Sporosarcina antarctica Yu et al., 2008 as Paenisporosarcinaantarctica comb. nov. and emended description of the genus Paenisporosarcina. Int J Syst Evol Microbiol. 2013;63(8):2927–33.

40. Krishnamurthi S, Bhattacharya A, Mayilraj S, Saha P, Schumann P, Chakrabarti T. Description of Paenisporosarcina quisquiliarum gen. nov., sp. nov., and reclassification of Sporosarcina macmurdoensis Reddy et al. 2003 as Paenisporosarcina macmurdoensis comb. nov. Int J Syst Evol Microbiol. 2009;59(6):1364–70.

41. Brune A, Ludwig W, Schink B. Propionivibrio limicola sp. nov., a fermentative bacterium specialized in the degradation of hydroaromatic compounds, reclassification of Propionibacter pelophilus as Propionivibrio pelophilus comb. nov. and amended description of the genus Propionivibrio. Int J Syst Evol Microbiol. 2002;52(2):441–4.

42. WARWICK S, BOWEN T, McVEIGH H, EMBLEY TM. A Phylogenetic Analysis of the Family Pseudonocardiaceae and the Genera Actinokineospora and Saccharothrix with 16S rRNA Sequences and a Proposal To Combine the Genera Amycolata and Pseudonocardia in an Emended Genus Pseudonocardia. Int J Syst Evol Microbiol. 1994;44(2):293–9.

43. Sakai M, Hosoda A, Ogura K, Ikenaga M. The Growth of Steroidobacter agariperforans sp. nov., a Novel Agar-Degrading Bacterium Isolated from Soil, is Enhanced by the Diffusible Metabolites Produced by Bacteria Belonging to Rhizobiales. Microbes Environ. 2014;29(1):89–95.

44. Fahrbach M, Kuever J, Remesch M, Huber BE, Kämpfer P, Dott W, et al. Steroidobacter denitrificans gen. nov., sp. nov., a steroidal hormone-degrading gammaproteobacterium. Int J Syst Evol Microbiol. 2008;58(9):2215–23.

45. van GYLSWYK NO, HIPPE H, RAINEY FA. Schwartzia succinivorans gen. nov., sp. nov., Another Ruminal Bacterium Utilizing Succinate as the Sole Energy Source. Int J Syst Evol Microbiol. 1997;47(1):155–9.

46. Richardson H. The Medieval Fairs and Markets of York. Borthwick Publications; 1961. 44 p.

47. Cohen CL. Technological advancements in microbial analyses of periodontitis patients: Focus on illumina® sequencing using the Miseq system on the 16s rrna gene: A clinical and microbial study [Internet] [M.S.]. [United States -- California]: University of Southern California;

2013 [cited 2016 Jul 13]. Available from: http://search.proquest.com/docview/1459428381/abstract/DE2DBCF305BC43AAPQ/1

48. Callens M, Macke E, Muylaert K, Bossier P, Lievens B, Waud M, et al. Food availability affects the strength of mutualistic host–microbiota interactions in Daphnia magna. ISME J. 2016 Apr;10(4):911–20.

49. Hogan DA, Willger SD, Dolben EL, Hampton TH, Stanton BA, Morrison HG, et al. Analysis of Lung Microbiota in Bronchoalveolar Lavage, Protected Brush and Sputum Samples from Subjects with Mild-To-Moderate Cystic Fibrosis Lung Disease. PLOS ONE. 2016 Mar 4;11(3):e0149998.

50. Vartoukian SR, Palmer RM, Wade WG. Cultivation of a Synergistetes strain representing a previously uncultivated lineage. Environ Microbiol. 2010 Apr;12(4):916–28.

51. Sarkar G, Sommer SS. [28] Removal of DNA contamination in polymerase chain reaction reagents by ultraviolet irradiation. In: Wu R, editor. Methods in Enzymology [Internet]. Academic Press; 1993 [cited 2016 Jul 14]. p. 381–8. (Recombinant DNA Part I; vol. 218). Available from: http://www.sciencedirect.com/science/article/pii/007668799318030G

52. Kemp BM, Smith DG. Use of bleach to eliminate contaminating DNA from the surface of bones and teeth. Forensic Sci Int. 2005 Nov 10;154(1):53–61.

53. Young JM, Weyrich LS, Breen J, Macdonald LM, Cooper A. Predicting the origin of soil evidence: High throughput eukaryote sequencing and MIR spectroscopy applied to a crime scene scenario. Forensic Sci Int. 2015 Jun;251:22–31.

54. Poretsky R, Rodriguez-R LM, Luo C, Tsementzi D, Konstantinidis KT. Strengths and Limitations of 16S rRNA Gene Amplicon Sequencing in Revealing Temporal Microbial Community Dynamics. PLoS ONE [Internet]. 2014 Apr 8 [cited 2016 Sep 7];9(4). Available from: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3979728/

55. Lin Z, Zhen Z, Wu Z, Yang J, Zhong L, Hu H, et al. The impact on the soil microbial community and enzyme activity of two earthworm species during the bioremediation of pentachlorophenol-contaminated soils. J Hazard Mater. 2016 Jan 15;301:35–45.

56. Joergensen RG, Küntzel H, Scheu S, Seitz D. Movement of faecal indicator organisms in earthworm channels under a loamy arable and grassland soil. Appl Soil Ecol. 1998 May;8(1–3):1–10.

57. Weyrich LS, Duchane S, Soubrier J, Arriola L, Llamas B, Breen J, et al. Reconstructing Neandertal behavior, diet, and disease using ancient DNA from dental calculus. Nature. 2016;

Chapter III

Diet driven differentiation of oral microbiota in ancient Britain

Statement of Authorship

Title of Paper	Unappreciated microbiota clade in ancient Britain linked to complex dietary drivers
Publication Status	Unpublished and Unsubmitted work written in manuscript style
Publication Details	In preparation for: Nature Communications

Principal Author

Name of Principal Author (Candidate)	Andrew G. Farrer		
Contribution to the Paper	Performed decontamination protocols, DNA extraction, PCR amplification, and sequencing library preparation. Analysed the data and interpreted the results. Wrote the manuscript.		
Overall percentage (%)	75%		
Certification:	This paper reports on original research I conducted during the period of my Higher Degree by Research candidature and is not subject to any obligations or contractual agreements with a third party that would constrain its inclusion in this thesis. I am the primary author of this paper.		
Signature		Date	6th October 2016

Co-Author Contributions

By signing the Statement of Authorship, each author certifies that:

i. the candidate's stated contribution to the publication is accurate (as detailed above);

ii. permission is granted for the candidate in include the publication in the thesis; and

iii. the sum of all co-author contributions is equal to 100% less the candidate's stated contribution.

Name of Co-Author	Laura S. Weyrich		
Contribution to the Paper	Advised on data analysis and interpretation of results, and discussed and edited the manuscript		
Signature		Date	7th October 2016

Name of Co-Author	Neville Gully		
Contribution to the Paper	Discussed the results and edited the manuscript		
Signature		Date	10th October 2016

Name of Co-Author	Keith Dobney		
Contribution to the Paper	Discussed the results and edited the manuscript		
Signature		Date	10th October 2016

Name of Co-Author	Alan Cooper		
Contribution to the Paper	Discussed the results and edited the manuscript		
Signature		Date	11th October 2016

Diet driven differentiation of oral microbiota in ancient Britain

A. G. Farrer[1*], K. Dobney[2], N. Gully[3], L. S. Weyrich[1+] and A. Cooper[1+]

Affiliations:

[1]Australian Centre for Ancient DNA, School of Biological Sciences, University of Adelaide, Adelaide, South Australia, Australia

[2]Department of Archaeology, Classics and Egyptology, School of Histories, Languages and Cultures, University of Liverpool, Liverpool, United Kingdom

[3]School of Dentistry, Faculty of Health Sciences, University of Adelaide, Adelaide, South Australia, Australia

[+]These authors contributed equally to this work

Corresponding author:

[*]andrew.farrer@adelaide.edu.au

Abstract

Modern oral microbiota can trigger severe oral and systemic disease. Developing medical treatments to alter the microbiota and combat these diseases requires identification of the environmental factors that drive microbiota structure in everyday life. Ancient DNA studies of dental calculus allow direct analysis of historical oral microbiota, providing insight into the key drivers of microbiota structure through time. We demonstrate that a complex of dietary factors differentiated the historical populations of Great Britain into two major groups using functional analysis of dental calculus samples from across England and Scotland. We suggest that these groups represent socio-economic differences based on the access to high status foods such as meat and dairy products. These results demonstrate potential archaeological applications to study multiple dietary components in ancient populations and suggest that altering available nutrients may be a simple but powerful method of altering oral microbiota in modern populations.

Today, the oral microbiota of industrialized populations are associated with both local and systemic disease[1]. Oral microbes have been linked to oral disease, such as dental caries and periodontitis[2], and to wider health impacts including cardiovascular disease[3], cancer[4], and mental health[5]. Microbiota across the body are known to be defined and altered by environmental pressures[6]. Modern studies have indicated that microbiota can be significantly altered by the living environment, interactions between people, disease, and diet[7–11]. Consequently, it is critical to understand the diversity, specific species, and functions of oral microbiota at both the individual and the population level to develop medical treatments for microbiota related diseases and to identify the aspects of the living environment that define these properties.

Comparisons of microbiota in industrialized and non-industrialized groups have been used to identify differences between modern, industrialized life and the assumed ancestral state with the aim of identifying the cultural and environmental factors that have resulted in modern, disease associated microbiota[12–14]. These comparative studies have revealed that non-industrialized groups typically do not suffer from the same local or systemic disease loads as industrialized peoples, and that significant variation exists between gut microbiota from different non-industrialized peoples[12], which is likely explained by the unique cultural and environmental histories of each group. As different modern, non-industrialized groups have unique microbiota, it is unclear which, if any, are representative of the ancestral state of industrialized peoples. Consequently, the direct study of historical populations is necessary to specifically identify the historical factors that have defined and altered the microbiota in individual populations. This information is critical to identify which modern cultural behaviours might be involved in altering microbiota and leading to disease.

Ancient DNA studies have the ability to examine historical microbiota and test hypotheses about the impacts of changes in environmental and cultural factors through time. Oral microbiota are preserved in calcified dental plaque deposits (calculus)[15] which are routinely found on archaeological human

remains[16] and can be removed without damaging the tooth surface. Studies examining ancient bacterial DNA trapped within calculus in European specimens have identified that both the agricultural (Neolithic) revolution (~8,000 BCE) and the industrial revolution (~1750 CE) were associated with major changes in the oral microbiota which are likely linked to increases in the overall intake of dietary carbohydrates[17]. However, the individual and population level structure of oral microbiota directly prior to these revolutions and the lifestyle factors that defined them, remain unknown. Changes in the historical environment potentially exposed people to densely populated cities, poor waste removal, human movement and trade while impacting dietary access and exposure to diseases, each of which may have defined the microbiota of individuals and, potentially, the population[18].

Diet is a major driving factor of microbiota diversity[19], and dietary changes impact microbiotia composition and function at both the population and individual level. For example, the community structure of gut microbiota has been observed to alter within 24 hours of a change in diet (i.e. high protein and fat to high fibre) indicating the large impact of nutrient availability[11]. Furthermore, multiple studies have shown that dietary groups can be differentiated by microbial function; for example, carnivorous or herbivorous mammals can be separated by key microbial functions associated with amino acid metabolism[19]. The humanization of diet (i.e. a reduction in fibre) from wild to captive non-human primates is also associated with a reduction in the microbial fibre-digestion pathways[20].

Given that diet will have been a key driver of human microbiota through history[17], it is likely that changes in microbial functions linked to diet (i.e. meat, fibre, and carbohydrate consumption) will be observable in historical oral microbiota and indicate altered dietary patterns between and during the agricultural and industrial revolutions. However, amino acid, fibre, and carbohydrate metabolism of microbiota have not yet been explored within any ancient population. Therefore, detailed studies of ancient microbiota and their functions from different populations must be undertaken to understand

how cultural and dietary factors may have impacted the microbiota through time.

Here, we analyse a detailed transect of ancient dental calculus samples from Great Britain to identify specific factors that altered microbiota prior to industrialization. Sixty-two samples were collected from Pre-Roman (prior to 43 CE) to Medieval (~1700 CE) periods from twelve archaeological sites across England and Scotland, including: Breedon on the Hill (Leicestershire, Anglo-Saxon), Hinxton (Cambridgeshire, Iron Age and Anglo-Saxon), Jewbury (York, Medieval), Kirk Hill (Fife), Kirkwall Cathedral (Orkney, Viking), Linton (Cambridgeshire, Anglo-Saxon), Newark Bay (Orkney, Viking), Oakington (Cambridgeshire, Anglo-Saxon), Raunds (Northamptonshire, Anglo-Saxon), St. Brides (London, post-Medieval), St. Helens-on-Walls (York, Medieval), St Ninnians Isle (Orkney, Viking), Bronze Age Yorkshire, and Pictish (Scotland). These samples have thorough archaeological reports that provide detailed information on culture, paleopathology, age, and sex, allowing us to interrogate the cultural and dietary changes that influenced microbiota. This data set provides the first detailed resolution of microbial species and function in an ancient human population.

Results

Robust oral microbiota are retrieved from ancient British specimens

Careful consideration was given to the risk of laboratory and environmental contamination, as endogenous signals can easily be obscured or misinterpreted due to contaminating microbial DNA[21–24]. To minimize contamination, samples were processed in an ultra-sterile, specialized ancient DNA laboratory and underwent a decontamination protocol to remove environmental contamination present on the outer surface (i.e. UV radiation for 30 minutes on each side followed by a three-minute submersion in 5% bleach)[17]. Optimized silica-based ancient DNA extraction and metagenomic library preparation protocols were applied[17,25] and negative controls were incorporated throughout to monitor laboratory and reagent contamination. We also obtained modern plaque (n = 5)[26] and dental calculus (n = 1)[27] data for comparison.

Ancient DNA samples contain low amounts of endogenous DNA (< 1%)[28]. Consequently, dental calculus samples may not contain sufficient preserved DNA to provide a representation of the ancient oral microbiota. In addition, when handling such low biomass samples, a large proportion of the DNA recovered may come from contamination by the low concentration of microbial DNA present within “sterile” reagents and laboratories[22]. To minimize the risk of misinterpretation due to DNA degradation and contaminating DNA, we applied four assessment procedures to identify high quality samples and remove potential contaminant DNA in the dataset. First, Bray-Curtis pairwise dissimilarity was calculated between all samples (n = 62) and negative controls (n = 28). Of the negative controls, 21 had an average Bray-Curtis dissimilarity of < 0.6 when compared to each other. The minimal variation indicated that these controls were a consistent representation of the laboratory environment (i.e. were not cross-contaminated with samples). Dental calculus samples were excluded if they lacked endogenous signal (i.e. had a dissimilarity less than four standard deviations from the 21 laboratory representative negative controls). This procedure only excluded two samples, reflecting the robust signal derived

from dental calculus specimens. Secondly, any species detected within the negative controls was removed from all calculus samples[27]. This removed 26.2% of all species level identifications from the data. Thirdly, we normalized the data set to remove samples with fewer than 100,000 reads resulting in a normalized dataset with 129,760 reads per sample. Fourthly, we examined whether there were any correlations in the data with obvious potential confounding factors to test for bias introduced by experimental procedure. Data were not significantly correlated with sampling site within the mouth (upper or lower jaw, left or right side of the mouth, and tooth type) or laboratory handling (extraction, library, and sequencing groups) (ANOSIM, $R < 0.3$ and $p > 0.05$, Table S1). However, oral microbiota are known to vary with oral geography[29]; therefore, only samples from molar teeth were analysed to limit any potential bias within the ancient data. Ultimately, 33 ancient calculus samples, representing 10 archaeological sites from the Pre-Roman to Medieval periods, and six published modern samples were deemed robust and analysed further.

Three distinct microbiota groups are observed in ancient Britain

We first assessed the similarities and differences between the ancient and modern samples. Bray Curtis dissimilarities of the rarefied data were UPGMA clustered, revealing three statistically different groups (Figure 1; ANOSIM, $R = 0.785$ $p = 0.001$, Table S1). Two groups consisted exclusively of ancient samples (Group 1 (n=12) and Group 2 (n=20)). These clusters did not have significantly different diversities (observed species, Figure 2a) and had no correlation with time (ANOSIM, $R < 0.3$, $p > 0.05$, Table S1). However, the dominant primary colonising microorganism (i.e. an oral species that binds the tooth surface[6]) was distinct within each of the two ancient groups. Streptococcus species were the dominant primary colonizer in Group 1, while Methanobrevibacter species were dominant in Group 2b (Figure 2). These differences in primary colonizers were also linked to community level changes in the microbiota (Kruskall Wallis, $p < 0.05$. Table S2). 35 taxa had increased abundance in Group 1 relative to Group 2, while

68 taxa were increased in Group 2 relative to Group 1. This demonstrates that the two distinct groups consisted of different microbial communities.

Interestingly, modern samples clustered together with a single ancient sample (St. Helens on the Walls 2, York, 1100 – 1550 CE) in Group 3 (Modern). This group of generally modern samples was significantly different from both ancient groups (ANOSIM, Modern vs. Group 1: $R=0.797$, $p=0.001$; Modern vs. Group 2: $R=0.989$, $p=0.001$, Table S1), even though Streptococcus species was also the primary colonizer of samples within the Modern group, similar to Group 1. Indeed, samples within the Modern Group showed no significantly different taxa with Group 1 (Kruskall Wallis, $p < 0.05$, Table S3a), while 57 taxa were significantly different from Group 2 (Kruskall Wallis, $p < 0.05$, Table S3b). The similarities between the Modern Group and Group 1 suggest that Group 1 may be the precursor to modern microbiota in Great Britain.

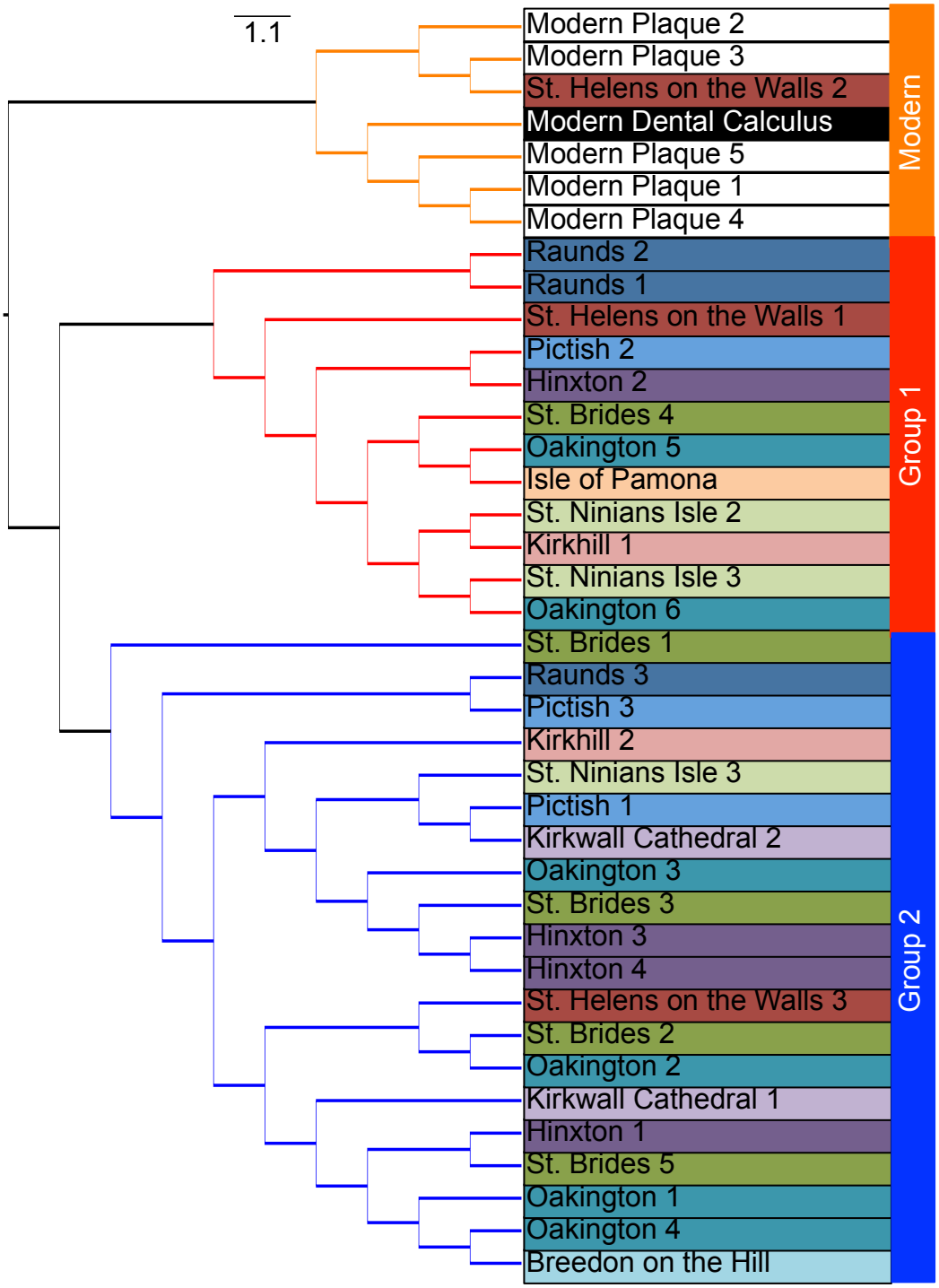


Figure 1: *Ancient dental calculus samples clustered using the UPGMA method based on Bray-Curtis dissimilarity. Samples fall into three major groups as denoted on the right hand bar (Modern – Orange, Group 1 – Red, Group 2 – Blue). Individual samples a coloured according to archaeological site, which do not associate with the three groups.*

Location and disease do not define ancient microbiota diversity

We assessed archaeological, palaeopathological, and functional information to search for factors that may have driven the formation of the distinctive groups of oral microbiota within ancient Britain. Importantly, the samples did not correlate with archaeological sites or with broader geographical or cultural classifications (Region, Country, or Urban vs. Rural groups (ANOSIM, $R < 0.3$ and $p > 0.05$, Table S1)), which were patterns that had previously been observed in other datasets[17]. This would suggest that the clusters observed here are based on individualistic factors linked to a person's behaviour rather than factors that exist at the population level. Therefore, we tested personal factors such as sex and age estimates. However, these factors also did not explain the groupings (ANOSIM, $R < 0.3$ and $p > 0.05$, Table S1).

Next, we investigated the impacts of disease by correlating the 3 oral microbiota groups with palaeopathological data, where available. While no larger significant trends were identified, patterns within single archaeological sites were suggestive of trends. For example, children from the Anglo-Saxon site at Oakington ($n = 2$) were within Streptococcus-dominated Group 1, while the adults ($n = 4$) from the same site fell within Methanobrevibacter-dominated Group 2 (Figure 1a, Turquoise). Dental calculus is unusual in children and normally only forms from the late teens[30] onwards, suggesting that disease may have contributed to the formation of microbiota within Group 1. Similarly, Group 1 also contained a diseased male from St. Brides Church suffering from an abnormal growth in the mouth and a benign bony swelling (torus palatinus), London (Figure 1a, Green). However, the lack of disease related metadata from all individuals inhibited our ability to test disease related links throughout Group 1.

We next examined the link between the groups and the presence of known human pathogens preserved within the calculus as a marker of oral and systemic disease[27] (Table S4). The Red Complex of oral pathogens (*Porphymonas gingivalis*, *Treponema denticola*, and *Tannerella forsynthia*), which is associated with severe periodontal disease[31], was present in all

ancient samples (Table S4). In addition, the two well-known oral pathogens *Streptococcus mutans* and *Fusobacterium nucleatum* were present (n = 23 and n = 33, respectively) although they did not differentiate the two clusters. However, a single non-oral pathogen (*Clostridium botulinum*, a food-borne pathogen and environmental microbe[32]) was significantly increased in Group 2 (Kruskall-Wallis, $p < 0.05$, Table S1) and was the only known pathogen statistically linked to a group. However, the low proportion of this pathogen suggests that it was probably not driving overall microbiota diversity (present in 27 samples, read count < 70 per sample). However, its presence does suggest that historical individuals were regularly exposed to *C. botulinum*, which could have been a healthy part of their microbiota acquired from food sources or through environmental exposure. Interestingly, primary food vectors at this time would have been fermented, salted, or smoked fish and meat products. Consequently, to be increased within Group 2 is unexpected but may be as a result of increased environmental exposure due to the social status of individuals in Group 2. It is unlikely that this is environmental contamination as it was preferentially observed in one group over another across a mixture of archaeological sites. Despite these small-scale observations with paleopathology and *C. botulinum*, large trends in oral health and systemic disease do not appear to drive the different groups of microbiota in ancient Britain.

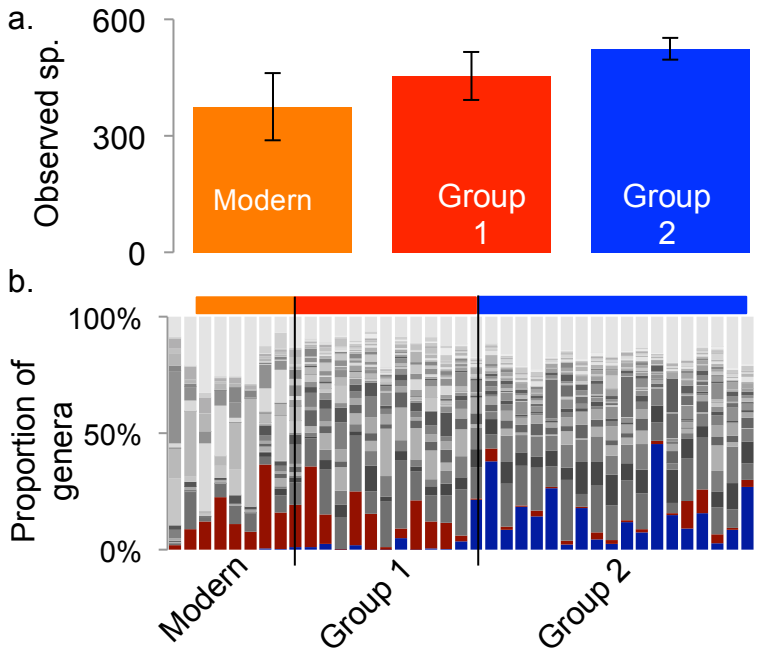


Figure 2: *a. Mean alpha diversity (observed species) and one standard deviation of each group when all taxonomic levels are rarefied to 120,000 reads. b. Indication of the two major primary colonizers in the three groups. Stacked bar chart of genera with Streptococcus species (Red) and Methanobrevibacter species (Blue).*

Meat consumption is more prevalent in the Streptococcus-dominated Group 1

While no eukaryotic sequences corresponding to dietary food sources were identified within our low-coverage data, bacterial community structure and functional differences can divulge large-scale dietary differences[19,20,27,33]. Eukaryotic sequences were examined. However, the data available in low-coverage data was insufficient to compare between samples. Therefore, we compared previously published dental calculus samples with associated dietary information to our data; these published datasets include Neandertals with either meat eating or foraging habits, a foraging chimpanzee[27], and meat eating medieval European farmers[34] (Figure 3a). We examined similarities in species between these published samples and Group 1 and Group 2 individuals using SourceTracker (Figure 3b). Using our data as

references, Bayesian modelling estimated the proportion of microbiota diversity from the published samples with known diets as coming from Group 1, Group 2, or an unknown source. Individuals that had evidence for high meat consumption (Spy II Neandertal and the Medieval farming samples) contained more taxa from Group 1, while individuals evidenced to have lower meat consumption (the El Sidron 1 Neandertal and chimpanzee)[27], had the highest proportions of Group 2 taxa. However, the El Sidron 2 Neandertal, previously designated as a low meat eating forager, had mostly Group 1 microbes (96%). Overall, these patterns suggest that the bacteria within Group 2 might be typically indicative of a diet that contains little to no meat while Group 1 members had meat in their diet.

We examined this link to meat-eating by exploring amino acid metabolism functions within the metagenomic diversity, as specific amino acid metabolism pathways are linked to either carnivore or herbivore diets[19]. Within our dataset, we identified 84.6% (33 of 39) of the amino acid metabolism pathways previously identified within mammalian microbiota as indicators of high or low meat consumption. 70 different individual functions were identified within these 33 pathways. Of these, 24.3% (17 functions) were significantly different between the three groups (Kruskal-Wallis, $p < 0.05$) and had a mean abundance > 10 reads in one of the three groups (Figure 4, Table S5). All of the amino acid functions linked with meat eating (n=4) were enriched in Group 1 and the Modern Group. The Streptococcus-enriched Group 1 and Modern Group also contained four of the 13 functions associated with plant eating, suggesting that these individuals were eating both meat and plants. This reflects the dietary inferences previously generated for the Spy II Neandertal and the medieval farming individuals[27]. In contrast, only functions associated with plant eating were enriched in Group 2, suggesting a primarily plant-based diet. These data indicate that different dietary strategies linked to the consumption of meat appear to have led to differentiation of ancient oral microbiota in Britain.

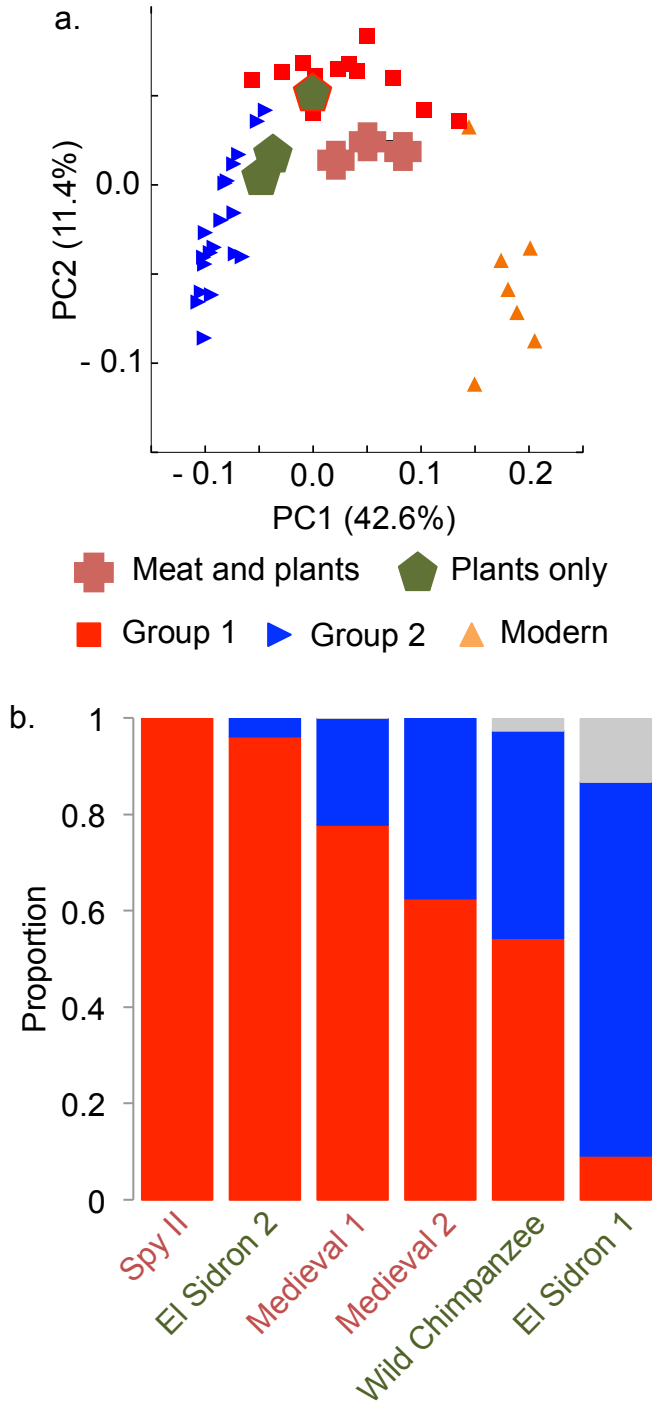


Figure 3: *a. PCoA plot of three groups with previously published data. Individuals previously identified as meat and plant eaters fall within Group 1 and exclusive foragers within Group 2. El Sidron 2 Neandertal (Red border) unexpectedly falls within Group 1. b. Proportion of previously published samples that match the ancient Group 1 or Group 2 communities.*

High fibre and carbohydrate consumption delineates ancient oral microbiota groups

To examine the level of plant consumption between the groups, we explored metabolic functions associated with high and low dietary fibre. We identified 22 of 23 (97.5%) previously recognized marker functions for fibre digestion, such as butanoate metabolism and lipopolysaccharide biosynthesis [20], and 68.2% (15 functions) (Kruskal-Wallis, $p < 0.05$, mean abundance > 10 reads in one of the three groups) (Figure 4, Table S5). All the genes associated with a high fibre diet (n=5) and few of those linked to low-fibre (3 of 10) were in greater abundance within the Methanobrevibacter-enriched Group 2. The higher archaeal content of Group 2 samples further suggests a herbivorous diet[19]. This indicates that individuals within Group 2 consumed a higher fibre, plant-based diet than the two Streptococcus dominated groups (Group 1 and Modern). Individuals within the Modern group also possessed an increase in genes indicative of low fibre in comparison to Group 1, reflecting a further decrease in high-fibre food sources in the modern diet.

To further examine the low-fibrous dietary food sources we also assessed carbohydrate metabolism. Carbohydrates have been shown to be an important driver of ancient oral microbiota diversity[17]. We examined carbohydrate metabolism functions identified within the samples from Level 4 of the SEED database of microbial function, providing insight into gene pathways associated with carbohydrates such as maltose and sucrose. 87 carbohydrate functions were present in our dataset, and 31 (35.6%) significantly varied between the three groups (Kruskal Wallace $p < 0.05$). Significantly higher levels of sugar metabolism were identified in meat eating, low-fibre Group 1 and Modern Group individuals including sequences associated with fructose, sucrose, trehalose, mannose, beta-glucosides, and maltose metabolism (Figure 4, Table S5). These sugars are common dietary sugars and are found in plants, particularly fruits and vegetables[35], indicating a diet rich in “high quality” plant material. In contrast, Group 2 was enriched for metabolic functions that use other carbon sources for energy,

such as formaldehyde and pyruvate (Figure 4, Table S5). Functions linked to the production of methane (methanogenesis) were also observed in Group 2, as expected given the presence of the methanogenic Methanobrevibacter species (Table S5). These functions suggest a limited availability of carbohydrates in the diet. Overall, this suggests that Streptococcus-dominated Group 1 and Modern individuals consumed more dietary carbohydrates than Group 2.

Metabolism of dairy products was different between ancient groups

We also analysed sugars present in milk as disparate consumption of milk in England has been previously identified in dental calculus research[36]. Gene groups involved in the utilization of the dairy sugars, lactose and galactose, were enriched in the Streptococcus-dominated Group 1 and Modern Group individuals. In contrast, Group 2 lacked the microbial functions linked to milk digestion (Figure 4, Table S5). One third of individuals from medieval England had milk proteins present in their dental calculus[36]. Here, Group 1, whose members have microbiota enriched for diary digestion, also represent one third of the ancient individuals. Concordance with protein analysis suggests that microbiota functional analysis is providing accurate indications of historical diet. This functional analysis suggests that differences in milk consumption contribute to the different microbiota observed in ancient Britain. Together, these patterns indicate that variation in meat, carbohydrate, and dairy consumption shifted ancient oral microbiota and may have selected for the modern industrialized microbiota observed today.

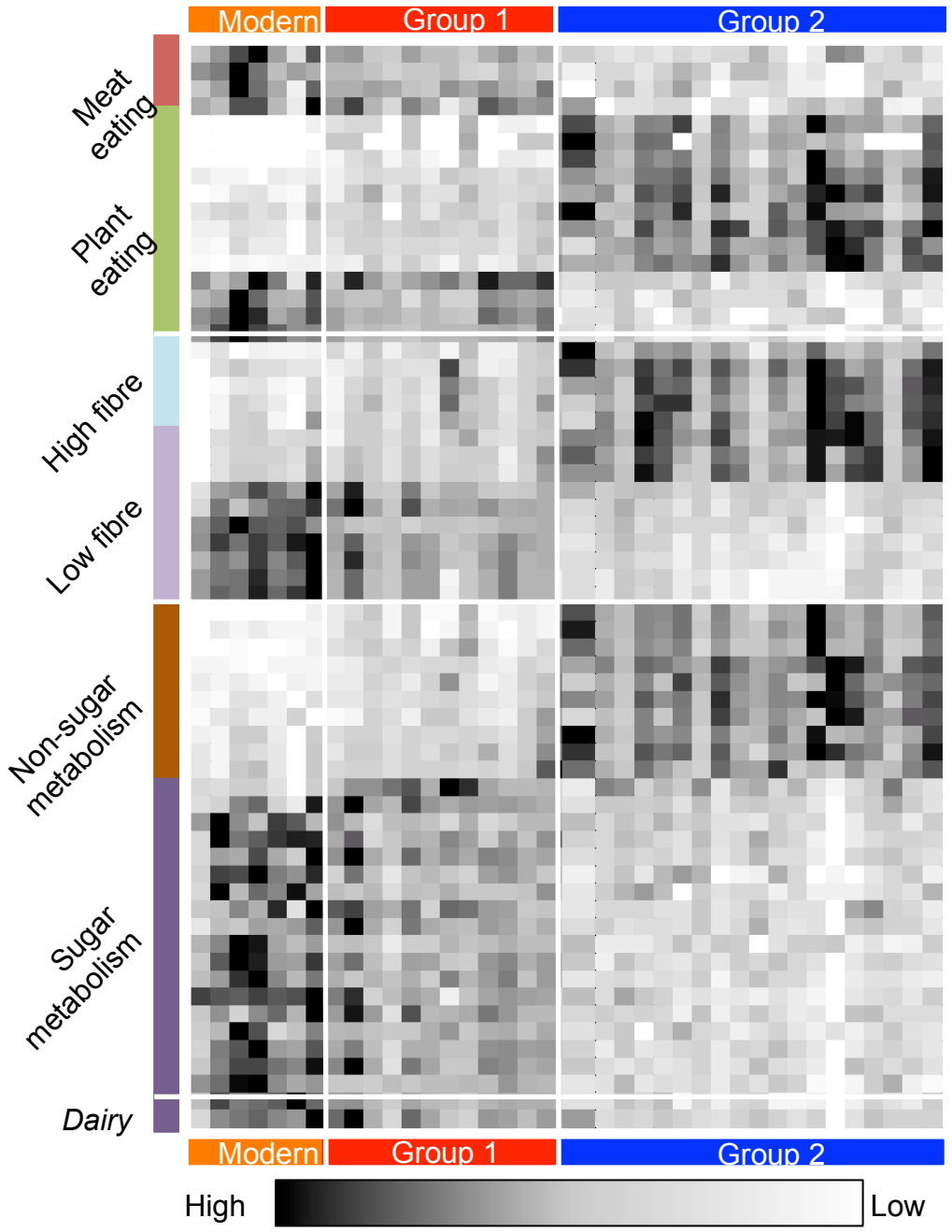


Figure 4: *Heat-map indicating abundance of significantly different metabolic functions in each sample (Kruskall-Wallis, $p < 0.05$). Samples are sorted into Groups (columns) and different metabolic functions (rows). Modern and Group 1 are enriched for meat associated amino acid metabolism, fibre digestion functions associated with low fibre diets, and sugar digestion (including dairy sugars). Group 2 show an opposing pattern.*

Discussion

Ancient DNA studies of dental calculus provide a detailed assessment of oral microbiota from historical individuals and represent a new tool to answer detailed archaeological questions around population substructure, diet, and disease. Here we have revealed a previously unappreciated level of diversity within the ancient oral microbiota of Great Britain. Two major groups of ancient microbiota existed and were stable over at least 1,800 years, suggesting that the many major events throughout this period, including multiple invasions, cultural developments, and disease epidemics failed to markedly alter oral microbiota. The apparent loss of one group also highlights the dramatic impact that recent history has had on the composition of oral microbiota. A suite of dietary factors was likely the key driver of microbiota diversity within ancient British individuals. Diet may also have provided the selection pressure for the community structure of modern microbiota. We suggest that the two-group pattern is likely linked to individualistic factors of lifestyle, wealth, and socio-economic rank. This study demonstrates that ancient DNA analysis of both bacterial species and their functions can provide unprecedented information about the past.

Socio-economic rank was associated with differential access to food items throughout history. The Streptococcus-dominated Group 1 individuals contain metabolic functions linked to meat, low-fibre carbohydrate, and dairy consumption. This is in stark contrast to the Methanobrevibacter-dominated Group 2, which was linked to a high-fibre, plant based diet. These observations may be linked with diets that are related to different lifestyles or socioeconomic ranks. For example, lower classes in Britain during the period analysed had a diet dominated by cereals and vegetables, while wealthier individuals (aristocracy, upper clergy, and wealthy townspeople) had additional access to meat and fish[37]. Our data likely reflects this socioeconomic split, with Group 1 indicative of the upper classes. Interestingly, dairy consumption was tightly linked with the group indicative of an upper class society. Although dairy products were ubiquitous in medieval Britain[38], this suggests that upper class had a higher rate of consumption.

Similarities between microbiota in Group 1 and modern individuals, such as having Streptococcus as a dominant primary colonizer[6]., suggest that the microbial species associated with Group 1 individuals are the pre-cursor community structure to the modern, industrial oral microbiota. The similarities in community and function indicate a continued selection for this microbiota structure over time that has ultimately resulted in the loss of the Group 2 microbiota in modern, industrialized individuals. We observe Group 2 individuals throughout our data, which finishes ~1700 CE. Consequently, our data is compatible with major changes in oral microbiota occurring during and following the Industrial Revolution. The Industrial Revolution saw marked increases in sugar consumption, notably sucrose and fructose. Processing of cereal grains became capable of removing germ and bran from the milled flour, reducing the fibre content. Consumption of dairy products and fatty meats also grew[39]. Therefore, the Industrial Revolution provided a refined version of the Group 1 associated diet, which eventually was accessible to the whole population. The functional suite of Group 1 microbiota is appropriate for this diet, suggesting the shift towards a complex of dietary factors that was once a high status diet has moulded the modern, industrialized oral microbiota.

Ancient groups of microbiota in this study do not match the microbiota of modern non-industrialized populations. To date, only one study has investigated oral microbiota of a non-industrialized population. This study identified an decrease in Prevotella, Fusobacterium, and Gemella in the non-industrialized Amerindians of South America compared to the industrial controls (USA) [13]. We do not observe a similar shift in bacterial taxa in this study as those taxa remained stable or increased in modern individuals in relation to the ancient samples. This demonstrates that ancient British and Amerindians do not share oral microbial community structure, which suggests that using modern non-industrial populations as ancestral proxies for Western populations is not accurate. Indigenous groups, such as the Amerindians, have unique cultural and environmental pressures and histories that differed from those in ancient Britain. In addition, the ancient

British populations included in this dataset had lived with agriculture for ~7,000 years [40], resulting in opportunities for unique co-evolution between humans and their microbiota. This experience and its consequences do not exist in modern individuals without any history of farming, such as modern, non-industrialized groups.

Previous analysis of diet using ancient dental calculus has required very deep sequencing (>5 million reads per individual) to allow direct identification of preserved DNA from food items. However, this depth of sequencing is costly, limiting the number of samples that can be analysed. In this study, we describe the use of low-coverage shotgun sequencing and microbial function analysis as an alternative means to infer dietary information. This approach can be used as part of a growing toolbox to infer dietary information. To date, biomolecular studies have generally analysed diet using isotopic analysis. These studies use naturally occurring isotopes (e.g. ^{14}N and ^{13}C) to infer the use of marine and terrestrial protein, and C_3 vs. C_4 plants[40]. Multiple studies have indicated the importance of marine fish in Britain and their transport inland[41]. However, unlike the microbiota data, isotopic studies have not identified a single split within the population split. The distinct differentiation in microbiota seen here may be biological, in that while diet might vary more continuously across the whole population, thresholds exist at which nutrient balances support one of the two microbiota community structures. Alternatively, this initial study may lack the resolution to discern shifts between further sub-divisions of dietary factors. The combination of isotopic and microbiota analysis should provide further resolution of diet in ancient populations. In addition, proteomics is now being applied to dental calculus[36] and has the power to indicate tissue specific identifications of dietary items. While isotopic and proteomic analysis permits the detailed analysis of specific elements of diet (e.g. C_3 vs. C_4 plants or milk proteins[36]), microbial DNA offers an opportunity to analyse multiple dietary elements simultaneously. Microbiota may be used to identify broad-scale patterns of dietary change within a population, one that is based on total available bionutrients (i.e. amino acids, fibre, and carbohydrates), which

cannot be observed with other dietary methods. Future studies will be required to explore the impact of change in one dietary factor on functional and taxonomic profiles. Such analyses will also need to ensure high quality data to avoid misinterpretation.

Contamination is a major issue for low biomass samples[22] such as ancient DNA specimens. Studies of microbial DNA are particularly vulnerable to bias from contamination due to the ubiquity of microbes and microbial DNA in the environment. In this study, 441 species (26.2%) of all species level identifications were removed by filtering laboratory contamination. This indicates the need for careful implementation and further development of laboratory and bioinformatic methods to obtain high quality data for future studies. In this study, we address this issue by performing a field standard decontamination protocol to limit environmental DNA contamination and by presenting a novel filtering strategy to remove low quality samples that have low endogenous DNA and/or have become contaminated with DNA from the laboratory environment and reagents. We use a comparison of all samples using Bray-Curtis dissimilarity metrics to exclude samples that are similar to negative controls. This method is a simple way to further increase confidence in the quality of samples to be analysed. In addition, this method can highlight a sample batch for further investigation of cross-contamination. A control sample that has experienced cross-contamination with dental calculus microbial DNA will fall away from other control samples, likely outside the 0.6 Bray-Curtis dissimilarity threshold, indicating that the samples processed with that negative control should be additionally assessed. This tool provides a novel approach to analysing ancient microbial communities and ensures that studies of dietary related functions can be effectively conducted in other ancient human populations.

The observation that diet appears to have driven population level differentiation of microbiota in historic times suggests that consideration of diet in modern populations may be a simple but effective way to manage microbiota. Further, specific seeding of nutrients may promote specific taxa and allow manipulation of community composition to prevent or cure

microbiota-based disease. Additional study is required to identify the fine scale role of nutrients and to examine the competition between microbial taxa. However, the stark pattern of microbiota differentiation seen here indicates the prospective use of microbiota analysis as an archaeological tool for diet and social structure, as well as the potential for effective pre-biotic medical treatments.

Methods

Sample collection

Samples were handled using sterile procedures as previously outlined[15]. Gloves were worn to limit contamination from the researcher and a sterile dental pick was used to remove the dental calculus deposit from the tooth surface. Enamel damage was avoided by applying pressure applied in parallel to the tooth. Calculus fragments were stored in labeled, sterile, zip seal bags. Sample metadata, including the oral location of the fragment, was recorded. Samples were transported to the ancient DNA facility at the Australian Centre for Ancient DNA, University of Adelaide, Australia. Access to samples was provided by the Natural History Museum (NHM), Royal College of Surgeons, Cambridge Archaeology East, and the Aberdeen Museum.

Decontamination protocols

Individual dental calculus fragments were exposed to UV radiation for 30 minutes on each side, followed by submerging the calculus sample in 3 mL of 5% bleach in a sterile petri dish for 3 minutes [17]. Following decontamination, all samples were washed in 1 mL of sterile 80% ethanol for one minute to remove residual chemicals (i.e. bleach) prior to DNA extraction.

DNA extraction, shotgun library preparation, and sequencing

Each sample underwent an in-house, silica based DNA extraction, as previously described [25]. Total volumes of lysis and guanidine DNA binding buffer were altered to account for small sample sizes as follows: 1.8 mL lysis buffer (1.6 mL 0.5 M EDTA (0.5M), 200 µL SDS (10%), and 20 µL proteinase K (20 mg/ml)) and 3 mL guanidinium DNA binding buffer. Two negative controls were included for every seven dental calculus samples. Shotgun libraries were generated without enzymatic damage repair using the protocol described previously[27]. Briefly, 20 µL of DNA extract was had single-strand overhangs removed prior to ligation of truncated, barcoded Illumina adaptors

and filling of the adaptor sequences. Samples had enzymes removed between each step (Qiagen MiniElute Reaction Clean-up Kits). Truncated Illumina adaptors were bound and the sample amplified in a 13 cycle, triplicate PCR amplification (HiFi). Agencourt AMPure XP system was used to purify pooled PCR products. Full length indexed sequencing adaptors were incorporated in a second 13 cycle, triplicate amplification[42]. Samples were pooled, purified, and diluted to form a 2 nM sequencing library. Quantification was completed using a TapeStation (Agilent) and quantitative PCR (KAPA Illumina quantification kit). The Illumina NextSeq, High sensitivity 2 x 150 bp kit was used for sequencing.

Bioinformatic analysis

Indices were used to demultiplex FastQ files using sabre (available here: https://github.com/najoshi/sabre). Reads were merged and samples fully demultiplexed and the 5' and 3' barcodes and adaptor sequence removed using bbmerge (available here: http://sourceforge.net/projects/bbmap/) and Adaptor Removal[43]. Taxonomy was derived from collapsed reads using MALTX[44]. Only collapsed reads were used because fragments greater than 300 bp were considered to be contamination[27]. MEGAN5[45] was used to calculate Bray-Curtis dissimilarity matrix of samples and negative controls based on genera level taxa identifications. Negative controls with Bray-Curtis dissimilarity > 0.6 in comparison to any other negative control was excluded. Four times the standard deviation of the remaining negative controls was calculated, and any sample within this range removed from the analysis. Samples were normalized, ignoring unassigned reads. All species identifications within negative control samples were removed from all samples. Finally, calculus from non-molar teeth was removed from the dataset.

Statistical analysis

To explore population level structure of microbiota, a UPGMA tree was generated using Bray-Curtis dissimilarity in MEGAN5. To identify

correlations with metadata and identification of taxa shifts, modern and ancient samples were exported from MEGAN5 and converted into QIIME format. Correlations with multiple metadata fields were completed in QIIME (V1.8)[46]. The dataset was filtered for singletons, Bray-Curtis dissimilarity calculated at a rarefaction depth of 120,000 reads, and the ANOSIM test applied (9999 permutations). If a sample did not have a known value for the metadata field being tested, that sample was excluded from that specific test. For metadata fields that were significantly correlated, Kruskall-Wallis was calculated from the singleton-filtered dataset. Alpha diversity was assessed within the identified groups by calculating the observed species from the singleton-filtered dataset at a rarefaction depth of 120,000 reads. Pathogens[27] were identified within the dataset and presence or absence metrics calculated.

To identify links with diet, published samples were compared to the dataset by generating a PCoA analysis of Bray-Curtis values and assignment of previously published data to ancient British samples using SourceTracker (V0.9.6)[47]. For analysis, ancient samples were denoted as sources, based on the two ancient groups identified with the UPGMA tree. Previously published data were labelled as sinks. SourceTracker was run with default parameters (1,000 subsampling, 10 iterations per sink sample) in R (V3.1.0)[48] using the QIIME wrapper. The suitability of the source populations was confirmed using the "take-one-out" method. To explore functional profiles of microbiota, functional tables were exported from MEGAN5 into QIIME and analysed as the taxonomic data. Amino Acid and Carbohydrate functions were extracted from the SEED database. For Amino Acids, all functions matching the Enzyme Commission numbers identified as distinguishing of carnivores and herbivores[19] were exported. For carbohydrates, all Level 4 functional groups were exported. Fibre metabolism functions[20] were exported from the KEGG database within MEGAN5. The 26 specific functions identified as relating to high and low fibre were exported to QIIME.

Acknowledgements

The authors would like to thank Chris Stringer and Robert Kruszynski at the Natural History Museum, London; Stephen Schiffels and Cambridge Archaeology East, Milly Farrell at the Royal College of Surgeons; and the Aberdeen Museum for access to samples.

References

1. Rautemaa, R., Lauhio, A., Cullinan, M. P. & Seymour, G. J. Oral infections and systemic disease—an emerging problem in medicine. Clin. Microbiol. Infect. 13, 1041–1047 (2007).
2. Pedersen, A. M. L. Oral Infections and General Health. (Springer International Publishing, 2016).
3. Kumar, P. S. Oral microbiota and systemic disease. Anaerobe 24, (2013).
4. Farrell, J. J. et al. Variations of oral microbiota are associated with pancreatic diseases including pancreatic cancer. Gut 61, 582–588 (2012).
5. Grau, A. J. et al. Periodontal Disease as a Risk Factor for Ischemic Stroke. Stroke 35, 496–501 (2004).
6. Avila, M., Ojcius, D. M. & Yilmaz, O. The Oral Microbiota: Living with a Permanent Guest. DNA Cell Biol. 28, 7 (2009).
7. Ruiz-Calderon, J. F. et al. Walls talk: Microbial biogeography of homes spanning urbanization. Sci. Adv. 2, e1501061 (2016).
8. Chase, J. et al. Geography and Location Are the Primary Drivers of Office Microbiome Composition. mSystems 1, e00022-16 (2016).
9. Goodrich, J. K. et al. Human Genetics Shape the Gut Microbiome. Cell 159, 789–799 (2014).
10. Weyrich, L. S. et al. Resident Microbiota Affect Bordetella pertussis Infectious Dose and Host Specificity. J Infect Dis (2013). doi:10.1093/infdis/jit597

11. David, L. A. et al. Diet rapidly and reproducibly alters the human gut microbiome. Nature (2013). doi:10.1038/nature12820
12. Schnorr, S. L. et al. Gut microbiome of the Hadza hunter-gatherers. Nat. Commun. 5, (2014).
13. Clemente, J. C. et al. The microbiome of uncontacted Amerindians. Sci. Adv. 1, e1500183 (2015).
14. Martínez, I. et al. The Gut Microbiota of Rural Papua New Guineans: Composition, Diversity Patterns, and Ecological Processes. Cell Rep. 11, 527–538 (2015).
15. Weyrich, L. S., Dobney, K. & Cooper, A. Ancient DNA analysis of dental calculus. J. Hum. Evol. 79, 119–124 (2015).
16. Dobney, K. & Brothwell, D. in Teeth and Anthropology (eds. Cruwys, E. & Foley, R. A.) 55–81 (1986).
17. Adler, C. J. et al. Sequencing ancient calcified dental plaque shows changes in oral microbiota with dietary shifts of the Neolithic and Industrial revolutions. Nat. Genet. 45, 450–455 (2013).
18. Weyrich, L. S. in Reticulate Evolution (ed. Gontier, N.) 231–253 (Springer International Publishing, 2015). doi:10.1007/978-3-319-16345-1_9
19. Muegge, B. D. et al. Diet drives convergence in gut microbiome functions across mammalian phylogeny and within humans. Science 332, 970–974 (2011).
20. Clayton, J. B. et al. Captivity humanizes the primate microbiome. Proc. Natl. Acad. Sci. 201521835 (2016). doi:10.1073/pnas.1521835113

21. Boers, S. A., Jansen, R. & Hays, J. P. Suddenly everyone is a microbiota specialist. Clin. Microbiol. Infect. 22, 581–582 (2016).
22. Salter, S. J. et al. Reagent and laboratory contamination can critically impact sequence-based microbiome analyses. BMC Biol. 12, 87 (2014).
23. Cooper, A. & Poinar, H. N. Ancient DNA: Do It Right or Not at All. Science 289, 1139–1139 (2000).
24. Yang, D. Y. & Watt, K. Contamination controls when preparing archaeological remains for ancient DNA analysis. J. Archaeol. Sci. 32, 331–336 (2005).
25. Brotherton P, Haak W, Templeton J, Brandt G, Soubrier J, Adler CJ, Richards SM, Der Sarkissian C, Ganslmeier R, Friederich S, Dresely V, van Oven M, Kenyon R, Van der Hoek M, Korlach J, Luong K, Ho SYW, Quintana-Murci L, Behar DM, Meller H, Alt KW, Cooper A & The Genographic Project. Neolithic mitochondrial haplogroup H genomes and the genetic origins of Europeans. Nat. Commun. Accepted 02/2013,
26. Belda-Ferre, P. et al. The oral metagenome in health and disease. ISME J. 6, 46–56 (2012).
27. Weyrich, L. S. et al. Reconstructing Neandertal behavior, diet, and disease using ancient DNA from dental calculus. Nature (2016).
28. Carpenter, M. L. et al. Pulling out the 1%: Whole-Genome Capture for the Targeted Enrichment of Ancient DNA Sequencing Libraries. Am. J. Hum. Genet. 93, 852–864 (2013).
29. Simón-Soro, Á. et al. Microbial Geography of the Oral Cavity. J. Dent. Res. 92, 616–621 (2013).

30. Bhat, M. Periodontal Health of 14-17-year-old US Schoolchildren. J. Public Health Dent. 51, 5–11 (1991).
31. Rocas, I. N., Siqueira, J. F., Jr., Santos, K. R. & Coelho, A. M. 'Red complex' (Bacteroides forsythus, Porphyromonas gingivalis, and Treponema denticola) in endodontic infections: a molecular approach. Oral Surg Oral Med Oral Pathol Oral Radiol Endod 91, 468–71 (2001).
32. Ting, P. T. & Freiman, A. The story of Clostridium botulinum: from food poisoning to Botox. Clin. Med. 4, 258–261 (2004).
33. Ley, R. E. et al. Evolution of Mammals and Their Gut Microbes. Science 320, 1647–1651 (2008).
34. Warinner, C. et al. Pathogens and host immunity in the ancient human oral cavity. Nat. Genet. (2014). doi:10.1038/ng.2906
35. Lee, C. (Cy), Shallenberger, R. & Vittum, M. Free sugars in fruits and vegetables. (1970).
36. Warinner, C. et al. Direct evidence of milk consumption from ancient human dental calculus. Sci. Rep. 4, 7104 (2014).
37. Müldner, G. & Richards, M. P. Fast or feast: reconstructing diet in later medieval England by stable isotope analysis. J. Archaeol. Sci. 32, 39–48 (2005).
38. Woolgar, C. M., Serjeantson, D. & Waldron, T. Food in Medieval England: Diet and Nutrition. (OUP Oxford, 2006).
39. Cordain, L. et al. Origins and evolution of the Western diet: health implications for the 21st century. Am. J. Clin. Nutr. 81, 341–354 (2005).

40. Richards, M. P., Schulting, R. J. & Hedges, R. E. M. Archaeology: Sharp shift in diet at onset of Neolithic. Nature 425, 366–366 (2003).
41. Mays, S. A. Carbon Stable Isotope Ratios in Mediaeval and Later Human Skeletons From Northern England. J. Archaeol. Sci. 24, 561–568 (1997).
42. Meyer, M. & Kircher, M. Illumina Sequencing Library Preparation for Highly Multiplexed Target Capture and Sequencing. Cold Spring Harb. Protoc. 2010, pdb.prot5448 (2010).
43. Lindgreen, S. AdapterRemoval: easy cleaning of next-generation sequencing reads. BMC Res. Notes 5, 337 (2012).
44. Herbig, A. et al. MALT: Fast alignment and analysis of metagenomic DNA sequence data applied to the Tyrolean Iceman. bioRxiv 50559 (2016). doi:10.1101/050559
45. Huson, D. H., Auch, A. F., Qi, J. & Schuster, S. C. MEGAN analysis of metagenomic data. Genome Res. 17, 377–386 (2007).
46. Caporaso, J. G. et al. QIIME allows analysis of high-throughput community sequencing data. Nat. Methods 7, 335–336 (2010).
47. Knights, D. et al. Bayesian community-wide culture-independent microbial source tracking. Nat. Methods 8, 761–763 (2011).
48. R Core Team. R: A language and environment for statistical computing. R Found. Stat. Comput. Vienna Austria (2014).

Chapter IV

Biological and cultural drivers of oral microbiota in Medieval and Post-Medieval London

Statement of Authorship

Title of Paper	Biological and cultural drivers of oral microbiota in Medieval and Post-Medieval London
Publication Status	Unpublished and Unsubmitted work written in manuscript style
Publication Details	In preparation for: Proceedings of the Royal Society B

Principal Author

Name of Principal Author (Candidate)	Andrew G. Farrer		
Contribution to the Paper	Conceived the experiments. Collected samples. Performed decontamination protocols, DNA extraction, PCR amplification, and sequencing library preparation. Analysed the data and interpreted the results. Wrote the manuscript.		
Overall percentage (%)	80%		
Certification:	This paper reports on original research I conducted during the period of my Higher Degree by Research candidature and is not subject to any obligations or contractual agreements with a third party that would constrain its inclusion in this thesis. I am the primary author of this paper.		
Signature		Date	2nd October 2016

Co-Author Contributions

By signing the Statement of Authorship, each author certifies that:

i. the candidate's stated contribution to the publication is accurate (as detailed above);

ii. permission is granted for the candidate in include the publication in the thesis; and

iii. the sum of all co-author contributions is equal to 100% less the candidate's stated contribution.

Name of Co-Author	Laura S. Weyrich		
Contribution to the Paper	Advised on data analysis and interpretation of results, and discussed and edited the manuscript		
Signature		Date	7th October 2016

Name of Co-Author	Neville Gully		
Contribution to the Paper	Discussed and edited the manuscript		
Signature		Date	10th October 2016

Name of Co-Author	Keith Dobney		
Contribution to the Paper	Discussed the results and edited the manuscript		
Signature		Date	10th October 2016

Name of Co-Author	Alan Cooper		
Contribution to the Paper	Edited the manuscript		
Signature		Date	11th October 2016

Name of Co-Author	Jelena Bekvalac		
Contribution to the Paper	Assisted with experimental design and collection of samples. Discussed the results		
Signature		Date	3rd October 2016

Name of Co-Author	Rebecca Redfern		
Contribution to the Paper	Assisted with experimental design and collection of samples. Discussed the results		
Signature		Date	6th October 2016

Biological and cultural drivers of oral microbiota in Medieval and Post-Medieval London, UK

A. G. Farrer[1*], J. Bekvalac[2], R. Redfern[2], N. Gully[3], K. Dobney[4], A. Cooper[1+], and L. S. Weyrich[1+]

Affiliations:

[1]Australian Centre for Ancient DNA, School of Biological Sciences, University of Adelaide, Adelaide, South Australia, Australia

[2]Centre for Human Bioarchaeology, Museum of London, London, United Kingdom

[3]School of Dentistry, Faculty of Health Sciences, University of Adelaide, Adelaide, South Australia, Australia

[4]Department of Archaeology, Classics and Egyptology, School of Histories, Languages and Cultures, University of Liverpool, Liverpool, United Kingdom

[+]These authors contributed equally to this work

Corresponding author:

[*]andrew.farrer@adelaide.edu.au

Abstract

There is limited understanding of the cultural and environmental factors that defined ancient microbiota. The oral microbiota has been analysed in previous ancient DNA studies of dental calculus. However, these studies lacked the resolution to identify the factors that defined the microbiota within individuals. Here, we present 128 oral microbiota from Medieval and Post-Medieval London, UK (1066 CE – 1853 CE). Using this detailed survey of a single ancient population through time, we explore microbiota association with extensive experimental, archaeological, and historical metadata. We identify a significant association between microbiota and oral geography, which has potentially confounded microbiota studies to date. However, by controlling for tooth type and surface, we are able to demonstrate the first associations between ancient microbiota and systemic health. In addition, we observe a change through time that, we suggest, correlates with the changing human demographic of London through time. This study indicates the potential of ancient microbiota to infer detailed health and socio-economic information and provides the baseline data to explore further populations.

Introduction

Microbial communities on and in the human body (microbiota) fulfill key functional roles including releasing otherwise inaccessible nutrients from food, removing dead epithelial cells, restoring tooth enamel, and interacting with the immune system (1–5). However, many diseases have now been linked to alterations in microbiota composition and/or function including oral disease, arthritis, respiratory disorders, cancer, obesity, and mental disorders (6–10). Therefore, understanding the microbiota and its roles in the body are of major importance to the future of medicine. Consequently, examining how microbial communities relate to health and respond to cultural and environmental factors is an important facet to understanding human health.

Studies examining how microbiota can be altered in modern, living people are generally unable to track the long-time impacts over multiple generations. Alternatively, archaeological studies provide a natural experiment to examine the adaption of microbiota within human populations and can reveal the history of the bacterial communities found in modern populations. Studies of historical microbiota could also monitor the impacts of changing lifestyle, such as different living environments, social economic status, disease, and diets, on specific bacteria and microbiota community structure. To explore this potential and to examine how microbiota can be related to lifestyle characteristics, studies of well-documented historical individuals are needed. Such analysis can develop the groundwork for understanding how much historical information can be inferred directly from an individual's microbiota composition and function (11).

The analysis of historical microbiota is routinely possible through ancient DNA (aDNA) analysis of calcified dental plaque (calculus), which is widespread in the archaeological record (12). Dental plaque consists of a microbial biofilm that grows on the surface of the teeth and forms part of the oral microbiota. Dental calculus forms during life as a result of calcium ions precipitating from the saliva and crystallizing within the plaque layer (13). This fossilization process means that dental calculus remains the only known

accurate record of archaeological microbiota (12,14–19). Oral bacterial communities respond to changes in lifestyle and environment and play key roles in oral and systemic disease making dental calculus an excellent marker to examine how microbiota and health were altered in past populations. Bacterial DNA from ancient dental calculus has been used to reveal major changes in human oral microbiota during the Neolithic (agricultural) and Industrial Revolutions (~10,000 ybp and ~200 ybp, respectively) (12), to perform highly detailed analyses of diet (17,18), and to track pathogen evolution through time (17,18). However, there are still no studies of microbiota within a single human population through time, and such a record is critical to understanding how microbiota differ and change through time in relation to life history.

Several technical factors have also not yet been resolved in aDNA analysis of dental calculus. For example, the approaches to contamination control and detection need to be constantly addressed and advanced, as non-endogenous DNA entering the analytical pathway will continue to be a major issue for the analysis of microbiota (20). Another key area is the influence of oral geography (i.e. the specific location in the mouth) because microbial composition is known to vary between different tooth surfaces in modern individuals (21). Surprisingly, this has not been considered in detail in studies of ancient oral microbiota to date. Variation resulting from oral geography might potentially provide false positive results or mask true, historical patterns, and this issue is set to become more intrusive as studies increase resolution. Failure to appreciate the potential variation or risks inherent in the research of dental calculus samples is likely to result in misinterpretation or errors and needs to be explored in further depth.

Here, we use 128 dental calculus samples with extensive biological and cultural metadata to gain fine scale insight into microbiota alterations within a population of Medieval and Post Medieval Londoners (1066 – 1853). The Museum of London collections have been extensively studied, providing detailed dating, paleopathology, and cultural information for each sample. In addition, London is an important city in the development of western,

industrialized civilization, and the history of the area reflects that of many other cities across Europe. Focusing on a single city removes the impact of geographical variation that may mask the finer scale microbiota changes relating to the environment of an individual. It also allows us to examine how microbiota within historical London were linked to age, sex, diet, location, oral and systemic disease, and socio-economic rank. We also explore technical factors that can potentially confound ancient dental calculus research, such as oral geography, sample size, and sampling and laboratory protocols.

Methods

Archaeological Context and Site Information

161 samples were collected from nine archaeological sites in a 10-square mile section of London (16km^2), forming a continuous transect from 1066 – 1853 CE (Figure 1). This included monastic (Bermondsey Abbey, Spital Square, Merton Priory, St. Benet Sharehog, St. Mary Graces, and St. Brides), laymen (Guildhall Yard), low class (Cross Bones), and upper class (Chelsea Old Church) burial sites and cemeteries. All individuals were over 18 years old and had extensive metadata collated in the Wellcome Osteological Research Database (WORD), which is hosted by the Museum of London (Table S1). This included information about sex and age estimates, blood disorders, dental and vertebral anomalies and pathologies, and joint disease.

Sample collection

All sampling was completed using sterile procedures as previously published (11). Briefly, a facemask and gloves were worn to limit contamination from the researcher, and the gloves were changed between each sample to limit cross contamination between samples. A sterile dental pick was used to remove the dental calculus deposit from the tooth surface. Pressure was applied in parallel to the tooth surface to ensure that no damage was done to the enamel. Calculus fragments were collected in

sterile aluminum foil and stored in sterile plastic bags for transport to the ancient DNA facility at the Australian Centre for Ancient DNA, University of Adelaide, Australia.

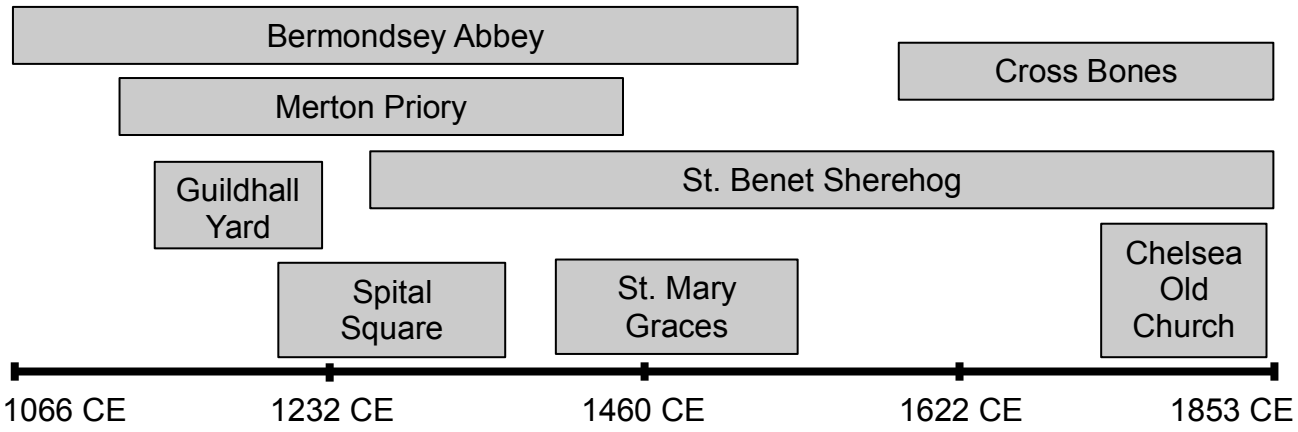


Figure 1: *The eight archaeological sites included in this study and the length of time each site was actively used for burials. Samples throughout these periods are included in this study.*

Sample Decontamination

All laboratory work was conducted in a specialized aDNA facility at the University of Adelaide that was specifically designed and built to conduct aDNA research. Prior to DNA extraction, each sample underwent a decontamination protocol to limit environmental microbial DNA on the calculus surface, as previously described (12). Samples were exposed to UV radiation for 30 minutes on each side in sterile sample trays, followed by emersion in three mL of 5% bleach in a sterile petri dish. Following decontamination, all samples were rinsed in one mL of 80% ethanol for one minute to remove residual bleach prior to extraction.

DNA Extraction, Shotgun Library preparation, and Sequencing

To recover preserved DNA, dry samples were powdered in a sterile tube, immediately following decontamination. A modified silica-based DNA extraction was used on all samples, as previously described (18). The total

volumes of lysis and DNA binding buffers were modified to account for the small sample size: 1.7 mL lysis buffer (1.6 mL 0.5 M EDTA (0.5M); 100 µL SDS (10 %); and 20 µL proteinase K (20mg/ml)) and 3 mL guanidinium DNA binding buffer. Each DNA extraction also included two negative controls at a ratio of one negative control per seven calculus samples. Next, shotgun libraries were generated without enzymatic damage repair as previously applied to dental calculus samples (18). Briefly, 20 µL of DNA extract was prepared by enzymatic polishing to produce blunt ended fragments before ligation of truncated barcoded Illumina adaptors, and filling of gaps in adaptor sequences. Qiagen MiniElute Reaction Clean-ups were completed after each step, and each sample underwent triplicate PCR amplification for 13 cycles (HiFi) with full length indexed Illumina adaptor (22) to increase concentration while maintaining complexity. PCR products were pooled and purified with the Agencourt AMPure XP system. A 2 nM sequencing library was produced following a second round of pooling, purification, and quantification using a TapeStation (Agilent) and quantitative PCR (KAPA Illumina quantification kit). In total, 128 out of the 161 dental calculus samples yielded DNA after sequencing using a high output 2 x 150 bp kit on the Illumina NextSeq.

Bioinformatic analysis and quality filtering

To identify the microbial communities preserved within samples, raw FastQ files were demultiplexed using sabre (available here: https://github.com/najoshi/sabre). Bbmerge (available here: http://sourceforge.net/projects/bbmap/) was used to merge reads (5 bp overlap), and Adaptor Removal (23) was applied to identify and remove the 5' and 3' barcode and adaptor sequences. Reads greater than 300 bp were discarded, as they likely represent modern contamination (15). Next, microbial species and functions were identified using MALTX (18,24) against the NCBI nr database (2014), and the resulting information was uploaded and filtered using default LCA parameters in MEGAN5 (25). To identify samples that lack endogenous signal, Bray-Curtis dissimilarities were

calculated in MEGAN5 for all in-house metagenomic samples and compared with negative controls. Negative controls that formed a tight cluster with an average dissimilarity less than 0.6 were considered representative of the laboratory environment, and any sample that fell within four times the mean standard deviation of this group was removed from the analysis on the basis of the lack of endogenous oral microbiota signal. Samples with < 100,000 reads were also removed from downstream analysis to ensure accurate reconstruction of the ancient bacterial community, as these were clearly poorly preserved. The identified reads within all remaining samples were normalized to 129,760 sequences, which was the lowest number of reads observed in any sample. Lastly, laboratory contaminant signal was removed from all samples by filtering any species observed in the negative controls from the calculus samples. All 128 sequenced samples from the Museum of London were retained throughout the bioinformatics analysis and were considered suitable for further analysis.

Analysis of bias associated with oral geography or sample processing

To assess the impact of oral geography and sample handling, the physical location of the sample and the workflow metadata were compared to the microbiota. All reads assigned to cellular organisms for each sample were exported from MEGAN5 and transformed for use in QIIME (V1.8) (26). Within QIIME, Bray-Curtis dissimilarity was calculated for all samples at a rarefaction of 120,000 reads per sample. ANOSIM tests (9999 permutations) were used to correlate oral sampling location (e.g. upper or lower jaw, tooth type and surface), sample information (i.e. fragment size and sub- or supra-gingival), and processing details (i.e. date of sampling, extraction, sequencing) with the Bray-Curtis matrix (Table S1). Following this analysis, samples were filtered based on tooth type, tooth surface, and sub- or supra-gingival calculus. The three largest datasets were taken forward for further analysis (Molar, Lingual/Palatal, Supra-gingival (n = 36); Premolar, Lingual/Palatal, Supra-gingival (n = 18); and Incisor, Lingual/Palatal, Supra-gingival (n = 18)).

Linking microbiota with health, culture, and environment of historical individuals

To identify the health, cultural, and environmental factors that correlated with the species identified within calculus, Bray Curtis dissimilarities were calculated (120,000 read rarefaction) from the filtered QIIME datasets. ANOSIM tests (9999 permutations) were used to identify metadata (Table S1) that were significantly associated with microbiota diversity. To ensure statistical tests were valid, a minimum of five samples per group was needed when comparing only two metadata categories, while a minimum of three samples per group was enforced when three metadata categories were compared. Sample without metadata for a certain category were excluded from specific tests. Correlations were deemed significant with a $p < 0.05$. For metadata categories that were significant, Kruskall-Wallis was used to identify specific species that contributed to the differences observed in each metadata category ($p < 0.05$).

Linking microbiota functional variation with individual metadata

To explore shifts in any function identified within these oral microbiota, abundance of annotated genes observed in each sample were exported from MEGAN5 to QIIME. Gene functions were annotated in level 5 of the SEED database (27). A Goodness of fit (G) test was used to identify microbial functions that significantly varied in frequency ($p < 0.05$) between the groups identified from taxonomic analysis.

Identifying dietary differences that contribute to microbiota variation

To assess if alterations in diet were related to the microbiota groups associated with lifestyle, amino acid and fibre metabolism functions previously identified to delineate carnivory vs. herbivory and high or low dietary fibre, respectively (28,29). In addition, carbohydrate metabolism pathways from the SEED database (level 4) were exported from MEGAN5 to

QIIME (19). The specific functions linked to each metadata-category were filtered by oral geography as above, and assessed with the G-test ($p < 0.05$).

Results

Robust oral microbiota are recovered from historic London samples

Samples were assessed for quality and bias to increase the resolution of changes in endogenous DNA between samples. First, comparison of samples to negative controls provided an assessment of sample quality and the limits of any potential contamination, as previously published (19). All sequenced samples (n = 128) fell more than four standard deviations from the cluster of negative controls, indicating that they all contained endogenous signal. We then removed any species identified in negative controls from the calculus samples, removing 464 taxa (26.4% of all species level identifications) from the overall dataset. Lastly, ANOSIM tests against laboratory methods were conducted and indicated that there was no correlation between the data and date of sampling, DNA extraction, or library preparation (ANOSIM, $p > 0.01$, Table S1). Consequently, no experimental biases were apparent.

Oral geography accounts for some variation between individuals

Previous studies of the modern oral microbiota have indicated that oral geography (i.e. where the bacterial community is within the mouth) can drastically impact microbial diversity (21). We correlated the complete ancient dataset (n = 128) with the oral sampling location of the dental calculus samples to assess if oral geography impacts the microbiota present in historical samples. ANOSIM results identified that tooth type significantly correlated with the observed microbiota in each sample (ANOSIM p=0.0001, Table S2). The variation between tooth type was also the variable that explained the first axis on PCoA plot calculated from Bray-Curtis values of all samples (Figure 2) accounting for 44.9 % of the variation in the data. This is an unexpectedly high proportion of total variation, and brings into question

previous studies that have not controlled for oral geography. To control for tooth type, we processed samples from each tooth type independently. Within tooth types, we observed that tooth surface (buccal, lingual/palatal, interproximal) was a significant driver of diversity ($p < 0.01$, Figure 3, Table S2) except for canines (n=14). However, as the canine data set contained the fewest samples this lack of correlation may be associated lack of statistical power. As tooth surface had impacted the data from other tooth types, lack of association between microbiota from canine teeth and tooth surface could not be seen as a confirmation of no association. Consequently, canine samples were not included in further analysis. Lastly, differences between sub- and supra-gingival microbiota community physiology have also been reported in modern populations (30), so we examined this parameter as well. Sub- and supra-gingival calculus samples were not statistically different when controlled for tooth type and tooth surface ($p > 0.05$, Table S3). However, very few samples were representative of sub-gingival dental calculus (5 of 128 samples, Table S1) and correct classification of sub-gingival calculus is complex in archaeological samples because of the lack of soft tissue (gum tissue), potentially biasing the analysis. However, to minimize contributions of this potential bias these five sub-gingival samples were removed from downstream analysis. Overall, these data indicate that oral geography has the potential to drive microbiota diversity within ancient calculus analyses and will need to be accounted for to avoid introducing bias.

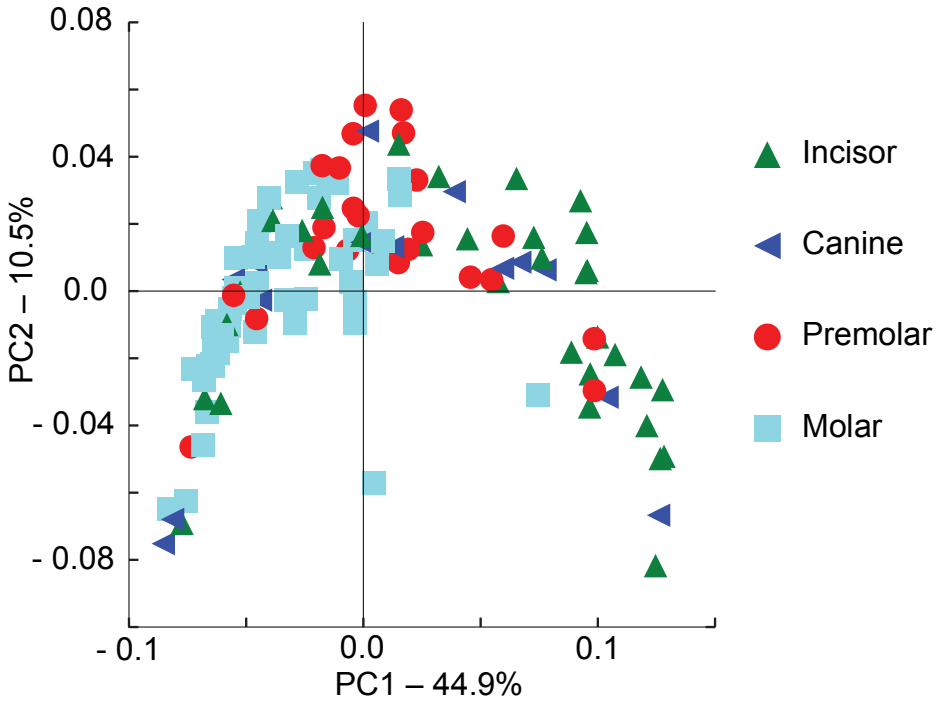


Figure 2: *PCoA plot of Bray-Curtis dissimilarity for all dental calculus samples (n = 128). Individual samples are coded by the tooth type from which the sample was taken. The first axis indicates the separation of tooth types, notably molar and incisor*

Microbiota are correlated with disease and time

While alterations in modern microbiota have been linked to disease and lifestyle (15), these factors have not yet been explored in ancient samples. Therefore, we explored the difference in microbiota between individuals with known oral and systemic diseases and a wide range of lifestyle factors. We also explored microbiota variation through time. First, we examined potential relationships between sample metadata and microbiota from supra-gingival dental calculus samples within three data sets: the lingual/palatal side of the molar (n=36), premolar (n=18), and incisor (n=18) teeth. An ANOSIM test was performed to examine the ability of any of 66 different metadata fields (Table S1) to explain some of the variation within each data set, including archaeological site, paleopathology, cultural factors,

disease, and period. While no metadata category significantly explained variation within the incisor group, disease factors were related to variation within the molar and premolar groups. Abscesses (oral disease) and various joint pathologies (porosity, osteophytic lipping, and non-specific periostitis – examples of systemic disease) were significantly associated with variation within the molar data set. Similarly, variation within the premolar teeth data set was correlated with chronic bone pathologies of the spine (Schmorls nodes). In addition, when the data was split into 300-year blocks (1000-1300, 1300-1600, and 1600-1900 CE) the date of the sample also explained variation in both premolar and molar teeth. Notably, there was no significant correlation with other time factors (change between 100-year, 200-year, 400-year, or 500-year blocks, or 100-year blocks within each of the 300-year divisions) (ANOSIM $p < 0.01$, Table S4). These results indicate that historical microbiota in London were impacted by disease, and the microbial communities shifted in increments recognizable with 300-year time bins, likely around two time points which potentially relate to the beginning and end of the presence of plague in Britain (1300 and 1600 CE). Each of these factors was investigated further to examine why these factors altered microbiota in the past.

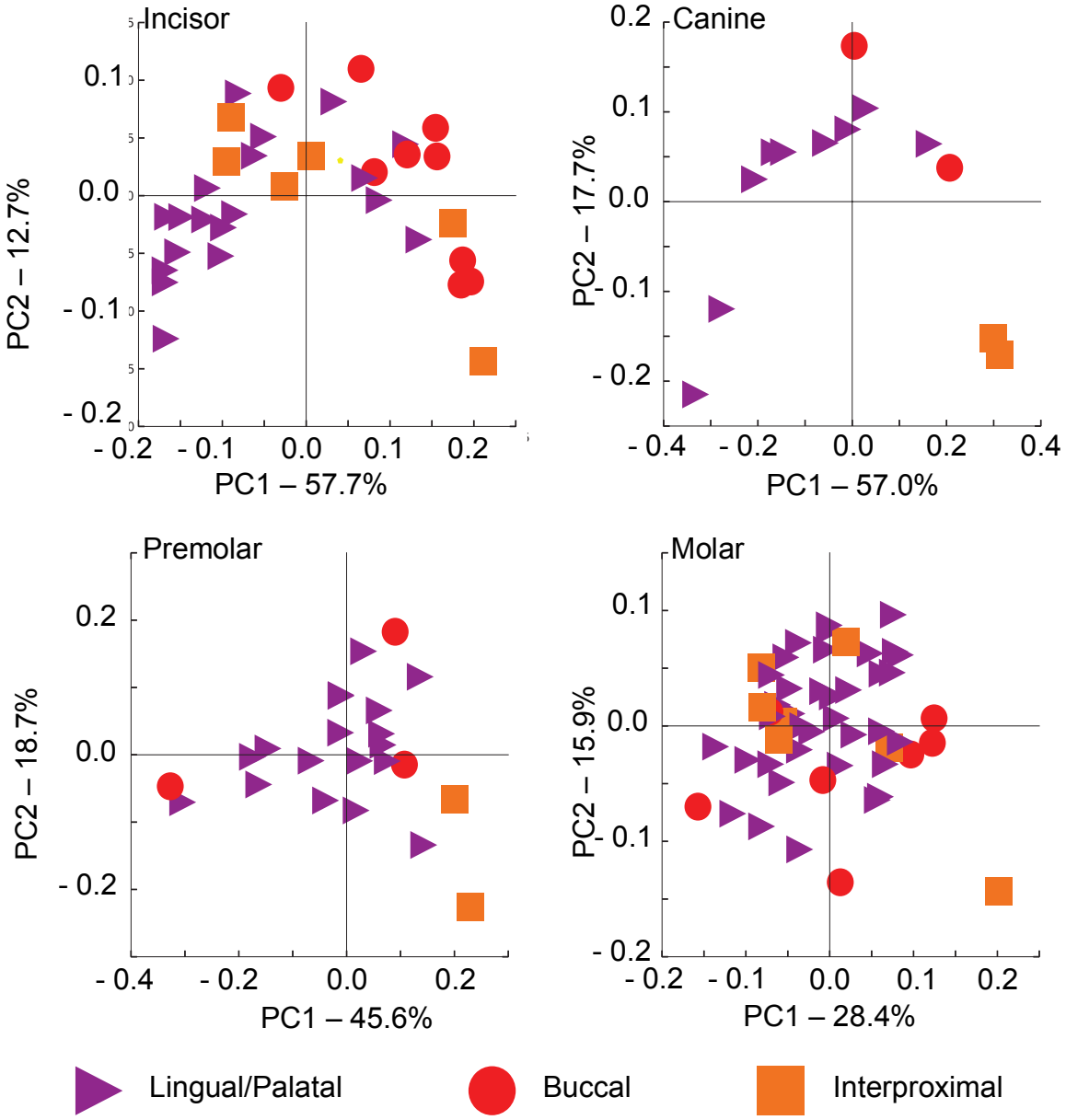


Figure 3: *PCoA plot of Bray-Curtis dissimilarity for all dental calculus samples divided by tooth type. Individual samples are coded by the tooth surface from which the sample was taken.*

Common oral pathogens were associated with microbiota alteration

The presence/absence of abscesses within the oral cavity correlated with microbiota within supra-gingival, lingual/palatal dental calculus of molar teeth. Therefore, we explored the taxonomic and functional traits associated with dental abscesses in historical London. The taxonomic variation

associated with dental abscesses was polymicrobial, and 30 bacteria species were increased in individuals with dental abscesses, while 20 species were decreased (Kruskall-Wallis, $p < 0.05$, Table 1, Table S5). Notably, Prevotella species and Streptococcus species (bacteria associated with dental abscesses) (31) were significantly increased in individuals with abscesses. However, the abscesses associated *Porphymonas* species and *Treponema* species were significantly lower in individuals with abscesses. In addition, other potential microbes involved in abscess formation were not significantly different (e.g. *Clostridium*, *Fusobacterium*, and *Bacteroides*). However, individuals with dental abscesses did have increased proportions of Methanobrevibacter species. Methanobrevibacter species have been associated with severe periodontal disease in modern populations (32), but they have also been shown to be a healthy component of ancient oral microbiota (19) serving as a primary colonizer.

Despite the significantly different species, there was little change in the microbial functions linked with ancient dental abscesses. A single genetic function was identified in the metagenomic sequence data, which is involved in production of flagella on bacteria (Flagella hook length protein (FliK)) (G-test, $p < 0.05$). Flagella allow bacteria to be motile (33), and their presence can be associated with bacterial pathogenicity, allowing bacteria to move through mucus layers (34) and gain access below the gum line. Flagella can also act as anchors, adhering bacteria to surfaces (34), which could allow pathogenic bacteria to adhere to the tooth surface and trigger abscess formation. Motility has previously been linked to oral disease (35,36) providing a potential link between higher flagella production and disease (dental abscess) presence.

Relationships exist between bone pathologies and microbiota composition

Microbiota variation in the calculus samples from molar teeth was also found to be associated with two joint diseases (porosity and osteophytic lipping) and a bone disorder (non-specific periostitis). Observations of

porosity indicate excessive opening of the pores in bone to allow blood, nerve, and other soft tissue to enter (37) and can be associated with age of the individual (38) as well as many diseases including anemia and rickets (39). Osteophytic lipping occurs when bone spurs (outgrowths of bone) form around the joint surfaces; they are often present alongside arthritis (40). Within our data set, the group of individuals with porosity and osteophytic lipping overlapped (Table S1). All individuals with porosity (n = 9) also had osteophytic lipping while five additional individuals only had osteophytic lipping. Consequently, taxonomic and functional differences may not be specifically associated with either trait individually. The final disease association is periostitis, an inflammation of tissue surrounding the bone, which is often indicative of disease in the underlying bone (41). In living patients, it can be associated with various diseases (including syphilis and skin ulcers) and is linked with trauma (41). In contrast to the molar samples, within premolar teeth microbiota were correlated with Schmorl's nodes, a vertebral bone pathology. While the nodes are common, their pathological status is not understood. However, it is thought that they appear in response to weakening of the vertebra by other pathologies or stress (42,43). While each of these bone pathologies is classified independently, it is possible that they have arisen from a similar cause, such as an increase in manual labor within this population.

Porosity

Microbiota variation was explored between those samples with and without signs of bone porosity. Porosity was linked to an increase in 28 microbiota taxa and a decrease in 28 taxa (Kruskall-Wallis, $p < 0.05$, Table S6). Increases included Prevotella species, which have previously been associated with bone disease such as arthritis (44). However, as with abscesses, the three Red Complex taxa were decreased in individuals with the disease state. The repeated reduction of the Red Complex in association with different disease is likely a genuine pattern as different individuals contain the highest levels of each taxon. Consequently, while the presence

of a single individual may skew the statistical interpretation of one member, it is unlikely that this would be consistent across all three members. Only two inferred microbiota functions showed significant differences between those samples with and without porosity, both increasing with the disease (G-test $p < 0.05$). The flagella associated protein was also associated with porosity, as it was for abscesses. An archaeal DNA polymerase (EC.2.7.7.7) representing a core cellular function was also associated with this disease and is likely linked to the increase in the archaeal Methanobrevibacter species. Next, an assessment of dietary functions demonstrated an increase in a high fibre associated metabolic function (G-test $p < 0.05$), suggesting that individuals with bone porosity ate a diet with high levels of dietary fibre. In addition, two single-carbon carbohydrate metabolism functional groups were also identified (methanogenesis and ethanolamine utilization, G-test $p < 0.05$). Methanogenesis is a key function of Methanobrevibacter species, and the significant increase in frequency is likely linked to an increase of these taxa with these individuals. Ethanolamine is a non-sugar carbohydrate found in animal and vegetable foods (45) and can be used as a carbon and/or nitrogen source by multiple bacteria, including Clostridium, which was at higher abundance in diseased individuals (46). The presence of these specific carbohydrate utilization functions suggests there may have been a lack of rich carbohydrates, such as sugar, in the diet.

Osteophytic lipping

Osteophytic lipping was associated with an increase in 19 microbiota taxa (Kruskall-Wallis, $p < 0.05$) compared to specimens without the pathology. These taxa are similar to those increased with porosity, which was to be expected given the overlap of individuals with these diseases. For example, Prevotella species and Methanobrevibacter species are also present in higher abundance in individuals with osteophytic lipping. Both the flagella associated protein and the archaeal DNA polymerase were enriched in the diseased individuals, as with those suffering from porosity. However, a second archaeal methanogen function that is part of the folate pathway was

also enriched (CoB--CoM heterodisulfide reductase (EC 2.8.4.1.)). Folate is a B vitamin necessary for DNA and RNA synthesis and is therefore fundamental to cellular processes (47). Notably, microbiota that produce B vitamins have been identified in the gut (47). Therefore, it is possible, that Methanobrevibacter species may be an unappreciated source of B vitamins for the host. Lastly, increased abundance of dietary related functions also identified methanogenesis (carbohydrates) and a high fibre marker function (Table S7). Together, these findings indicate that microbiota alterations linked to porosity and osteophytic lipping are similar, and may be associated with groups experiencing a high fibre diet.

Periostitis

Within calculus samples from individuals with periostitis, Prevotella species and Methanobrevibacter species were again increased, as observed with porosity and osteophytic lipping. However, several unique bacterial species were linked with this disease, including Capnocytophaga, Clostridia, Eubacterium, Firmicutes, Mogibacterium, Neisseria, and Pyramidobacter (Kruskall-Wallis, $p < 0.05$, Table S8). This indicates that periostitis may be linked to a unique shift in the human microbiota, which is different to those observed for porosity and osteophytic lipping. Functional metabolic changes linked to this disease included the archaeal folate pathway (CoB-CoM heterodisulfide reductase (EC 2.8.4.1.)). The repeated observation of archaeal-associated functional alterations in these disease states is likely to reflect the shared increase in Methanobrevibacter species prevalence. A single high fibre marker function was also associated with periostitis, again indicating that bone disease is linked to high-fibre diets.

Schmorls nodes

Variation in premolar microbiota was correlated with Schmorl's nodes, a vertebral bone pathology. However, no individual microbiota species or functions were significantly different between diseased and non-diseased individuals (Kruskal-Wallis, $p > 0.05$, Table S9). Schmorl's nodes

can originate from developmental disease, pathologies that weaken the vertebral disks, or physical stress on the spine (42). The lack of a specific causal factor for this paleopathology would explain why no association with the oral microbiota could be inferred. Together, these results indicate that systemic bone diseases and the oral microbiota are correlated. Theincrease in metabolic functions linked to the digestion of high fibre foods and non-sugar carbohydrates suggests that alterations in diet and lifestyle may also be indirectly contributing to this disease, and may be associated with lower socio-economic status (19).

Microbiota taxa and functions indicate a decrease in diet quality over time

Alterations in oral microbiota across multiple generations are key to understanding the impacts of cultural and environmental factors. Variation in both molar and premolar data sets seemed to relate to 300-year periods. Within these periods, 43 taxa decreased and 19 increased (Kruskall-Wallis, $p < 0.05$) in the calculus from molar teeth. This included an increase in taxa such as Methanobrevibacter species, while Streptococcus species and Prevotella species decreased. However, the Kruskal-Wallis test could detect no differences in premolar teeth over these periods. This suggests that Methanobrevibacter species-associated microbiota communities increased over time in London. These communities have been previously linked to high-fibre, low meat diets (19) and are likely representative of a low socioeconomic class. The change through time in molar teeth was also linked with an increase in the disease associated flagella hook length control protein (FliK) (g-test, $p < 0.05$) and two hypothetical proteins (D-serine permease (DsdX) and FIG019733 possible DNA binding protein). In relation to dietary functions, both calculus samples from molar and premolar teeth indicated a reduction in meat and sugar metabolism (specific genes; Table S10). This was particularly evident in the premolar communities, where metabolism of sucrose and dairy sugars (lactose and galactose) was decreased in later time points (Figure 4). In premolar teeth, an increase in

coarse fibre metabolism was observed through time, while three low fibre related functions were decreased through time (Table S10). Together, this functional analysis suggests a reduction in diet quality over the 800 years studied.

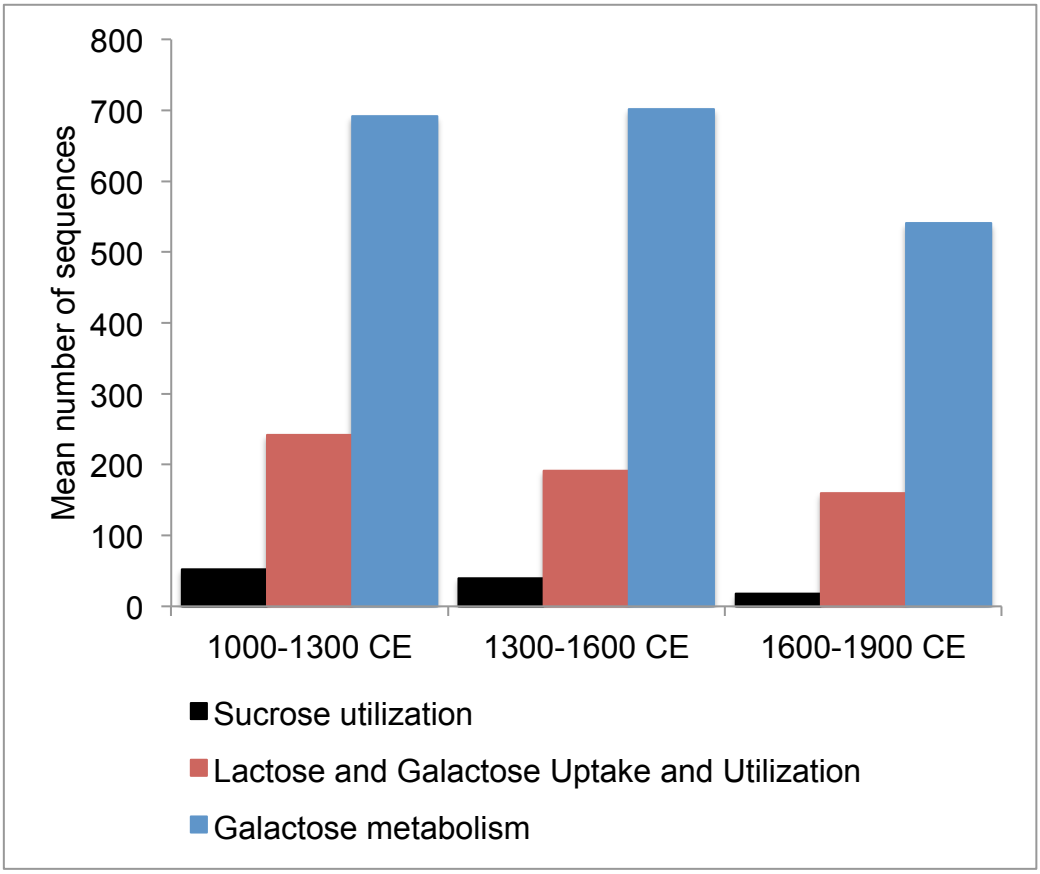


Figure 4: *Bar Chart showing change in sucrose and dairy sugar through time. Bars represent the mean number of sequences for each time period associated with each functional trait. These functional traits were significantly different between individuals of different time periods (Kruskal-Wallace, $p < 0.05$) Premolar-Lingual/Palatal-Supragingival sub-group*

Discussion

In the largest ancient dental calculus study to date, we reveal that the oral geography of the mouth is an important factor to be considered in order

to identify fine scale factors that impact the human microbiota. To complete the analysis in this study, we controlled for tooth type and tooth surface, as each were shown to be significantly associated with variation in the data. We additionally accounted for sub- vs. supra-gingival dental calculus to remove potential bias that was previously reported (48). Once these factors were considered, variation in the microbiota associated with oral health, systemic disease, alterations through time, and diet could be observed. This indicates the potential of dental calculus as a high-resolution archaeological tool to reveal new information about past human lifestyles and health.

Modern studies have previously identified different microbiota communities in different locations of the mouth (21). Despite this, it had not yet been examined in studies of ancient dental calculus. Within this study, the oral geography of dental calculus proved to be a significant source of variation within the data, raising questions about the potential validity of conclusions drawn from previous ancient calculus research. In future, such studies must account for tooth type and surface when analysing dental calculus samples to prevent bias. To address this issue, we performed statistical analysis on calculus samples from the same oral location, which allowed the detection of previously masked signals even though it markedly reduced the total sample size per group. Notably, different sampling locations within the mouth revealed different associations with metadata. This may indicate that the communities from different teeth respond differently and that there are unknown factors that impact specific areas of the mouth differently - or that there are still too few samples available to consistently recover trends. It is important to note that the different tooth samples come from different individuals, which limits the ability to directly compare the results. Consequently, these results represent the first associations of ancient microbiota with disease, but replication and further high-resolution studies are required to determine the extent and accuracy of these observations. To combat this, future studies should sample from a single oral location across individuals where possible, as well as examine calculus from multiple locations within the mouth of a single individual to

assess within-individual variation. The expansion of current bioinformatic methods to analyse ancient microbiota data should also be modified to account for or remove variation that results from oral geography. In addition, dental calculus samples from different regions within the mouth should not be combined into a bulk sample for analysis.

Although oral pathogens have been previously identified in other studies of ancient dental calculus (12,16–19), the microbiota community structure has not yet been linked to the presence of oral diseases. In this study, we examine alterations in community structure that are associated with the presence of oral pathogens. Association of abscesses with microbiota structure included the significant increase in some of the known, abscess-associated pathogens, including Prevotella species and Streptococcus species. Suggesting that the association seen was linked to the disease state and not an artifact. Notably, both Streptococcus species and Methanobrevibacter species were both increased in individuals with abscesses, even though these taxa have been found to be mutually exclusive in healthy individuals (19), which suggests that signals of lower-socio-economic status (increased Methanobrevibacter species) and abscesses (increased Streptococcus species) can be identified simultaneously in high resolution studies. While further study is required to confirm microbiota and cultural and environmental associations, this demonstrates the potential of microbiota studies to explore a wide range of lifestyle factors simultaneously.

In modern populations, oral microbiota have been linked to a wide range of systemic diseases, as diverse as arthritis, heart disease, and mental disorders (5,6,9). In this study, several indicators of different bone diseases were related to microbiota structure. This is the first study to indicate a correlation between systemic disease and the oral microbiota in ancient individuals. Microbiota from molar teeth showed correlations with two joint diseases, porosity and osteophytic lipping, and a link with periostotitis (inflammation of the tissue surrounding bone). In all cases, oral Prevotella species were enriched when the diseases were present. In modern

populations, Prevotella species are associated with arthritis in modern individuals (6) although the link between this species and disease is not yet known. Two possibilities may explain the association of Prevotella species and bone disease. First, there may be a direct link between this species and the disease, as oral bacteria can escape the mouth and cause inflammation and lesions elsewhere in the body (49). This option is consistent with the increased presence of bacterial sequences related to motility, which directly provides oral bacteria with a method of moving into the gum line and more easily entering the blood stream. Second, Prevotella species may serve as a marker for overall community alterations that are linked to socioeconomic status or lifestyles. This may explain the association between diseases and increases in Prevotella taxa within the Methanobrevibacter species-dominated microbiota community. The latter is an indicator of a high-fibre, low-carbohydrate diet that is expected in individuals without access to high-status foods, such as meat (19). Individuals with low socioeconomic rank are likely to suffer more bone disease through manual labor, which may result in the correlation between disease and microbiota.

In the premolar data set, oral microbiota composition was linked to the presence of bone spurs in the vertebra (Schmorl's nodes). However, the causative microorganisms for this disease are less clear, as no taxa were significantly altered when tested with the Kruskall-Wallis test and no functional differences were observed. However, the lack of clarity about the causes of these nodes in modern individuals suggests this is not an unexpected result and that the consistent patterns of taxa and functional change associated with socio-economic rank (*i.e.* poor diet and working conditions) throughout the data support the formation of such pathologies.

In this study, microbiota community structure and function were associated with different periods in ancient London and appeared to reflect a decrease in diet quality over time. Previous studies of rural populations have identified major microbiota shifts that likely correspond to dietary change during or following the Industrial Revolution (~1750) (12). This has been related to the increasing availability of processed flour and sugar across

society. The Industrial Revolution was presumed to be the second of two major alterations (the first being the Neolithic Revolution) that resulted in the modern oral microbiota. The modern oral microbiota is dominated by Streptococcus species as primary colonizers (50). In our study, individuals examined from the 1800s are likely to have experienced the effects of the industrial revolution, suggesting that more modern-like microbiota should have been observed. However, the oral microbiota shows a trend toward Methanobrevibacter species-dominated microbiota becoming more prevalent in the 1600-1900 CE period compared to earlier times (1000-1300 CE and 1300-1600 CE). Microbiota with high proportions of Methanobrevibacter species are associated with low socio-economic rank (19). This shift toward a population highly dominated by lower socio-economic groups could be explained by the alterations in the demographic structure of London. London's population almost doubled from 600,000 in 1700 to 1,000,000 in 1800 (51). Much of this growth came from an influx of rural people, particularly workers moving into the city for work. Consequently, the shift in microbiota may represent the increasing proportion of low socio-economic classes, rather than a population-level dietary shift.

This study reveals individualistic disease patterns that impacted the human oral microbiota and explores the use of dental calculus as an archaeological tool to study lifestyle alterations in ancient populations. We also identified significant impacts of different oral geographies, highlighting that future ancient calculus samples and detailed metadata must be carefully collected, documented, and analysed to accurately interpret these disease patterns. Once these biases were controlled for, we were able to identify potential links between oral microbiota and an array of oral and bone-related diseases. Controlling for bias can allow changes in community structure and function to be linked to unique factors, such as ancient disease and socio-economic rank and the associated lifestyle and diet. Large-scale sampling of a single population allows detailed information about our past to be revealed. In addition, this study lays the groundwork to build a comprehensive

understanding of oral microbiota, allowing researchers to gain individualistic information when other information is unavailable.

References

1. Rosenbaum M, Knight R, Leibel RL. The gut microbiota in human energy homeostasis and obesity. Trends Endocrinol Metab. 2015 Sep;26(9):493–501.

2. Trivedi B. Microbiome: The surface brigade. Nature. 2012;492(7429):2.

3. Li X, Wang J, Joiner A, Chang J. The remineralisation of enamel: a review of the literature. J Dent. 2014 Jun;42, Supplement 1:S12–20.

4. Iwase T, Uehara Y, Shinji H, Tajima A, Seo H, Takada K, et al. Staphylococcus epidermidis Esp inhibits Staphylococcus aureus biofilm formation and nasal colonization. Nature. 2010 May 20;465(7296):346–9.

5. Noverr MC, Huffnagle GB. Does the microbiota regulate immune responses outside the gut? Trends Microbiol. 2004 Dec;12(12):562–8.

6. Rautemaa R, Lauhio A, Cullinan MP, Seymour GJ. Oral infections and systemic disease—an emerging problem in medicine. Clin Microbiol Infect. 2007 Nov 1;13(11):1041–7.

7. Weyrich LS, Feaga HA, Park J, Muse SJ, Safi CY, Rolin OY, et al. Resident Microbiota Affect Bordetella pertussis Infectious Dose and Host Specificity. J Infect Dis [Internet]. 2013 Dec 15; Available from: http://www.ncbi.nlm.nih.gov/pubmed/24227794

8. Farrell JJ, Zhang L, Zhou H, Chia D, Elashoff D, Akin D, et al. Variations of oral microbiota are associated with pancreatic diseases including pancreatic cancer. Gut. 2012 Apr 1;61(4):582–8.

9. Grau AJ, Becher H, Ziegler CM, Lichy C, Buggle F, Kaiser C, et al. Periodontal Disease as a Risk Factor for Ischemic Stroke. Stroke. 2004 Feb 1;35(2):496–501.

10. Turnbaugh PJ, Ley RE, Mahowald MA, Magrini V, Mardis ER, Gordon JI. An obesity-associated gut microbiome with increased capacity for energy harvest. Nature. 2006 Dec 21;444(7122):1027–31.

11. Weyrich LS, Dobney K, Cooper A. Ancient DNA analysis of dental calculus. J Hum Evol. 2015 Feb;79:119–24.

12. Adler CJ, Dobney K, Weyrich LS, Kaidonis J, Walker AW, Haak W, et al. Sequencing ancient calcified dental plaque shows changes in oral microbiota with dietary shifts of the Neolithic and Industrial revolutions. Nat Genet. 2013 Apr;45(4):450–5.

13. White DJ. Processes contributing to the formation of dental calculus. Biofouling. 1991 Aug 1;4(1–3):209–18.

14. Dobney K, Brothwell D. Dental Calculus: Its Relevance to Ancient Diet and Oral Ecology. In: Cruwys E, Foley RA, editors. Teeth and Anthropology. 1986. p. 55–81. (British Archaeological Reports).

15. Weyrich LS. Evolution of the Human Microbiome and Impacts on Human Health, Infectious Disease, and Hominid Evolution. In: Gontier N, editor. Reticulate Evolution [Internet]. Springer International Publishing; 2015 [cited 2015 Oct 27]. p. 231–53. (Interdisciplinary Evolution Research). Available from: http://link.springer.com/chapter/10.1007/978-3-319-16345-1_9

16. De La Fuente C, Flores S, Moraga M. DNA from Human Ancient Bacteria: A Novel Source of Genetic Evidence from Archaeological Dental Calculus. Archaeometry. 2013;55(4):767–78.

17. Warinner C, Rodrigues JFM, Vyas R, Trachsel C, Shved N, Grossmann J, et al. Pathogens and host immunity in the ancient human oral cavity. Nat Genet. 2014;

18. Weyrich LS, Duchane S, Soubrier J, Arriola L, Llamas B, Breen J, et al. Reconstructing Neandertal behavior, diet, and disease using ancient DNA from dental calculus. Nature. 2016;

19. Farrer AG, Dobney K, Gully NJ, Weyrich LS, Cooper A. Diet driven differentiation of oral microbiota in ancient Britain. 2016;

20. Salter SJ, Cox MJ, Turek EM, Calus ST, Cookson WO, Moffatt MF, et al. Reagent and laboratory contamination can critically impact sequence-based microbiome analyses. BMC Biol. 2014 Nov 12;12(1):87.

21. Simón-Soro Á, Tomás I, Cabrera-Rubio R, Catalan MD, Nyvad B, Mira A. Microbial Geography of the Oral Cavity. J Dent Res. 2013 Jul 1;92(7):616–21.

22. Meyer M, Kircher M. Illumina Sequencing Library Preparation for Highly Multiplexed Target Capture and Sequencing. Cold Spring Harb Protoc. 2010 Jun 1;2010(6):pdb.prot5448.

23. Lindgreen S. AdapterRemoval: easy cleaning of next-generation sequencing reads. BMC Res Notes. 2012;5:337.

24. Herbig A, Maixner F, Bos KI, Zink A, Krause J, Huson DH. MALT: Fast alignment and analysis of metagenomic DNA sequence data applied to the Tyrolean Iceman. bioRxiv. 2016 Apr 27;50559.

25. Huson DH, Auch AF, Qi J, Schuster SC. MEGAN analysis of metagenomic data. Genome Res. 2007 Mar 1;17(3):377–86.

26. Caporaso JG, Kuczynski J, Stombaugh J, Bittinger K, Bushman FD, Costello EK, et al. QIIME allows analysis of high-throughput community sequencing data. Nat Methods. 2010 May;7(5):335–6.

27. Overbeek R, Begley T, Butler RM, Choudhuri JV, Chuang H-Y, Cohoon M, et al. The Subsystems Approach to Genome Annotation and

its Use in the Project to Annotate 1000 Genomes. Nucleic Acids Res. 2005 Jan 1;33(17):5691–702.

28. Muegge BD, Kuczynski J, Knights D, Clemente JC, González A, Fontana L, et al. Diet drives convergence in gut microbiome functions across mammalian phylogeny and within humans. Science. 2011 May 20;332(6032):970–4.

29. Clayton JB, Vangay P, Huang H, Ward T, Hillmann BM, Al-Ghalith GA, et al. Captivity humanizes the primate microbiome. Proc Natl Acad Sci. 2016 Aug 29;201521835.

30. Zijnge V, Leeuwen MBM van, Degener JE, Abbas F, Thurnheer T, Gmür R, et al. Oral Biofilm Architecture on Natural Teeth. PLOS ONE. 2010 Feb 24;5(2):e9321.

31. Shweta, Prakash SK. Dental abscess: A microbiological review. Dent Res J. 2013;10(5):585–91.

32. Horz H-P, Conrads G. Methanogenic Archaea and oral infections — ways to unravel the black box. J Oral Microbiol [Internet]. 2011 Feb 23 [cited 2016 Feb 17];3. Available from: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3086593/

33. Waters RC, O'Toole PW, Ryan KA. The FliK protein and flagellar hook-length control. Protein Sci Publ Protein Soc. 2007 May;16(5):769–80.

34. Josenhans C, Suerbaum S. The role of motility as a virulence factor in bacteria. Int J Med Microbiol. 2002 Jan 1;291(8):605–14.

35. Listgarten MA, Levin S. Positive correlation between the proportions of subgingival spirochetes and motile bacteria and susceptibility of human subjects to periodontal deterioration. J Clin Periodontol. 1981 Feb 1;8(2):122–38.

36. Armitage GC, Dickinson WR, Jenderseck RS, Levine SM, Chambers DW. Relationship Between the Percentage of Subgingival Spirochetes and the Severity of Periodontal Disease. J Periodontol. 1982 Sep 1;53(9):550–6.

37. Cardoso L, Fritton SP, Gailani G, Benalla M, Cowin SC. Advances in assessment of bone porosity, permeability and interstitial fluid flow. J Biomech. 2013 Jan 18;46(2):253–65.

38. Gabet Y, Bab I. Microarchitectural Changes in the Aging Skeleton. Curr Osteoporos Rep. 2011 Sep 8;9(4):177.

39. Pinhasi R, Mays S. Advances in Human Palaeopathology. John Wiley & Sons; 2008. 410 p.

40. Ak J. Osteophytes and the osteoarthritic femoral head. J Bone Joint Surg Br. 1975 1975;57(3):314–24.

41. Ortner DJ. Identification of Pathological Conditions in Human Skeletal Remains. Academic Press; 2003. 663 p.

42. Kyere KA, Than KD, Wang AC, Rahman SU, Valdivia–Valdivia JM, Marca FL, et al. Schmorl's nodes. Eur Spine J. 2012 Apr 28;21(11):2115–21.

43. Sonne-Holm S, Jacobsen S, Rovsing H, Monrad H. The epidemiology of Schmorl's nodes and their correlation to radiographic degeneration in 4,151 subjects. Eur Spine J. 2013 Mar 16;22(8):1907–12.

44. Scher JU, Abramson SB. The microbiome and rheumatoid arthritis. Nat Rev Rheumatol. 2011 Aug 23;7(10):569–78.

45. Mangine GT, Gonzalez AM, Wells AJ, McCormack WP, Fragala MS, Stout JR, et al. The effect of a dietary supplement (N-oleyl-phosphatidyl-ethanolamine and epigallocatechin gallate) on dietary compliance and

body fat loss in adults who are overweight: A double-blind, randomized control trial. Lipids Health Dis. 2012 Oct 4;11:127.

46. Garsin DA. Ethanolamine Utilization in Bacterial Pathogens: Roles and Regulation. Nat Rev Microbiol. 2010 Apr;8(4):290–5.

47. LeBlanc JG, Milani C, de Giori GS, Sesma F, van Sinderen D, Ventura M. Bacteria as vitamin suppliers to their host: a gut microbiota perspective. Curr Opin Biotechnol. 2013 Apr;24(2):160–8.

48. Segata N, Haake SK, Mannon P, Lemon KP, Waldron L, Gevers D, et al. Composition of the adult digestive tract bacterial microbiome based on seven mouth surfaces, tonsils, throat and stool samples. Genome Biol. 2012;13:R42.

49. Billing F. Focal Infection: Its broader application in the etiology of general disease. J Am Med Assoc. 1914;63(11):5.

50. Kolenbrander PE, London J. Adhere today, here tomorrow: oral bacterial adherence. J Bacteriol. 1993 Jun;175(11):3247–52.

51. Black J. London: a history. Carnegie; 2009.

Chapter V

Incorporating science communication into standard group research practice

Statement of Authorship

Title of Paper	Incorporating science communication into standard group research practice
Publication Status	Unpublished and Unsubmitted work written in manuscript style
Publication Details	In preparation for: PLoS Biology: Community Page

Principal Author

Name of Principal Author (Candidate)	Andrew G. Farrer		
Contribution to the Paper	Co –instigated and maintained the outreach programme, including editing of content and preparing the promotional posts analysed in the study. Conceptualized the framework. Collated and analysed data. Wrote the manuscript and generated the figures.		
Overall percentage (%)	75%		
Certification:	This paper reports on original research I conducted during the period of my Higher Degree by Research candidature and is not subject to any obligations or contractual agreements with a third party that would constrain its inclusion in this thesis. I am the primary author of this paper.		
Signature		Date	29th August 2016

Co-Author Contributions

By signing the Statement of Authorship, each author certifies that:

i. the candidate's stated contribution to the publication is accurate (as detailed above);

ii. permission is granted for the candidate in include the publication in the thesis; and

iii. the sum of all co-author contributions is equal to 100% less the candidate's stated contribution.

Name of Co-Author	Maria Lekis		
Contribution to the Paper	Co –instigated and maintained the outreach programme, including editing of content and preparing the promotional posts analysed in the study Discussed and edited the manuscript.		
Signature		Date	29th August, 2016

Name of Co-Author	Laura S. Weyrich		
Contribution to the Paper	Discussed and edited the manuscript		
Signature		Date	7th October 2016

Name of Co-Author	Alan Cooper		
Contribution to the Paper	Edited the manuscript		
Signature		Date	2nd September, 2016

Incorporating science communication into standard group research practice

Authors:

Andrew G. Farrer[1*], Maria Lekis[1], Laura S. Weyrich[1+], & Alan Cooper[1+]

Affiliations:

[1]Australian Centre for Ancient DNA, School of Biological Sciences, University of Adelaide, Adelaide, South Australia, Australia

[+]These authors contributed equally to this work

Corresponding author:

[*]andrew.farrer@adelaide.edu.au

Abstract

Science communication provides society with direct and comprehensible access to research, and academics profit from an increased public profile that provides professional and personal benefits. Here, we provide a framework to establish a free and sustainable communication programme within a research group, which minimizes an individual researcher's time investment. We draw upon our experience at the Australian Centre for Ancient DNA, University of Adelaide, Australia to outline key considerations in planning, producing, and promoting communication efforts. We also demonstrate outcomes with direct research relevance to indicate the power of science communication for engagement and as a research tool.

Introduction

Science communication is the core of informing and engaging non-specialists on the state and relevance of research (1), and has been shown to have professional and personal benefits to researchers (2,3). However, time spent producing, promoting, and evaluating content is commonly bemoaned as lost research time (4). Here, we present a framework for sustainably incorporating science communication into the standard operating practice of a research group using free resources and team contributions to minimize workload (Fig. 1). This method has directly benefitted our research at the Australian Centre for Ancient DNA (ACAD), University of Adelaide, Australia through the establishment of research and public collaborations, attraction of new staff and students, and promotion of research. By involving the whole research group, this framework also offers supported opportunities for students to develop presentation and editing skills, and, through volunteer roles, skills in team and digital management. Based on our experience, we outline three areas of consideration: planning and building a science communication programme, producing and promoting content, and evaluating success.

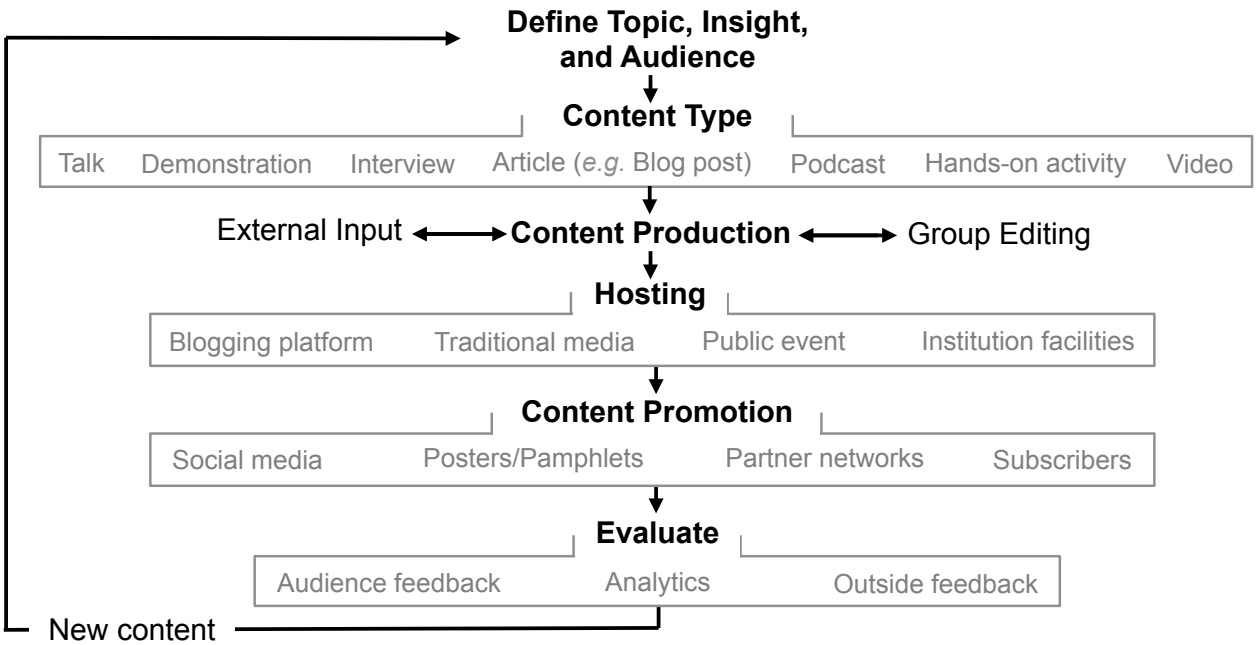


Figure 1: *Workflow for producing, promoting, and evaluating content.*

1. Planning and building a science communication programme

It is necessary to consider a trio of objectives when setting up a science communication programme: desired outcomes, target audience, and content hosting. Our initial phase of development focused on group discussion of these objectives, before a group of volunteers set up the required online infrastructure. Volunteer roles then focused on researching and promoting communication opportunities to group members, and establishing and maintaining online platforms. Volunteer roles particularly suited members interested in developing skills in team and media management. However, this framework relied on whole group participation, which was made an expectation, to achieve the distribution of workloads.

Discussion of desired outcomes focused on previous positive experience with media and citizen science projects to identify successful strategies. Following this, we defined our primary goal as maintenance and growth of an existing profile to engage the public with our research, justify government funding, and encourage citizen science. ACAD is a 90% grant-funded research group of 25 – 30 active researchers with dedicated administrative support. The majority of members are PhD candidates and early career researchers. Consequently it was key that the programme allowed all members to contribute in order to develop their communication skills: a key attribute required of academics.

Appreciation of the target audience is a vital aspect of successful communication. As a group, we agreed that a high school level of science understanding was suitable, given our broad target audience. Skills in assessing audience knowledge level and interest were introduced to the group by partnering with science communication initiatives, such as the international Children's University programme. Many institutions have science communication initiatives that can provide expert advice, and offer opportunities to present in a range of content styles to pre-existing audiences. We recommend partnering with such initiatives to gain skills and to form and expand an audience base.

Hosting platforms, where content can be made public, should also be selected to limit effort in curating and posting content and to maximize ease of use by the audience. We initially provided written content on a blogging platform (Wordpress.com). Free blogging platforms, such as Wordpress.com and Blogger, provide customizable templates, store text and image, and allow simple embedding of audio and video players from other hosting sites (e.g. YouTube or Vimeo). This flexibility allowed us to incorporate multiple media without our audience needing to follow us across multiple websites. ACAD platforms were made relatable by using group photographs over more generic laboratory or logo visuals, and allowed reader subscription to promote audience retention. Social media were used to promote our blog. Different social media networks are home to different audiences and are effective at targeting and expanding an audience base (5). Since our group aims for broad appeal, ACAD promoted its content on two of the most popular social networks: Facebook and Twitter (6). However, if the target audience is more specific, other networks may be more suitable (e.g. LinkedIn, for business). We integrated our communication programme into our research presence by embedding social media feeds and links into our university-based website using code automatically generated by Facebook and Twitter.

2. Producing and promoting content

Content should be a self-contained narrative that communicates novel, relatable insights into research and researchers (1). People do not relate to abstract concepts effectively, and are more engaged by discussion which is emotionally engaging (7). This does not mean that science communication should be excessively emotive, but that content must show a human element (excitement, concern, etc.) for the topic, alongside factual information (8). To achieve this, content can describe studies and the personal inspiration behind them, use specialist knowledge to address social themes, or promote events, awards, and graduations (indicating the human and emotional element to research groups). To limit production commitment

and enforce careful consideration of how to present all necessary concepts, we used limits of 1,000 words or 10 minutes (with accompanying images and graphics). This also made content accessible and engaging to a browsing audience. Authors may include third-party information, graphics, photographs, or videos to help convey concepts. However, when doing so, both author and the group must be aware of the privacy, copyright, and plagiarism laws associated with reusing third-party material. Obtaining permission or identifying CreativeCommons licenses should be completed in advance and the required attribution given. Failure to do so harms the reputation of the group, so appropriate acknowledgment of intellectual property must be taken seriously. Group, volunteer, and/or partner editing is a key element of this framework, providing a knowledge base and test audience for guidance when producing content (Fig. 1).

Group editing ensures maintenance of relevance, interest value, and quality. Communication with a lay audience is difficult (9), and appropriate editing for effective communication is typically the greatest hurdle for both authors and editors. Opportunities to give and receive edits allow members to gain skills from one another, which is particularly relevant when some members have experience with partner initiatives (e.g. Children's University). Content editing by a putative member of the target audience is ideal, highlighting difficult areas that fellow researchers may not appreciate (9). At ACAD, our administrative coordinator volunteered for this role, and noted that editing requirements lessened as members submitted more content, indicating authors' improved confidence and efficiency over time.

Once production is complete, and the content released on a hosting platform, it should be promoted in order to maximize viewership and impact. Social media are a powerful tool for content promotion, providing access to a large range of social and interest groups. However, social media networks are highly structured, and groups of individual users can become distinct and isolated from posts and conversations occurring outside their primary interests (10). These groups can be engaged using existing functionality and by tailoring the discussion of science to apply to specific interests or needs.

Hashtags (keywords), tagging (alerting specific users to the content), and incorporation of key questions, call-to-action statements, images, and video should be used to target and inform audience members of the relevance of the promoted content to their interests. For example, ACAD used hashtags to direct posts to an interest group (e.g. #forensics, a group that may not recognise the relevance of our work from the centre title), and used tagging to alert people or groups (e.g. our institution, @UniAdelaide) of relevant posts and to converse with audience members. However, while bringing content to the attention of a variety of interest groups is beneficial and necessary for audience growth, the rate of posting should be limited so your content does not become spam. Engagement rates should be monitored to identify posts that are ignored, and additional consideration should be given to determine why specific posts fail to engage the target audience. Additionally, tagging an institution should not replace alerting the media office to new research or events in advance. Finally, use of social media management tools, such as Hootsuite and Tweetdeck, can reduce monitoring times by allowing simultaneous organization of multiple accounts and scheduling of posts in advance.

3. Evaluating success

Assessment of success in meeting agreed-upon objectives of the outreach programme can be carried out using freely available metrics on audience engagement. Quantifying audience interaction with platforms and content gives demonstrable measures of success and provides important feedback to improve the production cycle. It is important to convey such analytics to the whole group, so that everyone can be motivated by the impacts of their work (9). At ACAD, our main objective was to engage a broad audience. Consequently, we tracked how all audience members were viewing and interacting with our content. Here, we present data from ACAD during 2015 to demonstrate the reach of our programme and lessons from our experience.

Contrary to popular opinion, use of social media does not have to be intense to grow and engage an audience. Throughout 2015, ACAD made 42 posts ($\bar{x}$ = 0.8 per week) on Facebook and 104 posts ($\bar{x}$ = 2 per week) on Twitter. As shown in Fig. 2, this was accompanied by a steadily growing number of followers (i.e. audience members signing up to the ACAD profile). Increasing the total number of followers also increases the potential number of people who can engage with the content (e.g. click to expand the post, view image, or follow a URL). This signal of active consideration is important to identify content that engages the audience. However, observations have shown that engagement rates drop as follower count increases (11). At ACAD, we set our goal engagement rate at > 5% of the number of followers, though > 1% would have been appropriate given our follower count (11). Analytics revealed that 98% of ACAD Facebook posts and 89% of ACAD Twitter posts were above the 1% threshold in 2015. However, only 29% of Twitter posts passed our 5% goal, in contrast to 86% on Facebook (Fig. 2). Consequently, we are working to improve our Twitter engagement, and are using analytics to identify the audience preferences specific to this platform.

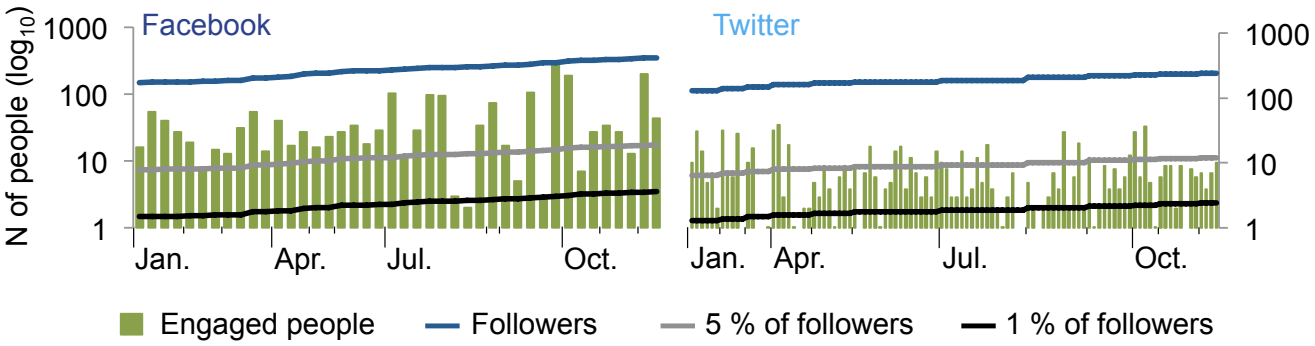


Figure 2: *Growth of follower count (blue line) and engagement with the 42 Facebook posts (left) and 104 Twitter posts (right) placed during 2015 in comparison to the benchmarks of 5 % (grey line) and 1 % (black line) of follower count. Each green bar represents the number of people engaged with a single post.*

Targeting content to the relevant audience can maximize engagement. However, audience preferences may contrast with the goals of the group. Analysis of preferred post types on ACAD's Facebook page

showed that this audience preferred social post types (Fig. 3). While we wanted to engage the public with our researchers, our primary goal was to focus on research topics, and deviating from this goal would not have achieved the desired research image. Thus, we did not alter the balance of our post types, despite the potential to increase our engagement metrics by including more purely social posts. In contrast, we altered our strategy on Twitter, as it became apparent that the audience did not favour social posts, and preferred information regarding opportunities (e.g. the chance to work or study at ACAD) and alerts of new blog posts. Such analysis improved management and planning of posts for each platform, particularly by saving time on formatting ineffective posts for Twitter. While social media was a key area of active promotion, we noted that individuals who found content via internet searches and promotion on non-ACAD websites also engaged with ACAD content.

Figure 3: *Normalized audience engagement with different post categories on ACAD Facebook (left) and Twitter (right) profiles in 2015. Larger font size relates to an increased average engagement per post. Post categories: "Blog": Link to ACAD hosted content, "Social": ACAD social events, graduations, and member recognition, "Opportunity": Chance to work or study at ACAD, "Publicity": Appearance in media, and public or school talks, "Research": Published paper, conference presentation, or grant success, and "Resource": Service provided by ACAD (e.g. OAGR, workshops, and training seminars).*

Once released, outreach content forms an online legacy, which can continue to engage the public some time after the date of content release. In total, the ACAD Wordpress.com site gained more views from web searches than from social media referrals. Although the highest views per day were associated with ACAD social media posts, audience engagement via social media was short-lived. We observed a three-day lifespan (from the time of posting) for views coming from our social media links. However, views from internet searches continued indefinitely. This legacy allows for future resurgence of interest when a new user redistributes content. For example, a YouTube video received 1,000 views in a single week, nearly seven months after it was initially posted, because a new user promoted the content on a third-party website. Hence, the legacy of online content continually draws new viewers, contributing to science communication with no additional effort from the group.

Finally, we note a range of research outcomes that resulted from ACAD's increased visibility online that were not captured by analytics. These included public donation of over 400 human DNA samples for forensic war dead identification projects, and invitations to speak in schools and feature on popular science television and radio shows. It is also important to note that our communication efforts also engaged our peers, being cited as an element that inspired applications for research and student positions. It also directly resulted in an international collaboration to share genomic data for the new, open access Online Ancient Genome Repository (OAGR) (12). These outcomes, alongside reporting of analytics, are now incorporated into research grant applications as an effective way to promote research. Further, researchers have noted that involvement in science communication was both informative and enjoyable.

Conclusion

Academic researchers are increasingly engaging with science communication tools and sharing their research passions with the public. Here, we have presented a framework that incorporates science

communication into the standard operating practice of a research group, and demonstrated that engagement and enthusiasm from public and peers has resulted in significant research benefits. Sharing of responsibilities across multiple members of a research group, and monitoring and analysing audience engagement, reduces workload on any single individual. Furthermore, the framework provides researchers and students with opportunities to actively build communication and editing skills with the support and advice of the research group and partner initiatives. In recognition of the research benefits gained, science communication should not be viewed as an addition to academic work, but as a powerful tool within it.

References

1. Burns TW, O'Connor DJ, Stocklmayer SM. Science Communication: A Contemporary Definition. Public Underst Sci. 2003 Apr 1;12(2):183–202.

2. Phillips DP, Kanter EJ, Bednarczyk B, Tastad PL. Importance of the Lay Press in the Transmission of Medical Knowledge to the Scientific Community. N Engl J Med. 1991 Oct 17;325(16):1180–3.

3. Eysenbach G. Can Tweets Predict Citations? Metrics of Social Impact Based on Twitter and Correlation with Traditional Metrics of Scientific Impact. J Med Internet Res. 2011 Dec 16;13(4):e123.

4. Ecklund EH, James SA, Lincoln AE. How Academic Biologists and Physicists View Science Outreach. PLOS ONE. 2012 May 9;7(5):e36240.

5. Osterrieder A. The value and use of social media as communication tool in the plant sciences. Plant Methods. 2013;9:26.

6. Bik HM, Goldstein MC. An Introduction to Social Media for Scientists. PLoS Biol. 2013 Apr 23;11(4):e1001535.

7. Roeser S. Risk Communication, Public Engagement, and Climate Change: A Role for Emotions. Risk Anal. 2012 Jun 1;32(6):1033–40.

8. Fiske ST, Dupree C. Gaining trust as well as respect in communicating to motivated audiences about science topics. Proc Natl Acad Sci. 2014 Sep 16;111(Supplement 4):13593–7.

9. Brownell SE, Price JV, Steinman L. Science Communication to the General Public: Why We Need to Teach Undergraduate and Graduate Students this Skill as Part of Their Formal Scientific Training. J Undergrad Neurosci Educ. 2013 Oct 15;12(1):E6–10.

10. Bruns A, Burgess J, Highfield T. A "Big Data" Approach to Mapping the Australian Twittersphere. In: Arthure P, Bode K, editors. Advancing

Digital Humanities: Research, Methods, Theories. Palgrave Macmillan UK; 2014.

11. Leander M. What is a good Facebook engagement rate? [Internet]. Michael Leander - Award Winning Marketing Speaker & Trainer. 2011 [cited 2016 May 16]. Available from: http://www.michaelleander.me/blog/facebook-engagement-rate-benchmark/

12. Online Ancient Genome Repository [Internet]. [cited 2016 Jul 28]. Available from: https://www.oagr.org.au/

Discussion

Overall summary and significance

This thesis aimed to identify the structure of oral microbiota within a historical population and the aspects of historical life that defined and altered this microbial community. In addition, it presents methods to optimize the recovery of microbiota from dental calculus, allowing for the high-resolution studies of a population utilized here. The data presented in this paper provides important insights into the association of microbiota and human lifestyle. These results are relevant to medical and archaeological fields for progressing the understanding and application of microbiota.

In Chapter I, I explored the use of aDNA analysis of DC as an archaeological tool. I discussed the currently known factors that define microbiota structure and indicate how microbiota composition can be used to identify past lifestyle traits. This review suggested how studies of populations with detailed historical records, such as the British population analysed in the following chapters, could be used to calibrate microbiota studies for application to poorly understood populations. In addition, this chapter discussed the considerations needed for aDNA studies of dental calculus to be completed successfully, including assessments of environmental and laboratory contamination, DNA marker type, and sequencing technology.

Chapter II addressed the issue of environmental contamination on the outside of ancient dental calculus samples. I explored multiple decontamination protocols to remove microbes and microbial DNA from the surface of dental calculus that had been introduced from the preservation matrix (*i.e.* soil), handling during excavation and storage, or the storage environment. This is a critical issue, as incorporation of contaminant DNA into analysis can mask endogenous signal, or generate a signal that results in misinterpretation of the data. In addition, removal of contamination prior to sequencing maximizes sequencing capacity and thus sequencing quality of endogenous DNA. This is the first study to test decontamination methods on dental calculus, including a comparison of the two published protocols to date. In this analysis, I indicated that UV irradiation (30 minutes, each side)

and bleach submersion (5%, 3 minutes) was the most efficient at removing identifiable contamination. This method of decontamination was then applied in Chapters III and IV.

Chapter III applies the principles and observations in Chapters I and II to reveal two distinct microbiota present in an ancient population. In addition to the decontamination method introduced in Chapter II, I also develop a novel bioinformatics filtering step that compares samples to negative controls and applys stringent, standard controls for microbiota. This study of 33 ancient samples from Great Britain revealed two distinct groups of unidentified diversity that were present through British history from pre-Roman (pre 43 CE) to Early Victorian (~1850 CE) periods. This was also the first ancient DNA study to examine microbial function within the human microbiota and identify that metabolic functional data related to diet primarily drive variation between the two groups. The impact of diet on the microbiota also suggests that socio-economic rank, which defined access to food in past populations, may play a significant role in determining microbiota community structure within the population. The higher socio-economic rank group was defined by an increased microbial metabolism for meats, sugar, and dairy foods, while the second group was associated with dietary food sources linked to lower socioeconomic ranks.

Chapter IV was able to examine changes beyond diet and identify microbiota associations with oral and systemic disease and an alteration through time from the 1000's CE to the 1800s CE. This study was conducted in collaboration with the Museum of London where I collected 161 dental calculus samples from a single city, producing the largest study of historical microbiota from a single location to date. The decontamination protocols put in place during Chapter II and III allowed me to analyse 128 robust samples. This large sample size allowed me to examine and identify the contributions of oral geography to sample variation. This is a major issue that needs to be addressed in all future studies to ensure genuine interpretation of ancient and oral microbiota data. Once oral geography was considered during data analysis, a detailed study examining factors that alter

microbiota was completed. This high-resolution data identifies several novel facets of disease, diet, and socio-economic status that drive microbiota diversity. Changes in these facets were linked to a demographic shift in London, as the population grew with a bias towards an increase in lower socio-economic classes, rather than an overall shift in microbiota among the population.

Finally, Chapter V addresses the issue of making research accessible to academics and the public. This chapter outlines and discusses a framework for (large) research groups to produce a science communication programme that will benefit the group, the individual researchers, and the research, while minimizing the time input for individual researchers. This kind of framework is critical for the future of science outreach to maximize the benefits of science communication to research. I demonstrate that public facing communication has provided the Australian Centre for Ancient DNA, University of Adelaide, Australia with multiple direct benefits to research by creating awareness that sparked collaboration between the Centre and its peers and the public. This paper indicates the potential of public facing science communication to breach the barriers between disciplines and initiate academic communication and collaboration.

In summary, this thesis demonstrates the capacity, and thus potential, for dental calculus to be used to detect fine scale changes in the bacterial community in relation to diet, disease, and time. It provides the first insights into the ancient microbiota within a single population through time, producing the initial observations and data sets from which future studies can expand and use as a baseline for comparison. In the early stages of historical microbiota research, this thesis specifies three key recommendations for study methods to optimize future microbiota recovery and avoid bias. In this discussion chapter, I summarise and explore the significant outcomes of this research and look forward to future research avenues and the practices that promote communication between academic fields and with the public.

Oral microbiota are defined by individualistic factors

In this thesis, the diet and disease state of an individual were identified as the key defining factors of ancient oral microbiota. Previous studies of ancient microbiota have suggested that geography and population culture drove microbiota diversity. For example, Adler *et al.* demonstrated that individuals from the same archaeological site and culture were more similar (1). However, these observations were made when studying large time frames and geographies with a limited number of samples. In Chapter III, I demonstrate that samples do not follow this pattern when the geographical variation is reduced to a single country (Great Britain) and the time frame is reduced to ~ 2,000 years, rather than ~ 8,000 years. This was repeated again within the higher resolution data set of Chapter IV, which came from seven nearby, yet different, archaeological sites, in London, UK. This indicates that geography or shared environment of living location were not the major definers of microbiota in ancient populations.

Microbiota function is associated with diet

In living individuals, changes in diet can alter the microbiota rapidly (2). These links between diet and microbiota have led some researchers to suggest that dietary alterations may have played a key role during human evolutionary history by shaping the modern microbiota. For example, the gut microbiota of non-human primates becomes human-like following a shift from their natural diet to one more closely resembling that of modern humans (3). Previous studies of ancient DC have also identified diet as a significant driver of microbiota structure (1,4,5). However, limited sample availability in ancient studies could only predict large dietary changes by comparing different cultures (in different geographies) over time, which would also have been linked to a wide range of cultural and social differences. While direct inference of diet was possible in more recent studies due to the recovery of eukaryotic DNA (4,5), these studies still compared samples across expansive periods and lacked resolution to understand how a single population would be affected by the observed variations in diet. In Chapter III

and IV, examination of microbial metabolic functions demonstrated that several specific dietary components varied among the population. These functions included signals of meat, fibre, and carbohydrate (including dairy) metabolism. Diet explained the differentiation of two taxonomically distinct microbiota groups present within the ancient British population. This is the first ancient DNA study to use microbial metabolic function as a proxy of diet (rather than direct evidence from eukaryotic DNA) and provides new insights into the nutrition available in ancient diets. Notably, this information correlated with known dietary differences between socio-economic groups. Of the two groups of microbiota that were identified, one group was shown to have increased metabolic functions associated with meat and sugar metabolism, including an increased capacity to break down sugars indicative of dairy products. In the past, diets of this nature were associated with higher status individuals (6). The second microbiota group contained a functional profile indicative of the poor quality diet of a lower class individual, which included higher fibre and non-sugar carbohydrate metabolism.

Identification of these specific dietary elements as drivers of microbiota structure has not yet been demonstrated in modern humans. While the impact of diet has been studied, many of these studies used 16S rRNA gene sequencing and thus cannot examine microbial functions directly. In the few studies where functional information has been linked to diet, focus has been placed on changes as a result of experimental dietary shifts in relation to the production of disease associated metabolites (7,8). However, functional profiles associated with plant-based and meat-based diets have been identified in modern individuals following short term, major, experimental dietary shifts in humans (2). In both previously published modern human data and the ancient humans studied here, the functions used to differentiate between the levels of meat and fibre in the diet came from studies of mammalian, non-human carnivores and herbivores, and captive and wild non-human primates, respectively. This indicates how little is known about the specific impacts of diet on human microbiota diversity and function. In this thesis, I directly identify, and demonstrate distinct

functional profiles related to dietary factors, providing the first study to infer diet from metabolic functions. However, future studies should investigate functional profiles further, with the aim of identifying other distinguishing, human specific, functional factors across multiple populations and host species. Additional functions linked to diet that should be explored further in archaeological individuals include functions associated with marine resources, the presence of fermentation products, and characteristic compounds of differing grain crops (such as flavonoids (9)). These factors need to be assessed in controlled experiments and could not be assessed within the British population because the level of dietary information is not sufficient in archaeological samples. However, once identified, the functions could be assessed in historical populations to provide detailed inferences of diet.

Inference of diet using microbial metabolic function required a much lower sequencing depth than assessment of eukaryotic DNA from dental calculus (~ 120,000 reads per sample compared to 45 million reads (5)). This reduces the cost of sequencing per sample, allowing larger-scale surveys to be completed. Savings in sequencing costs could be directed toward isotopic and proteomic analyses of dental calculus. Together, these methods could provide broad and specific dietary information (such as animal tissue or C_3/C_4 plant information (10,11)), indication of the geographical origin of the human host (12), and human health (by identification of immune proteins (5)). However, isotopic and proteomic studies of dental calculus are even more recent fields than microbiota analysis. Proteomic analysis of diet has, to date, only been applied to milk protein (11), while isotopic analysis requires calibrating against traditional methods (e.g. collagen) to assess what information can be inferred (13). Further studies are required to refine and streamline the analysis, and create technical protocols that allow easy integration of isotopic, proteomic, and DNA assessment within the same individual or sample.

Oral Disease

Previous studies identified key oral pathogens in ancient individuals(1,4,5). However, they did not look at overall community shifts in association with oral disease. In Chapter III, microbiota composition was not associated with the presence or absence of oral pathogens such as the Red Three Complex (*Porphyromonas gingivalis*, *Tannerella forsythia*, and *Treponema denticola*), *Streptococcus mu*tans, or *Fusobacterium nucleatum*. However, Chapter IV identified links between abscesses and microbiota community structure, including increases in the known abscess pathogens: *Prevotella* species and *Streptococcus* species. This is the first time that an ancient oral pathology has been linked to large-scale changes of the microbiota structure. Dental abscesses are known to be a polymicrobial disease (14), and several modern studies using metagenomic tools have begun to report the polymicrobial incidence of dental caries (cavities) or periodontal disease (gum disease) (15–17). These studies and the observations made here, support whole community composition being related to pathology It may be that novel oral pathogens, or unique combinations of microbiota, are also linked to abscesses in ancient populations. Future studies should look to identify functional traits present in modern and ancient abscesses and other oral diseases to identify potential microbes that caused ancient diseases.

Systemic Disease

The relationship between systemic disease and the oral microbiota is likely to be complex, as associations can be due to both direct and indirect factors. Direct associations may be due to the entry of oral microbes into the bloodstream, which has been associated with various systemic disease, such as arthritis, cancer, and mental disorders (18–20). Alternatively, systemic disease within the body may trigger or exacerbate other diseases (*i.e.* inflammation in the mouth or gut (19)), altering the body and, thus, the microbiota. However, a cultural or environmental variable may impact both without any direct link between the disease and the microbiota. For example,

a person of high socio-economic rank with a rich diet is more susceptible to diseases such as gout (21). Similarly, a rich diet will alter the microbiota structure (as demonstrated in Chapter III) independently of the presence of the gout. To date, ancient DNA analyses of microbiota have not identified a link between oral microbiota and systemic disease. However, the data set used in Chapter IV provided the resolution and metadata to test these associations. From the paleopathological data, osteoarthritis and identifiable blood disorders were considered the most likely diseases to be linked to ancient microbiota. Osteoarthritis has been associated with modern oral microbiota (22), and blood disorders are likely to impact the oral environment and influence inflammation in bone tissue. However, in Chapter IV, I identify that ancient microbiota were linked to the presence or absence of two joint diseases (porosity and lipping), periostosis (inflammation of the outer bone tissues), and Schmorl's nodes on the vertebra. This is the first evidence of an association between ancient microbiota and systemic disease. While the specific mechanism that links the microbiota and these diseases remains unknown, the microbial functional shifts associated with these diseases indicate that individuals had poor diets and were likely of lower socio-economic class. It is plausible that the link between microbiota and these diseases may not be causal, and may be correlative. For example, stress on bones, which may have contributed to these diseases, are more likely in the lives of poorer working class individuals. Previous studies lacked the resolution to identify these factors, likely due to the large geographically spread, the limited sample number, and the biasing contributions of oral geography. However, the observations of Chapter IV are initial and require further study to confirm. It is of particular importance to examine multiple dental calculus samples from a single individual to explore whether different teeth respond differently to the same disease state. Despite this, the link between oral microbiota and systemic health indicates the potential to examine the health status of past individuals using dental calculus and demonstrates the complex array of factors that drive microbiota variation that need to be considered in modern medical research.

Overarching drivers of microbiota in ancient Britain

The links between microbiota and both diet and disease suggests that socio-economic status is the overall driver of microbiota structure in ancient Britain. This is a finer level of social structure than has previously been identified using ancient dental calculus, and the appearance of microbiota groups representing upper and lower socio-economic groups at multiple archaeological sites highlights the significance of a suite of individual factors, such as lifestyle and disease, on microbiota diversity. This has drastic implications for modern medical studies; all factors of a person's life must be addressed when looking to assess or restore a healthy microbiota state. The results of this thesis are also an important archaeological finding that indicates the detailed lifestyle information that can be derived from ancient dental calculus. Future studies should explore historical individuals with extensive information regarding their social rank to assess the associations identified in this study and to identify the socio-economic status at which one type of microbiota is maintained over the other. A similar study could also utilize known family groups to assess how microbial inheritance from the mother obscures the signals linked to socio-economic status.

It is clear that microbiota are shaped by a complex series of drivers, as demonstrated in this thesis, and that further study is required to fully understand the precise nature of each interaction and to identify further drivers. However, several drivers can be ruled out. Notably, there are no significant correlations with age or sex. In modern hunter-gatherer individuals, microbiota have been associated with sex (23). However, this has not been seen in modern individuals or here, in Chapters III and IV. Sex divisions were associated with division of labour and the resulting differential access to resources. The results here indicate that resources were not sufficiently different between men and women in ancient Britain to alter the microbiota. It remains unknown whether sex differences were present in Europe and, if so, when they disappeared in European populations. Microbiota have also been shown to alter with age (24,25), although the age groupings focus on more extreme age differences than could be applied with

the archaeological data (*i.e.* individuals in their 30's compared to individuals in their 60's). Consequently, age related changes might have been too fine to be detected in these studies.

Changes through time

Microbiota are known to alter over time. Human microbiota have also diverged dramatically from the Great Ape ancestral community, potentially linked to dietary changes (3,26), and human microbiota have differentiated significantly within themselves, likely due to cultural differences between industrialised and non-industrialised lifestyles (1,4,23,27,28). However, there were previously no studies with sufficient resolution to identify the specific drivers of change through time. This thesis addresses this issue by presenting a continuous transect from the pre-Roman to Early Victorian period (pre-43 CE to ~1850 CE) in Great Britain. The combined data set of Chapter III and IV examines the impacts of events outside the agricultural and industrial revolutions. The data of this thesis suggest that the oral microbiota and its population level diversity in Great Britain were stable through time. This is unexpected, as the history of this period had many events that were predicted to impact the microbiota, as outlined in Chapter I. These included invasion, war, human migration, trade, and disease epidemics. Within Britain, I hypothesised that the Viking invasions during the 900s CE and the Norman invasion of 1066 CE would have caused change due to the introduction of novel microbiota into the population. I also hypothesized that major disease events, most notably the Black Death (1348 CE), also would have influenced microbiota diversity at the population level. However, these factors did not correlate with microbiota change. The two groups of microbiota identified in Chapter III were present in all periods, including the additional periods covered by Chapter IV. There are several possibilities as to why the hypothesized events did not produce detectable signatures within the microbiota data presented here. Firstly, there may be a lack of samples or sequencing data to resolve the shifts that occurred.

Thousands of samples or more detailed species identifications may be needed to identify specific events in the microbiota, such as the Viking or Norman invasions. With increased sampling, Chapter IV was able to resolve some of the factors (individual disease) that were undetectable in Chapter III, suggesting that alterations linked to invasions may be detectable with greater sampling depths. Despite this, the analysis generally suggests that there were minimal differences between the invaders and locals during invasions, and that any impacts were minimal in comparison to diet and disease. To further investigate the impact of invasion, large studies need to focus on single periods and target documented burial sites used before, during, and after the major invasion period. However, it is also plausible that these events did not have an impact on the microbiota. For example, Chapter IV covers the period of the Black Death, yet no microbiota alterations linked to this infectious disease were observed. While filtering may have reduced the sensitivity to an extent that the shift could not be detected, the acute disease may have also occurred too rapidly to impact microbiota composition within the slowly forming calculus sample. However, significant shifts in microbiota through time were observed in 300 year blocks, which split the data into pre-1300, 1300-1600, and post-1600. Given that the bubonic plague was present in Great Britain from the first incidence in 1348 CE (Black Death) until 1666 CE (Great Plague) (29), it is tempting to link this shift to a pre-plague microbiota, a plague exposed microbiota, and a post-plague microbiota. However, this time frame covers many other historical events. For example, the 1300 CE shift may be linked to the poor harvests and food shortages associated with the Little Ice Age, and the 1600 CE shift to the cultural impacts of the Reformation or the Great Fire of London. Equally, as suggested in Chapter IV, this change through time may be an artefact of the changing demography of London as the population grew. These potential explanations highlight the dangers of correlating observed biomolecular shifts with historical events. To ascertain the impact of the Black Death and the other historical events noted, specific studies must be

conducted, targeting specific sites and utilising samples and analyses to directly analyse these events.

The microbiota structure in ancient British calculus was maintained until the end of the time transect in Chapter IV. This was unexpected, as the industrial revolution was previously reported to have shifted the modern microbiota community structure (1). While the benefits of the industrial revolution would have taken time to reach the majority of the population (29), this data suggests that the impacts of industrialization did not affect the microbiota throughout the population until after 1850, and potentially the 1900s. Consequently, the data in this thesis indicate that the transition to the modern microbiota structure was later, and therefore even more rapid than expected. The stability of the microbiota through the preceding 2,000 years, and potentially longer, indicates how extreme the cultural changes must have been to trigger the formation of a distinct microbiota structure in modern individuals. Given that 66% of the population were part of the second microbiota group in Chapter III, and 90% in Chapter IV, the majority of the population experienced a rapid alteration, which makes the modern microbiota dysbiosis even less surprising. The 1900s include a range of developments that likely impacted the microbiota, such as antibiotics, food additives, further heavy metal contamination (lead), the first and second world wars, and globalization. While the data presented here provides a baseline for further studies and indicates the degree of change, future studies need to assess microbiota throughout the 1900s to identify how and why the transition to the modern microbiota began.

This thesis demonstrated that the population level structure of the microbiota was stable through time and was largely robust to major changes like invasion, disease epidemics, or other factors during the ~2,000 year period studied. This work also indicates that the transition from the historical microbiota structure to the modern structure began later than previously reported. Medically, this demonstrates how significant recent events were in the history of our microbiota, and how the modern “healthy” microbiota is unlikely to represent a healthy, natural state. However, this data cannot

indicate the health status of historical microbiota, and it should not be assumed that the ancestral (or historical state) was necessarily healthier.

Working with ancient DNA

Throughout this thesis, three key technical issues and protocols were discussed. First, Chapter II examined the control of environmental contamination. This was followed by Chapter III, which described a method to utilise negative controls to identify low quality samples prior to filtering of contaminant species. Finally, Chapter IV identified the risks of sampling bias (*i.e.* oral geography) by creating false signal within the data set. In future studies, each of these issues will require consideration, both at the time of sampling (*e.g.* oral geography) and during laboratory procedures (*e.g.* environmental and laboratory contamination.

Oral microbiota are known to vary across the different tooth types and surfaces (30). However, this has not previously been analysed or controlled in ancient DNA studies. In chapter IV, the association between the microbiota population structure and oral sampling site was tested. This revealed that a large percentage of the variation present within the data (~44%) was linked to tooth type (*i.e.* molar, premolar, canine, or incisor) and tooth surface (*i.e.* buccal, lingual/palatal, or interproximal). Without accounting for this variation, associations with disease and time were masked. Consequently, this is a major issue that must be accounted for in future studies, particularly in studies exploring fine scale variation between individuals, as done in this thesis. Importantly, observations in Chapter III were present when only molar teeth were analysed, demonstrating that the patterns identified are genuine. Failure to account for oral geography in future studies will include, at best, variation that masks signal or, at worst, the introduction of a false signal that appears to be cultural or environmental. To account for this variation, calculus samples should be collected from a single oral location, and where possible, detailed metadata should be kept on the specific location the sample was taken from (*i.e.* tooth type, number,

surface, mandible/maxillia, left/right, etc.). Further, previous studies have pooled dental calculus taken from multiple teeth (from the same individual) prior to DNA extraction (5). The findings within Chapter IV indicate that this method should not be used as it mixes different biological signals. If multiple oral locations are examined, the data needs to be compared to sampling metadata and filtered to remove confounding variation. Consequently, obtaining comparable microbiota from different individuals requires care during sampling, laboratory, and analysis stages.

Environmental contamination must also be accounted for to limit false positive results. Environmental contamination is any microbial DNA that is present on the dental calculus having being deposited post-mortem, and includes DNA from the microbial community that degrades the human body, the matrix the remains are preserved within (*e.g.* water or soil), handling by excavators, museum curators, other researchers, and the storage environment. These factors apply to all ancient DNA samples, and decontamination of samples is a standard practice in the field. However, the most common contaminant are microbes (31), which pose little threat to large mammalian genomic studies, as microbial sequences are disregarded during mapping to reference genomes. However, when the target DNA is microbial, the successful physical removal of environmental prior to DNA extraction is critical to avoid these biases. Despite this, there is not a field standard for decontaminating dental calculus. This is addressed in Chapter II, when I assessed a range of decontamination methods for efficiency, including two published protocols. As a result of this comparison, I recommend a 30-minute UV exposure to each side of the dental calculus fragment, followed by submergence in 5% bleach for three minutes. This method was the most successful at removing soil microorganisms and was consistent across a test group of real-world, archaeological dental calculus samples. Implementation of a decontamination standard across the field will allow for more confident meta-analysis of data as it becomes available. However, further studies should continue to pose and test novel decontamination methods to improve the removal of contaminant DNA,

while limiting the impacts on endogenous DNA. For example, variations in UV exposure and bleach concentration could be assessed to identify optimal conditions for the removal of contaminating DNA.

Reliance on databases and bioinformatics tools as the sole method to identify contamination post sequencing should be used with caution. In Chapter II, bioinformatic filtering methods were applied, yet SourceTracker analysis still assigned large proportions of reads as unknown, despite having human oral and skin, and soil microbe comparison data. This proportion was highest in untreated samples, suggesting that many 'unknown' species likely arise from contamination. This also signifies that the reference samples provided to SourceTracker are not extensive enough to account for all the species present in ancient dental calculus. This could be improved by using soil data from the specific sites where the samples were collected, and by using skin microbiota data obtained from swabs of the team that handled the samples after excavation. However, constructing a comprehensive comparison data set is difficult given the variation in preservation, microbial content of soil (32), and handling of different samples. Consequently, practical procedures to minimize contamination are critical.

Laboratory contamination cannot be removed or accounted for using the protocols outlined in Chapter II, as this contaminating DNA is introduced into the samples from the laboratory environment during processing. Consequently, negative controls (*e.g.* samples processed in tubes with no sample added) are critical to monitor this type of contamination. Negative controls are, again, an ancient DNA field standard (33). However, laboratory contamination mostly consists of microbial DNA (34), which again does not pose issues for mammalian genomic studies but is problematic for ancient microbiota analysis. In Chapters III and IV, species identified within negative controls accounted for ~25% of the total detected species. This highlights the critical need to sequence and assess negative controls when working with ancient samples that have low endogenous DNA levels. In addition, poor quality samples with low endogenous DNA will more closely resemble a negative control than a well-preserved sample. Consequently, I introduced a

simple analysis in Chapter III that uses Bray-Curtis dissimilarity values to identify microbial communities that were not distinct from the negative controls. These samples were excluded from further analysis, further removing bias from my data sets. Stringent methods such as these must be applied to obtain a robust data set, and it is recommended that this analysis be incorporated as a field standard in future studies to ensure that laboratory microorganisms are not driving the signals observed in ancient microbiota studies.

It is concerning that ~25 % of identified species must be removed from the data during the filtering of laboratory contamination. However, 16S rRNA gene analysis of ancient microbiota studies have identified up to 97% of species as laboratory contaminations (4). The rate of contamination in the two methods may be different, but shotgun sequencing also provides a single read count for each species identified, rather than multiple blocks of reads assigned to unique operational taxonomic units (OTUs) that are assigned to a single taxon. Therefore, reads without species level identifications cannot be removed from shotgun data sets as the entire taxon would be lost. However, a tool commonly used in ancient eukaryotic analysis is assessment of the damage patterns on the DNA fragments. Future development of bioinformatics for calculus analysis could exploit contamination to better identify modern microbial contamination within sequencing. DNA fragments are already being aligned to reference genomes in programs such as MALT and MALTX (35), and an assessment of mismatches along the fragment based on these existing alignments could identify potentially endogenous, ancient DNA from modern contaminants. While the exclusion of modern DNA would not remove all remaining contaminants, this method would improve confidence in the microbiota community being analysed.

The analysis of DNA from ancient dental calculus is a young field and consequently there is not a standardised analytical framework for analysis of these data. Info here = things that must be considered for future work and

will be key in developing such a standard. It is also, then, critical that data is freely available such that future studies can

Working with ancient DNA is a delicate process, particularly when the target DNA is microbial. In this thesis, I have explored and presented several methods to account for confounding factors during the analysis of ancient dental calculus. It is critical that these factors (oral geography, sample decontamination, and assessment of negative controls) are accounted for in future studies to maintain and improve data quality and allow correct inferences of historical microbiota.

Future work

Ancient DNA analysis of dental calculus is increasingly becoming an accepted tool within medical and archaeological research. This thesis expands on this premise and identifies detailed alterations in historical microbiota. However, during this research, several issues have arisen that need to be addressed in future analyses. In addition, the high-resolution data obtainable from dental calculus suggests potential future research directions. In this section, I present eight future directions that can build upon the work presented here.

1. Detailed assessment of calculus from different regions of mouth

The impacts of oral geography cannot be full understood until the variation of microbiota at different oral locations is fully examined. In Chapter IV, different oral locations significantly impacted microbiota diversity and masked signals linked with disease. Therefore, calculus samples from different oral locations were not utilized to examine individuals within the time transect. Further analyses of living individuals will aid in this issue, particularly by detailing the taxonomic and functional differences of microbiota from different oral locations. Community differences also need to be examined in relationship to temporal changes within the oral environment (*e.g.* level of saliva present). If these variables can be quantitatively described, then there

is a potential to apply statistical or bioinformatic methods to account for this variation (see point 4). If the communities are unique and no large-scale trends in microbiota can be linked with oral geography (*i.e.* the microbiota in different areas of the mouth respond differently to external factors), oral geography can only be addressed by sampling single locations. Future ancient DNA studies should also compare different samples from the same individual to confirm that the trends observed in modern individuals are true in the past. These studies are critical to control for oral geography variation in ancient studies.

2. Integrated sampling procedures

This thesis has presented the first high-resolution analysis of ancient oral microbiota and past living environments. However, the human genome can also play key roles in shaping the microbiota. Ancient DNA studies have the potential to observe the co-adaption of the human genome and the microbiota. For example, microbiota could be analysed through time as lactase persistence was being selected for in the human genome. Building on the observations within Chapter III, dental calculus would provide an alternate measure of the proportion of the population consuming dairy and an indication of the quantity, providing insight into the use of dairy and thus the level of selection upon the genome. Such comparisons could be extended to all dietary associated genes, including carbohydrates and fatty acids (36). The results of Chapter IV indicate that microbiota shifts could be identified in relation to evolution of disease resistance in the human genome. Consequently, future studies should assess ancient microbiota, genomes, and environments from the same individual. While human DNA has been recovered from dental calculus, multiple rounds of enrichment are required to obtain mitochondrial genome (37). Therefore, a tooth or bone sample should be collected alongside the dental calculus and metadata. It may also be possible to revisit archaeological remains that have previously been sampled (for genomic or microbiota analysis) and take the partner (dental calculus or bone/tooth) sample. However, new studies and excavations

should integrate human and microbiota sampling to gain the maximum amount of data with minimal damage to the remains. As the genomic era expands, microbiota analyses should be routinely integrated now to avoid retroactively accessing material.

3. Integrated extraction methods

To maximise the information recovered from samples of dental calculus, integrated extraction protocols to recover multiple biomolecules, such as proteins and DNA should be applied. Proteomic analysis of dental calculus is a powerful tool to assess microbial and eukaryotic proteins (*i.e.* immune proteins and dietary constituents (2,9)). For example, the presence of dairy items in the diet has been demonstrated by identification of the β-lactoglobulin protein (11). Further analyses can target proteins from a wide range of dietary items. The tissue specificity of proteins may allow identifications of species and the tissue type consumed. Additionally, identification of human immune proteins can reveal the immune responses occurring within the mouth, providing an insight into oral disease (5). Isotopic analysis of dental calculus has isolated $\delta^{13}C$ and $\delta^{15}N$ isotopes to study diet (13), and may be used to indicate marine resources and C_3/C_4 plant use. However, this method is controversial. It has been stated that analysis of isotopes from dental calculus is flawed and does not correlate with data from collagen. The complex and variable formation of dental calculus also confounds standardisation of isotopic results (38). Future studies may demonstrate that isotopes from dental calculus provide important insights, which may not be the same as can be acquired from other biological isotope sources (*e.g.* bone, teeth, hair, nails). However, if isotopic analyses are to be undertaken, considerable theoretical and developmental work is required. Until such methods can be defended, archaeological use of isotopes from dental calculus should not be undertaken. In addition, microfossils, which were the original target of research (39), are important sources of dietary and environmental information. To allow the integration of these methods, extraction protocols should be optimised to recover DNA, proteins,

microfossils, and (potentially) isotopes simultaneously from the same dental calculus fragment. Multiple extraction methods have been developed for recovering DNA and proteins from a single sample, which also can capture RNA and metabolites (40–43). However, further development is required to include microfossil and isotopic recovery. Currently, these extraction methods have not been applied to ancient samples where the concentrations of the biomolecules are in low concentrations. Optimization for ancient samples will be critical to recover high quality data. However, analysis of multiple biomolecules from a single sample will broaden the array of lifestyle factors, such as diet and disease, that can be analysed and allow complementation of methods to provide greater confidence of inferences.

4. Improvement of microbial identification

Current microbiota studies use comprehensive databases of microbial genomes to identify the microbes preserved with the dental calculus sample. However, the total diversity of ancient microbes is poorly understood, which limits the ability to identify new taxa or novel functions, and may result in misidentification of sequences due to unappreciated bacterial evolution. Studies should continue to assess the microbes within the human body and the environment to expand the reference databases, and ensure that taxonomic and functional assignments are more robust (44). In addition, ancient DNA studies should also continue the recovery of ancient microbial genomes from metagenomic samples. This has been done in modern samples using methods such as GroopM. GroopM compares related metagenomes and uses coverage similarity to reconstruct individual microbial genomes. This method reduces the need for sequencing depth to identify genomes, which is necessary for ancient DNA analyses. However, further development will be required to optimize for ancient DNA, which contains short fragments of damaged DNA potentially confounding current methods. While many ancient DNA studies have focused on ancient pathogens, commensals represent the microbiota present throughout the life and provide most of the functions critical to health and disease. Current

studies of ancient microbiota may be missing unique microbial and functional diversity that has been lost in modern microbiota and thus is not present in the current databases. Incorporating ancient genomes into reference databases would provide greater resolution when identifying ancient microbial species and their functions and may allow the appreciation of unique species within ancient microbiota. Recovering and reconstructing ancient microbial genomes on a large-scale is difficult and costly, but should be a long-term goal that remains a widespread theme within the field.

5. Development of multivariate analyses for ancient microbiota

Chapters III and IV demonstrate that microbiota are altered in response to multiple factors simultaneously. Current analytical methods applied in the ancient DNA field can only assess the association of the microbiota with a single metadata field. However, a clear understanding of the hierarchy of external factors that alter the microbiota can only be established through multivariate analyses. In this thesis, the contributions of oral geography and disease metadata had to be analysed individually, due to the lack of multivariate approaches. In Chapter IV, I created sub-sets of data from the total 128 samples to account for variation associated with oral geography. Once the data was split, I analysed the links between microbiota and disease. While sampling can be utilised to avoid bias from oral geography, other overlapping metadata characteristics cannot be analysed in this way (*e.g.* consumption of meat and oral disease). Consequently, multivariate analysis could be applied to discern individual contributions of individual metadata fields, removing variation due to oral geography (or any other factor). This would allow data sets, such as those in Chapters IV, to utilize all samples for analysis, increasing the statistical power to detect fine scale alterations. Statistical methods capable of indicating the relative impact of multiple metadata fields have been applied to data from modern microbiota analyses. These include direct and indirect gradient analyses and display the drivers of sample distribution on ordination plots (such as a plot from a Principal Coordinates Analysis) (45). While these analyses indicate the

major drivers of the inter-sample variation, they do not allow for variation of one factor to be removed from the data. An analysis method to overcome this is the MaAsLin (Multivariate Association with Linear Models) package (46). This pipeline can account for variation within the microbial taxonomy or function as described by one or more metadata fields and remove this trend, allowing observation of finer scale associations. However, this method has not been applied to ancient DNA samples and further analysis is required to determine if it can be used with ancient DNA data sets with confidence. Simulated ancient DNA reads and microbiota should be used to test the MaAsLin pipeline under controlled conditions. However, Chapter IV provides a large data set upon which these trials with real data could be applied.

6. Examination of additional populations and geographies

This thesis has focussed on Great Britain as a model population to examine historical microbiota. However, Great Britain cannot be assumed to harbour microbiota representative of Europe, other colonised Western countries (*e.g.* United States or Australia), or uncontacted regions. Therefore, the analyses conducted in this thesis should be repeated in different geographical regions around Europe and the world. It remains unknown if the presence of two microbiota groups is present in other European populations. Although, I would predict that this would be the case. The level of trade and movement between countries prior to the Industrial Revolution (29,47) ,the fact that two German samples (5) fell within the diversity of the British samples (Chapter III, Figure 3a), and that there were no significant associations with influx of people or goods into Britain (via invasion, migration, or trade) suggests that there was sufficient interconnection between populations throughout time that the microbiota would be consistent across Europe.. The basic diets of social groups are also similar across Europe (48), which would support two similar microbiota groups. The data present in this thesis provides a baseline for comparison, and will allow researchers to qualify and quantify the variation that exists between different regions and assess these assumptions.

7. Identifying the micriobiota of colonization

The microbiota data from Britain and Europe provide an opportunity to appreciate the microbiota that were introduced during the Colonial period in the New World and elsewhere. Native peoples of colonised countries likely possessed unique microbiota compared to those examined in Britain, and the resultant shifts in native peoples' microbiota as a result of colonization are partially dependent on the novel microbes that were introduced by colonialists. As Chapters III and IV show, the microbiota of Britain during the colonial period was significantly different from that observed in modern, industrialised populations, indicating that modern Western microbiota are not a suitable proxy for colonialists. This is particularly true for early colonial periods prior to the Industrial Revolution. Ideally, a comparison of the European population, the individuals who arrived in the new colony, and the native people of that region should be assessed. This will allow indication of what microbes were brought by the colonialists, whether this marked a random sampling or specific sub-group of the European population, and how these microbiota differed from the local, native microbial communities. Appreciating this will impact our understanding of how colonialists initially impacted the health and microbial diversity within indigenous people, providing insight into how treatments can be tailored to overcome the pattern of poorer health in colonised, native populations.

8. Using microbiota to track migrations

Human migration has resulted in humans that inhabit a wide-range range of different environments with unique resources. How microbiota respond to an altered diet and the specific pattern of microbes that are retained as humans pass through different environments, may leave signatures that allow populations to be compared and traced through time. It may be possible to use these signatures in the microbiota to reconstruct human migratory patterns or paths. Examining the evolution of core microbial species could be applied to identify these patterns. Future studies of dental calculus should assess this possibility. The rapid adaptive ability of the

microbiota community and microbial genomes suggests that microbiota may offer a resolution that allows tracking of rapid migrations, such as the peopling of Polynesia (49,50).

Around the world, ancient DNA studies of dental calculus are still in the early stages. In the future, techniques must be refined to obtain robust, reproducible results from dental calculus. These techniques must then be used to produce a field standard so that microbiota reconstructions are not biased by different assumptions in independent studies. Ancient DNA studies must also look to combine their findings with those from other molecular analyses (*e.g.* proteomics), bioarchaeology methods (*e.g.* isotopes), and archaeological and historical information. Extensive collaborations will provide unique context to the data, and foster new questions to further develop dental calculus as an archaeological tool. Partnerships with medical researchers will also ensure that the data provided from ancient dental calculus can become medically relevant and reveal causal factors of modern diseases.

Science communication

In the final chapter of this thesis, I outlined a framework for science communication within a research group. Analysis of science communication efforts at the Australian Centre for Ancient DNA demonstrated that both the public and the Centre's academic peers were part of the audience engaged with blog posts, videos, and social media output. Further, both the public and peers entered into collaborations with the Centre. The public donated DNA samples to build a DNA database for war dead identification, while the academic collaboration resulted in the sharing of data, which was made publically available via the Online Ancient Genome Repository (OAGR, (51)). These two outcomes directly demonstrate the potential of science communication to benefit research.

The future of science communication at the Centre now has a solid base, both in terms of the infrastructure though which science communication can be conducted and an audience. Consistency is critical in maintaining an audience (52). Consequently, the focus for future communication is to maintain a regular output of science content. Novel formats will be explored and, where suitable, utilised, but as an addition to the current framework, not as a replacement for existing elements. By doing this, the centre can build upon its current audience and focus on increasing interaction with the existing audience as well as gaining new audience groups.

As noted in Chapter V, social media platforms are not homogenous, and individual users form interconnected groups who share information (53). Information posted by users outside of this group is unlikely to be visible to the group members. Academic disciplines are likely to also form such groups, as individual researchers follow ("keep up to date with") other members of their field who share relevant information. The followers of the Australian Centre for Ancient DNA social media accounts tend to be other ancient DNA researchers. This communication within the field is important and, as noted above, has resulted in a collaboration. However, for a research group to expose members of other disciplines to their research, it is important that the group identify the tags and language of various groups in order to place content into the forum of the other group. On Twitter this may mean the use of hashtags (key words). Examples of hashtags that could be used in relation to the research presented here include: #twitterstorians (historians), #archaeology (archaeology), #medicalresearch (medical research), and #microbiome (microbiota and microbiome research). These key words are likely to capture both public and academic users and offer a way to breach the barriers within social media and between academic disciplines.

The framework presented in Chapter V, places science in a public forum in an engaging manner. It does not address the complexities of engaging with specific ideas and looking to remove culturally engrained

opinions or pseudo-science. Academics involvement in such debates (*e.g.* vaccinations) is cautioned unless sufficient time can be devoted to understanding the argument and creating well though-out responses and examples. An understanding of the psychology behind changing opinions is also required for such debates to be effective. It is critical that scientists engage with these pseudo-science issues, and to work toward a science-literate society. The collaborations noted within the framework can include groups that provide guidance on entering this element of science communication. However, it should be noted that high quality science communication is important and provides a resource for professional science communicators and the interested public.

In Chapter V, the targeted audience included academic peers and the public. However, the time and effort put into making science comprehensible to a lay audience, including academics in separate and different fields, can also be directed at a third group – high school and undergraduate students. Students receive little insight into active research during their formal education. However, science communication content provides an accessible source of information on active research pitched at a suitable level for students toward the end of high school and in undergraduate degrees. Science communication outputs should be advertised to students and, potentially, incorporated into lessons/tutorials to provide an introduction to a topic, to inspire critical thinking and to further research. In addition, this insight could be a method to inspire a passion for science and inform future career pathways

In the context of this thesis, science communication has been used to promote its research across and beyond the multiple research fields it incorporates. Public facing communication was an aspect of the Museum of London application to sample (Chapter IV). Communication can benefit this research as described above. However, it is also important to the Museum. Science communication allows the Museum to promote its collections and the research benefits that they provide. To provide benefits to the research presented in Chapter IV and the Museum, three activities were planned.

Firstly, at the time of sampling, a small display was placed in the Museum galleries for an afternoon. The Museum osteology curators and I (as the researcher) were present to describe and discuss the project with Museum visitors. This was followed by a series of talks to academics and school students in the UK and Australia. The final communication project is to produce a series of short videos describing the research project from inception to results. These videos will be pitched toward high school students and made available via YouTube to promote microbiota research and provide insight into the elements and timeframe of a research project. These videos are still in production, although, filming has occurred throughout the research process to capture dental calculus sampling at the Museum, DNA extraction, library preparation, and data analysis at the Australian Centre of Ancient DNA. These films will allow viewers to watch the research at each stage, rather than only see a report of the results. Consequently, planning of science communication from the beginning of the project helped to support the sampling application, allowing this research to be undertaken. Through multiple communication events this research has been (and will continue to be) distributed.

Conclusion

Microbiota are a critical part of the human body, carrying out key functions of which the human body is incapable itself. However, modern, industrialized microbiota are in a disease state, which modern medicine needs to address. Unfortunately, studies of living people cannot trace the development of this modern disease state, identify the key cultural and environmental changes that drove this development, or explore how alterations impact the microbiota of subsequent generations. Ancient DNA studies of dental calculus allow direct insight into the structure of oral microbiota through time, allowing these questions to be addressed. Dental calculus is a calcified microbial biofilm, a preserved representation of the oral

microbiota, and the only known accurate representative of historical microbiota structure. Studies of ancient dental calculus have revealed that the microbial community was significantly altered during the agricultural and industrial revolutions. However, the population level structure of microbiota and the specific factors that drove this change remained unknown.

This thesis presents the largest study of dental calculus to date, totalling 281 samples, to identify the individualistic factors that defined and altered the microbiota in a historical population. Findings indicate that historical Britain contained two distinct groups of microbiota that were defined by diet and linked to the socio-economic rank of the individual. In addition, these groups were present throughout the ~2,000-year time transect, indicating the minimal impact of the multiple invasions and disease epidemics that struck the population. Further detailed study also revealed associations with oral and systemic disease for the first time. Together, these results indicate that the key drivers of microbiota structure are diet and disease and suggest that nutrient control may be a way to medically manipulate the microbiota. However, the potential for dental calculus to be used as an archaeological tool to study diet, disease, and socio-economic rank – areas of major archaeological, anthropological, and historical research – is also indicated.

To provide this high-resolution data, several key elements of sample processing were identified and addressed to provide robust data. These experimental and analytical methods should be used in further analyses of dental calculus and be key considerations in developing a standardised research method. Standardisation should identify the optimal methods to produce accurate and precise data, and account for the practicalities of archaeological, laboratory, and analytical work and biological patterns that need to identified or controlled for. In this thesis it is demonstrated that future studies must account for the oral geography of samples, apply a UV and bleach decontamination treatment to remove environmental decontamination, and sequence negative controls to assess sample quality and filter laboratory contamination from samples. High-resolution analyses,

as conducted here, are reliant on the recovery of high quality data to confidently make claims about cultural and environmental impacts on microbiota.

This thesis also addressed the need for science communication, outlining a group framework to maximise communication outputs and minimise investment from individual researchers. This framework should be exploited to target three major audience groups: academic peers, the public, and students. The benefits of doing so provide collaborations with peers and the public, help improve science literacy in society, and can inspire future researchers. Critically, in the context of this thesis, public facing science communication is a method to engage with academics in disparate fields (*i.e.* molecular biologists informing historians, archaeologists, and medical researchers of the research concepts and results). Ultimately this aims to create an awareness of research methods that may create the environment for new, trans-disciplinary research collaborations, and the advancement of scientific research.

References

1. Adler CJ, Dobney K, Weyrich LS, Kaidonis J, Walker AW, Haak W, et al. Sequencing ancient calcified dental plaque shows changes in oral microbiota with dietary shifts of the Neolithic and Industrial revolutions. Nat Genet. 2013 Apr;45(4):450–5.

2. David LA, Maurice CF, Carmody RN, Gootenberg DB, Button JE, Wolfe BE, et al. Diet rapidly and reproducibly alters the human gut microbiome. Nature. 2014 Jan 23;505(7484):559–63.

3. Clayton JB, Vangay P, Huang H, Ward T, Hillmann BM, Al-Ghalith GA, et al. Captivity humanizes the primate microbiome. Proc Natl Acad Sci. 2016 Aug 29;201521835.

4. Weyrich LS, Duchane S, Soubrier J, Arriola L, Llamas B, Breen J, et al. Reconstructing Neandertal behavior, diet, and disease using ancient DNA from dental calculus. Nature. 2016;

5. Warinner C, Rodrigues JFM, Vyas R, Trachsel C, Shved N, Grossmann J, et al. Pathogens and host immunity in the ancient human oral cavity. Nat Genet. 2014;

6. Spencer C. The British Isles. In: Kiple KF, Ornelas KC, editors. The Cambridge World History of Food [Internet]. Cambridge: Cambridge University Press; 2012 [cited 2016 Oct 14]. p. 1217–26. Available from: http://universitypublishingonline.org/ref/id/histories/CBO9781139058643A017

7. Hildebrandt MA, Hoffmann C, Sherrill–Mix SA, Keilbaugh SA, Hamady M, Chen Y, et al. High-Fat Diet Determines the Composition of the Murine Gut Microbiome Independently of Obesity. Gastroenterology. 2009 Nov;137(5):1716–1724.e2.

8. Albenberg LG, Wu GD. Diet and the Intestinal Microbiome: Associations, Functions, and Implications for Health and Disease. Gastroenterology. 2014 May;146(6):1564–72.

9. Dykes. Phenolic Compounds in Cereal Grains and Their Health Benefits. Cereal Foods World [Internet]. 2007 [cited 2016 Oct 15]; Available from: http://www.aaccnet.org/publications/plexus/cfw/pastissues/2007/abstracts/CFW-52-3-0105.html

10. Müldner G, Richards MP. Fast or feast: reconstructing diet in later medieval England by stable isotope analysis. J Archaeol Sci. 2005 Jan;32(1):39–48.

11. Warinner C, Hendy J, Speller C, Cappellini E, Fischer R, Trachsel C, et al. Direct evidence of milk consumption from ancient human dental calculus. Sci Rep. 2014 Nov 27;4:7104.

12. Wunder MB. Determining geographic patterns of migration and dispersal using stable isotopes in keratins. J Mammal. 2012 Apr 30;93(2):360–7.

13. Scott GR, Poulson SR. Stable carbon and nitrogen isotopes of human dental calculus: a potentially new non-destructive proxy for paleodietary analysis. J Archaeol Sci. 2012;39(5):1388–93.

14. Shweta, Prakash SK. Dental abscess: A microbiological review. Dent Res J. 2013;10(5):585–91.

15. Sheng WS, Hsueh PR, Hung CC, Teng LJ, Chen YC, Luh KT. Clinical features of patients with invasive Eikenella corrodens infections and microbiological characteristics of the causative isolates. Eur J Clin Microbiol Infect Dis Off Publ Eur Soc Clin Microbiol. 2001 Apr;20(4):231–6.

16. Bakaletz LO. Developing animal models for polymicrobial diseases. Nat Rev Microbiol. 2004 Jul;2(7):552–68.

17. Belda-Ferre P, Alcaraz LD, Cabrera-Rubio R, Romero H, Simón-Soro A, Pignatelli M, et al. The oral metagenome in health and disease. ISME J. 2012 Jan;6(1):46–56.

18. Benton D, Williams C, Brown A. Impact of consuming a milk drink containing a probiotic on mood and cognition. Eur J Clin Nutr. 2006 Dec 6;61(3):355–61.

19. Zitvogel L, Galluzzi L, Viaud S, Vétizou M, Daillère R, Merad M, et al. Cancer and the gut microbiota: An unexpected link. Sci Transl Med. 2015 Jan 21;7(271):271ps1-271ps1.

20. Scher JU, Abramson SB. The microbiome and rheumatoid arthritis. Nat Rev Rheumatol. 2011 Aug 23;7(10):569–78.

21. Nuki G, Simkin PA. A concise history of gout and hyperuricemia and their treatment. Arthritis Res Ther. 2006;8(1):S1.

22. Loyola-Rodriguez JP, Martinez-Martinez RE, Abud-Mendoza C, Patiño-Marin N, Seymour GJ. Rheumatoid arthritis and the role of oral bacteria. J Oral Microbiol [Internet]. 2010 Dec 21 [cited 2016 Sep 12];2. Available from: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3084578/

23. Schnorr SL, Candela M, Rampelli S, Centanni M, Consolandi C, Basaglia G, et al. Gut microbiome of the Hadza hunter-gatherers. Nat Commun [Internet]. 2014 Apr 15 [cited 2015 Feb 21];5. Available from: http://www.nature.com/ncomms/2014/140415/ncomms4654/full/ncomms4654.html?message-global=remove

24. Hopkins MJ, Sharp R, Macfarlane GT. Variation in human intestinal microbiota with age. Dig Liver Dis. 2002 Sep 1;34:S12–8.

25. Tiihonen K, Ouwehand AC, Rautonen N. Human intestinal microbiota and healthy ageing. Ageing Res Rev. 2010 Apr;9(2):107–16.

26. Moeller AH, Li Y, Ngole EM, Ahuka-Mundeke S, Lonsdorf EV, Pusey AE, et al. Rapid changes in the gut microbiome during human evolution. Proc Natl Acad Sci. 2014 Nov 18;111(46):16431–5.

27. Clemente JC, Pehrsson EC, Blaser MJ, Sandhu K, Gao Z, Wang B, et al. The microbiome of uncontacted Amerindians. Sci Adv. 2015 Apr 1;1(3):e1500183.

28. Martínez I, Stegen JC, Maldonado-Gómez MX, Eren AM, Siba PM, Greenhill AR, et al. The Gut Microbiota of Rural Papua New Guineans: Composition, Diversity Patterns, and Ecological Processes. Cell Rep. 2015 Apr 28;11(4):527–38.

29. Black J. London: a history. Carnegie; 2009.

30. Simón-Soro Á, Tomás I, Cabrera-Rubio R, Catalan MD, Nyvad B, Mira A. Microbial Geography of the Oral Cavity. J Dent Res. 2013 Jul 1;92(7):616–21.

31. Gilbert MTP, Bandelt H-J, Hofreiter M, Barnes I. Assessing ancient DNA studies. Trends Ecol Evol. 2005 Oct;20(10):541–4.

32. Young JM, Weyrich LS, Breen J, Macdonald LM, Cooper A. Predicting the origin of soil evidence: High throughput eukaryote sequencing and MIR spectroscopy applied to a crime scene scenario. Forensic Sci Int. 2015 Jun;251:22–31.

33. Cooper A, Poinar HN. Ancient DNA: Do It Right or Not at All. Science. 2000 Aug 18;289(5482):1139–1139.

34. Salter SJ, Cox MJ, Turek EM, Calus ST, Cookson WO, Moffatt MF, et al. Reagent and laboratory contamination can critically impact sequence-based microbiome analyses. BMC Biol. 2014 Nov 12;12(1):87.

35. Herbig A, Maixner F, Bos KI, Zink A, Krause J, Huson DH. MALT: Fast alignment and analysis of metagenomic DNA sequence data applied to the Tyrolean Iceman. bioRxiv. 2016 Apr 27;50559.

36. Voight BF, Kudaravalli S, Wen X, Pritchard JK. A Map of Recent Positive Selection in the Human Genome. PLOS Biol. 2006 Mar 7;4(3):e72.

37. Ozga AT, Nieves-Colón MA, Honap TP, Sankaranarayanan K, Hofman CA, Milner GR, et al. Successful enrichment and recovery of whole mitochondrial genomes from ancient human dental calculus. Am J Phys Anthropol. 2016 Mar 1;n/a-n/a.

38. Salazar-García DC, Richards MP, Nehlich O, Henry AG. Dental calculus is not equivalent to bone collagen for isotope analysis: a comparison between carbon and nitrogen stable isotope analysis of bulk dental calculus, bone and dentine collagen from same individuals from the Medieval site of El Raval (Alicante, Spain). J Archaeol Sci. 2014 Jul;47:70–7.

39. Dobney K, Brothwell D. Dental Calculus: Its Relevance to Ancient Diet and Oral Ecology. In: Cruwys E, Foley RA, editors. Teeth and Anthropology. 1986. p. 55–81. (British Archaeological Reports).

40. Tolosa JM, Schjenken JE, Civiti TD, Clifton VL, Smith R. Column-based method to simultaneously extract DNA, RNA, and proteins from the same sample. BioTechniques. 2007 Dec;43(6):799–804.

41. Mathieson W, Thomas GA. Simultaneously extracting DNA, RNA, and protein using kits: Is sample quantity or quality prejudiced? Anal Biochem. 2013 Feb 1;433(1):10–8.

42. Xie J, Yun H, Dong H, Zhao W, Wang G, Qiu G, et al. Simultaneous extraction, separation and purification of microbial genomic DNA and

total RNA from acidic habitat samples. Anal Methods. 2015 Jan 29;7(3):909–17.

43. Nakayasu ES, Nicora CD, Sims AC, Burnum-Johnson KE, Kim Y-M, Kyle JE, et al. MPLEx: a Robust and Universal Protocol for Single-Sample Integrative Proteomic, Metabolomic, and Lipidomic Analyses. mSystems. 2016 Jun 28;1(3):e00043-16.

44. Hugenholtz P. Exploring prokaryotic diversity in the genomic era. Genome Biol. 2002;3:reviews0003.

45. Ramette A. Multivariate analyses in microbial ecology. Fems Microbiol Ecol. 2007 Nov;62(2):142–60.

46. MaAsLin: Multivariate Association with Linear Models | The Huttenhower Lab [Internet]. [cited 2016 Oct 19]. Available from: https://huttenhower.sph.harvard.edu/maaslin

47. Hinton DA, Crawford S, Hamerow H, editors. The Oxford Handbook of Anglo-Saxon Archaeology [Internet]. 1st ed. Oxford University Press; 2011 [cited 2016 Aug 26]. Available from: http://www.oxfordhandbooks.com/view/10.1093/oxfordhb/9780199212149.001.0001/oxfordhb-9780199212149

48. Carlin M, Rosenthal JT. Food and Eating in Medieval Europe. Bloomsbury Publishing; 1998. 203 p.

49. Kirch PV. Peopling of the Pacific: A Holistic Anthropological Perspective. Annu Rev Anthropol. 2010;39(1):131–48.

50. Moodley Y, Linz B, Yamaoka Y, Windsor HM, Breurec S, Wu J-Y, et al. The Peopling of the Pacific from a Bacterial Perspective. Science. 2009 Jan 23;323(5913):527–30.

51. Online Ancient Genome Repository [Internet]. [cited 2016 Jul 28]. Available from: https://www.oagr.org.au/

52. Wong F. The Secrets of YouTube Success - 2012 Edition [Internet]. RocketJump. 2012 [cited 2015 Jan 5]. Available from: http://www.rocketjump.com/blog/the-secrets-of-youtube-success-updated

53. Bruns A, Burgess J, Highfield T. A “Big Data” Approach to Mapping the Australian Twittersphere. In: Arthure P, Bode K, editors. Advancing Digital Humanities: Research, Methods, Theories. Palgrave Macmillan UK; 2014.

Appendix I

Chapter III: Supplementary Information

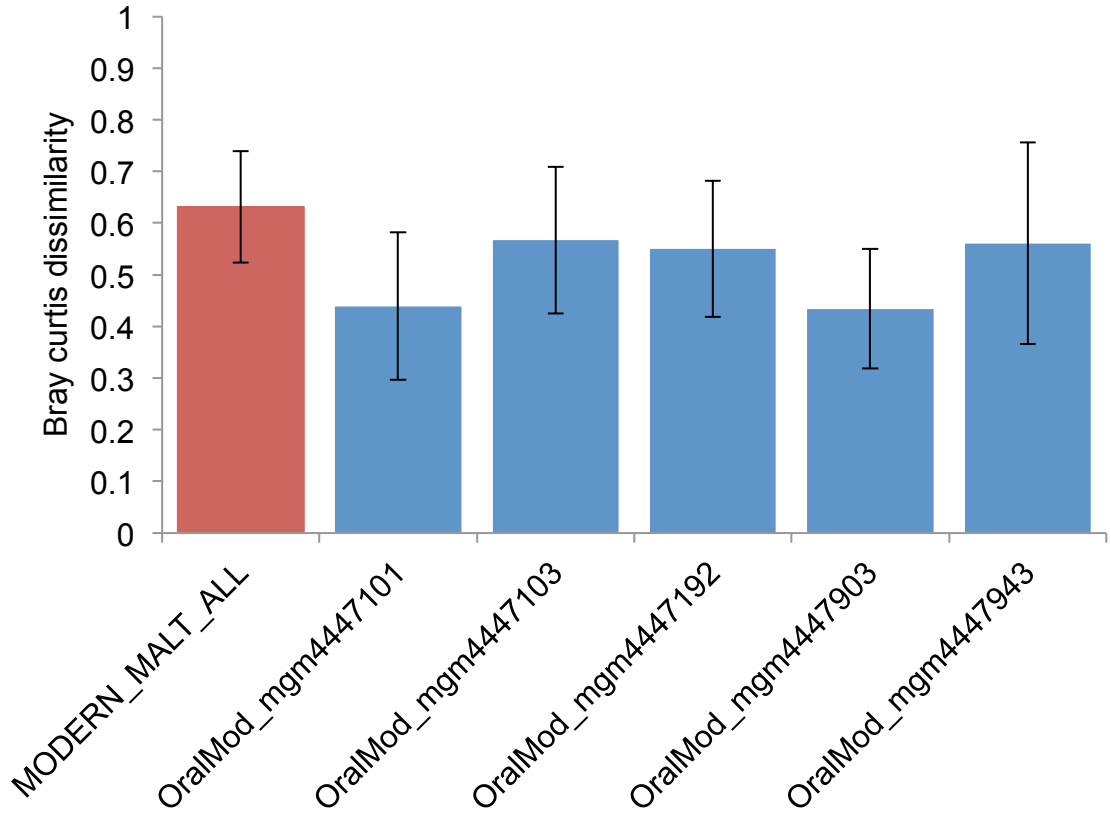


Figure S1: Average Bray curtis of each modern sample compared to the others. Calculus sample in red, plaque samples in blue. No significant difference between samples, which indicates that plaque and calculus samples are directly comparable.

Table S1

Test	R statistic	p-value
Group 1 vs.Group 2 vs.Modern	0.8153	0.0001
Group 1 vs.Group 2	0.7861	0.0001
Group 1 vs. Modern	0.7965	0.0001
Group 2 vs .Modern	0.9886	0.0001
## 3 Groups ##		
BlockPeriod	-0.0291	0.5752
## Ancient Groups (Group 1 & Group 2) ##		
Cemetry	-0.0276	0.5914
Clade	0.7861	0.0001
Country	0.0033	0.408
Region	0.0689	0.2293
RuralUrban	0.1224	0.1358
## 3 Groups ##		
Acinetobacter calcoaceticus baumannii complex	0.616	0.0017
Bordetella parapertussis	-0.0989	0.6241
Campylobacter jejuni	0.7186	0.053
Clostridium botulinum	0.6817	0.0001
Clostridium tetani	-0.1558	0.8128
Corynebacterium diphtheriae	0.6807	0.0003
Enterococcus faecium	0.2617	0.0012
Fusobacterium nucleatum	NA	NA
Helicobacter pylori	0.7186	0.0545
Mycobacterium tuberculosis complex	-0.0517	0.461
Neisseria gonorrhoeae	0.7277	0.0001
Porphyromonas gingivalis	NA	NA
Prevotella denticola	-0.024	0.6489
Pseudomonas aeruginosa group	0.0245	0.3769
Staphylococcus aureus	0.0346	0.3593
Streptococcus mutans	-0.1111	0.9024
Streptococcus pyogenes	-0.0147	0.4739
Tannerella forsythia	NA	NA
Treponema denticola	0.7967	0.0228

ANOSIM results showing significant differences ($R < 0.3$ and $p > 0.05$) between the three groups but not with period, geography, or pathogens (except for *Clostridium botulinum*)

Table S2a

	Test Statistic	Bonferroni p value	Group 1 mean	Group 2 mean
Aggregatibacter segnis	17.57	0.03	51.42	0.75
Campylobacter	19.69	0.01	1902.83	357.6
Campylobacter showae	20.39	0.01	547.50	52.4
Campylobacteraceae	18.36	0.02	51.92	15.95
Campylobacterales	18.71	0.01	89.67	21.8
Capnocytophaga	18.00	0.02	1364.33	143.75
Capnocytophaga gingivalis	16.79	0.04	59.00	1.2
Capnocytophaga sp CM59	18.02	0.02	137.08	2.65
Capnocytophaga sp oral taxon 326	23.59	0.00	46.83	0.8
Capnocytophaga sp oral taxon 329	17.57	0.03	82.33	0.85
Capnocytophaga sp oral taxon 332	16.76	0.04	32.75	0
Capnocytophaga sputigena	22.43	0.00	36.50	0
Cardiobacteriaceae	19.41	0.01	21.17	2.6
Cardiobacterium	19.36	0.01	280.67	27
Cardiobacterium hominis	19.63	0.01	148.92	12.5
Cardiobacterium valvarum	19.39	0.01	709.17	53.9
Eikenella corrodens	19.44	0.01	385.25	38.15
Epsilonproteobacteria	17.10	0.03	62.42	16.55
Flavobacteriaceae	17.36	0.03	263.42	78.8
Flavobacteriales	18.88	0.01	51.08	14.35
Flavobacteriia	19.39	0.01	47.42	7.75
Haemophilus	19.51	0.01	34.33	0
Kingella	19.52	0.01	10.00	0
Leptotrichia goodfellowii	19.52	0.01	75.83	0
Neisseria	19.70	0.01	1564.25	62.1
Neisseria elongata	20.69	0.01	277.00	0.75
Neisseria mucosa	16.76	0.04	76.92	0
Neisseria sp oral taxon 014	18.85	0.01	40.50	2
Neisseriaceae	19.70	0.01	533.75	49.95
Pasteurellaceae	17.68	0.03	309.58	46.95
Peptoclostridium	18.47	0.02	20.75	7.6
Prevotella saccharolytica	18.37	0.02	183.92	9.8
Prevotella sp oral taxon 472	17.37	0.03	368.50	11.45
Rothia	17.72	0.02	35.25	1.6
Selenomonas sp oral taxon 138	17.58	0.03	24.33	0.85

Part a: Kruskall-Wallis results indicating taxa that differ significantly ($p < 0.05$) between the two ancient groups of microbiota

Table S2b

	Test Statistic	Bonferroni p value	Group 1 mean	Group 2 mean
Clostridium clariflavum	17.72	0.02	0.00	13.7
Clostridium hiranonis	20.32	0.01	1.25	20.4
Clostridium hylemonae	20.42	0.01	5.17	25
Clostridium methylpentosum	17.59	0.03	0.00	20.5
Clostridium stercorarium	20.11	0.01	7.50	36.65
Eubacterium cellulosolvens	19.38	0.01	18.00	46.65
Ruminococcus obeum	21.72	0.00	11.58	35.7
Alkaliphilus metalliredigens	23.07	0.00	0.00	25.85
Alkaliphilus oremlandii	19.93	0.01	1.33	22.7
Anaerococcus hydrogenalis	17.59	0.03	0.00	16.3
Anaerofustis stercorihominis	20.08	0.01	8.33	55.1
Anaerotruncus colihominis	21.38	0.00	10.25	34.65
Archaea	19.03	0.01	12.17	65.45
Bacillaceae	20.24	0.01	27.83	51.45
Bacillales	20.23	0.01	69.58	120.55
Bacillus	21.87	0.00	45.75	82.65
Bacteria	20.39	0.01	13706.17	24454.95
Blautia hydrogenotrophica	17.26	0.03	13.42	38.95
Butyrivibrio fibrisolvens	20.26	0.01	1.33	30.1
Cellulosilyticum lentocellum	21.96	0.00	3.67	27.3
Clostridia	21.82	0.00	316.17	1032.35
Clostridiaceae	20.25	0.01	22.25	38.25
Clostridiales	21.10	0.00	2889.33	6741
Clostridiales Family XIII Incertae Sedis	19.01	0.01	114.17	442.6
Clostridium	21.82	0.00	136.17	277.45
Clostridium botulinum	21.98	0.00	8.17	37.2
Clostridium sp D5	20.06	0.01	15.92	38.65
Clostridium ultunense	19.56	0.01	18.67	58.55
Coprococcus catus	21.44	0.00	3.33	26.05
Desulfosporosinus	18.38	0.02	21.25	47.1
Desulfotomaculum	21.14	0.00	23.50	49.1
Dorea formicigenerans	18.39	0.02	9.17	23.1
Eggerthia catenaformis	20.39	0.01	33.00	168.9
Ethanoligenens harbinense	16.71	0.04	1.08	19.55
Eubacteriaceae	16.60	0.04	15.00	86.35
Eubacterium limosum	21.96	0.00	2.25	34.15
Eubacterium ventriosum	18.46	0.02	2.50	18.75
Euryarchaeota	18.33	0.02	109.25	855.15
Faecalibacterium prausnitzii	16.41	0.05	29.50	60.95

Part b: Kruskall-Wallis results indicating taxa that differ significantly ($p < 0.05$) between the two ancient groups of microbiota

Table S2c

	Test Statistic	Bonferroni p value	Group 1 mean	Group 2 mean
Firmicutes	21.82	0.00	3195.00	8422.5
Gottschalkia acidurici	22.16	0.00	2.83	29.85
Halanaerobium	19.46	0.01	0.00	13.8
Jonquetella anthropi	17.51	0.03	102.08	239
Listeria	20.10	0.01	16.67	30.55
Methanobacteriaceae	18.67	0.02	184.67	1852.6
Methanobacteriales	18.34	0.02	50.00	551.2
Methanobrevibacter	18.67	0.02	198.67	2133.15
Methanobrevibacter ruminantium	18.70	0.01	73.42	719.2
Methanobrevibacter smithii	18.67	0.02	586.17	6045
Methanococcales	19.38	0.01	0.00	14.85
Methanomicrobiales	18.61	0.02	1.17	19.4
Mogibacterium sp CM50	20.39	0.01	346.75	1315.8
Oscillibacter valericigenes	17.20	0.03	36.33	80
Pelosinus fermentans	19.47	0.01	3.58	22.35
Peptoclostridium difficile	22.10	0.00	6.17	39.6
Peptococcaceae	18.88	0.01	33.00	56.05
Peptoniphilus indolicus	19.94	0.01	8.00	129.75
Peptoniphilus sp oral taxon 386	21.10	0.00	47.58	256.5
Peptostreptococcaceae bacterium OBRC8	17.21	0.03	24.83	55.3
Pseudoflavonifractor capillosus	16.57	0.05	15.75	36.9
Pseudoramibacter alactolyticus	18.17	0.02	73.25	2072.75
Ruminococcaceae	19.05	0.01	48.50	80.5
Ruminococcaceae bacterium D16	18.38	0.02	18.17	46.55
Ruminococcus bromii	21.06	0.00	3.58	37.1
Ruminococcus flavefaciens	18.25	0.02	1.25	22.25
Shuttleworthia satelles	19.88	0.01	22.17	95.35
Thermoanaerobacterium	21.13	0.00	0.00	21
Thermotogaceae	16.76	0.04	18.58	37.2

Part c: Kruskall-Wallis results indicating taxa that differ significantly ($p < 0.05$) between the two ancient groups of microbiota

Table S3a

	Test Statistic	Bonferroni p value	Group 1 mean	Modern mean
Prevotella sp oral taxon 473	15.28	0.09	1.17	437.29
Prevotella sp oral taxon 299	13.60	0.22	0.00	62.00
Capnocytophaga canimorsus	13.60	0.22	0.00	22.14
Granulicatella elegans	13.60	0.22	0.00	46.57
Veillonella atypica	13.60	0.22	0.00	94.43
Veillonella dispar	13.60	0.22	0.00	82.29
Veillonella parvula	13.60	0.22	0.00	236.71
Veillonella sp 3 1 44	13.60	0.22	0.00	50.86
Veillonella sp 6 1 27	13.60	0.22	0.00	42.43
Veillonella sp oral taxon 158	13.60	0.22	0.00	38.71

Kruskall-Wallis results indicating that no taxa differ significantly ($p < 0.05$) between Group 1 and Modern microbiota groups

Table S3b - 1

	Test Statistic	Bonferroni p value	Group 1 mean	Group 2 mean
Rothia aeria	25.26	0.00	0.00	174.00
Fusobacterium hwasookii	25.26	0.00	0.00	70.00
Prevotella melaninogenica	25.25	0.00	0.00	630.29
Gemella haemolysans	25.25	0.00	0.00	88.71
Streptococcus infantis	25.25	0.00	0.00	61.14
Streptococcus parasanguinis	25.25	0.00	0.00	93.29
Leptotrichia goodfellowii	25.25	0.00	0.00	360.86
Rothia	21.30	0.00	1.60	428.57
Caudovirales	20.84	0.01	0.00	5.43
Capnocytophaga sp oral taxon 380	20.81	0.01	0.00	22.43
Streptococcus sp oral taxon 058	20.81	0.01	0.00	16.86
dsDNA viruses no RNA stage	20.81	0.01	0.00	3.57
unclassified Siphoviridae	20.81	0.01	0.00	22.71
Rothia dentocariosa	20.80	0.01	0.00	597.14
Prevotella sp oral taxon 299	20.80	0.01	0.00	62.00
Capnocytophaga canimorsus	20.80	0.01	0.00	22.14
Capnocytophaga sp oral taxon 324	20.80	0.01	0.00	35.57
Capnocytophaga sp oral taxon 332	20.80	0.01	0.00	120.57
Capnocytophaga sputigena	20.80	0.01	0.00	122.86
Streptococcus tigurinus	20.80	0.01	0.00	24.29
Veillonella atypica	20.80	0.01	0.00	94.43
Veillonella dispar	20.80	0.01	0.00	82.29
Veillonella parvula	20.80	0.01	0.00	236.71
Veillonella sp 3 1 44	20.80	0.01	0.00	50.86
Veillonella sp 6 1 27	20.80	0.01	0.00	42.43
Veillonella sp oral taxon 158	20.80	0.01	0.00	38.71
Kingella	20.80	0.01	0.00	31.71
Kingella oralis	20.80	0.01	0.00	221.00
Neisseria mucosa	20.80	0.01	0.00	827.71
Neisseria sicca	20.80	0.01	0.00	230.14
Neisseria sp oral taxon 020	20.80	0.01	0.00	42.71
Haemophilus	20.80	0.01	0.00	1084.00
Haemophilus parainfluenzae	20.80	0.01	0.00	1137.00
Streptococcus oralis	20.10	0.01	2.00	75.43
Streptococcus mitis	19.41	0.01	3.15	99.14
Prevotella sp oral taxon 473	18.86	0.02	2.25	437.29
Fusobacterium periodonticum	18.85	0.02	3.25	158.29
Capnocytophaga ochracea	18.25	0.02	0.65	34.14
Porphyromonas catoniae	18.24	0.02	1.25	384.14

Kruskall-Wallis results indicating taxa that differ significantly ($p < 0.05$) between Group 2 and Modern microbiota groups

Table S3b - 2

	Test Statistic	Bonferroni p value	Group 1 mean	Group 2 mean
Porphyromonas sp oral taxon 279	18.24	0.02	1.60	1522.86
Capnocytophaga gingivalis	18.24	0.02	1.20	301.71
Capnocytophaga sp oral taxon 326	18.24	0.02	0.80	181.57
Capnocytophaga sp oral taxon 329	18.24	0.02	0.85	756.57
Neisseria elongata	18.24	0.02	0.75	807.71
Veillonella ratti	16.67	0.05	0.00	44.57
Campylobacter concisus	16.67	0.05	0.00	24.43
Myoviridae	16.67	0.05	0.00	8.86
Prevotella oulorum	16.66	0.05	0.00	490.86
Prevotella veroralis	16.66	0.05	0.00	87.29
Capnocytophaga sp oral taxon 335	16.66	0.05	0.00	17.29
Streptococcus gordonii	16.66	0.05	0.00	30.00
Streptococcus sp oral taxon 056	16.66	0.05	0.00	14.14
Neisseria bacilliformis	16.66	0.05	0.00	85.86
Neisseria flavescens	16.66	0.05	0.00	57.14
Neisseria gonorrhoeae	16.66	0.05	0.00	67.43
Neisseria macacae	16.66	0.05	0.00	73.86
Neisseria sp GT4A CT1	16.66	0.05	0.00	48.71
Granulicatella adiacens	16.58	0.05	5.45	556.86
Prevotella sp oral taxon 472	15.92	0.07	11.45	494.43
Gemella morbillorum	15.60	0.08	15.05	166.57
Rhodobacterales	15.50	0.09	6.95	0.00
Alkaliphilus	15.42	0.09	8.20	0.14
Thermoanaerobacterales Family III Incertae Sedis	15.41	0.09	9.10	0.00
Methanosarcinales	15.40	0.09	5.35	0.00
Desulfitobacterium	15.39	0.09	19.30	0.00
Halanaerobiaceae	15.39	0.09	7.90	0.00
Halanaerobiales	15.38	0.09	8.25	0.00

Kruskall-Wallis results indicating taxa that differ significantly ($p < 0.05$) between Group 2 and Modern microbiota groups

Table S4a

	Porphyromonas gingivalis	Tannerella forsythia	Fusobacterium nucleatum	Treponema denticola	Streptococcus mutans	Clostridium botulinum	Enterococcus faecium	Prevotella denticola
N Present in G1	12	12	12	12	9	7	6	7
N Present in G2	19	19	19	19	12	19	16	10
N Present in Mod	8	8	8	7	8	1	2	6
Percentage G1	100	100	100	100	75	58.33	50	58.33
Percentage G2	100	100	100	100	63.16	100	84.21	52.63
Percentage Mod	100	100	100	87.5	100	12.5	25	75

Table S4b

	Neisseria gonorrhoeae	Pseudomonas aeruginosa group	Staphylococcus aureus	Clostridium tetani	Corynebacterium diphtheriae	Acinetobacter calcoaceticus baumannii complex	Bordetella parapertussis
N Present in G1	1	3	0	0	0	1	1
N Present in G2	0	1	4	4	0	0	2
N Present in Mod	6	1	1	0	4	3	0
Percentage G1	8.33	25	0	0	0	8.33	8.33
Percentage G2	0	5.26	21.05	21.05	0	0	10.53
Percentage Mod	75	12.5	12.5	0	50	37.5	0

Part 1 & 2: Presence/absence of pathogens tested for in the three microbiota groups. Number of samples with pathogen present, and percentage of samples shown for each group

Table S4c

	Streptococcus pyogenes	Mycobacterium tuberculosis complex	Campylobacter jejuni	Helicobacter pylori	Borrelia burgdorferi group	Parachlamydia	Corynebacterium amycolatum
N Present in G1	1	2	0	0	0	0	0
N Present in G2	2	0	0	0	0	0	0
N Present in G3	0	0	1	1	0	0	0
Percentage G1	8.33	16.67	0	0	0	0	0
Percentage G2	10.53	0	0	0	0	0	0
Percentage G3	0	0	12.5	12.5	0	0	0

Part 3: Presence/absence of pathogens tested for in the three microbiota groups. Number of samples with pathogen present, and percentage of samples shown for each group

Table S5a

Carbohydrate (SEED, Level 4)	Test-Statistic	Bonferroni_P	Group 1 mean	Group 2 mean	Modern mean
Pyruvate ferredoxin oxidoreductase	31.96	0.00	167.83	398.68	59.13
Pyruvate metabolism I anaplerotic reactions PEP	29.71	0.00	339.92	544.16	245.50
Lactose utilization	29.67	0.00	51.08	24.00	95.25
Lactate utilization	28.97	0.00	83.83	50.32	176.63
Methanogenesis	28.01	0.00	81.00	608.37	7.75
Isobutyryl CoA to Propionyl CoA Module	27.73	0.00	25.08	47.95	12.63
Formaldehyde assimilation Ribulose monophosphate pathway	27.67	0.00	3.50	28.26	1.00
Di Inositol Phosphate biosynthesis	26.48	0.00	12.17	29.16	1.88
Mannose Metabolism	26.42	0.00	97.25	53.05	149.63
L rhamnose utilization	25.93	0.00	51.83	29.79	80.13
Fructose utilization	25.27	0.00	63.00	36.05	99.75
Glycolysis and Gluconeogenesis including Archaeal enzymes	25.11	0.00	390.92	532.89	318.13
Acetyl CoA fermentation to Butyrate	24.89	0.00	131.00	203.74	104.38
Alpha acetolactate operon	24.73	0.00	3.83	0.53	7.00
Mannitol Utilization	23.59	0.00	54.08	29.05	83.00
Methylglyoxal Metabolism	22.94	0.00	29.92	17.16	50.25
Serine glyoxylate cycle	21.88	0.00	637.83	771.84	513.25
Maltose and Maltodextrin Utilization	21.80	0.00	305.75	186.37	308.75
Carboxysome	21.47	0.00	17.50	13.58	42.75
Sucrose utilization	20.79	0.00	42.08	17.37	63.25
Beta Glucoside Metabolism	19.86	0.00	68.25	33.42	75.50
Glycolysis and Gluconeogenesis	19.71	0.00	532.92	611.32	447.13
Chitin and N acetylglucosamine utilization	19.23	0.01	68.25	49.00	92.63
Glycerol fermenation to 1 3 propanediol	17.84	0.01	19.25	14.95	33.25
L fucose utilization temp	17.73	0.01	17.92	11.79	27.13
Fructooligosaccharides FOS and Raffinose Utilization	17.07	0.02	161.17	83.00	151.13
Trehalose Uptake and Utilization	16.91	0.02	31.92	18.16	32.38
Predicted carbohydrate hydrolases	15.61	0.03	49.33	38.84	76.13
Trehalose Biosynthesis	15.38	0.04	132.25	99.89	90.63
Lactose and Galactose Uptake and Utilization	15.08	0.04	220.33	145.11	261.50
Glycogen metabolism	14.90	0.05	284.08	196.58	247.38

Kruskal-Wallis output of the significantly different ($p < 0.05$) carbohydrate functions (SEED, Level 4) between the three microbiota groups

Table S5b

Fibre (KEGG)	Test-Statistic	Bonferroni P	Group 1 mean	Group 2 mean	Group 3 mean
Other glycan degradation	30.52	0.00	200.42	80.15	386.86
Glycosphingolipid biosynthesis globo series	29.39	0.00	98.92	33.70	160.00
Glycosaminoglycan degradation	27.20	0.00	48.83	11.45	134.00
Lipopolysaccharide biosynthesis	26.25	0.00	169.25	100.80	468.14
Glycosphingolipid biosynthesis ganglio series	25.77	0.00	47.92	10.95	120.00
Various types of N glycan biosynthesis	25.13	0.00	5.75	23.40	2.57
Galactose metabolism	24.05	0.00	706.83	481.10	721.71
Phenylalanine metabolism	22.85	0.00	214.00	343.55	192.86
Alanine aspartate and glutamate metabolism	22.83	0.00	1410.58	1732.75	1362.71
Butanoate metabolism	22.80	0.00	940.25	1146.95	733.00
Propanoate metabolism	19.93	0.00	793.17	937.40	680.14
Purine metabolism	19.91	0.00	3939.92	4435.90	3766.57
Pyrimidine metabolism	19.69	0.00	3161.33	3581.50	3015.71
D Arginine and D ornithine metabolism	19.32	0.00	2.25	0.65	6.43
Pyruvate metabolism	16.86	0.00	1718.25	1985.65	1517.86
Amino sugar and nucleotide sugar metabolism	13.93	0.02	1457.25	1293.95	1558.43

Kruskal-Wallis output of the significantly different ($p < 0.05$) fibre functions (Clayton *et* al. 2016) between the three microbiota groups

Table S5c

Amino Acids (SEED, Level 5)	Test-Statistic	Bonferroni P	Group 1 mean	Group 2 mean	Group 3 mean
Tryptophan synthase alpha chain EC 4 2 1 20	31.06	0.00	5.83	1.60	14.86
Adenosylhomocysteinase EC 3 3 1 1	30.48	0.00	37.00	79.35	13.14
Glutamine synthetase clostridia type EC 6 3 1 2	29.81	0.00	7.50	32.20	1.86
Coproporphyrinogen III oxidase oxygen independent EC 1 3 99 22	29.42	0.00	9.58	1.95	20.71
Serine hydroxymethyltransferase EC 2 1 2 1	28.76	0.00	76.00	104.30	46.43
Ketoisovalerate oxidoreductase subunit VorB EC 1 2 7 7	28.25	0.00	2.33	17.15	0.14
Ketoisovalerate oxidoreductase subunit VorA EC 1 2 7 7	26.64	0.00	2.08	21.85	0.14
S adenosylmethionine synthetase EC 2 5 1 6	26.62	0.00	62.75	88.75	45.71
NADP specific glutamate dehydrogenase EC 1 4 1 4	24.57	0.00	49.58	81.40	32.00
UDP 3 O 3 hydroxymyristoyl N acetylglucosamine deacetylase EC 3 5 1	23.13	0.00	7.50	2.95	13.57
Tryptophan synthase beta chain EC 4 2 1 20	22.85	0.00	40.92	23.05	39.14
L serine dehydratase EC 4 3 1 17	20.91	0.00	9.75	6.75	21.71
Aromatic amino acid aminotransferase EC 2 6 1 57	20.09	0.00	7.17	2.15	4.00
Threonine dehydratase biosynthetic EC 4 3 1 19	18.90	0.00	10.08	4.60	13.14
D 3 phosphoglycerate dehydrogenase EC 1 1 1 95	17.79	0.01	31.33	51.55	29.43
4 hydroxybenzoate polyprenyltransferase EC 2 5 1	17.64	0.01	5.00	1.80	5.29

Part 1: Kruskal-Wallis output of the significantly different ($p < 0.05$) amino acid functions (Muegge et al, 2014) between the three microbiota groups

Table S5d

Amino Acids (SEED, Level 5)	Test-Statistic	Bonferroni P	Group 1 mean	Group 2 mean	Group 3 mean
2 octaprenyl 3 methyl 6 methoxy 1 4 benzoquinol hydroxylase EC 1 14 13	16.70	0.01	2.17	0.20	7.57
Glutamate decarboxylase EC 4 1 1 15	16.59	0.01	6.75	6.20	0.43
Threonine dehydratase catabolic EC 4 3 1 19	15.94	0.02	0.83	1.50	0.00
Archaeal S adenosylmethionine synthetase EC 2 5 1 6	15.89	0.02	1.25	12.25	0.00
N acetylglutamate synthase EC 2 3 1 1	15.88	0.02	4.42	1.10	4.71
Threonine dehydratase EC 4 3 1 19	15.25	0.03	14.33	9.10	21.29
5 methyltetrahydropteroyltrigluta mate homocysteine methyltransferase EC 2 1 1 14	14.21	0.05	35.33	15.05	42.00
Glutamate synthase NADPH small chain EC 1 4 1 13	13.83	0.06	10.08	12.85	5.57
Ketoisovalerate oxidoreductase subunit VorC EC 1 2 7 7	13.38	0.07	0.17	1.65	0.00
Arginine deiminase EC 3 5 3 6	12.93	0.09	32.75	20.10	14.29
Coproporphyrinogen III oxidase oxygen independent EC 1 3 99 22 divergent putative	12.52	0.11	0.58	0.00	0.00
Biosynthetic Aromatic amino acid aminotransferase beta EC 2 6 1 57	12.41	0.12	1.50	0.65	2.14
Biosynthetic Aromatic amino acid aminotransferase alpha EC 2 6 1 57	11.32	0.21	1.17	0.15	4.29
Glutamate synthase NADPH large chain EC 1 4 1 13	11.16	0.22	10.33	21.50	20.57
Branched chain amino acid aminotransferase EC 2 6 1 42	10.33	0.34	31.67	43.80	28.86
Peptidoglycan N acetylglucosamine deacetylase EC 3 5 1	9.73	0.45	5.33	2.30	3.86

Part 2: Kruskal-Wallis output of the the significantly different ($p < 0.05$) amino acid functions (Muegge et al, 2014) between the three microbiota groups

Appendix II

Chapter IV: Supplementary Information

	All samples		Incisor only		Canine only		Premolar only		Molar only	
	R statistic	p-value	R statistic	p-value	R statistic	p-value	R statistic	p-value	R statistic	p-value
Date sampled	0.0057	0.3316	-0.0162	0.5624	-0.138	0.7894	0.063	0.2014	0.0229	0.2623
ExtractionID	0.0469	0.0233	0.0002	0.4747	0.546	0.0286	0.0547	0.2757	0.0776	0.0848
Library ID	-0.0083	0.6758	-0.0005	0.4469	0.4304	0.0105	0.0897	0.1672	-0.0094	0.5673
Tooth Type	0.285	0.0001	NA	NA	NA	NA	NA	NA	NA	NA
Tooth Surface	0.0476	0.1659	0.2478	0.0014	0.3684	0.0251	0.4926	0.0046	0.2512	0.005
Right/Left	0.023	0.0496	-0.0575	0.8437	0.143	0.1387	-0.0263	0.6286	0.0283	0.2447
Maxilla/Mandible	0.0384	0.1974	0.0724	0.1417	0.0143	0.4161	0.1872	0.0946	0.0638	0.1867
Sub/Supragingival	0.2461	0.0544	0.2356	0.1386	0.284	0.2132	0.1796	0.2887	0.4645	0.0931
Fragment Size	0.015	0.2862	-0.0155	0.5562	0.166	0.1286	0.0655	0.2494	0.1185	0.0193
Sample N =	128		36		14		24		54	

Table S2: Whole data set and sub-groups (by tooth type) assessed for association with sampling and laboratory protocols to identify potential confounding factors that explain variation within the data. ANSOIM test (9999 permutations) calculated from Bray-Curtis dissimilarity. Significant p values are highlighted green. Number of samples in each sub-group noted at the bottom.

	Incisor				Premolar				Molar			
	Buccal		Lingual		Buccal		Lingual		Buccal		Lingual	
	R statistic	p-value	R statistic	p-value	R statistic	p-value	R statistic	p-value	R statistic	p-value	R statistic	p-value
Date sampled	0.1111	0.3021	-0.1099	0.8672	NA	NA	-0.0603	0.7249	-0.0222	0.5115	0.0186	0.3523
ExtractionID	0.4531	0.0826	-0.038	0.5604	NA	NA	0.2143	0.1005	0.1	0.4856	0.1086	0.0975
Library ID	-0.0617	0.5953	0.0586	0.2782	0	0.6584	0.0758	0.2833	-0.2105	0.6825	0.0898	0.0722
Right/Left	-0.3587	1	0.0101	0.3714	NA	NA	-0.0137	0.5144	-0.3265	0.9144	-0.0003	0.4644
MaxillaMandible	-0.1125	0.6922	0.1667	0.1609	0	0.6699	0.1489	0.2425	0.1818	0.287	-0.0512	0.6416
Sub/Supragingival	NA	NA	0.2476	0.2129	NA	NA	NA	NA	NA	NA	0.6451	0.0541
Fragment Size	-0.2709	0.849	-0.0573	0.6274	NA	NA	0.0359	0.3797	0.7895	0.0457	0.0976	0.0838
Sample N =	9		18		3		18		7		36	

Table S3: Sub-groups of data assessed for association with sampling and laboratory protocols to identify potential confounding factors that explain variation within the data. ANSOIM test (9999 permutations) calculated from Bray-Curtis dissimilarity. With this level of filtering, there were no significant associations. Number of samples in each sub-group noted at the bottom.

	R statistic	p-value	N of samples	Group ratio		
Abscess Y/N	0.3089	0.0052	25	Diseased=7	Healthy=18	
Joint Porosity	0.1784	0.0308	30	Diseased=9	Healthy=29	
Vertebral CleftNeuralArch	0.7765	0.0342	30	Diseased=1	Healthy=29	
Joint Osteophytic Lipping	0.0926	0.0394	30	Diseased=14	Healthy=16	
Date 300	0.1172	0.0443	32	1000-1300=16	1300-1600=12	1600-1900=4
Non specific periostitis	0.1661	0.0466	36	Diseased=9	Healthy=27	

Table S4a: Molar-Lingual/Palatal-Supragingival sub-group of data assessed for association with Museum of London WORD database entries to identify potential factors that explain variation within the data. ANSOIM test (9999 permutations) calculated from Bray-Curtis dissimilarity. Only significant results are shown ($p < 0.05$).

	R statistic	p-value	N of samples	Group ratio	
VertebralPathology_SchmorlsNodes	0.3453	0.0115	18	Diseased=12	Healthy=6
Tooth_Rotation	0.6271	0.028	18	Diseased=2	Healthy=16
GeneralComments	0.6064	0.0333	18	Enthesopathy=2	Healthy=16
Tooth_Crowding	0.3513	0.0409	18	Diseased=4	Healthy=14
Overall_Tooth_Anomaly	0.3513	0.0415	18	Diseased=4	Healthy=14

Table S4b: Premolar-Lingual/Palatal-Supragingival sub-group of data assessed for association with Museum of London WORD database entries to identify potential factors that explain variation within the data. ANSOIM test (9999 permutations) calculated from Bray-Curtis dissimilarity. Only significant results are shown ($p < 0.05$).

Table S5a: Kruskall-Wallis results indicating taxa that were significantly different between individuals with and without dental abscesses ($p < 0.05$)

	Test-Statistic	Bonferroni P	Diseased mean	Healthy mean	Increased in:
Actinomyces	21.31	0.00	6760.86	6234.67	Diseased
Actinomyces_sp__oral_taxon_849	18.26	0.02	181.14	108.72	Diseased
Desulfomicrobium_baculatum	81.04	0.00	1190.86	791.39	Diseased
Desulfovibrionales	25.44	0.00	462.43	321.56	Diseased
Euryarchaeota	119.77	0.00	874.86	475.61	Diseased
Gemella	53.46	0.00	126.14	35.67	Diseased
Gemella_morbillorum	56.31	0.00	88.86	15.67	Diseased
Johnsonella_ignava	77.98	0.00	646.14	366.83	Diseased
Lactobacillales	16.50	0.04	182.14	112.67	Diseased
Lautropia_mirabilis	38.10	0.00	247.57	128.78	Diseased
Leptotrichia	29.20	0.00	186.43	96.28	Diseased
Leptotrichia_buccalis	29.02	0.00	83.71	27.94	Diseased
Methanobacteriaceae	460.63	0.00	2482.43	1194.67	Diseased
Methanobacteriales	197.85	0.00	744.00	297.33	Diseased
Methanobrevibacter	438.19	0.00	3125.86	1685.00	Diseased
Methanobrevibacter_smithii	876.99	0.00	8831.71	5326.28	Diseased
Neisseria	30.65	0.00	237.86	132.06	Diseased
Oribacterium_sp__oral_taxon_078	27.01	0.00	201.14	110.06	Diseased
Parvimonas	43.18	0.00	357.43	202.83	Diseased
Peptostreptococcus_stomatis	21.37	0.00	128.14	64.50	Diseased
Prevotella	65.13	0.00	523.00	293.83	Diseased
Prevotella_sp__oral_taxon_472	43.02	0.00	97.14	26.33	Diseased
Pseudoramibacter_alactolyticus	47.61	0.00	945.14	668.61	Diseased
Pyramidobacter_piscolens	41.49	0.00	577.14	378.67	Diseased
Selenomonas	32.29	0.00	643.00	455.11	Diseased
Streptococcus	1638.24	0.00	4168.43	1263.00	Diseased
Streptococcus_cristatus	139.00	0.00	323.14	90.28	Diseased
Streptococcus_mitis	18.63	0.01	20.29	1.56	Diseased
Streptococcus_pneumoniae	21.54	0.00	52.29	15.11	Diseased
Streptococcus_sanguinis	65.98	0.00	120.14	25.78	Diseased

Table S5b: Kruskall-Wallis results indicating taxa that were significantly different between individuals with and without dental abscesses ($p < 0.05$)

	Test-Statistic	Bonferroni P	Diseased mean	Healthy mean	Increased in:
Actinomyces_massiliensis	21.41	0.00	563.43	729.61	Healthy
Actinomycetales	25.32	0.00	2392.14	2752.94	Healthy
Aminomonas_paucivorans	21.72	0.00	66.86	132.06	Healthy
Bacteria	409.84	0.00	22675.57	27193.39	Healthy
Clostridiales	19.96	0.01	7264.86	7813.39	Healthy
Coriobacteriaceae	305.73	0.00	1927.14	3169.17	Healthy
Eggerthella	16.40	0.04	78.57	137.83	Healthy
Filifactor_alocis	22.76	0.00	244.57	361.72	Healthy
Firmicutes	58.42	0.00	8383.57	9402.67	Healthy
Fretibacterium_fastidiosum	152.55	0.00	2709.71	3696.33	Healthy
Olsenella	65.66	0.00	541.71	842.00	Healthy
Olsenella_sp__oral_taxon_809	64.86	0.00	429.71	699.06	Healthy
Olsenella_uli	29.64	0.00	497.14	683.89	Healthy
Porphyromonas_endodontalis	19.89	0.01	4.86	30.00	Healthy
Porphyromonas_gingivalis	46.87	0.00	148.14	290.28	Healthy
Propionibacterium_propionicum	348.59	0.00	1085.43	2135.44	Healthy
Synergistaceae	181.88	0.00	1236.43	2000.11	Healthy
Tannerella_forsythia	18.57	0.01	1148.29	1364.17	Healthy
Treponema_denticola	152.57	0.00	1066.71	1715.28	Healthy
uncultured_Termite_group_1_bacterium	52.56	0.00	53.43	156.28	Healthy

Table S5a: Kruskall-Wallis results indicating the functions (including dietary functions) that were significantly different between individuals with and without dental abscesses ($p < 0.05$)

	Test-Statistic	Bonferroni P	Diseased mean	Healthy mean	Increased in:
Flagellar hook length control protein (FliK)	82.88	0.00	185.29	49.83	Diseased
Carbohydrate					
	Test-Statistic	Bonferroni_P	Diseased mean	Healthy mean	Increased in:
Methanogenesis	61.43293267	3.99E-13	582.2857143	344.8888889	Diseased
Fibre					
NA					
Amino Acids					
NA					

Table S6a: Kruskall-Wallis results indicating taxa that were significantly different between individuals with and without porosity of the joints ($p < 0.05$) Molar-Lingual/Palatal-Supragingival sub-group

	Test-Statistic	Bonferroni P	Healthy mean	Diseased mean	Increased in:
Actinobacteria	24.43	0.00	2087.44	1780.19	Diseased
Actinomyces	21.27	0.00	6712.44	6188.67	Diseased
Bacteroidetes_oral_taxon_274	28.87	0.00	1003.11	776.71	Diseased
Campylobacter_gracilis	37.07	0.00	270.89	147.24	Diseased
Coriobacteriaceae	291.79	0.00	3302.78	2057.76	Diseased
Corynebacterium_matruchotii	93.38	0.00	376.67	156.71	Diseased
Euryarchaeota	297.30	0.00	950.44	342.52	Diseased
Gemella	22.43	0.00	92.22	38.71	Diseased
Gemella_morbillorum	21.12	0.00	59.33	19.38	Diseased
Leptotrichia	19.33	0.01	161.44	91.86	Diseased
Leptotrichia_buccalis	24.81	0.00	72.11	24.19	Diseased
Methanobacteriaceae	940.33	0.00	2632.33	861.57	Diseased
Methanobacteriales	320.45	0.00	750.22	211.10	Diseased
Methanobacterium	20.11	0.01	52.56	16.19	Diseased
Methanobrevibacter	1107.63	0.00	3435.89	1212.90	Diseased
Methanobrevibacter_ruminantium	407.40	0.00	1409.78	535.38	Diseased
Methanobrevibacter_smithii	2960.65	0.00	10075.44	3786.00	Diseased
Methanosphaera_stadtmanae	29.08	0.00	104.00	40.24	Diseased
Olsenella	123.18	0.00	933.33	514.05	Diseased
Olsenella_sp__oral_taxon_809	99.09	0.00	761.89	421.76	Diseased
Olsenella_uli	54.82	0.00	715.56	462.38	Diseased
Oribacterium_sp__oral_taxon_078	26.43	0.00	182.89	97.43	Diseased
Prevotella	86.44	0.00	512.89	257.24	Diseased
Prevotella_sp__oral_taxon_472	42.51	0.00	85.89	20.76	Diseased
Selenomonas	41.63	0.00	614.67	408.90	Diseased
Streptococcus	156.23	0.00	2486.00	1681.57	Diseased
Thermoplasmata	24.35	0.00	241.33	144.81	Diseased
Veillonellaceae	29.02	0.00	737.67	545.05	Diseased

Table S6b: Kruskall-Wallis results indicating functions (including dietary) that were significantly different between individuals with and without porosity of the joints ($p < 0.05$) Molar-Lingual/Palatal-Supragingival sub-group

	Test-Statistic	Bonferroni_P	Diseased mean	Healthy mean	Increased in:
Flagellar hook length control protein FliK	86.29	0.00	169.11	39.62	Diseased
Archaeal DNA polymerase II large subunit EC 2 7 7 7	19.78	0.02	64.89	23.71	Diseased
Thermosome subunit	18.15	0.06	59.56	21.76	Diseased
CoB CoM heterodisulfide reductase subunit A EC 1 8 98 1	18.05	0.06	80.78	35.48	Diseased
DNA directed RNA polymerase subunit A EC 2 7 7 6	15.06	0.29	50.67	18.81	Diseased
Formate dehydrogenase alpha subunit EC 1 2 1 2	13.23	0.76	44.44	16.48	Diseased
Cell division control protein 6	12.86	0.92	39.67	13.86	Diseased
OTU	Test-Statistic	Bonferroni_P	Diseased mean	Healthy mean	
Methanogenesis	184.55	0.00	566.00	198.13	
Fibre					
OTU	Test-Statistic	Bonferroni_P	Diseased mean	Healthy mean	
Various_types_of_N_glycan_biosynthesis	10.31	0.03	29.86	9.94	
Amino Acids					
NA					

Table S7a: Kruskall-Wallis results indicating taxa that were significantly different between individuals with and without osteophytic lipping of the joints ($p < 0.05$) Molar-Lingual/Palatal-Supragingival sub-group

	Test-Statistic	Bonferroni P	Diseased mean	Healthy mean	Increased in:
Actinobacteria	26.96	0.00	2042.21	1723.75	Diseased
Actinomyces	120.17	0.00	7006.07	5768.06	Diseased
Actinomyces_massiliensis	40.20	0.00	811.43	575.81	Diseased
Actinomycetales	66.53	0.00	2922.71	2332.06	Diseased
cellular_organisms	16.30	0.05	4023.29	3669.25	Diseased
Coriobacteriaceae	122.41	0.00	2844.07	2070.06	Diseased
Corynebacterium_matruchotii	30.35	0.00	284.93	168.25	Diseased
Euryarchaeota	261.81	0.00	803.29	281.31	Diseased
Methanobacteriaceae	793.17	0.00	2180.43	703.63	Diseased
Methanobacteriales	269.06	0.00	609.14	166.06	Diseased
Methanobrevibacter	963.14	0.00	2889.71	996.13	Diseased
Methanobrevibacter_ruminantium	314.18	0.00	1174.64	467.88	Diseased
Methanobrevibacter_smithii	2721.25	0.00	8624.29	3090.31	Diseased
Methanosphaera_stadtmanae	25.16	0.00	88.50	33.88	Diseased
Olsenella	52.73	0.00	778.86	518.19	Diseased
Olsenella_sp__oral_taxon_809	44.99	0.00	640.00	422.13	Diseased
Olsenella_uli	24.24	0.00	624.79	462.69	Diseased
Prevotella	23.50	0.00	401.00	275.25	Diseased
Prevotella_sp__oral_taxon_472	24.13	0.00	63.71	19.81	Diseased

Table S7b: Kruskall-Wallis results indicating taxa that were significantly different between individuals with and without osteophytic lipping of the joints ($p < 0.05$) Molar-Lingual/Palatal-Supragingival sub-group

	Test-Statistic	Bonferroni P	Diseased mean	Healthy mean	Increased in:
Bacteria	138.40	0.00	24501.07	27174.75	Healthy
Bacteroidetes_oral_taxon_274	20.73	0.00	745.29	931.56	Healthy
Betaproteobacteria	25.13	0.00	237.00	359.00	Healthy
Burkholderiales	27.44	0.00	122.14	218.19	Healthy
Campylobacter	38.73	0.00	352.93	538.06	Healthy
Campylobacter_rectus	24.25	0.00	68.57	138.88	Healthy
Clostridia	19.93	0.01	1240.79	1473.25	Healthy
Clostridiales	108.14	0.00	7140.00	8437.19	Healthy
Comamonadaceae	19.79	0.01	95.00	166.56	Healthy
Eubacterium	25.02	0.00	1964.79	2290.94	Healthy
Eubacterium_saphenum	341.16	0.00	1904.79	3219.69	Healthy
Filifactor_alocis	35.91	0.00	283.64	444.75	Healthy
Firmicutes	88.58	0.00	8577.21	9854.50	Healthy
Fretibacterium_fastidiosum	243.06	0.00	3070.86	4416.25	Healthy
Johnsonella_ignava	23.98	0.00	422.57	577.13	Healthy
Lachnospiraceae	16.22	0.05	699.71	858.56	Healthy
Neisseria_elongata	24.76	0.00	16.00	57.56	Healthy
Porphyromonas_gingivalis	170.61	0.00	197.86	547.63	Healthy
Proteobacteria	39.22	0.00	1015.79	1317.94	Healthy
Pseudoramibacter_alactolyticus	192.55	0.00	718.07	1343.19	Healthy
Pyramidobacter_piscolens	22.38	0.00	333.43	467.00	Healthy
Slackia	17.49	0.02	252.64	355.56	Healthy
Streptococcus	35.96	0.00	1725.36	2095.75	Healthy
Streptococcus_anginosus_group	50.38	0.00	52.93	152.69	Healthy
Synergistaceae	146.11	0.00	1460.86	2188.69	Healthy
Tannerella_forsythia	22.97	0.00	1363.43	1625.31	Healthy
Treponema	32.51	0.00	442.14	628.25	Healthy
Treponema_denticola	189.68	0.00	1213.14	1988.63	Healthy

Table S7c: Kruskall-Wallis results indicating functions (including dietary) that were significantly different between individuals with and without osteophytic lipping of the joints ($p < 0.05$) Molar-Lingual/Palatal-Supragingival sub-group

	Test-Statistic	Bonferroni_P	Diseased mean	Healthy mean	Increased in:
Flagellar hook length control protein FliK	56.52	0.00	128.21	34.94	Diseased
CoB CoM heterodisulfide reductase subunit A EC 1 8 98 1	19.90	0.02	72.64	28.44	Diseased
Archaeal DNA polymerase II large subunit EC 2 7 7 7	19.03	0.04	55.79	18.81	Diseased
Thermosome subunit	16.51	0.13	50.71	17.69	Diseased
DNA directed RNA polymerase subunit A EC 2 7 7 6	14.34	0.42	43.57	15.06	Diseased
Carbohydrate					
			Diseased		
OTU	Test-Statistic	Bonferroni_P	mean	Healthy mean	
Methanogenesis	179.45	0.00	646.00	251.43	
Ethanolamine_utilization	12.02	0.05	127.67	189.24	
Fibre					
			Diseased		
OTU	Test-Statistic	Bonferroni_P	mean	Healthy mean	
Various_types_of_N_glycan_biosynthesis	10.58	0.03	34.67	12.62	
Amino Acids					
NA					

Table S8a: Kruskall-Wallis results indicating taxa that were significantly different between individuals with and without periostitis ($p < 0.05$) Molar-Lingual/Palatal-Supragingival sub-group

	Test-Statistic	Bonferroni P	Healthy mean	Diseased mean	Increased in
Bacteroidales	27.79	0.00	852.67	1084.44	Diseased
Bacteroides	17.21	0.03	270.52	375.78	Diseased
Bacteroidetes	18.35	0.02	686.33	854.33	Diseased
Capnocytophaga	69.38	0.00	199.44	401.78	Diseased
Capnocytophaga_gingivalis	28.47	0.00	3.70	34.67	Diseased
Capnocytophaga_sp__CM59	41.14	0.00	17.26	77.44	Diseased
Clostridia	31.02	0.00	1241.56	1534.78	Diseased
Clostridiales	79.78	0.00	7333.15	8455.00	Diseased
Clostridiales_Family_XIII__Incertae_Sedis	42.94	0.00	534.07	770.11	Diseased
Clostridium_sp__HGF2	32.89	0.00	67.26	151.00	Diseased
Eubacterium	63.63	0.00	2014.19	2552.67	Diseased
Euryarchaeota	130.32	0.00	401.37	792.22	Diseased
Firmicutes	77.94	0.00	8734.81	9940.89	Diseased
Methanobacteriaceae	358.07	0.00	1047.96	2099.44	Diseased
Methanobacteriales	61.81	0.00	291.48	513.11	Diseased
Methanobrevibacter	613.13	0.00	1373.30	2989.44	Diseased
Methanobrevibacter_ruminantium	305.77	0.00	575.04	1327.44	Diseased
Methanobrevibacter_smithii	2397.74	0.00	3975.74	9594.33	Diseased
Methanosphaera_stadtmanae	37.32	0.00	37.67	110.56	Diseased
Mogibacterium_sp__CM50	36.45	0.00	816.33	1078.78	Diseased
Neisseria_elongata	33.75	0.00	23.22	81.00	Diseased
Oribacterium_sp__oral_taxon_078	26.05	0.00	96.00	180.22	Diseased
Prevotella	29.45	0.00	317.26	469.00	Diseased
Pyramidobacter_piscolens	97.57	0.00	323.85	625.67	Diseased
Thermoplasmata	21.89	0.00	170.30	267.89	Diseased

Table S8b: Kruskall-Wallis results indicating taxa that were significantly different between individuals with and without periostitis ($p < 0.05$) Molar-Lingual/Palatal-Supragingival sub-group

	Test-Statistic	Bonferroni_P	Healthy mean	Diseased mean	Increased in
Actinobacteria	62.55	0.00	2009.59	1539.11	Healthy
Actinomyces	59.71	0.00	6301.44	5463.67	Healthy
Actinomycetales	25.49	0.00	2617.04	2264.44	Healthy
Bacteria	64.60	0.00	26112.07	24307.56	Healthy
Betaproteobacteria	24.75	0.00	348.81	229.56	Healthy
Chloroflexi	22.54	0.00	168.37	92.22	Healthy
Coriobacteriaceae	116.04	0.00	2770.70	2026.11	Healthy
Deltaproteobacteria	35.25	0.00	434.37	276.67	Healthy
Desulfobulbaceae	26.13	0.00	199.33	110.00	Healthy
Desulfobulbus_propionicus	88.82	0.00	610.52	324.56	Healthy
Desulfomicrobium_baculatum	31.53	0.00	1034.37	794.56	Healthy
Desulfovibrio	19.47	0.01	284.07	188.44	Healthy
Eubacterium_saphenum	482.44	0.00	2780.96	1378.22	Healthy
Fretibacterium_fastidiosum	322.80	0.00	4165.93	2684.67	Healthy
Johnsonella_ignava	85.65	0.00	588.93	313.11	Healthy
Olsenella	36.35	0.00	754.26	538.00	Healthy
Olsenella_sp__oral_taxon_809	28.55	0.00	613.96	440.78	Healthy
Parvimonas	41.45	0.00	341.15	193.22	Healthy
Peptoniphilaceae	20.39	0.01	162.48	91.00	Healthy
Peptostreptococcus_stomatis	29.59	0.00	131.11	57.33	Healthy
Porphyromonas_gingivalis	38.52	0.00	393.33	238.11	Healthy
Propionibacterium_propionicum	209.42	0.00	2212.11	1352.33	Healthy
Proteobacteria	73.51	0.00	1294.78	894.67	Healthy
Selenomonas	35.72	0.00	651.22	453.11	Healthy
Slackia	43.70	0.00	379.11	218.44	Healthy
Streptococcus	180.95	0.00	2898.81	1963.67	Healthy
Streptococcus_anginosus_group	18.03	0.02	128.56	69.22	Healthy
Synergistaceae	96.64	0.00	1936.07	1372.00	Healthy
Treponema_denticola	50.61	0.00	1743.63	1348.56	Healthy

Table S8c: Kruskall-Wallis results indicating functions (including dietary) that were significantly different between individuals with and without periostitis ($p < 0.05$) Molar-Lingual/Palatal-Supragingival sub-group

	Test-Statistic	Bonferroni_P	Healthy mean	Diseased mean	Increased in:
CoB CoM heterodisulfide reductase subunit A (EC.1.8.98.1)	13.05	0.84	37.37	75.44	Diseased
Carbohydrate					
OTU	Test-Statistic	Bonferroni_P	Healthy mean	Diseased mean	
Methanogenesis	133.19	0.00	266.33	602.11	
Fibre					
OTU	Test-Statistic	Bonferroni_P	Healthy mean	Diseased mean	
Alanine aspartate and glutamate metabolism	10.57	0.03	1667.56	1860.67	
Amino Acids					
NA					

Table S9a: Kruskall-Wallis results indicating taxa that were significantly different between individuals with and without Scmorl's nodes ($p < 0.05$) Premolar-Lingual/Palatal-Supragingival sub-group

	Test-Statistic	Bonferroni P	Diseased mean	Healthy mean	Increased in:
Actinomyces_massiliensis	19.23	0.01	1066.50	873.50	Diseased
Aminobacterium_colombiense	26.83	0.00	168.33	86.33	Diseased
Bacteria	344.02	0.00	19960.00	16424.83	Diseased
Campylobacter	17.61	0.02	1390.75	1178.17	Diseased
Campylobacter_gracilis	68.09	0.00	650.00	385.83	Diseased
Campylobacter_rectus	33.18	0.00	175.00	83.33	Diseased
candidate_division_TM7_single_cell_isolate_TM7c	17.34	0.02	736.58	585.33	Diseased
cellular_organisms	58.28	0.00	2924.08	2369.17	Diseased
Clostridia	86.72	0.00	813.92	480.67	Diseased
Clostridiales	226.57	0.00	5096.50	3688.67	Diseased
Clostridiales_Family_XIII__Incertae_Sedis	52.05	0.00	347.25	182.50	Diseased
Clostridium_sp__HGF2	20.20	0.01	24.33	2.50	Diseased
Comamonadaceae	32.04	0.00	466.92	309.67	Diseased
Corynebacterium	34.64	0.00	378.25	233.33	Diseased
Corynebacterium_matruchotii	292.73	0.00	2093.92	1129.83	Diseased
Deltaproteobacteria	43.43	0.00	396.50	232.17	Diseased
Desulfobulbaceae	23.44	0.00	190.83	107.67	Diseased
Desulfobulbus_propionicus	68.97	0.00	560.50	316.17	Diseased
Desulfomicrobium_baculatum	151.57	0.00	801.42	382.33	Diseased
Desulfovibrio	58.94	0.00	333.17	163.67	Diseased
Desulfovibrio_sp__A2	19.19	0.01	25.25	3.17	Diseased
Desulfovibrionaceae	18.69	0.01	87.50	39.33	Diseased
Desulfovibrionales	59.99	0.00	328.75	159.33	Diseased
Eubacterium_saphenum	243.87	0.00	1064.33	462.33	Diseased
Filifactor_alocis	57.98	0.00	306.50	146.17	Diseased
Firmicutes	309.34	0.00	5896.83	4139.33	Diseased
Fretibacterium_fastidiosum	720.25	0.00	4015.50	1961.67	Diseased
Jonquetella_anthropi	29.58	0.00	170.83	84.67	Diseased
Methanobrevibacter	17.68	0.02	606.58	468.83	Diseased
Methanobrevibacter_smithii	54.46	0.00	1894.50	1467.17	Diseased

Table S9b: Kruskall-Wallis results indicating taxa that were significantly different between individuals with and without Scmorl's nodes ($p < 0.05$) Premolar-Lingual/Palatal-Supragingival sub-group

	Test-Statistic	Bonferroni P	Diseased mean	Healthy mean	Increased in:
Mogibacterium_sp__CM50	78.68	0.00	562.17	303.17	Diseased
Parvimonas	16.52	0.04	284.33	195.50	Diseased
Porphyromonas_gingivalis	373.75	0.00	692.58	152.33	Diseased
Propionibacterium_propionicum	229.70	0.00	5767.83	4253.50	Diseased
Pseudoramibacter_alactolyticus	162.81	0.00	500.25	175.33	Diseased
Slackia	26.43	0.00	254.17	151.17	Diseased
Synergistaceae	228.32	0.00	1826.75	1025.17	Diseased
Syntrophomonas_wolfei	24.80	0.00	33.58	4.50	Diseased
Tannerella_forsythia	352.47	0.00	2111.08	1063.17	Diseased
Treponema	44.47	0.00	647.67	429.50	Diseased
Treponema_denticola	145.23	0.00	1819.83	1164.17	Diseased
uncultured_Termite_group_1_bacterium	172.84	0.00	311.17	66.00	Diseased
Burkholderiales	16.16	0.05	692.25	550.67	Diseased

Table S9c: Kruskall-Wallis results indicating taxa that were significantly different between individuals with and without Scmorl's nodes ($p < 0.05$) Premolar-Lingual/Palatal-Supragingival sub-group

	Test-Statistic	Bonferroni P	Diseased mean	Healthy mean	
Actinobacteria	69.15	0.00	2096.42	2669.83	Healthy
Actinomyces	138.61	0.00	6772.08	8212.17	Healthy
Actinomyces_naeslundii	22.25	0.00	57.58	119.83	Healthy
Actinomyces_sp__ICM39	20.12	0.01	2.58	24.50	Healthy
Actinomyces_sp__oral_taxon_848	53.39	0.00	83.42	205.83	Healthy
Aggregatibacter	52.16	0.00	76.75	193.67	Healthy
Aggregatibacter_aphrophilus	42.91	0.00	25.08	94.67	Healthy
Capnocytophaga	555.35	0.00	905.67	2198.83	Healthy
Capnocytophaga_gingivalis	188.09	0.00	41.17	270.17	Healthy
Capnocytophaga_sp__CM59	47.73	0.00	110.67	238.33	Healthy
Capnocytophaga_sp__oral_taxon_326	28.17	0.00	27.75	82.33	Healthy
Capnocytophaga_sp__oral_taxon_329	90.55	0.00	27.17	147.17	Healthy
Capnocytophaga_sp__oral_taxon_332	29.42	0.00	16.25	63.17	Healthy
Capnocytophaga_sputigena	46.00	0.00	22.17	92.33	Healthy
Chlorobiaceae	441.82	0.00	10.25	377.17	Healthy
Chlorobium	47.22	0.00	13.33	75.00	Healthy
Chlorobium_Pelodictyon_group	174.04	0.00	2.67	142.50	Healthy
Chryseobacterium	18.30	0.01	1.08	18.50	Healthy
Coriobacteriaceae	86.87	0.00	1660.92	2242.17	Healthy
Eikenella_corrodens	29.09	0.00	151.08	259.83	Healthy
Eubacterium	280.31	0.00	1858.92	3023.17	Healthy
Flavobacteriaceae	156.81	0.00	198.08	529.83	Healthy
Flavobacteriales	23.91	0.00	32.75	85.00	Healthy
Flavobacteriia	20.12	0.01	30.17	75.67	Healthy
Fusobacteriales	17.26	0.03	31.00	72.83	Healthy
Fusobacterium	151.39	0.00	366.92	778.83	Healthy
Gemella	46.31	0.00	101.75	223.00	Healthy
Gemella_morbillorum	40.72	0.00	72.92	171.33	Healthy
Granulicatella_adiacens	24.36	0.00	31.33	83.33	Healthy
Haemophilus	22.31	0.00	26.33	72.50	Healthy

Table S9d: Kruskall-Wallis results indicating taxa that were significantly different between individuals with and without Scmorl's nodes ($p < 0.05$) Premolar-Lingual/Palatal-Supragingival sub-group

	Test-Statistic	Bonferroni P	Diseased mean	Healthy mean	
Haemophilus_parainfluenzae	24.21	0.00	1.83	25.67	Healthy
Johnsonella_ignava	151.17	0.00	692.92	1228.33	Healthy
Lachnospiraceae	18.10	0.02	897.00	1086.33	Healthy
Lactobacillales	32.38	0.00	195.25	324.33	Healthy
Lautropia_mirabilis	694.22	0.00	2326.17	4481.50	Healthy
Leptotrichia	61.96	0.00	256.08	466.17	Healthy
Leptotrichia_goodfellowii	62.60	0.00	74.00	203.33	Healthy
Leptotrichiaceae	18.98	0.01	46.42	98.33	Healthy
Neisseria	815.42	0.00	854.33	2465.00	Healthy
Neisseria_macacae	17.84	0.02	3.75	25.50	Healthy
Neisseria_mucosa	79.72	0.00	39.75	162.50	Healthy
Neisseria_sicca	36.87	0.00	15.83	70.17	Healthy
Neisseriaceae	72.26	0.00	287.92	529.17	Healthy
Olsenella	30.68	0.00	471.33	657.00	Healthy
Olsenella_sp__oral_taxon_809	25.47	0.00	383.58	536.33	Healthy
Pasteurellaceae	429.70	0.00	234.83	915.00	Healthy
Peptostreptococcaceae	130.27	0.00	180.75	466.17	Healthy
Porphyromonas	43.69	0.00	84.42	193.17	Healthy
Porphyromonas_catoniae	72.33	0.00	45.50	165.50	Healthy
Porphyromonas_sp__oral_taxon_279	17.21	0.03	26.50	65.83	Healthy
Prevotella	35.37	0.00	751.92	1000.50	Healthy
Prevotella_sp__oral_taxon_472	42.12	0.00	151.92	286.83	Healthy
Streptococcus	1311.62	0.00	4702.58	8890.83	Healthy
Streptococcus_sanguinis	214.65	0.00	173.08	559.83	Healthy
Veillonella	30.17	0.00	39.83	104.83	Healthy
Veillonellaceae	26.03	0.00	929.08	1162.17	Healthy

Table S9e: Kruskall-Wallis results indicating functions (including dietary) that were significantly different between individuals with and without Scmorl's nodes ($p < 0.05$) Premolar-Lingual/Palatal-Supragingival sub-group

	Test-Statistic	Bonferroni P	Diseased mean	Healthy mean	Increased in:
Duplicated ATPase component BL0693 of energizing module of predicted ECF transporter	13.88	0.55	148.42	91.00	Diseased
Carbohydrate NA					
Fibre NA					
Amino Acids NA					

	Test-Statistic	Bonferroni P	1000 1300 mean	1300 1600 mean	1600 1900 mean
Atopobium rimae	38.02	0.00	134.13	93.75	51.50
Bacteria	238.09	0.00	26070.94	25805.25	22982.00
Bacteroidales	36.31	0.00	929.13	783.42	689.00
Bacteroides	29.17	0.00	333.56	245.67	212.00
Bacteroidetes	56.36	0.00	763.69	634.00	498.00
Bacteroidetes oral taxon 274	536.85	0.00	1083.00	858.25	277.75
Campylobacter	121.32	0.00	603.50	411.50	281.50
Campylobacter gracilis	28.04	0.00	216.13	156.83	120.50
Campylobacter showae	82.29	0.00	152.06	70.17	35.50
Candidatus Methanomethylophilus alvus	50.53	0.00	135.19	66.67	46.00
Centipeda periodontii	32.13	0.00	38.19	15.25	4.25
Coriobacteriaceae	414.47	0.00	2933.19	2433.50	1588.50
Corynebacterium matruchotii	117.30	0.00	365.94	169.25	150.00
Desulfobulbaceae	52.60	0.00	198.00	171.67	83.00
Desulfobulbus propionicus	218.27	0.00	615.94	521.50	214.75
Eubacterium	65.57	0.00	2298.88	2055.25	1782.50
Facklamia ignava	31.72	0.00	38.69	34.83	5.50
Firmicutes	236.73	0.00	9306.13	9226.50	7532.25
Gemella	69.00	0.00	114.13	58.50	22.00
Gemella morbillorum	71.67	0.00	78.00	35.58	6.50
Lachnospiraceae	110.39	0.00	888.81	755.67	504.75
Lachnospiraceae bacterium oral taxon 082	22.45	0.01	31.75	25.75	5.25
Lactobacillales	49.08	0.00	191.31	119.33	79.25
Leptotrichia	45.98	0.00	135.81	93.83	46.25

Table S10a: Kruskall-Wallis results indicating taxa that were significantly different between individuals of different time periods ($p < 0.05$) Molar-Lingual/Palatal-Supragingival sub-group

	Test-Statistic	Bonferroni P	1000 1300 mean	1300 1600 mean	1600 1900 mean
Leptotrichia buccalis	42.41	0.00	53.56	24.83	6.50
Neisseria	44.51	0.00	192.56	100.58	93.75
Olsenella	194.24	0.00	834.19	617.75	359.75
Olsenella sp oral taxon 809	151.78	0.00	680.75	497.75	300.00
Olsenella uli	100.72	0.00	688.00	495.17	366.75
Oribacterium sp oral taxon 078	23.29	0.01	113.88	87.83	52.50
Peptoniphilus indolicus	33.14	0.00	109.44	98.75	43.50
Peptostreptococcus stomatis	95.20	0.00	141.50	112.92	25.25
Prevotella	61.31	0.00	384.94	276.58	198.25
Prevotella saccharolytica	25.17	0.00	38.25	15.58	7.00
Pyramidobacter piscolens	115.26	0.00	509.19	329.92	225.25
Selenomonas	457.67	0.00	830.69	368.42	193.50
Selenomonas sputigena	64.97	0.00	85.63	35.58	12.00
Streptococcus	3718.28	0.00	3994.56	1506.92	358.25
Streptococcus cristatus	434.71	0.00	349.44	87.83	12.50
Thermoplasmata	112.23	0.00	274.69	135.00	86.00
Thermoplasmatales archaeon BRNA1	22.55	0.01	57.19	28.75	18.25
Treponema denticola	116.60	0.00	1697.81	1657.17	1175.50
Veillonellaceae	251.99	0.00	862.81	497.58	336.00
Propionibacterium propionicum	384.55	0.00	1948.19	1773.58	3012.50
Capnocytophaga	86.04	0.00	323.13	150.42	156.25
Neisseria elongata	77.94	0.00	74.94	8.08	15.00
Capnocytophaga sp CM59	53.21	0.00	58.69	4.92	32.75
Methanobacterium	45.69	0.00	25.06	15.83	72.25
Cardiobacterium valvarum	33.47	0.00	154.25	69.67	100.75

Table S10b: Kruskall-Wallis results indicating taxa that were significantly different between individuals of different time periods ($p < 0.05$) Molar-Lingual/Palatal-Supragingival sub-group

	Test-Statistic	Bonferroni P	1000 1300 mean	1300 1600 mean	1600 1900 mean
Parvimonas	25.11	0.00	366.94	258.00	261.25
Burkholderiales	24.39	0.00	225.19	139.17	150.75
Actinomyces	460.55	0.00	5397.69	6744.08	7857.75
Actinomyces massiliensis	326.90	0.00	509.81	698.00	1223.75
Actinomyces sp oral taxon 448	49.51	0.00	298.94	417.83	494.75
Actinomyces urogenitalis	41.44	0.00	435.44	597.25	633.00
Actinomycetales	190.12	0.00	2295.81	2681.92	3315.75
cellular organisms	41.92	0.00	3682.88	3790.17	4225.75
Desulfomicrobium baculatum	1519.84	0.00	600.75	1034.00	2611.50
Desulfovibrio sp A2	126.39	0.00	7.56	21.83	108.75
Desulfovibrionales	523.94	0.00	251.81	408.42	987.75
Eggerthia catenaformis	37.10	0.00	121.44	152.83	231.50
Eubacterium saphenum	483.31	0.00	1878.31	3237.58	3269.75
Euryarchaeota	365.33	0.00	417.56	474.42	1061.00
Methanobacteriaceae	1357.79	0.00	1034.44	1243.08	3108.25
Methanobacteriales	627.97	0.00	254.06	298.92	1001.00
Methanobacterium paludis	21.84	0.02	7.25	12.92	34.75
Methanobrevibacter	1180.61	0.00	1460.38	1772.00	3677.00
Methanobrevibacter smithii	2065.92	0.00	4511.88	5679.50	9590.75
Methanothermobacter thermautotrophicus	21.12	0.02	1.94	7.33	22.25
Peptoniphilaceae	40.99	0.00	117.75	148.42	232.25
Pseudoramibacter alactolyticus	253.13	0.00	707.25	1436.17	1050.25
Johnsonella ignava	165.94	0.00	480.38	639.50	260.50

Table S10c: Kruskall-Wallis results indicating taxa that were significantly different between individuals of different time periods ($p < 0.05$) Molar-Lingual/Palatal-Supragingival sub-group

	Test-Statistic	Bonferroni P	1000 1300 mean	1300 1600 mean	1600 1900 mean
Porphyromonas gingivalis	165.26	0.00	262.94	516.92	199.00
Clostridium sp HGF2	139.24	0.00	96.06	126.33	6.50
Fretibacterium fastidiosum	110.11	0.00	3272.56	4174.50	3665.50
Clostridiales	84.07	0.00	7710.25	7826.17	6811.25
Slackia	73.27	0.00	347.81	386.17	192.50
uncultured Termite group 1 bacterium	68.49	0.00	106.25	129.25	31.50
Synergistaceae	68.31	0.00	1563.38	2030.17	1657.00
Streptococcus anginosus group	58.17	0.00	111.88	152.75	48.00
Tannerella forsythia	31.73	0.00	1250.56	1533.83	1470.50
Synergistes sp 3 1 syn1	20.21	0.04	267.19	333.33	228.25

Table S10d: Kruskall-Wallis results indicating taxa that were significantly different between individuals of different time periods ($p < 0.05$) Molar-Lingual/Palatal-Supragingival sub-group

	Test-Statistic	Bonferroni P	1000 1300 mean	1300 1600 mean	1600 1900 mean
Streptococcus	6527.03	0.00	7960.86	2727.00	1069.50
Lautropia mirabilis	2111.31	0.00	2749.57	2313.67	416.50
Neisseria	789.00	0.00	1428.00	554.67	358.50
Capnocytophaga	590.17	0.00	1407.29	927.00	403.00
Streptococcus cristatus	542.71	0.00	536.86	111.67	62.50
Bacteroidetes oral taxon 274	450.82	0.00	2421.29	1704.67	1167.00
Streptococcus sanguinis	327.47	0.00	354.43	140.00	31.50
Actinomyces turicensis	250.32	0.00	408.71	270.00	77.00
Pasteurellaceae	239.26	0.00	370.86	270.00	67.00
Capnocytophaga gingivalis	210.28	0.00	163.14	32.33	6.50
Flavobacteriaceae	203.80	0.00	355.14	289.33	79.00
Bacteroidetes	201.55	0.00	1463.86	892.00	869.00
Lachnospiraceae	116.99	0.00	1104.71	834.00	656.50
Lactobacillales	113.96	0.00	290.43	137.67	96.50
Leptotrichia goodfellowii	105.36	0.00	104.00	88.67	8.50
Paludibacter propionicigenes	103.00	0.00	151.00	79.00	24.00
Prevotella	94.10	0.00	737.71	669.00	424.00
Neisseria mucosa	89.46	0.00	91.00	18.00	12.00
Peptostreptococcaceae	82.43	0.00	270.14	133.67	110.00
Aggregatibacter	80.10	0.00	122.57	117.33	26.50
Neisseriaceae	73.60	0.00	363.57	233.00	171.50
Fusobacterium	73.08	0.00	558.14	340.67	340.00
Granulicatella adiacens	72.17	0.00	66.00	9.33	8.50
Capnocytophaga sp oral taxon 329	71.91	0.00	80.57	30.67	8.00

Table S10e: Kruskall-Wallis results indicating taxa that were significantly different between individuals of different time periods ($p < 0.05$) Premolar-Lingual/Palatal-Supragingival sub-group

	Test-Statistic	Bonferroni P	1000 1300 mean	1300 1600 mean	1600 1900 mean
Leptotrichia	71.87	0.00	293.29	278.67	134.50
Capnocytophaga sp CM59	71.32	0.00	161.86	123.00	45.50
Gemella	66.42	0.00	160.14	95.00	46.50
Neisseria elongata	63.65	0.00	145.57	133.67	45.50
Streptococcus pneumoniae	57.88	0.00	76.14	23.33	13.00
Lachnospiraceae oral taxon 107	54.43	0.00	127.71	78.33	36.00
Campylobacter showae	42.23	0.00	398.14	276.33	243.00
Johnsonella ignava	36.09	0.00	926.43	900.33	704.00
Neisseria sp oral taxon 014	31.78	0.00	91.86	55.67	31.00
Actinomyces odontolyticus	25.44	0.00	38.86	24.33	6.50
Flavobacteriales	24.80	0.00	54.29	39.33	14.50
Capnocytophaga sp oral taxon 326	24.33	0.00	46.43	27.33	10.50
Bacilli	24.17	0.00	181.43	114.33	106.50
Facklamia	23.78	0.01	25.71	15.33	2.00
Streptococcaceae	22.71	0.01	39.29	14.33	9.50
Flavobacteriia	21.06	0.02	47.57	33.67	13.00
Leptotrichiaceae	19.91	0.04	57.14	50.33	20.50
Actinomyces sp oral taxon 181	19.79	0.04	32.29	31.00	7.50
Porphyromonas gingivalis	3831.16	0.00	274.14	155.33	2711.00
Tannerella forsythia	1423.64	0.00	1787.71	1160.67	3601.50
Eubacterium saphenum	1260.85	0.00	905.43	425.67	2078.00
Fretibacterium fastidiosum	1235.25	0.00	3754.43	2503.33	5609.00
Firmicutes	997.75	0.00	5932.14	4514.67	8005.00
Clostridiales	817.58	0.00	5063.43	4043.33	6981.00
Pseudoramibacter alactolyticus	478.95	0.00	391.29	289.67	986.00
Synergistaceae	428.43	0.00	1703.71	1236.33	2472.50

Table S10f: Kruskall-Wallis results indicating taxa that were significantly different between individuals of different time periods ($p < 0.05$) Premolar-Lingual/Palatal-Supragingival sub-group

	Test-Statistic	Bonferroni P	1000 1300 mean	1300 1600 mean	1600 1900 mean
Bacteria	419.53	0.00	20055.86	18989.67	23034.50
Desulfomicrobium baculatum	377.40	0.00	852.14	262.00	372.00
Eubacterium	276.13	0.00	2130.29	1259.00	2066.50
Campylobacter rectus	258.20	0.00	123.71	82.67	386.50
Clostridia	214.58	0.00	783.29	583.33	1181.50
Bacteroidales	181.70	0.00	1508.71	912.67	1468.50
Filifactor alocis	162.52	0.00	298.86	167.67	485.00
Clostridiales Family XIII Incertae Sedis	152.56	0.00	295.14	248.67	568.50
Treponema denticola	137.76	0.00	1719.71	1362.00	2044.50
Slackia	136.33	0.00	220.57	203.67	466.00
Campylobacter	123.34	0.00	1344.71	921.33	1420.50
Clostridium sp HGF2	120.92	0.00	12.43	4.33	89.50
Desulfovibrionales	119.57	0.00	343.00	118.67	187.50
Veillonellaceae	82.15	0.00	963.43	671.00	1017.50
Mogibacterium sp CM50	74.31	0.00	526.00	448.67	733.50
Treponema	73.55	0.00	579.00	501.67	798.00
Porphyromonadaceae	65.28	0.00	174.29	114.67	269.50
Selenomonas	58.84	0.00	591.00	366.00	543.00
Jonquetella anthropi	58.79	0.00	154.86	107.00	247.00
Aminobacterium colombiense	58.42	0.00	153.29	105.67	244.50
Veillonella	58.01	0.00	93.14	25.33	26.00
Subdoligranulum sp 4 3 54A2FAA	56.59	0.00	36.29	27.00	102.00
Desulfovibrio	54.69	0.00	322.86	198.00	366.50
candidate division TM7 genomosp GTL1	50.31	0.00	172.43	158.00	293.00

Table S10g: Kruskall-Wallis results indicating taxa that were significantly different between individuals of different time periods ($p < 0.05$) Premolar-Lingual/Palatal-Supragingival sub-group

	Test-Statistic	Bonferroni P	1000 1300 mean	1300 1600 mean	1600 1900 mean
Porphyromonas endodontalis	40.71	0.00	14.86	11.00	55.00
Dethiosulfovibrio peptidovorans	39.78	0.00	91.00	63.33	152.50
Fusobacterium necrophorum	37.88	0.00	63.14	23.67	85.50
Clostridium	34.99	0.00	531.29	460.33	655.00
Streptococcus anginosus	31.91	0.00	58.00	17.00	18.50
Peptoniphilus sp oral taxon 386	31.05	0.00	102.00	73.33	155.50
Pyramidobacter piscolens	30.84	0.00	356.43	348.00	490.00
Bacteroides	29.53	0.00	420.43	302.33	439.00
Syntrophomonas wolfei	25.28	0.00	41.57	9.33	38.00
Methanobrevibacter smithii	399.98	0.00	1313.71	1819.00	2538.00
candidate division TM7 single cell isolate TM7c	151.65	0.00	596.00	617.00	1028.50
Methanobrevibacter	107.44	0.00	419.14	604.67	774.50
Methanobacteriaceae	79.55	0.00	305.57	442.00	567.00
cellular organisms	44.10	0.00	2788.14	2883.00	3275.00
Methanobrevibacter ruminantium	29.94	0.00	188.57	276.33	305.00
Campylobacter gracilis	21.66	0.02	423.43	532.00	562.50
Methanobacteriales	21.49	0.02	73.29	104.33	140.50
Actinomycetales	1943.65	0.00	2232.57	5802.00	2770.00
Actinomyces	5072.19	0.00	4810.71	13860.67	6866.00
Actinomyces massiliensis	1306.58	0.00	414.43	1992.33	1836.50
Coriobacteriaceae	817.70	0.00	1908.71	2631.33	956.50
Actinobacteria	776.77	0.00	2754.29	3126.67	1367.50
Actinomyces urogenitalis	667.93	0.00	300.29	1175.00	420.50
Actinomyces sp oral taxon 448	468.96	0.00	197.71	856.33	393.50
Cardiobacterium valvarum	455.22	0.00	414.43	721.67	129.00
Propionibacterium propionicum	397.88	0.00	4363.71	6291.00	4728.00
Corynebacterium matruchotii	384.75	0.00	1378.57	1789.33	806.50

Table S10h: Kruskall-Wallis results indicating taxa that were significantly different between individuals of different time periods ($p < 0.05$) Premolar-Lingual/Palatal-Supragingival sub-group

	Test-Statistic	Bonferroni P	1000 1300 mean	1300 1600 mean	1600 1900 mean
Olsenella	320.10	0.00	554.14	789.00	234.00
Olsenella sp oral taxon 809	286.34	0.00	443.57	647.00	177.00
Cardiobacterium	185.83	0.00	179.00	306.67	58.00
Burkholderiales	163.95	0.00	558.00	812.67	378.00
Olsenella uli	139.94	0.00	396.29	584.33	248.50
Actinobacteridae	139.01	0.00	191.43	449.00	206.00
Proteobacteria	131.00	0.00	1874.57	1983.33	1363.00
Peptostreptococcus stomatis	116.32	0.00	184.29	235.00	60.00
Actinomycetaceae	104.76	0.00	118.29	314.67	157.00
Betaproteobacteria	92.72	0.00	707.00	741.00	438.50
Actinomyces sp oral taxon 849	90.79	0.00	105.71	253.33	101.50
Cardiobacterium hominis	85.86	0.00	122.57	180.67	46.00
Comamonadaceae	80.41	0.00	357.00	546.33	299.50
Corynebacterium	66.00	0.00	264.86	357.67	172.50
Gemella morbillorum	65.79	0.00	119.14	67.67	25.50
Propionibacterium	52.86	0.00	202.14	324.00	173.00
Prevotella sp oral taxon 472	52.27	0.00	132.43	150.00	54.00
Actinomyces sp oral taxon 178	52.22	0.00	131.29	221.00	98.00
Parvimonas	49.65	0.00	283.43	294.67	159.00
Atopobium rimae	46.08	0.00	79.86	120.00	37.00
Propionibacterium freudenreichii	43.21	0.00	23.71	59.00	9.00
Actinomyces oris	35.79	0.00	37.86	98.33	42.00
Collinsella aerofaciens	32.57	0.00	50.71	71.33	19.00
Actinomyces sp oral taxon 170	31.88	0.00	45.57	97.67	39.00

Table S10i: Kruskall-Wallis results indicating taxa that were significantly different between individuals of different time periods ($p < 0.05$) Premolar-Lingual/Palatal-Supragingival sub-group

	Test-Statistic	Bonferroni P	1000 1300 mean	1300 1600 mean	1600 1900 mean
Actinomyces sp oral taxon 848	29.63	0.00	119.57	207.00	127.00
Actinomyces sp oral taxon 171	28.88	0.00	34.14	85.33	38.00
Parvimonas sp oral taxon 393	28.10	0.00	36.86	40.00	7.50
Streptococcus agalactiae	27.85	0.00	36.00	6.33	10.00
Micromonosporaceae	26.94	0.00	9.29	35.67	7.00
Alphaproteobacteria	26.77	0.00	103.29	158.00	81.00
Actinomyces naeslundii	26.66	0.00	52.43	104.33	48.50
Chloroflexi	23.90	0.01	129.57	170.00	91.50
Leptotrichia buccalis	22.35	0.01	84.29	97.67	44.00
Gammaproteobacteria	21.98	0.01	247.71	262.67	171.50
Atopobium	21.68	0.02	48.86	58.33	19.50
Prevotella saccharolytica	19.42	0.05	75.57	92.33	42.50

Table S10j: Kruskall-Wallis results indicating taxa that were significantly different between individuals of different time periods ($p < 0.05$) Premolar-Lingual/Palatal-Supragingival sub-group

	Test-Statistic	Bonferroni P	1000 1300 mean	1300 1600 mean	1600 1900 mean
FIG019733 possible DNA binding protein	22.45	0.04	0	0	0.75
D serine permease DsdX	22.45	0.04	0	0	0.75

Table S10j: Kruskall-Wallis results indicating functions that were significantly different between individuals of different time periods ($p < 0.05$) Molar-Lingual/Palatal-Supragingival sub-group

	Test-Statistic	Bonferroni P	1000 1300 mean	1300 1600 mean	1600 1900 mean
NA	NA	NA	NA	NA	NA

Table S10j: Kruskall-Wallis results indicating functions that were significantly different between individuals of different time periods ($p < 0.05$) Molar-Lingual/Palatal-Supragingival sub-group

Carbohydrate

OTU	Test-Statistic	Bonferroni P	1000 1300 mean	1300 1600 mean	1600 1900 mean
Methanogenesis	189.68	0.00	293.25	350.42	683.00
Fructooligosaccharides FOS and Raffinose Utilization	15.56	0.04	149.94	117.50	89.25

Fibre

NA

Amino Acids

OTU	Test-Statistic	Bonferroni P	1000 1300 mean	1300 1600 mean	1600 1900 mean
Archaeal S adenosylmethionine synthetase EC 2 5 1 6	25.47	0.00	1.94	2.75	21.25

Table S10k: Kruskall-Wallis results indicating dietary functions that were significantly different between individuals of different time periods ($p < 0.05$) Molar-Lingual/Palatal-Supragingival sub-group

Carbohydrate

OTU	Test-Statistic	Bonferroni P	1000 1300 mean	1300 1600 mean	1600 1900 mean
Pyruvate ferredoxin oxidoreductase	39.82	0.00	222.57	181.00	317.50
Methanogenesis	31.49	0.00	92.71	125.33	184.00
Fructooligosaccharides FOS and Raffinose Utilization	27.19	0.00	187.00	153.33	100.00
Maltose and Maltodextrin Utilization	26.22	0.00	337.14	326.33	226.50
Pyruvate metabolism I anaplerotic reactions PEP	20.20	0.00	454.71	396.67	533.00
Sucrose utilization	18.64	0.01	52.57	39.67	17.50
Beta Glucoside Metabolism	17.17	0.02	81.14	78.00	40.00
Lactose and Galactose Uptake and Utilization	16.97	0.02	241.57	191.33	160.00

Fibre

OTU	Test-Statistic	Bonferroni P	1000 1300 mean	1300 1600 mean	1600 1900 mean
Galactose metabolism	26.10	0.00	691.57	702.00	540.50
Amino sugar and nucleotide sugar metabolism	19.02	0.00	1562.86	1518.00	1340.50
Glycosphingolipid biosynthesis globo series	12.40	0.04	88.29	68.00	47.50

Amino Acid

OTU	Test-Statistic	Bonferroni P	1000 1300 mean	1300 1600 mean	1600 1900 mean
Glutamine synthetase clostridia type EC 6 3 1 2	16.34	0.02	20.14	15.67	44.00

Table S10l: Kruskall-Wallis results indicating dietary functions that were significantly different between individuals of different time periods ($p < 0.05$) Premolar-Lingual/Palatal-Supragingival sub-group

Appendix III

Reconstructing Neandertal behavior, diet, and disease using ancient DNA from dental calculus

(Weyrich *et al*. 2016)

This appendix presents a study of Neandertal dental calculus, to which I contributed a section of statistical analysis and provided edits and comments on the manuscript. The samples analysed in this study are used in Chapter III to provide dietary comparisons to the British data.

Reconstructing Neandertal behavior, diet, and disease using ancient DNA from dental calculus

Authors: Laura S Weyrich[1], Sebastian Duchene[2], Julien Soubrier[1], Luis Arriola[1], Bastien Llamas[1], James Breen[1], Alan G Morris[3], Kurt W Alt[4], David Caramelli[5], Veit Dresely[6], Milly Farrell[7], Andrew G Farrer[1], Michael Francken[8], Neville Gully[9], Wolfgang Haak[1], Karen Hardy[10], Katerina Harvati[8], Petra Held[11], Edward C. Holmes[2], John Kaidonis[9], Carles Lalueza-Fox[12], Marco de la Rasilla[13], Antonio Rosas[14], Patrick Semal[15], Arkadiusz Soltysiak[16], Grant Townsend[9], Donatella Usai[17], Joachim Wahl[18], Daniel H. Huson[19], Keith Dobney[20], and Alan Cooper[1*]

Affiliations:

[1]Australian Centre for Ancient DNA, School of Biological Sciences, University of Adelaide, Adelaide, South Australia, Australia

[2]Marie Bashir Institute for Infectious Diseases and Biosecurity, Charles Perkins Centre, School of Life and Environmental Sciences and Sydney Medical School, University of Sydney, Sydney, Australia

[3]Department of Human Biology, University of Cape Town, Cape Town, South Africa

[4]Danube Private University, Krems, Austria; State Office for Heritage Management and Archaeology, Saxony-Anhalt, Germany; Heritage Museum, Halle, Germany; and Institute for Prehistory and Archaeological Science, Basel University, Switzerland

[5]Department of Biology, University of Florence, Florence, Italy

[6]Archaeology Saxony-Anhalt and Heritage Museum, Halle, Germany

[7]Human Origins and Palaeo Environments Group, Oxford Brookes University, Oxford, United Kingdom

[8]Paleoanthropology, Senckenberg Centre for Human Evolution and Paleoenvironments, Eberhard Karls University of Tübingen

[9]School of Dentistry, The University of Adelaide, Adelaide, Australia

[10]ICREA (Catalan Institution for Research and Advanced Studies) Pg. Lluís Companys 23, 08010 Barcelona, Catalonia, Spain; Departament de Prehistòria, Facultat de Filosofia i Lletres, Universitat Autònoma de Barcelona, Barcelona, Catalonia, Spain.

[11]Institute of Anthropology, University of Mainz

[12]Institute of Evolutionary Biology, CSIC-Universitat Pompeu Fabra, Barcelona, Spain

[13]Área de Prehistoria, Departamento de Historia, Universidad de Oviedo, Oviedo, Spain

[14]Paleoanthropology Group, Department of Paleobiology, Museo Nacional de Ciencias Naturales, CSIC, Madrid, Spain

[15]Scientific Service Heritage, Royal Belgian Institute of Natural Sciences, Brussels, Belgium

[16]Department of Bioarchaeology, Institute of Archaeology, University of Warsaw, Warsaw, Poland.

[17]Istituto Italiano per l'Africa e l'Oriente (IsIAO), Rome, Italy

[18]State Office for Cultural Heritage Management Baden-Württemberg, Esslingen Germany

[19]Department of Algorithms in Bioinformatics, University of Tübingen, Tübingen, Germany

[20]Department of Archaeology, Classics and Egyptology, School of Histories, Languages and Cultures, University of Liverpool, Liverpool, United Kingdom

Introductory Paragraph

Recent genomic data has revealed multiple interactions between Neandertals and modern humans[1], but there is currently little genetic evidence about Neandertal behavior, diet, or disease. We shotgun sequenced ancient DNA from five Neandertal dental calculus specimens to characterize regional differences in Neandertal ecology. In Spy, Belgium, Neandertal diet was heavily meat based, and included woolly rhinoceros and wild sheep (mouflon), characteristic of a steppe environment. In El Sidrón, Spain, no meat was detected, and dietary components of mushrooms, pine nuts, and moss reflected forest gathering[2,3]. Differences in diet were also linked to an overall shift in the oral bacterial community (microbiota) in Neandertals, and suggested that meat consumption contributed to significant variation between Neandertal microbiota. Evidence for self-medication was recognized in one male El Sidrón Neandertal with a dental abscess[4], who also suffered from a chronic gastrointestinal pathogen (*Enterocytozoon bieneusi*). Metagenomic data from this individual also contained a nearly complete genome of the archaeal commensal *Methanobrevibacter oralis* – the oldest draft microbial genome (10.2x depth of coverage) generated to date at ~48,000 years old. DNA preserved within dental calculus represents an important new resource of behavioral and health information for ancient hominin specimens, as well as a unique long-term study system for microbial evolution.

Main Text

Neandertals remain our closest known extinct hominin relative, and co-existed and occasionally interbred with anatomically modern humans (AMHs) across Eurasia in the Late Pleistocene[1]. Neandertals became extinct in Europe around 40,000 years ago (40 Kyr), while the extinction process across the rest of Eurasia is less clear[5,6]. Isotopic and archaeological data through the last glacial cycle (~120-12 Kyr) suggest that Neandertals were as carnivorous as polar bears or wolves[7,8], with a diet heavily based on large terrestrial herbivores, such as reindeer, woolly mammoth, and woolly rhinoceros[9]. In contrast, microwear analysis of tooth surfaces from Neandertals in different ecological settings, such as wooded areas or open plains, suggests diets were guided by local food availability[3]. Furthermore, analysis of phytoliths, starch granules, and proteins preserved in calcified dental plaque (calculus) indicate that Neandertal diets included many plants, including some that were potentially cooked prior to consumption or used for medicinal purposes[10,11]. As a result, Neandertal diet remains a topic of considerable debate, with little known about the specific animals and plants consumed, or the potential impacts on Neandertal health or disease.

While genomic studies continue to reveal evidence of interbreeding between AMHs and Neandertals across Eurasia[12], little is known about the health consequences of these interactions. The genetic analysis of Neandertal dental calculus represents an opportunity to examine this issue, and to reconstruct Neandertal diet, behavior, and disease[13,14]. Here, we report the first genetic analysis of dental calculus from five Neandertals: from El Sidrón cave in Spain (n=2); Spy cave in Belgium (n=2); and Breuil Grotta in Italy (n=1), alongside a

historic chimpanzee (n=1) and a modern human (n=1) for comparison, as well as low coverage sequencing of calculus from a wide-range of ancient humans (Table S1). To provide increased resolution of the diseases that may have impacted Neandertals, we also deeply sequenced (>147 million reads) dental calculus from the best preserved specimen, El Sidrón 1, who was suffering from a dental abscess[4].

Recent reports have identified that significant size-based PCR amplification biases confound the standard approach used to study ancient dental calculus[15] (*i.e.* using 16S ribosomal RNA (rRNA) amplicon sequencing to identify bacterial taxa[13,14]). Consequently, we examined the Neandertal dental calculus specimens – the oldest examined to date – using metagenomic shotgun approaches with Illumina sequencing, in addition to standard 16S rRNA amplicons (V4 region)[16]. Decontamination protocols and multiple blank controls (EBCs) were used to monitor contamination. A stringent filtering strategy was also developed to remove laboratory contaminant sequences in both data sets using QIIME (amplicon) or MEGAN5 (shotgun)[17,18] limiting bias from laboratory contamination[19,20]. As anticipated, the 16S amplicon data sets were not representative of the biodiversity revealed by shotgun sequencing (Figures S3-S4, S12, and S15-S16; Tables S2 and S7), and clustered together irrespective of sample age, while containing disproportionately large amounts of non-oral and environmental contaminant microorganisms (Figure 1, Tables S2 and S7). These results confirm the importance of shotgun sequencing approaches for ancient specimens, as amplicon data can be biased by DNA degradation and lead to preferential amplification of contaminating modern DNA and/or short target sequences (*i.e.* prokaryotic species with short ribosomal sequences)[15]. As a result,

the 16S amplicon data sets were discarded in favour of metagenomic shotgun sequencing.

The shotgun data sets consisted of short DNA fragments (*e.g.* <70 bp) which complicated accurate bacterial species identification using standard software, such as MG-RAST and DIAMOND (Figure S14)[18,21]. To circumvent this problem, we used a novel metagenomic alignment tool that rapidly identifies species from shorter fragment lengths using a BLASTX like algorithm (MEGAN ALignment Tool with BLASTX-like alignments; MALTX) against the NCBI non-redundant reference sequence database (2014)[22]. We benchmarked this tool by comparing its performance to other accepted programs within the field (Figure S8-S10 and S14)[18]. We verified that the bacterial diversity accurately matched the subset of 16S rRNA fragments identified within the shotgun data set. We also confirmed this pattern for two previously published ancient calculus data sets (Figure 1 and S13; Table S6)[23]. Within the ancient dental calculus samples, the Grotta Breuil Neandertal failed to produce amplifiable sequences, and was therefore excluded from downstream analysis. Bioinformatic filtering of environmental contaminants revealed that the Spy Neandertals were more heavily impacted by environmental contamination (Figure S15-S17; Table S7)[18]. Indeed, shotgun sequences from Spy I clustered more closely to those from the modern individual than other Neandertals (Figure 1), contained similar diversity to environmental samples (Figure S20), and presented DNA damage patterns characteristic of contamination with modern DNA sequences (Figure S22). Therefore, this individual was also excluded from further analyses[18]. The three robust Neandertal shotgun metagenomic data sets (El Sidrón 1, El Sidrón 2, and Spy II) contained an average of 93.76% bacterial, 5.91% archaeal, 0.27%

eukaryotic, and 0.06% viral identifiable sequences, similar to previously published ancient and modern human dental calculus (Figure 2A and S17)[14]. The six dominant bacterial phyla in the modern human mouth (Actinobacteria, Firmicutes, Bacteroidetes, Fusobacteria, Proteobacteria, and Spirochaetes) were also dominant in each of the Neandertals, with an average of 222 bacterial species per individual (Figure 2A and S17)[18].

We first examined Neandertal diets using the eukaryotic diversity preserved within the dental calculus, after filtering spurious results[18]. Calculus from the Spy II individual contained high numbers of reads mapping to rhinoceros (*Ceratotherium simum*) and sheep (*Ovis aries*), as well as the edible 'grey shag' mushroom (*Coprinopsis cinerea*) (Table 1). Bones of woolly rhino, reindeer, mammoth and horses were present in Spy Cave[24,25], while wild mouflon sheep were known to inhabit Europe throughout the Pleistocene[18,26]. Woolly rhino has long been suspected to be part of the Spy Neandertal diet[27], and the genetic signals confirm the highly carnivorous diet inferred from the isotope data obtained from the Spy individuals[8,28].

The dietary profile in El Sidrón Neandertals was markedly different from Spy, and contained no sequences matching large herbivores or suggesting high meat consumption. However, reads mapping to edible mushrooms ('split gill'; *Schizophyllum commune*), pine nuts (*Pinus koraiensis*), forest moss (*Physcomitrella patens*), and poplar (*Populus trichocarpa*) were identified (Table 1). Sequences mapping to plant fungal pathogens were also observed (*Zymoseptoria tritici, Phaeosphaeria nodorum, Penicillium rubens,* and *Myceliophthora thermophila*), suggesting the El Sidrón Neandertals may have been consuming molded herbaceous material. There is limited zooarchaeological

evidence about the El Sidrón individuals, and protein analysis of dental calculus has provided the only known dietary information, identifying different plant starches and evidence of cooking[11]. Our genetic data create the first detailed description of El Sidrón diet, and together with the dental calculus sequences from Spy, indicate that Neandertal groups across Europe used multiple subsistence strategies[29] according to location and food availability[2,3]. Further analyses, such as DNA hybridization and proteomic sequencing, will be needed to verify and extend the dietary reconstructions[18].

Bitter compounds were previously identified in calculus from El Sidrón 1, suggesting that this individual may have been self-medicating a dental abscess[11]. Our findings support this suggestion, as this was the only individual whose calculus contained sequences corresponding to the natural antibiotic producing *Penicillium* and poplar, whose bark, roots, and leaves contains the natural pain killer, salicylic acid (*i.e.* the active ingredient in aspirin)[30]. Surprisingly, this individual also yielded sequences matching the intracellular eukaryotic pathogen, microsporidia (*Enterocytozoon bieneusi*), which causes acute diarrhea in humans[31]. This suggests that this Neandertal may have been medicating a gastrointestinal disease, alongside the chronic dental abscess.

To examine how oral microorganisms (microbiota) in Neandertals were linked to dietary composition, we compared the filtered shotgun data to a wide range of ancient calculus specimens from humans with varying diets including: ancient Later Stone Age (LSA) African gatherers; African Pastoralist Period individuals with high meat consumption[32]; European hunter-gatherers with a diet that included a wide range of protein sources; and early European farmers with diets largely based around carbohydrates and milk consumption (see SI for

archaeological descriptions of dietary information)[18]. We used UPGMA to cluster Bray Curtis distances obtained by comparing both complete and rarefied shotgun sequenced oral microbiota[18], which revealed four distinct groups: forager-gatherers with limited meat consumption (El Sidrón Neandertals, chimpanzee, and LSA African gatherers); hunter-gatherers (or pastoralists) with a frequent meat diet (Spy Neandertal, African pastoralists, and European hunter-gatherers); ancient agriculturalists (European farming individuals); and modern humans (Figure 2B). This analysis identifies a split between hunter-gatherers and agriculturalists, as previously observed[13], but also reveals two distinct hunter-gatherer groups, potentially differentiated by the quantity of meat consumed in their diet. Meat consumption appears to have impacted early hominin microbiota and health, similar to differences observed between carnivorous and herbivorous mammals[33,34]. This finding also indicates that dental calculus may be used to directly infer the dietary behavior of ancient hominins.

We then examined the Neandertal microbial diversity for signs of disease. Neandertal microbiota were more similar to the historic chimpanzee sample than the modern human, and contained less potentially pathogenic Gram-negative species, which are associated with secondary enamel colonization, increased plaque formation, and periodontal disease (18.9% Gram-negatives in Neandertals, compared to 77.6% in the modern human; Figure S21)[35]. Contrary to earlier reports[13], this does not appear to be due to taphonomic biases that remove Gram-negative species, as all types of microbial taxa were equally damaged and fragmented (Table 2)[18], suggesting this is a biological observation. The lower levels of these immunostimulatory Gram-negative taxa in Neandertals may be related to the reduced presence of Fusobacteria taxa (Figure S17), as this keystone

group facilitates the binding of Gram-negative microorganisms to the primary colonizers that bind tooth enamel (*e.g. Streptococcus*, *Actinomyces*, and *Methanobrevibacter* species)[36]. Importantly, the increased diversity of Gram-negative immunostimulatory taxa in modern humans are strongly linked to a wide-range of Western diseases[37].

Several oral pathogens could be identified within the shotgun data, although the short ancient sequences and diverse metagenomic background complicated identification. We established a number of exclusion criteria to verify the authenticity of short sequences of specific bacterial pathogens, including the assessment of ancient DNA damage, phylogenetic position, and bioinformatic comparisons to differentiate close relatives[18]. Pathogens that passed the exclusion criteria included the caries-associated taxa *Streptococcus mutans*, which was identified in all Neandertals (0.08% to 0.18%) (Tables S10-S11). All three members of the 'red complex' pathogens associated with modern periodontal disease[38] were identified in at least one Neandertal (*Porphyromonas gingivalis*: 0-0.52%; *Tannerella forsythia* 0.05-2.4%; and *Treponema denticola* 0-1.87%), although no single individual contained all three pathogens. These oral pathogens support the isolated evidence of Neandertals with dental caries, periodontal disease, and associated tooth-picking to relieve dental pain from specimens in Krapina (Croatia), Shanidar (Iraq) and Cova Forada (Spain)[39–41]. *Bordetella parapertusssis* and *Pasteurella multocida*, which together cause exacerbated whooping cough infections, were also detected in El Sidrón 1 (Table S9); however, only a limited number of *B. parapertussis*-specific reads were identified (i.e. only 212 reads, mostly in a region containing hypothetical proteins, mapped more efficiently to *B. parapertussis* than *Bordetella petrii*, an environmental

isolate). Similarly, a variety of other pathogens (*Neisseria gonorrhoeae*, *Streptococcus pyogenes*, and *Corynebacterium diptheriae*) were identified but could not be unambiguously distinguished from closely related commensal oral taxa (Figure S23). These difficulties highlight the need for rigorous criteria when identifying pathogenic strains from ancient metagenomic data[18].

We also examined the commensal microorganisms in Neandertals in greater detail. Within the deeply sequenced El Sidrón 1 oral microbiota, we were able to recover eight draft ancient microbial genomes with >1x depth of coverage, corresponding to the most prevalent microbial taxa (Table 2). Of particular interest was an archaeal species that dominated the oral metagenome of El Sidrón 1 (14.7%; Figure S17), but was present in lower proportions in the other Neandertal specimens (1.4% and 1.2% in El Sidrón 2, and Spy II, respectively). The large differences in G/C content between bacteria and archaea facilitated efficient read mapping of the archaeal sequences (Table 2), which grouped closest to the modern human-associated *Methanobrevibacter oralis* JMR01 strain. We were able to produce the oldest draft microbial genome to date at ~48 Kyr[42] and the first draft ancient archaeal genome, *Methanobrevibacter oralis* subsp. *neandertalensis* (44.7% of 2.1 Mbp covered at a 10.3x depth of coverage; Table 2 and Figure 3). The DNA damage profile (C-to-T at 5'ends (33%); G-to-A at 3' ends (36%); probability of terminating an overhang (lambda; 0.38); cytosine deamination probability in a double-strand context (d_D; 0.05); cytosine deamination probability in a single strand context (d_S; 1); Table 2) and fragment length distribution (average 58.67 bp) of *M. oralis* subsp. *neandertalensis* were consistent with ancient DNA damage.

Phylogenetic analysis of seven modern *Methanobrevibacter* genomes revealed that *M. oralis* subsp. *neandertalensis* is most closely related to the modern human *M. oralis* strain (JMR01), and together form a sister group to *Methanobrevibacter smithii* strains (commensal archaea found in the modern gut) (Figure 3B). Dating using a strict molecular clock[18] places the divergence between *M. oralis* subsp. *neandertalensis* and modern human *M. oralis* strains between 112-143 Kyr (95% highest posterior density interval; mean date of 126 Kyr) (Figure 3B), long after the genomic divergence of Neandertals and modern humans (450-750 Kyr)[43]. This suggests that microbial species were transferred during subsequent interactions, potentially during early interbreeding events between modern humans and Neanderthals, likely in the Near East[44].

Within the *M. oralis* subsp. *neandertalensis* genome, 1,929 coding sequences matched those present in the modern human *M. oralis*, while 136 in the latter appeared to be absent in *M. oralis* subsp. *neandertalensis* (6.5% of all coding sequences in *M. oralis*) (Table S16). The absent loci included analogs of genes encoding antiseptic resistance (*qacE*) and those required for regulation of maltose metabolism (*sfsA*), potentially reflecting the impacts of modern oral hygiene and relative dearth of carbohydrates in ancient Neandertal diets, respectively. As expected, bacterial immunity loci were also variable, with regions encoding for CRISPR Cas2 and Cas6 in modern *M. oralis* missing, while the Cas1 CRISPR system could only be partially assembled in *M. oralis* subsp. *neandertalensis* (Table S16). The ratio of non-synonymous to synonymous mutations per site (d_N/d_S) between translatable *M. oralis* subsp. *neandertalensis* protein coding sequences and modern human *M. oralis* suggested that 58% were under strong purifying selection ($d_N/d_S<0.1$) (Table 3)[18]. Only 4% appeared to be

under putative positive selection ($d_N/d_S>1$), and included the conjugal transfer gene, *traB*, which aids in the uptake of foreign DNA (i.e. plasmid transfer)[45] and *mutT*, which is involved in DNA mismatch repair[46]. Hence, much of the *M. oralis* genome appears to be subject to purifying selection, perhaps in part reflecting the relatively stable environmental conditions in the hominin mouth over time[47]. In addition, these data suggest that any adaptive evolution in the modern human *M. oralis* genome may have occurred primarily through the uptake of new DNA sequences, rather than the fixation of beneficial point mutations.

Preserved dental calculus represents an important new resource of behavioral, dietary, and health information for ancient hominin specimens, which is critical for understanding how these factors continue to impact modern humans today. The use of dental calculus provides detailed insight into the lifeways of Neandertals, linking and supplementing zooarchaeological, paleontological, and morphological records in a single approach, especially in specimens with limited contextual information (*i.e.* El Sidrón). In addition, dental calculus provides a unique long-term study system for microbial evolution in real time, allowing researchers to examine how hundreds of different microbial species have evolved, and were spread amongst different hominins.

References

1. Green, R. E. *et al.* A Draft Sequence of the Neandertal Genome. *Science* **328,** 710–722 (2010).
2. Fiorenza, L. *et al.* Molar Macrowear Reveals Neanderthal Eco-Geographic Dietary Variation. *PLOS ONE* **6,** e14769 (2011).
3. Zaatari, S. E., Grine, F. E., Ungar, P. S. & Hublin, J.-J. Neandertal versus Modern Human Dietary Responses to Climatic Fluctuations. *PLOS ONE* **11,** e0153277 (2016).
4. Rosas, A. *et al.* Paleobiology and comparative morphology of a late Neandertal sample from El Sidrón, Asturias, Spain. *PNAS* **103,** 19266–19271 (2006).
5. Meyer, M. *et al.* A High-Coverage Genome Sequence from an Archaic Denisovan Individual. *Science* **338,** 222–226 (2012).
6. Higham, T. *et al.* The timing and spatiotemporal patterning of Neanderthal disappearance. *Nature* **512,** 306–309 (2014).
7. Richards, M. P. *et al.* Neanderthal diet at Vindija and Neanderthal predation: The evidence from stable isotopes. *PNAS* **97,** 7663–7666 (2000).
8. Richards, M. P. & Trinkaus, E. Isotopic evidence for the diets of European Neanderthals and early modern humans. *PNAS* **106,** 16034–16039 (2009).
9. Bocherens, H., Drucker, D. G., Billiou, D., Patou-Mathis, M. & Vandermeersch, B. Isotopic evidence for diet and subsistence pattern of the Saint-Césaire I Neanderthal: review and use of a multi-source mixing model. *Journal of Human Evolution* **49,** 71–87 (2005).
10. Henry, A. G., Brooks, A. S. & Piperno, D. R. Microfossils in calculus demonstrate consumption of plants and cooked foods in Neanderthal diets

(Shanidar III, Iraq; Spy I and II, Belgium). *PNAS* (2010). doi:10.1073/pnas.1016868108

11. Hardy, K. *et al.* Neanderthal medics? Evidence for food, cooking, and medicinal plants entrapped in dental calculus. *Naturwissenschaften* **99,** 617–626 (2012).
12. Fu, Q. *et al.* An early modern human from Romania with a recent Neanderthal ancestor. *Nature* **524,** 216–219 (2015).
13. Adler, C. J. *et al.* Sequencing ancient calcified dental plaque shows changes in oral microbiota with dietary shifts of the Neolithic and Industrial revolutions. *Nat Genet* **45,** 450–455 (2013).
14. Warinner, C. *et al.* Pathogens and host immunity in the ancient human oral cavity. *Nat Genet* **46,** 336–344 (2014).
15. Ziesemer, K. A. *et al.* Intrinsic challenges in ancient microbiome reconstruction using 16S rRNA gene amplification. *Scientific Reports* **5,** 16498 (2015).
16. Caporaso, J. G. *et al.* Ultra-high-throughput microbial community analysis on the Illumina HiSeq and MiSeq platforms. *ISME J* **6,** 1621–1624 (2012).
17. Huson, D. H., Mitra, S., Ruscheweyh, H.-J., Weber, N. & Schuster, S. C. Integrative analysis of environmental sequences using MEGAN4. *Genome Res.* **21**, 1552–1560 (2011).
18. Materials and methods are available as supplementary materials on *Nature* Online.
19. Salter, S. J. *et al.* Reagent and laboratory contamination can critically impact sequence-based microbiome analyses. *BMC Biology* **12,** (2014).

20. Weiss, S. *et al.* Tracking down the sources of experimental contamination in microbiome studies. *Genome Biol* **15,** (2014).
21. Meyer, F. *et al.* The metagenomics RAST server – a public resource for the automatic phylogenetic and functional analysis of metagenomes. *BMC Bioinformatics* **9,** 386 (2008).
22. Herbig, A. *et al.* MALT: Fast alignment and analysis of metagenomic DNA sequence data applied to the Tyrolean Iceman. *bioRxiv* 50559 (2016). doi:10.1101/050559
23. Warinner, C. *et al.* Pathogens and host immunity in the ancient human oral cavity. *Nat Genet* **46,** 336–344 (2014).
24. Germonpré, M., Udrescu, M. & Fiers, E. The fossil mammals of Spy. *Anthropologica et Præhistorica* **123/2012,** 298–327 (2013).
25. Germonpré, M., Udrescu, M. & Fiers, E. Possible evidence of mammoth hunting at the Neanderthal site of Spy (Belgium). *Quaternary International* **337,** 28–42 (2014).
26. Nouvelles données paléogéographiques et chronologiques sur les Caprinae (Mammalia, Bovidae) du Pléistocène moyen et supérieur d'Europe :: EI-SEV publishing: :: Euskomedia. Available at: http://www.euskomedia.org/analitica/10263. (Accessed: 27th July 2016)
27. Patou-Mathis, M. Neanderthal subsistence behaviours in Europe. *International Journal of Osteoarchaeology* **10,** 379–395 (2000).
28. Naito, Y. I. *et al.* Ecological niche of Neanderthals from Spy Cave revealed by nitrogen isotopes of individual amino acids in collagen. *Journal of Human Evolution* **93,** 82–90 (2016).

29. Henry, A. G., Brooks, A. S. & Piperno, D. R. Plant foods and the dietary ecology of Neanderthals and early modern humans. *Journal of Human Evolution* **69,** 44–54 (2014).
30. Tsai, C. J. *et al.* Salicylate metabolism in Populus. *BMC Proceedings* **5,** 1–2 (2011).
31. Tumwine, J. K. *et al.* Enterocytozoon bieneusi among children with diarrhea attending Mulago Hospital in Uganda. *Am. J. Trop. Med. Hyg.* **67,** 299–303 (2002).
32. Leonard, W. R. & Crawford, M. H. *The Human Biology of Pastoral Populations*. (Cambridge University Press, 2002).
33. Muegge, B. D. *et al.* Diet Drives Convergence in Gut Microbiome Functions Across Mammalian Phylogeny and Within Humans. *Science* **332,** 970–974 (2011).
34. Moeller, A. H. *et al.* Rapid changes in the gut microbiome during human evolution. *PNAS* **111,** 16431–16435 (2014).
35. Marsh, D. P. D. & Bradshaw, D. J. Dental plaque as a biofilm. *Journal of Industrial Microbiology* **15,** 169–175 (1995).
36. Signat, B., Roques, C., Poulet, P. & Duffaut, D. Fusobacterium nucleatum in periodontal health and disease. *Curr Issues Mol Biol* **13,** 25–36 (2011).
37. Cho, I. & Blaser, M. J. The Human Microbiome: at the interface of health and disease. *Nat Rev Genet* **13,** 260–270 (2012).
38. Holt, S. C. & Ebersole, J. L. Porphyromonas gingivalis, Treponema denticola, and Tannerella forsythia: the ‘red complex’, a prototype polybacterial pathogenic consortium in periodontitis. *Periodontology 2000* **38,** 72–122 (2005).

39. Lozano, M., SubirÃ, M. E. lia, Aparicio, J., Lorenzo, C. & GÃ^{3}mez-Merino, G. Toothpicking and Periodontal Disease in a Neanderthal Specimen from Cova ForadÃ Site (Valencia, Spain). *PLoS ONE* **8,** e76852 (2013).

40. Trinkaus, E. Dental remains from the Shanidar adult Neanderthals. *Journal of Human Evolution* **7,** 369–382 (1978).

41. Topić, B., Raščić-Konjhodžić, H. & Cižek Sajko, M. Periodontal disease and dental caries from Krapina Neanderthal to contemporary man - skeletal studies. *Acta Med Acad* **41,** 119–130 (2012).

42. Wood, R. E. *et al.* A New Date for the Neanderthals from El Sidrón Cave (Asturias, Northern Spain). *Archaeometry* **55,** 148–158 (2013).

43. Stringer, C. The origin and evolution of Homo sapiens. *Phil. Trans. R. Soc. B* **371,** 20150237 (2016).

44. Sankararaman, S., Patterson, N., Li, H., Paabo, S. & Reich, D. The date of interbreeding between Neandertals and modern humans. *PLoS genetics* **8,** e1002947 (2012).

45. Kosońo, S., Kataoka, M., Seki, T. & Yoshida, T. The TraB protein, which mediates the intermycelial transfer of the Streptomyces plasmid pSN22, has functional NTP-binding motifs and is localized to the cytoplasmic membrane. *Mol. Microbiol.* **19,** 397–405 (1996).

46. Fowler, R. G. & Schaaper, R. M. The role of the mutT gene of Escherichia coli in maintaining replication fidelity. *FEMS Microbiol. Rev.* **21,** 43–54 (1997).

47. Dethlefsen, L., McFall-Ngai, M. & Relman, D. A. An ecological and evolutionary perspective on human–microbe mutualism and disease. *Nature* **449,** 811–818 (2007).

Figures and Legends
Figure 1

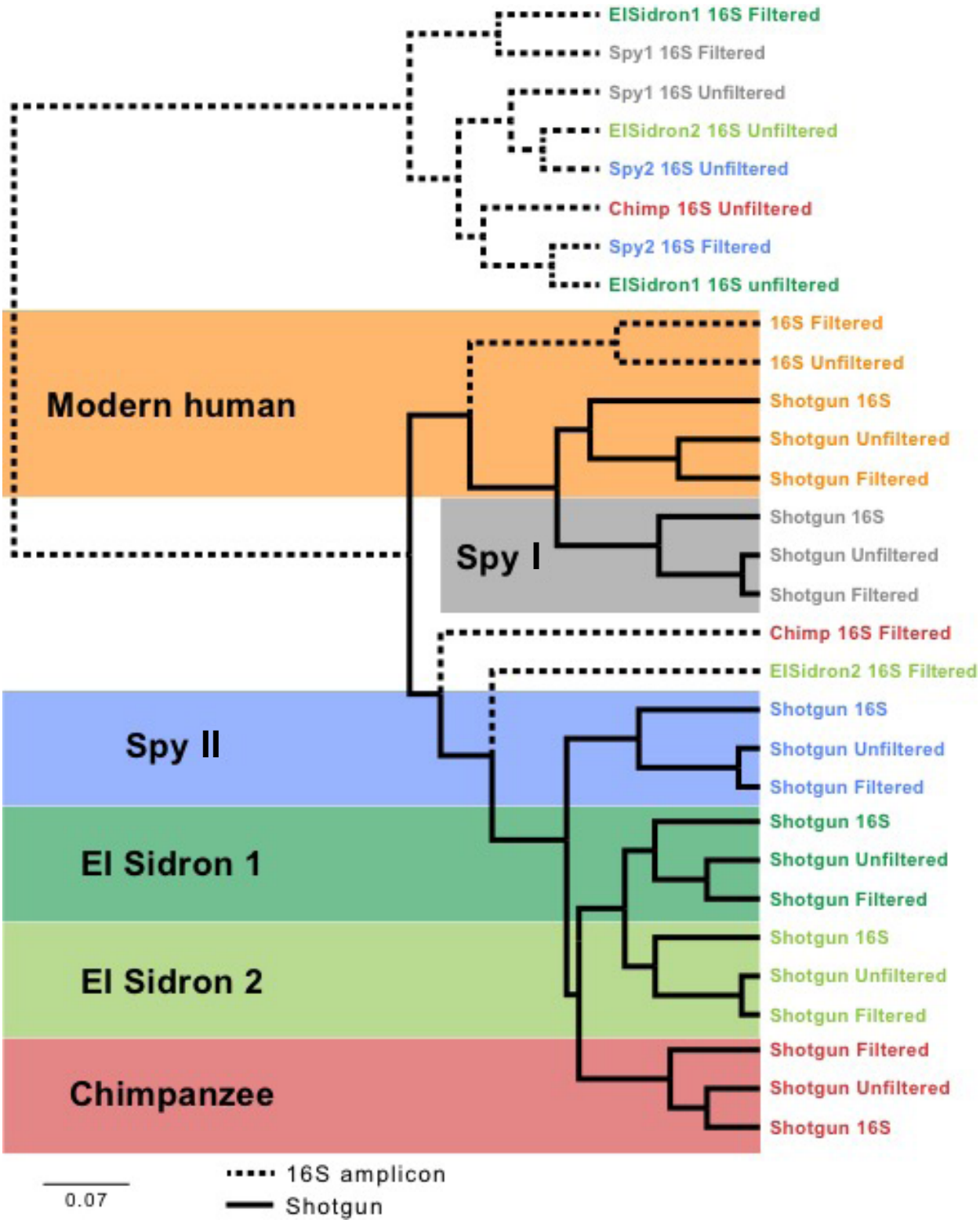
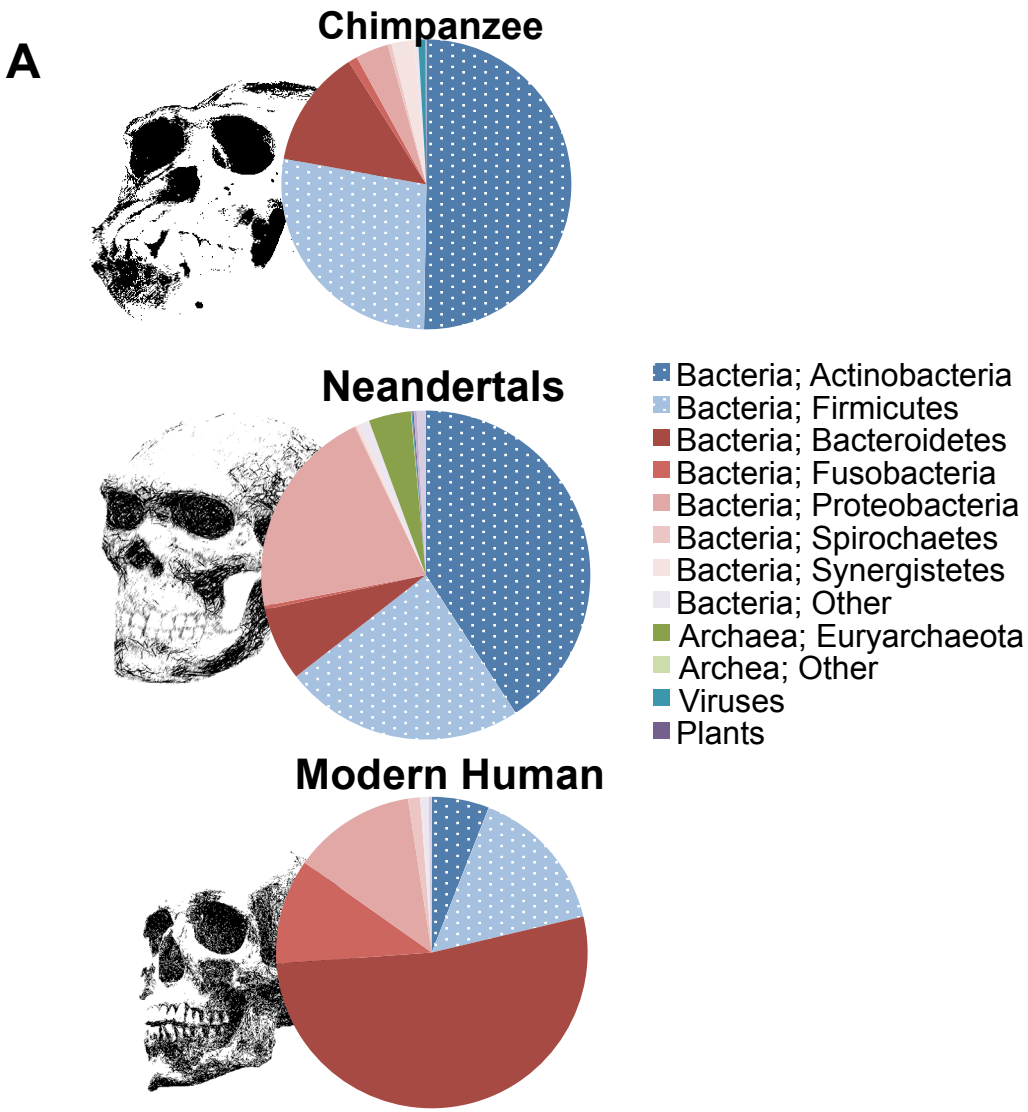


Figure 1: Comparison of 16S amplicon and shotgun data sets obtained from ancient, historic, and modern dental calculus samples. Filtered and unfiltered 16S rRNA amplicon and shotgun data sets, as well as 16S rRNA shotgun sequences identified using GraftM, were compared using UPGMA clustering of Bray Curtis distances from a chimpanzee (red), Neandertals (El Sidrón 1 (dark green), El Sidrón 2 (light green), Spy I (grey), Spy II (blue), and a modern human (orange).

Figure 2


A
Chimpanzee
Neandertals
Modern Human
Bacteria; Actinobacteria
Bacteria; Firmicutes
Bacteria; Bacteroidetes
Bacteria; Fusobacteria
Bacteria; Proteobacteria
Bacteria; Spirochaetes
Bacteria; Synergistetes
Bacteria; Other
Archaea; Euryarchaeota
Archea; Other
Viruses
Plants

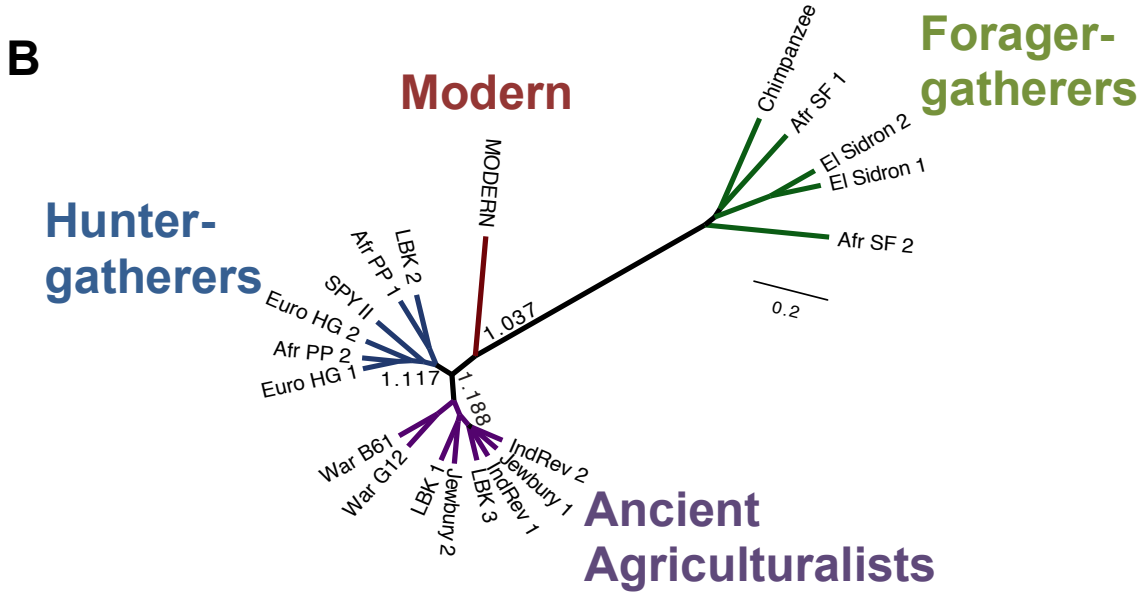

B
Modern
Forager-gatherers
Hunter-gatherers
Ancient Agriculturalists
MODERN
Chimpanzee
Afr SF 1
El Sidron 2
El Sidron 1
Afr SF 2
0.2
1.037
1.188
1.117
LBK 2
Afr PP 1
SPY II
Euro HG 2
Afr PP 2
Euro HG 1
War B61
War G12
LBK 1
Jewbury 2
LBK 3
IndRev 1
Jewbury 1
IndRev 2

Figure 2: Bacterial community composition at the phyla level of oral microbiota from chimpanzee, Neandertal, and modern human samples. Oral microbiota from shotgun data sets of a wild-caught chimpanzee (A), Neandertals (n=3; B), and a modern human (C) are presented at the phyla level. Phyla names were simplified for clarity, and unidentified reads were excluded. Gram-positive (blue) and Gram-negative (red) phyla are differentiated by color. (D) UPGMA clustering of Bray Curtis values obtained from 22 oral metagenomes is displayed. Definitions for abbreviations can be found in the SI.

Figure 3

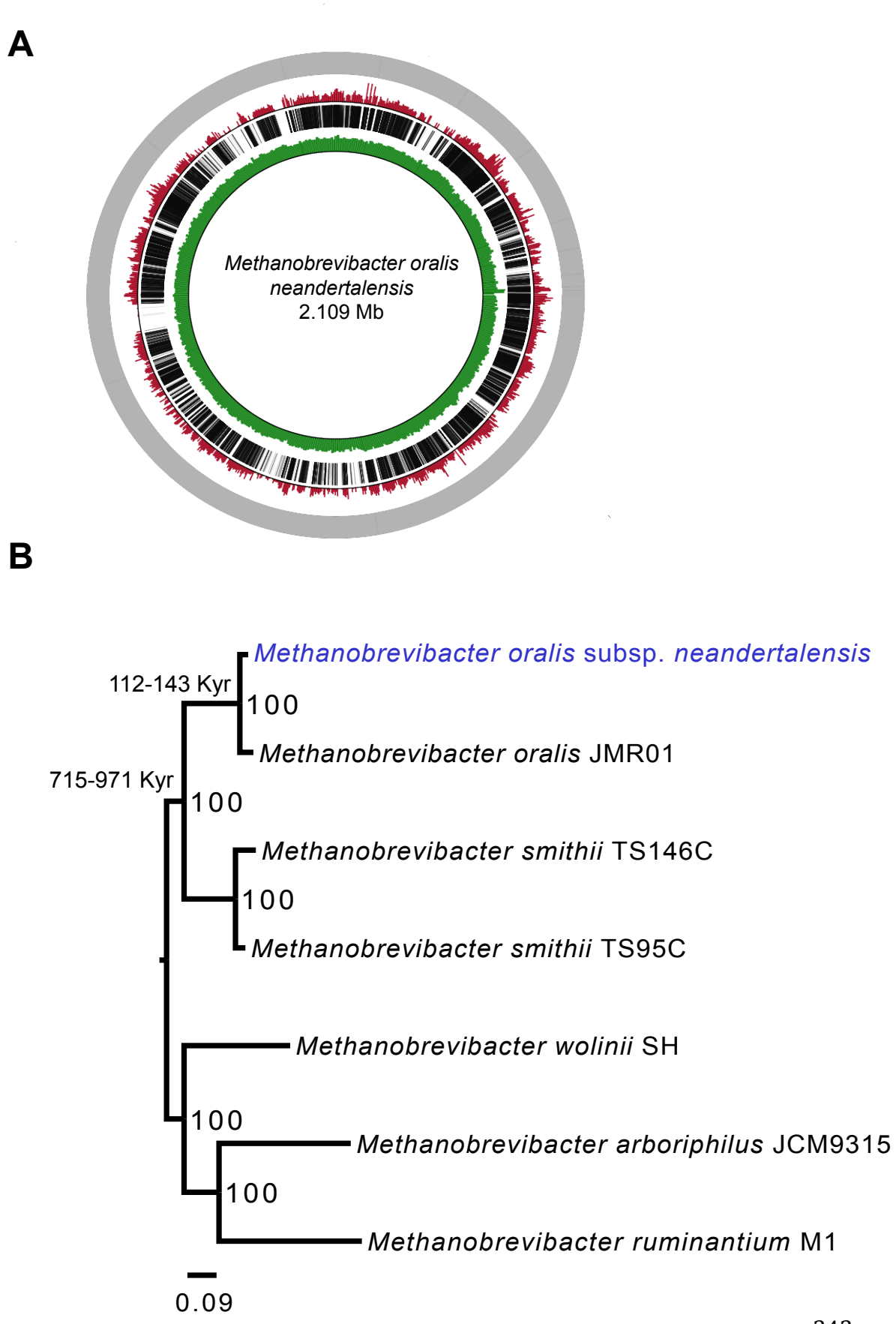

A
Methanobrevibacter oralis
neandertalensis
2.109 Mb
B
Methanobrevibacter oralis subsp. neandertalensis
112-143 Kyr
100
Methanobrevibacter oralis JMR01
715-971 Kyr
100
Methanobrevibacter smithii TS146C
100
Methanobrevibacter smithii TS95C
Methanobrevibacter wolinii SH
100
Methanobrevibacter arboriphilus JCM9315
100
Methanobrevibacter ruminantium M1
0.09

Figure 3: 48,000 year old archaeal draft genome and phylogeny of *Methanobrevibacter oralis neandertalensis*.

(A) Ancient sequences mapping to *Methanobrevibacter oralis* JMR01 are displayed in a Circos plot (black), alongside the depth of coverage obtained (red). The reference sequence is displayed (grey) with the GC content of the reference sequence calculated in 2500 bp bins (green). (B) A Methanobrevibacter phylogeny was constructed from whole genome alignments in RAxML with 100 bootstrap replicates, with the percent support shown in each node. The estimated dates placed onto this tree were calculated from a whole genome phylogeny using a Bayesian methodology (in BEAST) assuming a strict clock model.

Table 1

Scientific Name	Common Name of Likely Source	Hominid Pathogen (X) or Medicinal Uses (+)	El Sidron 1	El Sidron 2	Spy II	Chimpanzee	Modern Human	Lab Control (EBC)	Spy I*
Zymoseptoria tritici	Plant (wheat) pathogen		4.13%	0	0	0	0	*0*	*2.87%*
Phaeosphaeria nodorum	Plant (wheat) pathogen		12.22%	0	0	3.98%	0	*0*	*0.45%*
Penicillium rubens	Food fungus	+	3.97%	0	0	0	0	*0*	*1.35%*
Myceliophthora thermophila	Cellulose Fungus		0	0	0.56%	0	0	*0*	*0.13%*
Coprinopsis cinerea	Edible Mushrom (grey shag)		0	0	2.44%	0	0	*0*	*0*
Schizophyllum commune	Edible Mushroom (split gill)		3.65%	0	0	0	0	*0*	*0.10%*
Malassezia globosa	Human fungal commensal		3.65%	8.89%	0	0	19.92%	*0*	*5.49%*
Enterocytozoon bieneusi	Intracellular parasite (microsporidia)	X	8.10%	0	0	0	0	*0*	*0*
Ovis aries	Sheep (wild mouflon)		0.00%	0	62.03%	0	0	*0*	*1.17%*
Ceratotherium simum	White Rhino (woolly rhino)		0	0	34.40%	0	0	*0*	*0.11%*
*Ixodes scapularis**	Tick		0	0	0	0	2.15%	*0*	*0.15%*
Physcomitrella patens	Moss		2.06%	0	0	0	0	*0*	*0.09%*
Pinus koraiensis	Pine Tree		13.49%	19.60%	0	4.45%	0	*0*	*0.40%*
Populus trichocarpa	Poplar Tree	+	2.86%	0	0	0	0	*0*	*0.44%*
Total Eukaryotic Reads			630	551	532	427	3760	*5*	*25294*

Table 1: Dietary information preserved in calculus.

DNA sequences mapping to eukaryotic species are shown as a proportion of the total eukaryotic reads identified within each sample. Eukaryotic sequencing identified in the extraction blank controls and the Spy I Neandertal, which is heavily contaminated with modern DNA, are shown to the right. * denotes samples or taxa that are likely the results of contamination, as they do not represent biological processes[18].

Table 2

Reference Genome				Mapped Reads										
Reference Genome	Sequence Reference Number	Length (Mbps)	GC content (%)	Base Covered (Mbps)	% Sequence Covered	Unique Hits	Depth	Depth (Avg coverage)	Average Read Length	5p-C-T	3p-G-A	DeltaD	DeltaS	Lambda
Methanobrevibacter oralis JMR01	NZ_CBWS00000000	2.107	27.8	0.941	44.7	370115	10.3	15.16	58.67	0.33	0.36	0.05	1	0.38
Candidatus Saccharibacteria oral TM7	NZ_CP007496.1	0.705	44.5	0.131	18.6	108919	7.9	5.83	52.46	0.37	0.41	0.01	1	0.38
Campylobacter gracilis ATCC 33236	NZ_CP012196.1	2.282	46.6	1.199	52.5	94472	2.1	2.40	51.7	0.38	0.41	0.01	1	0.36
Propionibacterium propionicum F0230a	NZ_018142.1	3.449	66.1	2.083	60.4	130748	1.8	1.89	48.85	0.37	0.43	0	1	0.43
Fretibacterium fastidiosum	gi 296110870	2.728	55.5	1.466	53.7	121822	2.09	2.43	48	0.39	0.43	0	1	0.41
Eubacterium infirmum F0142	NZ_AGWI00000000	1.9	40.1	0.176	9.3	52170	1.4	10.73	51.53	0.33	0.38	0.02	1	0.41
Peptostreptococcus stomatis DSM 17678	GCF_000147675.1	1.988	36.7	1.222	61.5	94743	2.6	2.90	54.62	0.36	0.4	0.02	1	0.38
Eubacterium sphenum ATCC 49989	NZ_GG688422.1	1.084	40.6	0.261	24.1	23124	1.1	3.46	52.87	0.37	0.41	0.03	1	0.36

Table 2: Draft microbial genomes present in El Sidrón 1.

Eight draft microbial genomes from Gram-positive, Gram-negative, eubacterial, and archaea were obtained from the deeply sequenced El Sidrón 1 specimen by read mapping. The sequence coverage, GC content, sequencing depth, and damage profile (average fragment length and base pair modifications calculated from MapDamage2) are displayed for each genome.

Table 3

Gene Number	CDS	Genebank	dN/dS ratio	Gene Annotation	Coding Protein Function
Gene1211	1184	WP_042693702.1	0	NZ_HG796201.1	preprotein translocase subunit SecG
Gene291	283	WP_042691749.1	0	NZ_HG796199.1	SAM-dependent methyltransferase
Gene303	295	WP_042691777.1	0	NZ_HG796199.1	fibrillarin
Gene343	343	WP_042691868.1	0	NZ_HG796199.1	sugar fermentation stimulation protein SfsA
Gene394	394	WP_042691937.1	0	NZ_HG796199.1	30S ribosomal protein S2
Gene401	401	WP_042691950.1	0	NZ_HG796199.1	transcriptional regulator
Gene745	745	WP_042692741.1	0	NZ_HG796200.1	50S ribosomal protein L37
Gene757	757	WP_042693268.1	0	NZ_HG796200.1	acyltransferase
Gene766	766	WP_042692795.1	0	NZ_HG796200.1	DNA-directed RNA polymerase
Gene769	769	WP_042692805.1	0	NZ_HG796200.1	30S ribosomal protein S6
Gene772	772	WP_042692815.1	0	NZ_HG796200.1	50S ribosomal protein L24
Gene773	773	WP_042692817.1	0	NZ_HG796200.1	30S ribosomal protein
Gene810	810	WP_042692911.1	0	NZ_HG796200.1	transcriptional regulator
Gene836	836	WP_042692956.1	0	NZ_HG796200.1	endonuclease DDE
Gene880	880	WP_042693050.1	1.52	NZ_HG796200.1	uracil transporter
Gene724	724	WP_042692699.1	2.67	NZ_HG796200.1	acetylesterase
Gene269	269	WP_042691703.1	3.64	NZ_HG796199.1	conjugal transfer protein TraB
Gene1206	1206	WP_042693692.1	12	NZ_HG796201.1	DNA mismatch repair protein MutT

Table 3: Purifying and positive selection in *M. oralis neandertalensis.*

The ratio of non-synonymous to synonymous substitutions per site (d_N/d_S) was calculated for coding regions with sufficient coverage and that were conserved between *M. oralis* and *M. oralis* subsp. *neandertalensis*. Genes that have undergone strong purifying ($d_N/d_S < 0.1$) or positive ($d_N/d_S > 1$; grey) selection are displayed if the function of the gene was annotated. Hypothetical proteins and those not matching to the *M. oralis* genome during BLAST searches are not shown.

Acknowledgements

We thank Giorgio Manzi (University of Rome), the Odontological Collection of the Royal College of Surgeons, Royal Belgian Institute of Natural Sciences, Museo Nacional de Ciencias Naturales, and Adelaide Universities for access to dental calculus material. We thank Adam Croxford for DNA sequencing and Alan Walker, Johannes Krause, and Alexander Herbig for critical feedback.

Author Contributions

LSW, KD, and AC designed study; AGM, KWA, DC, VD, MF, MF, NG, WH, KH, KH, PH, JK, CLF, MR, AR, PS, AS, DU, and JW provided samples and interpretations of associated archaeological goods; LSW performed experiments; LSW, SD, EH, JS, BL, JB, LA, and AGF performed bioinformatics analysis and interpretation of the data; DHH developed bioinformatics tools; NG, JK, and GT analyzed medical relevance of data; LSW and AC wrote the paper; and all authors contributed to editing the manuscript.

Author Information

Raw and analyzed data sets and the scripts utilized for this analysis are available in the Online Ancient Genome Repository (OAGR) (currently available at: https://www.oagr.org.au/experiment/view/16/?token=9LCF0GKSL7DHO3FPR4YBSZCGYD0ASJ). The Australian Research Council supported this work, and the authors declare no competing financial interests. Requests for materials should be addressed to laura.weyrich@adelaide.edu.au.

Methods

Sampling handling and DNA extraction

Samples were stored and all molecular biology procedures prior to PCR amplification stages were carried out at the Australian Centre for Ancient DNA facility at The University of Adelaide. All experiments were performed within UV treated, still-air working hoods located in isolated, still-air working rooms that have been designed to allow highly technical ancient DNA research to be performed with ultra-low levels of background contamination (*i.e.* workflow is monitored, facilities are irradiated with ultraviolet light each night, and the general facility is under positive air pressure). To minimize environmental contamination, each dental calculus sample was UV treated for 15 minutes on each side, soaked in 2 mL of 5% bleach for 3 min, rinsed in 90% ethanol for 1 minute, and dried at room temperature for several minutes. Directly proceeding decontamination, DNA extraction was performed using an in-house silica-based method, as previously described[1] but with decreased buffer volumes (1.8 mL lysis buffer (1.6 mL EDTA; 200 uL SDS; 20 uL 20 mg/mL proteinase K) and 3 mL quanidine DNA binding buffer).

DNA library preparation and sequencing

Once DNA was extracted, 16S ribosomal RNA amplicon libraries of the V4 region were constructed by PCR amplification[2]. Each sample was amplified in triplicate, and samples were pooled, Ampure cleaned, and quantified using a TapeStation and quantitative PCR (KAPA Illumina quantification kit), prior to sequencing with an Illumina MiSeq 300 cycle kit (~40 samples/run). Frequent and repetitive extraction blank controls (EBCs) are used throughout all experimental procedures, *i.e.* extraction, amplification, and library preparation. Several key

samples were selected for shotgun metagenomic sequencing. Shotgun metagenomic libraries were constructed as previously described[3], with 5 bp forward and reverse barcodes. Metagenomic libraries were Ampure cleaned, quantified using a TapeStation and quantitative PCR (KAPA Illumina quantification kit), and pooled at equimolar concentrations prior to sequencing.

16S rRNA amplicon library analysis

To process the 16S amplicon data, sequences were de-multiplexed using the CASAVA pipeline and joined into amplicons using fastq_joiner (ea-utils)[4]. Quality filtering and trimming was completed using Cutadapt, and sequences were then imported into QIIME 1.6.0 for analysis[5]. In QIIME, OTUs were clustered in UCLUST at 97%, and representative sequences were taxonomically identified using the Greengenes (gg_12_10) database[6]. After OTU selection, strict filtering was applied to all samples (SI Section II). Diversity was analysed in QIIME, and phylogenetic analysis was visualized in FigTree (http://tree.bio.ed.ac.uk/software/figtree). Statistical analysis were performed by anosim in QIIME, and the calculation of the Jaccard or Bray Curtis indices, hierarchal clustering, and heatmap construction was completed in R using the vegan and gplots packages (http://cran.r-project.org).

Shotgun DNA Library Analysis

To process shotgun metagenomic data, reads were merged with a 5 bp overlap using bbmerge, and reads matching the forward and reserve barcodes with one mismatch were retained using AdapterRemoval[7]. Taxonomic identifications were made using protein alignments in MetaPhlAn [8], MG-RAST[9], DIAMOND[10], and the new Metagenome ALignment Tool with the BLASTX-like approach (MALTX) developed in the Huson laboratory at the University of Tubingen[11].

Taxonomic assignments were then filtered using default LCA parameters in MEGAN5[12], and data was exported at specific taxonomic classification levels (*i.e.* phyla, species, etc.) for downstream analysis. Reference genomes were excluded if they were known to have human DNA contamination[13]. Statistical analyses were done using a Mann-Whitney U test (comparisons of phyla in one taxa compared to other samples), a heteroscedastic t-test (direct comparisons between specific taxa in two samples/groups), or LefSe (identifying taxa that distinguish one group from another)[8]. Genomes were assembled by mapping to a reference genome using specific ancient DNA parameters in bwa[14], and authenticated using MapDamage 2.0[15]. Phylogenetic analyses were completed by mapping reads to reference genomes, aligning genomic sequences using progressiveMauve[16], and inferring trees in RAxML v8.1.21[17] using the GTRGAMMA model. A Bayesian approach was utilized to estimate dates of divergence between strains and clades. Detailed descriptions of each procedures are available in the Supplemental Information[18].

References

1. Brotherton, P. *et al.* Neolithic mitochondrial haplogroup H genomes and the genetic origins of Europeans. *Nat Commun* **4,** 1764 (2013).
2. Caporaso, J. G. *et al.* Ultra-high-throughput microbial community analysis on the Illumina HiSeq and MiSeq platforms. *ISME J* **6,** 1621–1624 (2012).
3. Meyer, M. & Kircher, M. Illumina sequencing library preparation for highly multiplexed target capture and sequencing. *Cold Spring Harb Protoc* **2010,** pdb.prot5448 (2010).
4. Aronesty, E. *Command-line tools for processing biological sequencing data*. Available at: https://github.com/ExpressionAnalysis/ea-utils (2011).
5. Caporaso, J. G. *et al.* QIIME allows analysis of high-throughput community sequencing data. *Nature Methods* **7,** 335–336 (2010).
6. DeSantis, T. Z. *et al.* Greengenes, a Chimera-Checked 16S rRNA Gene Database and Workbench Compatible with ARB. *Appl Environ Microbiol* **72,** 5069–5072 (2006).
7. Lindgreen, S. AdapterRemoval: easy cleaning of next-generation sequencing reads. *BMC Res Notes* **5,** 1–7 (2012).
8. Segata, N. *et al.* Metagenomic biomarker discovery and explanation. *Genome Biol.* **12,** R60 (2011).
9. Aziz RK *et al.* The RAST Server: Rapid Annotations using Subsystems Technology. *BMC Genomics* **9,** (2008).
10. Buchfink, B., Xie, C. & Huson, D. H. Fast and sensitive protein alignment using DIAMOND. *Nat Meth* **12,** 59–60 (2015).

11. Herbig, A. *et al.* MALT: Fast alignment and analysis of metagenomic DNA sequence data applied to the Tyrolean Iceman. *bioRxiv* 50559 (2016). doi:10.1101/050559
12. Huson, D. H., Mitra, S., Ruscheweyh, H.-J., Weber, N. & Schuster, S. C. Integrative analysis of environmental sequences using MEGAN4. *Genome Res* **21**, 1552–1560 (2011).
13. Longo, M. S., O'Neill, M. J. & O'Neill, R. J. Abundant Human DNA Contamination Identified in Non-Primate Genome Databases. *PLOS ONE* **6,** e16410 (2011).
14. Schubert, M. *et al.* Improving ancient DNA read mapping against modern reference genomes. *BMC Genomics* **13,** 178 (2012).
15. Ginolhac, A., Rasmussen, M., Gilbert, M. T., Willerslev, E. & Orlando, L. mapDamage: testing for damage patterns in ancient DNA sequences. *Bioinformatics* **27,** 2153–5 (2011).
16. Darling, A. E., Mau, B. & Perna, N. T. progressiveMauve: Multiple Genome Alignment with Gene Gain, Loss and Rearrangement. *PLoS ONE* **5,** e11147 (2010).
17. Stamatakis, A. RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics* **22,** 2688–90 (2006).
18. Materials and methods are available as supplementary materials on *Nature* Online.

Appendix IV

Ethical approval for analysis of human dental calculus (H-2012-108)

THE UNIVERSITY of ADELAIDE

RESEARCH BRANCH
OFFICE OF RESEARCH ETHICS, COMPLIANCE AND INTEGRITY

THE UNIVERSITY OF ADELAIDE
LEVEL 4, RUNDLE MALL PLAZA
50 RUNDLE MALL
ADELAIDE SA 5000 AUSTRALIA

TELEPHONE +61 8 8313 6028
FACSIMILE +61 8 8313 3700
EMAIL hrec@adelaide.edu.au
CRICOS Provider Number 00123M

23 February 2016

Dr L Weyrich
Dental School

Dear Dr Weyrich

ETHICS APPROVAL No: **H-2012-108**

PROJECT TITLE: **Using dental calculus to reveal the origins and evolution of human disease**

Thank you for the application and project status report dated 3.9.15 and the amended documents in response dated 11.2.2016, requesting extension of approval for the above project. The extension of ethics approval for an additional three years has been approved by the Low Risk Human Research Ethics Review Group (Faculty of Health Sciences).

The ethics expiry date for this project is: **31 August 2018**

Ethics approval is granted for three years and is subject to satisfactory annual reporting. The form titled *Project Status Report* is to be used when reporting annual progress and project completion and can be downloaded at http://www.adelaide.edu.au/ethics/human/guidelines/reporting. Prior to expiry, ethics approval may be extended for a further period.

Participants in the study are to be given a copy of the Information Sheet and the signed Consent Form to retain. It is also a condition of approval that you **immediately report** anything which might warrant review of ethical approval including:

- serious or unexpected adverse effects on participants,
- previously unforeseen events which might affect continued ethical acceptability of the project,
- proposed changes to the protocol; and
- the project is discontinued before the expected date of completion.

Yours sincerely

Sabine Schreiber
Secretary, Human Research Ethics Committee
Office of Research Ethics, Compliance and Integrity