Advances in the study of ancient biomolecules in archaeological dental calculus

Dissertation

In Partial Fulfilment of the Requirements for the Degree of
„Doctor of Philosophy“ (PhD)

Submitted to the Council of the Faculty
of Biological Sciences
of the Friedrich Schiller University Jena

by B.Sc. M.Sc. Zandra Fagernäs
born on 24 September 1992 in Kronoby

Gutachter:
1. Prof. Dr. Christina Warinner (Max-Planck-Institut für Evolutionäre Anthropologie, Leipzig)
2. Prof. Dr. Johannes Krause (Max-Planck-Institut für Evolutionäre Anthropologie, Leipzig)
3. Dr. Nicolás Rascovan (Institut Pasteur, Paris)

Beginn der Promotion: 18.10.2021
Dissertation eingereicht am: 23.09.2021
Tag der öffentlichen Verteidigung: 15.03.2022

TABLE OF CONTENTS

Introduction 4
The human microbiome 6
Ancient microbes and microbiomes 6
Why study ancient microbiomes? 9
Challenges with studying ancient microbiomes 10
Preservation assessment and authentication 13
Method development 14
Checking our assumptions 15
Preserving archaeological materials 16
Interdisciplinary studies and their strengths 17
Aims and objectives 19
Overview of manuscripts 20
Manuscript A 24
Manuscript B 56
Manuscript C 68
Manuscript D 97
Discussion 127
The stability of the dental calculus microbiome 127
Seeing the trees in the forest 130
Limitations of archaeological dental calculus 133
The importance of study design 136
Future directions 138
Summary 140
Zusammenfassung 142
References 144
Eigenständigkeitserklärung 155
Acknowledgements 156
Appendix 157
Appendix A: Summary of author contributions 157
Appendix B. Supplementary information for Manuscript B 159
Appendix C: Supplementary information for Manuscript C 163
Appendix D: Supplementary information for Manuscript D 165

INTRODUCTION

Studies of the human past are commonly thought of as being about skeletons and grave goods, pots and temples, scripts and statues. A large part of the available information about the human past is, however, hidden from the naked eye - in the form of microbes. Microbes can be found everywhere, and are a significant component of our life in many forms; as commensals, symbionts, pathogens, or even microbes whose activity we harness for our own benefit, such as food fermentation. These microbes, when preserved in the archaeological record, can tell us about several different aspects of our past, from health and disease, to diet and occupation.

Under suitable conditions, microbes can be preserved over thousands of years in archaeological substrates. Although the microbes are no longer viable, and their cellular structures slowly degrade, their biomolecules can be studied with methods from archaeological science fields, such as archaeogenetics and palaeoproteomics. The limits of ancient biomolecule preservation are still being explored and continuously extended, but for example, an authentic oral microbiome metagenome has been recovered from dental calculus from 100,000 years ago (Fellows Yates *et al.*, 2021). It can be expected that microbial proteins are preserved over even longer time periods, as protein preservation has been found to exceed the limits of DNA preservation (Chen *et al.*, 2019; Welker *et al.*, 2020). Over time, biomolecular evidence is, however, degraded and lost, as environmental factors such as heat and humidity lead to chemical alterations and breakdown of molecules, and environmental microbes colonize the archaeological material (Dabney, Meyer and Pääbo, 2013; Hendy *et al.*, 2018). What is recovered from archaeological substrates is therefore a mix of ancient and modern microbes, through biomolecules that are chemically different from the original assembly and have been fragmented into short pieces. From this, details about past cultures and human lives can be reconstructed, given that appropriate methods are used.

Ancient biomolecule research is a relatively young field, and ancient microbiome research is a recent addition to it. Ancient proteins were first studied through immunological assays (Boyd and Boyd, 1937), although the accuracy of this type of assay was later brought into question due to protein degradation (Downs and Lowenstein, 1995). The full potential of ancient protein studies started to be realized after the application of mass spectrometry (Ostrom *et al.*, 2000), and later tandem mass spectrometry (Cappellini *et al.*, 2012), allowing researchers to study the entire proteome of a substrate. The first ancient DNA study was published in 1984 (Higuchi *et al.*, 1984), but ancient metagenomes were inaccessible to researchers until about 20 years later, when molecular biology methods had developed enough to facilitate this type of research

(Tito *et al.*, 2008; Adler *et al.*, 2013). Other types of biomolecules are also studied in archaeological substrates, such as metabolites (e.g., Hardy *et al.*, 2012), RNA (e.g., Fordyce *et al.*, 2013), and lipids (e.g., Evershed, 1993), but in this thesis, the focus will be on DNA, and to a lesser extent, proteins.

Development of the field of ancient biomolecule research has contained challenges. Early studies struggled with separating authentic ancient biomolecules from contaminants, leading to some false interpretations. For example, DeSalle *et al.* (1992) claimed to have extracted 25-30 million years old DNA from a termite encased in amber. The validity of this find (and other similar ones) was, however, questioned quite soon after, as this is far beyond the theoretical preservation limits of DNA, and researchers started to realize how easily contamination could cause false positives when using PCR-technologies (polymerase chain reaction) (Austin, Smith and Thomas, 1997). Authentication of ancient DNA took a great leap forward after next-generation sequencing technologies were introduced, and when it was realized that the damage present in ancient DNA occurs preferentially in fragment ends and can be used to estimate whether the sequences are of ancient or modern origin (Dabney, Meyer and Pääbo, 2013; Jónsson *et al.*, 2013). Standards were also set for how to minimize contamination and assess authenticity, for both archaeogenetics and palaeoproteomics (Cooper and Poinar, 2000; Hendy *et al.*, 2018). Many valuable lessons have been learned through the growing pains of these new fields, leading to development of better methods and standards.

Currently, both archaeogenetics and palaeoproteomics are at a stage where ancient biomolecules are accessible, workflows are set to minimize contamination, and authentication criteria are established. However, there is still substantial room for development and optimization, and a need for broader application of these methods. New methods continue to expand the applications of ancient biomolecules to study the past in ways that were previously thought impossible, and existing methods are being fine-tuned to make sure that the finite archaeological substrates are used in the most productive way possible. In this thesis, I will demonstrate three different types of method development for biomolecular archaeological dental calculus research, which advance best practices in the field and bring forward new opportunities, beginning with a review of the biology of dental calculus and history and current state of the research field in Manuscript A. In Manuscript B, I present an optimized protocol allowing for extraction of both DNA and proteins from a single sample in order to limit destructive analyses; in Manuscript C I investigate the taxonomic variability across the dentition to improve study design and interpretation; and in Manuscript D I explore the application of metagenomic studies of dental calculus in tracing human migration patterns to expand the application of this type of studies.

The human microbiome

During the last few decades, archaeogenetic studies have illuminated human history and prehistory all over the world and through vast periods of time. However, humans also harbor a large variety of resident microbes, the estimated number of which ($3.8x10^{13}$), outnumbers the estimated number of human cells ($3.0x10^{13}$) in a body (Sender, Fuchs and Milo, 2016). The microbes that reside in and on our bodies are collectively called the human *microbiome*, a term first coined in 2001 and defined as "the ecological community of commensal, symbiotic, and pathogenic microorganisms that literally share our body space" (Lederberg and McCray, 2001). Together with our microbiome, we form a *holobiont*, functioning as an assemblage of the macrobe (the human) and our microbes (Youle *et al.*, 2013). The human microbiome includes major well-known and species-rich microbial assemblages, such as the oral and gut microbiomes, as well as smaller and less-known ones, such as the urogenital and eye microbiomes (NIH HMP Working Group *et al.*, 2009).

These microbes do much more than just live in and on our bodies. In recent years, many connections between microbial community composition and human health have been discovered, both positive and negative. For example, certain gut bacteria can synthesize vitamins *de novo,* and these vitamins can be used by the human host (LeBlanc *et al.*, 2013). On the other hand, the composition of the gut microbiome also has connections to the development of diabetes mellitus (Barlow, Yu and Mathur, 2015). Depending on the composition of the gut microbiome, it can both help us survive, and make us severely ill. But what does a 'healthy' microbiome look like? What did the ancestral human microbiome look like, before industrialization, globalization, and heavy sanitation and hygiene measures? How much have human lifestyle changes, such as the introduction of agriculture and the use of antibiotics, changed our microbiomes through time? To answer these kinds of questions, we need to look at what the human microbiome has looked like throughout our past, and turn to archaeological sciences.

Ancient microbes and microbiomes

Studies of ancient microbes have gained interest over the last years, as researchers have learned more about some of the deadly pathogens of the past, our commensal microbes, and the microbes whose properties we have harnessed to our favour, through studies of various archaeological substrates (Figure 1). Ancient bloodborne pathogens can mainly be recovered from parts of the skeleton that had blood flow while the individual was alive, such as the well-protected pulp cavity in a tooth, or from skeletal lesions caused by the pathogens (Spyrou *et al.*,

2019). Their spread and pathogenicity can be studied using biomolecular approaches, to date mainly archaeogenetics. Some examples of ancient pathogens that have been recovered from the and studied through their biomolecules are the bacteria *Yersinia pestis* (causing the plague) and *Mycobacterium tuberculosis* (tuberculosis), the viruses HBV (viral hepatitis) and VARV (smallpox), and the eukaryotic *Plasmodium* species causing malaria (Spyrou *et al.*, 2019). Although studying specific microbes can tell us about past diseases, microbiomes have the potential to give us knowledge about many other parts of the human past as well.

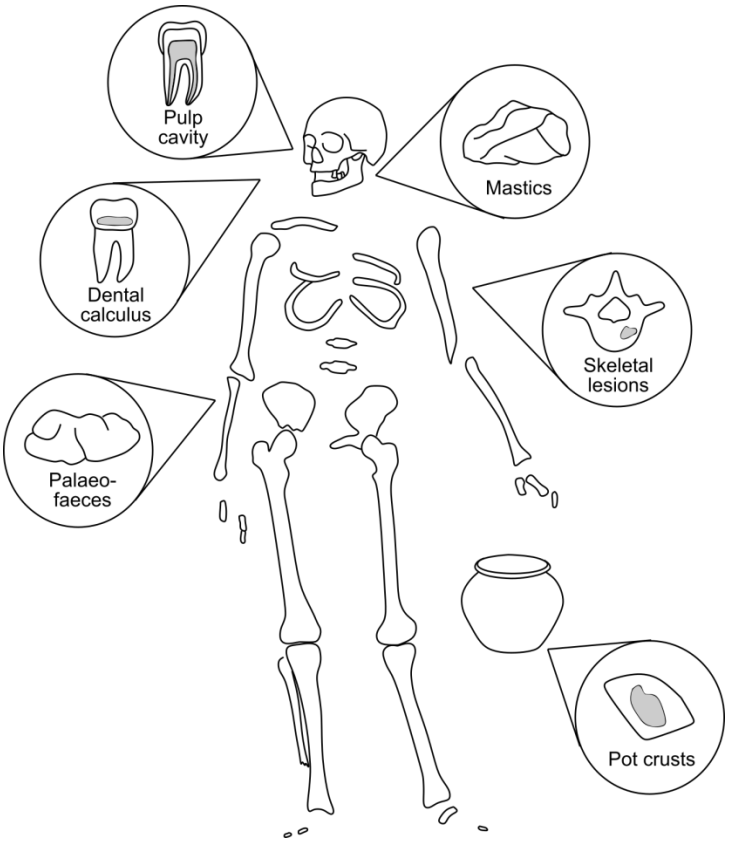


Figure 1. Examples of archaeological substrates which have the potential to contain ancient microbes in a burial site.

Both single-species palaeomicrobiological studies and ancient microbiome studies are informative of the human past, albeit different aspects of it. Studies of single species are mainly informative about disease, human interaction, and evolution of the species, whereas microbiomes can reflect factors such as diet, migration, evolution of the microbiome, and changes in lifestyle, all in one sample. Often, these two study types are combined, as interesting taxa are identified within a microbiome (e.g., Warinner *et al.*, 2014). What separates analyses of microbiomes from single-species analyses is mainly the lack of a set reference genome. Instead

of mapping DNA fragments from the study sample to one genome, the metagenomic DNA is mapped to hundreds of different genomes, in order to identify which taxa they originated from and reconstruct the microbial community. After this step, research is continued on either the microbial community as a whole, or on single species that are identified as informative for the study question. The species assignments are, however, complicated by factors inherent to microbes, such as incomplete species separation and horizontal gene transfer within microbiomes (Warinner *et al.*, 2017). Separation of contaminant microbes, stemming from sources such as the burial ground and storage location, from the original microbiome composition, may also be challenging. Further, taxa from the human microbiome can be closely related to environmental taxa, which may lead to DNA being incorrectly taxonomically assigned (Warinner *et al.*, 2017).

Most host-associated microbiomes will not be preserved in archaeological assemblages, as they decay and are overgrown by environmental microbes when the soft tissues in the human body decompose. In order for microbiomes to preserve in a state that resembles the original composition, protection from external influences, such as colonizing microbes and humidity, is essential. Only two of the human microbiomes are commonly preserved in the archaeological record: the gut and oral microbiomes. These are fortunately also among the most researched microbiomes in current day humans, allowing for direct comparisons between modern and ancient microbiomes. The gut microbiome is preserved in cases where faeces are protected from decomposition, either *in situ* in the body or after deposition, such as in very dry or cold environments (Warinner *et al.*, 2015), where microbes are preserved inside the fecal matter. The oral microbiome, on the other hand, is mainly preserved in the form of dental calculus, which calcifies during life, and this calcified matrix offers protection from decay and overgrowth (Warinner *et al.*, 2014).

Palaeofaeces are faeces from any species, which have undergone desiccation or partial mineralization. The term ‘coprolite’ is often used interchangeably with ‘palaeofaeces’, however, since fully fossilized faeces have traditionally been called coprolites, archaeological samples can more appropriately be termed palaeofaeces (Shillito *et al.*, 2020). Palaeofaeces have traditionally mainly been examined for evidence of diet and parasites, since remains of these are visible on a microscopic, or even macroscopic, scale. During recent years, the utility of palaeofaeces in researching the human gut microbiome has been explored, with several publications comparing ancient and modern gut microbiota (Shillito *et al.*, 2020). These studies have revealed differences between ancient and modern gut microbiomes in community composition (Tito *et al.*, 2012; Wibowo *et al.*, 2021), abundance of specific species (Tett *et al.*, 2019), and functional diversity (Jacobson *et al.*, 2020; Wibowo *et al.*, 2021). These differences, taken together with comparisons with modern populations with traditional lifestyles, may offer

clues as to the effects that different lifestyles and diets may have on the composition and function of the gut microbiome.

Dental calculus is formed during the lifetime of an individual, when the dental plaque biofilm periodically calcifies, encasing microbes and other particles in a mineral matrix. Dental calculus is ubiquitous in archaeological contexts across populations and time periods, and much more commonly found than palaeofaeces. It has been studied from individuals that are millions of years old, such as the Miocene ape Sivapithecus (12.5-8.5 million years before present (BP)) (Hershkovitz *et al.*, 1997). Early studies of dental calculus largely focused on microfossils, as they can be seen with a microscope and give clues about the diet of an individual. More recently, several studies have also looked at the genetics (Adler *et al.*, 2013; Warinner *et al.*, 2014; Weyrich *et al.*, 2017; Fotakis *et al.*, 2020), proteomics (Warinner *et al.*, 2014; Jersie-Christensen *et al.*, 2018; Fotakis *et al.*, 2020) and metabolomics (Velsko *et al.*, 2017) of the ancient oral microbiome, comparing it to modern oral microbiomes, and comparing ancient populations across time and space. The biology, archaeology, and current study of dental calculus are explored more thoroughly in Manuscript A.

Although palaeofaeces and dental calculus are the most commonly studied sources of ancient microbiomes, other substrates also exist, albeit much more scarcely. For example, mastics (Jensen *et al.*, 2019), historical medical specimens (Devault *et al.*, 2014) and frozen human remains (Lugli *et al.*, 2017) have yielded both single microbes and microbiomes. Although not part of the human microbiome, but still utilized by humans, microbes used for fermentation of food or beverages may be preserved in ancient pots (Aouizerat *et al.*, 2019), as food biomolecules have been recovered from pot sherds (Hendy *et al.*, 2018). Further discoveries of ancient microbes and microbiomes from archaeological substrates are likely to surface in the future, as archaeologists and archaeological scientists are increasingly aware of such potential findings.

Why study ancient microbiomes?

Studying health and disease of past human populations may give us important information with regard to modern day healthcare. By tracking the spread and dynamics of past pandemics, we can learn ways to predict the behaviour of future pandemics, allowing us to respond to them more efficiently. By studying what human microbiomes used to look like before the advent of antibiotics, we might find ways to combat antibiotic resistance. Many modern diseases are associated with a microbiome that is distinctly different from that found in states of health, a condition termed microbiome *dysbiosis* (Wilkins, Monga and Miller, 2019). Predicting what a

microbiome is ‘supposed to’ look like is non-trivial though, as our current microbiomes are already affected by factors like large-scale use of antibiotics (Gillings, Paulsen and Tetu, 2015) and an urbanized lifestyle (Segata, 2015). By studying the microbes of individuals that were alive thousands of years ago, we may gain a better understanding of what the ancestral human microbiome looked like. Although returning to an ancestral state of the microbiome may no longer be an option, given our changes in lifestyle and adaptations of the microbes over time (Carmody, Sarkar and Reese, 2021), elucidating which changes have taken place may offer connections to understanding the development of modern diseases.

Ancient human microbiomes contain information about several different aspects of the past and the present, in addition to mapping out differences between modern and ancient individuals. Throughout human history, humans have migrated over large areas and continuously changed their ways of living. Such changes, seen as e.g. alterations in living conditions, social interactions, diet, medication, and hygiene, are bound to have affected our accompanying microbiomes over time, as has been seen for modern microbiomes (Flandroy *et al.*, 2018). As an example, diet can be studied both through changes in the microbiome composition as a whole through altered nutrient sources (Harper and Armelagos, 2013), as well as adaptations of single species to altered nutrient sources (Fellows Yates *et al.*, 2021). Microbes may also directly reflect human migrations through patterns in their genetic diversity, and may be used as a proxy for human population genetic analyses when human genomes are not accessible (Mégraud, Lehours and Vale, 2016; Eisenhofer *et al.*, 2019).

Challenges with studying ancient microbiomes

Although the prospect of studying the past through ancient biomolecules is exciting and holds promise of discoveries about who we are and where we came from, it also comes with inherent challenges. For the general study of ancient DNA, there are two major issues that complicate genomic research as compared to modern DNA studies - *fragmentation* and *deamination*. Over time, chemical changes will cause nicks in the DNA strand backbones, eventually leading to breaks along the entire molecule length (Dabney, Meyer and Pääbo, 2013). These nicks occur when purines are hydrolysed and lost from the DNA strand, exposing the backbone of the DNA and leaving it susceptible to hydrolysis next to the depurination site, eventually leading to single-strand breaks (Briggs *et al.*, 2007). This fragmentation does not appear to be linearly related to sample age, but instead an initial fast fragmentation after the death of an organism is followed by a slow, stable fragmentation over a long period of time, leading to similar fragment length distributions in samples across a wide time range (Kistler *et al.*, 2017).

The second major modification of DNA over time is hydrolytic deamination of cytosine (C), into uracil (U) (Dabney, Meyer and Pääbo, 2013). During replication in laboratory processing, this uracil will be read as a thymine (T), leading to cytosine > thymine (C > T) transitions being detected in sequencing data during downstream data analyses. This means that during replication, these modified bases are paired with adenine (A) on the opposite strand (to complement the T), instead of the G that would complement the original C. Because of this, data from double-stranded DNA libraries show a guanine > adenine (G > A) transition on the complementary strand of the DNA molecule. These misincorporations are concentrated at the ends of the DNA strands (Briggs *et al.*, 2007), and are directly dependent on sample age and temperature at the burial site (Kistler *et al.*, 2017). Fragmentation and deamination make ancient DNA analysis challenging, but they also provide measures for estimating the authenticity of putative ancient DNA finds, which is essential in order to separate ancient DNA from modern contaminants (Jónsson *et al.*, 2013; Skoglund *et al.*, 2014).

For the analysis of ancient metagenomes, fragmentation and deamination provide some field-specific challenges. The short fragment length may increase the amount of spurious mismapping, and deamination will cause mismatches in alignment to reference genomes, which may erroneously be interpreted as nucleotide polymorphisms. The lack of a specific target organism makes the evaluation of these effects difficult. The effects of DNA degradation on the reconstruction of ancient microbiomes were a concern as the field started to grow. Using *in silico* datasets with simulated post-mortem degradation patterns, Velsko *et al.* (2018) tested the accuracy of several different taxonomic classifiers, and found that biases are mainly dependent on which classifier is used, rather than DNA degradation. This result was later confirmed by Eisenhofer and Weyrich (2019), who also noted that selection of a reference database can significantly alter the reconstructed microbiome. These systematic studies were essential, as they show that researchers can use standard taxonomic classifiers with ancient DNA, without DNA degradation biasing results. Although biases are introduced depending on the taxonomic classifier and database, the same is true for modern microbiome studies, and as long as the methodology is consistent across all samples of a study, the biases will be consistent and the samples can be compared to each other.

Studying ancient microbiomes comes with its own inherent challenges (Figure 2). The ancient microbes are no longer viable, and can therefore not be cultured and studied using traditional microbiological techniques. One of the major challenges is the presence of multiple species and strains, as microbiomes can consist of hundreds of taxa. Reconstructing the composition of the original microbiome is not trivial, especially in combination with damaged biomolecule fragments. Species that are missing from the database that is used cannot be identified in the ancient dataset, and biases towards model organisms in the amount of sequence data that is

present may artificially inflate their abundance in the dataset (Mann *et al.*, 2020). Microbial taxa or genomic variation that were present in ancient microbiomes, but have subsequently been lost from our microbiomes and therefore lack modern representatives, cannot be recovered through alignment to reference databases. Further, if so called 'dirty genomes' are present in the database, i.e. genomes that have adapters or other contaminant sequences included in them, false identifications can arise (Mann *et al.*, 2020; Steinegger and Salzberg, 2020).

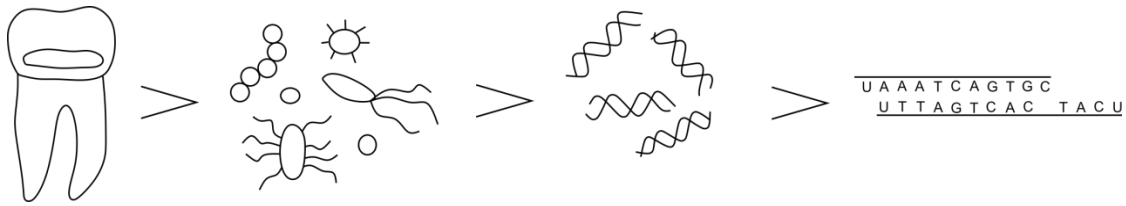


Figure 2. Challenges with studies of ancient microbiomes. What is recovered from archaeological dental calculus is a complex mixture of taxa, consisting of fragmented DNA molecules, with damage.

Finally, the composition of an ancient microbiome sample will change over time, both at the genomic and taxonomic level. The genomic composition of the sample is altered, as AT-rich fragments are preferentially lost due to the lower denaturation temperature of short AT-rich fragments (Mann *et al.*, 2018). This may bias the reconstructed taxonomic profile toward species with high GC-content genomes. Microbes will also colonize the archaeological substrates at several different stages post-mortem, with sources ranging from the burial soil to post-excavation handling, storage, and the laboratory environment. Through all these steps, the proportion of the original biomolecular composition that is present in the sample will progressively decrease, as the amount of contamination increases (Hendy *et al.*, 2018). However, through robust laboratory and bioinformatic method development, reconstructing and studying ancient microbiomes has been made feasible and reliable.

During the development of the field of ancient microbiome studies, different issues have been encountered and addressed, some of them not directly related to the microbiome. For example, during the early days of ancient dental calculus analysis, it was common to also report eukaryotic DNA identifications, as they may derive from the diet of an individual (Warinner *et al.*, 2014; Weyrich *et al.*, 2017). However, the validity of some of these finds have been questioned (Dickson, Oeggl and Stanton, 2017), due to the unlikeliness of the individuals eating some of the reported putative dietary items. A review of the study of diet through dental calculus metagenomic analyses pointed out several pitfalls, such as the low number of eukaryotic DNA fragments present generally hindering authentication, and databases biasing identifications

(Mann *et al.*, 2020). The authors of the review set guidelines for assessing validity of such dietary finds, in order to guide both researchers and reviewers in how eukaryotic DNA should be reported and which methods can be used to support the findings. Examples like this demonstrate the critical nature of method testing and development for raising standards in the field.

Although proteins generally preserve over longer time periods than DNA, they also degrade over time. Similar to DNA, bonds in the protein structure are damaged and proteins are fragmented into shorter peptides, making downstream protein identification more challenging (Hendy, 2021). Another major change in proteins over time is deamidation of the amino acids glutamine and asparagine into glutamic acid and aspartic acid, respectively. The deamidation of glutamine is often used as an indication of sample preservation, although it should be noted that deamidation is not directly correlated with age, but rather influenced by a range of environmental variables (van Doorn *et al.*, 2012; Schroeter and Cleland, 2016). Site-specific deamidation patterns have been harnessed to differentiate between authentic ancient proteins and modern contaminants, allowing e.g. separation between ancient dietary proteins and laboratory contaminants consisting of the same protein (Ramsøe *et al.*, 2020, 2021). Although other post-translational modifications of proteins related to degradation have been identified, these are not commonly used as signs of protein degradation (Hendy, 2021).

Preservation assessment and authentication

Over time, the composition of an ancient microbiome sample will change, as it gets colonized by environmental microbes from the burial soil and contaminated from different sources after excavation. These exogenous microbes may bias downstream analyses, unless care is taken to identify and exclude them. Currently, preservation is mainly assessed by comparing the ancient microbiome samples to modern microbiomes, through methods such as SourceTracker or principal component analysis (PCA) (Knights *et al.*, 2011; Warinner *et al.*, 2014; Weyrich *et al.*, 2017; Mann *et al.*, 2018; Fellows Yates *et al.*, 2021). Comparing ancient samples with modern ones can give researchers an idea of the proportion of the original microbiome that is still present. However, neither one of these methods gives researchers a threshold of when to discard samples due to poor preservation. SourceTracker estimates the proportion of taxa in a sample that stems from comparative samples, but only allows for preservation comparisons between the samples in a study. A PCA, on the other hand, shows how closely related the microbiomes in different samples are. Therefore, an analytic tool (called *cuperdec*) for assessing sample preservation was developed by Fellows Yates *et al.* (2021), which allows for estimation of the proportion of microbes that stems from a target source (e.g., the oral

microbiome), and can be directly compared to other sample types and modern microbiomes. This method thereby allows for calculations of a cut-off for well-preserved samples. A combination of these preservation assessment methods gives researchers both an estimation of the composition of the microbiome sample and contamination sources, as well as a set of criteria to determine which samples can be included in downstream analyses.

When studying a single species, criteria such as edit distance and damage patterns are often utilized to ensure that the organism is of ancient origin, and not a modern contaminant (Hübler *et al.*, 2019). The edit distance is a measure of how different the studied DNA sequences are from the reference genome, and can indicate if they indeed stem from the reference organism, or another closely related species. Most DNA sequences are assumed to have very few mismatches (some stemming from deamination), if aligned to the correct reference genome, whereas aligning to an incorrect species will lead to a higher number of mismatches per DNA fragment. Damage patterns show deamination of DNA fragments, as described earlier, and indicate if the studied DNA sequences are of ancient or modern origin. Damage patterns may be studied from selected representative species from a metagenomic sample, to ensure that the species assemblage is of ancient origin, as it is not feasible to authenticate every species in a metagenomic sample with current methods.

Method development

Method development is the backbone of scientific progress, whether stemming from laboratory or analytical advances. These are fundamental to progressing research strategies and expanding the boundaries of questions that can be answered in a field. Currently, ancient biomolecules are accessible to researchers, with established methods for recovery from many different types of archaeological substrates. It was, however, not until recently that this treasure trove of information was accessible to archaeological scientists. The field of archaeological sciences has grown tremendously during recent years, and methods have been developed to open up completely new horizons of research.

Method development takes place at multiple scales. There are of course the major developments, which revolutionize a field - in archaeogenetics they might be an extraction protocol allowing for recovery of very short DNA fragments (Dabney *et al.*, 2013), a new validation method (Jónsson *et al.*, 2013), or an entirely new source of ancient DNA being discovered (Slon *et al.*, 2017). However, smaller method developments take place in most studies, although they may not be the main topic of the study. For example, a method for isolating pigments from archaeological dental calculus, without damaging them, was developed

in a study looking at lapis lazuli fragments in a medieval German woman's dental calculus (Radini *et al.*, 2019). A novel method for estimating preservation of a metagenomic sample came from a study comparing ancient oral microbiomes of hominids (Fellows Yates *et al.*, 2021). At an even smaller scale, scripts that are published alongside manuscripts can be used as a base for other researchers developing appropriate analyses for their data. All these different types of developments are what together cause science to take giant leaps forward.

One example of astonishing development can be found in ancient human genetics. In 2006, approximately 100 mg of bone was used to generate ~255,000 unique DNA sequences from a Neanderthal (~38,000 BP) from Vindija cave in Croatia (Green *et al.*, 2006). The authors of this study state that "To achieve one-fold coverage of the Neanderthal genome (3 gigabases) without any further improvement in technology, about twenty grams of bone and 6,000 runs on the current version of the 454 sequencing platform would be necessary". However, there were considerable technological improvements in the following years, and in 2017, a 30-fold coverage genome was created from a Neanderthal (~50,000 BP) from the same archaeological site using just 41 mg of bone material (Prüfer *et al.*, 2017). The technological improvements were both specific for the field of archaeogenetics, such as the extraction protocol being optimized for retrieval of very short DNA fragments, and requiring less starting material, as well as for the field of genetics as a whole, such as improvements in sequencing technology (454 sequencing vs. Illumina).

Checking our assumptions

When a new field of research develops, things can move very fast. Most published papers will present 'the first' of something, and research groups are racing to publish the next novel findings. Method development and assumption checking are often forced to take a back seat in this flood of new results. Not until later, when the pace of novel discoveries has slowed down, are biases and assumptions checked - sometimes resulting in faulty methods being detected - a phenomenon that is also present within ancient DNA research (Jones and Bösl, 2021).

Examples of such delayed assumption checking can be found in metagenomic studies of archaeological dental calculus, a field which was founded in the 2010s and quickly led to several high-profile publications (Adler *et al.*, 2013; Warinner *et al.*, 2014; Weyrich *et al.*, 2017). Modern microbiome studies often use the 16S rRNA gene for species profiling, since the hypervariable regions it contains can be used to identify taxa in a microbial community without the need for extensive sequencing efforts. The same procedure was initially used to profile the microbial community in archaeological dental calculus, assuming that it would work in a similar

manner for archaeological samples as for modern samples (Adler *et al.*, 2013). Later it was, however, shown that microbial community composition profiling using 16S rRNA sequences is not suitable for ancient DNA, since it causes biases in the community composition (Ziesemer *et al.*, 2015). Ancient DNA sequences are generally shorter than the 16S rRNA hypervariable regions, and length polymorphisms in these regions can bias which species are detected, with species with long sequences dropping out from the dataset due to failed amplification. Any differences detected between ancient and modern samples using 16S rRNA may therefore be caused by a biased method, rather than biological differences. In a similar manner, dental plaque is commonly used as a comparative oral microbiome dataset, when evaluating the preservation and authenticity of reconstructed microbiomes from archaeological dental calculus. Velsko *et al.* (2019) have, however, shown that the community composition differs between dental plaque and dental calculus, reflecting the difference in maturation between the two substrates. Modern dental plaque is therefore not adequate as a comparative source for archaeological dental calculus, and instead modern dental calculus should be included in the study. These two examples illustrate that if care is not taken to check underlying assumptions of the study design, inappropriate conclusions may be drawn.

Preserving archaeological materials

Although it may seem like there is an endless amount of certain archaeological materials (such as dental calculus) available for studies, researchers still have an ethical responsibility to preserve the materials as far as possible (Austin *et al.*, 2019; Fox and Hawks, 2019). Once a sample has been used for biomolecular analysis, a part of it is often irreversibly destroyed, reducing the amount that is left for future analyses. If several types of destructive analyses are conducted on the same specimen, the majority of a sample may be used up. For example, a thorough study of an archaeological dental calculus sample may involve the extraction of DNA, proteins, microfossils, and metabolites. A small dental calculus deposit can therefore be used up completely, if every analysis is conducted independently. Seeing as methods are constantly evolving, and analyses that seemed unachievable ten years ago are routine today, we need to consider what methodological developments the future may bring. If a sample is completely used up today, and the attempted analysis fails, the new methods may never be used to test the original hypothesis.

There are several ways to minimize the amount of destructive analyses of valuable samples. First, combined extraction protocols make it possible to extract more than one type of information at a time. One such protocol is presented in Manuscript B, which combines extractions for DNA and proteins from ancient dental calculus., Other studies have combined

extraction methods for other types of studies, for example combining radiocarbon dating and DNA extraction from skeletal material (Korlević, Talamo and Meyer, 2018) or microfossil analysis and DNA extraction from dental calculus (Modi *et al.*, 2020). Second, samples with little or no value for morphological studies can preferentially be chosen for destructive analyses. For example, a Neanderthal-Denisova hybrid has been identified from just a small fragment of a long bone (Slon *et al.*, 2018). Screening of a large number of undiagnostic bone fragments from Denisova cave through collagen peptide mass fingerprinting allowed for this unique hominin to be discovered, without prior knowledge of the taxonomy, after which the sample could be studied further through archaeogenetic methods. Third, preservation can be predicted or samples pre-screened for molecular preservation prior to destructive analyses. As an example, a method has been published that allows for screening of archaeological specimens for presence of collagen or endogenous DNA before any extractions are performed, by using FTIR-ATR (Fourier-transform infrared spectroscopy in attenuated total reflectance), thereby lowering the risk of failed destructive analyses (Kontopoulos *et al.*, 2020).

One final, and often overlooked, way to avoid unnecessary destructive analyses is to define a study question prior to starting a study. Why do we want to perform analyses on this particular sample? What can we learn from it? And which method is the most appropriate one to answer our specific question? In this manner, unnecessary analyses that do not lead to results can be avoided. Clearly thought out aims and study designs also lower the risk of ending up with insufficient power to draw supported conclusions from a study, or introducing biases that will mislead conclusions. Research can be either hypothesis-driven or exploratory, both of which are commonly practiced ways of conducting research in archaeological sciences. With both approaches, there is a risk of not achieving any results, and samples going through destructive analyses without the data ever being published, but this risk may be higher for exploratory research. Therefore, having even a broad aim for the study prior to sampling may reduce the risk of performing unnecessary destructive analyses.

Interdisciplinary studies and their strengths

A key property of archaeological sciences is interdisciplinary work. This can lead to new methodologies being developed, as scholars from different disciplines share their expertise with each other. Ever since the start of the field, archaeology, linguistics, genetics, and proteomics have been combined to create multifaceted research on the past. As information about ancient times is often very fragmented - damaged DNA, extinct languages, crushed ceramics, etc. - a single discipline may only provide part of the history that is to be told. A combined approach of

several methods may lead to a more holistic view of the past, and avoid biases introduced by missing information from one discipline.

Disciplines may be combined at different parts of the research process. If different extraction methods are used to extract several types of biomolecules from a single sample, the interdisciplinarity starts early on in the process. For example, a dental calculus sample could be studied by four different disciplines - archaeogenetics, palaeoproteomics, metabolomics, and palaeobotany. Archaeogenetic analyses can provide information about the microbiome and the host, and through these about health, disease, lifestyle, and potentially migration (Warinner *et al.*, 2014; Eisenhofer *et al.*, 2019; Fotakis *et al.*, 2020; Fellows Yates *et al.*, 2021). Palaeoproteomic analyses can provide information about the diet of the host, as well as the function of the microbes and the host response to the microbes (Warinner *et al.*, 2014; Hendy *et al.*, 2018; Jersie-Christensen *et al.*, 2018). Metabolomic analyses may give clues about past medicines and food consumption (Velsko and Warinner, 2017; Velsko *et al.*, 2017). Finally, microfossils can provide information about the diet, lifestyle, and environment of past individuals (Hardy *et al.*, 2016; Geber *et al.*, 2019). Taken together, and combined with archaeological, anthropological, and linguistic information from the area, this interdisciplinary study would give a holistic view of the life of a past individual.

AIMS AND OBJECTIVES

The aim of this dissertation is to advance the study of ancient biomolecules, specifically DNA and proteins, from archaeological dental calculus at three different levels: i) Laboratory processing, ii) Study design and result interpretation, and iii) Expanding the use of dental calculus in studying the human past. Using state-of-the-art metagenomic and metaproteomic methods from the archaeological sciences, these aims are explored using carefully set up study designs, in order to ascertain that the results are reliable and reproducible.

In particular, the following questions are addressed:

- What is the history and current state of the research of dental calculus in archaeological sciences?
- Can proteins and DNA be simultaneously extracted from a single sample of dental calculus without introducing major biases?
- Does amount of processed dental calculus influence biomolecule yield and reconstructed metagenome and metaproteome?
- Does sampling location on the dental arcade significantly influence the reconstructed ancient dental calculus microbiome?
- Does DNA from other sources, aside from the microbiome, vary across the dental arcade in archaeological dental calculus?
- Does the dental calculus microbiome change with human migrations and cultural changes?
- Can human migrations be traced through phylogenetic analyses of microbes in archaeological dental calculus?

OVERVIEW OF MANUSCRIPTS

The dissertation includes the following four manuscripts:

Manuscript A

"Dental Calculus"

Zandra Fagernäs and Christina Warinner

In press for the *Handbook of Archaeological Science*, *2nd edition* (eds. Mark Pollard, Ruth-Ann Armitage and Cheryl Makarewicz)

Synopsis: Manuscript A is a review of the history, development, and current research on archaeological dental calculus. We begin by going through the current knowledge of how the dental plaque biofilm forms, and how it calcifies into dental calculus. Further, we discuss how the human immune system interacts with the dental calculus microbiome, and how dental plaque plays a part in the formation of periodontal diseases. We also go through the components of dental calculus, as this is essential knowledge for performing scientific analyses on the substrate. The history of archaeological dental calculus research is discussed in detail, starting with the first discoveries of information about the past being contained in this material, and touching upon how various different areas of archaeological science were applied to study dental calculus. We discuss the prevalence of dental calculus in the archaeological record - going back as far as the Miocene - and how biomolecules are preserved over time. Finally, we discuss the applications of archaeological dental calculus in the study of the oral microbiome, the host, diet, and craft activities.

This book chapter is intended for beginners in the field of archaeological science, and will give a thorough and multifaceted review of all the different ways dental calculus has been, and can be, studied to shed light on various aspects of the human past.

Manuscript B

"A unified protocol for simultaneous extraction of DNA and proteins from archaeological dental calculus"

Zandra Fagernäs, Maite I. García-Collado, Jessica Hendy, Courtney A. Hofman, Camilla Speller, Irina Velsko, Christina Warinner

Published in *Journal of Archaeological Science* (June 2020)

Synopsis: In Manuscript B, we present and thoroughly evaluate a protocol through which both DNA and proteins can be extracted from a single sample of archaeological dental calculus. We compare this protocol to previously published standard protocols for extraction of either DNA or proteins. The unified protocol is evaluated using dental calculus from six individuals from four different archaeological sites, from a range of time periods and estimated preservation states, in order to ensure robustness of the results. In addition, we assess the effects of sample mass on the results.

We find that DNA yield is reduced through the unified protocol; however, as dental calculus is a very rich source of ancient DNA, this is unlikely to affect the success rate of studies. No effects of the unified protocol can be detected on characteristics of the extracted DNA fragments, or any downstream analyses. Protein yield, on the other hand, depends both on the amount of starting material, and which protocol was used. A minor decrease in the recovery of peptides containing hydrophilic amino acids is detected when using the unified protocol, likely due to the majority of the aqueous fraction being used for DNA extraction. The biases introduced by the unified protocol are, however, very minor, and can be accounted for in study design. We conclude that it is possible to extract both DNA and proteins from a single sample of dental calculus with this protocol, without introducing any major biases, thereby minimizing the need for destructive analyses.

Manuscript C

"Understanding the microbial biogeography of ancient human dentitions to guide study design and interpretation"

Zandra Fagernäs, Domingo C. Salazar-García, Azucena Avilés Fernández, María Haber Uriarte, Amanda Henry, Joaquín Lomba Maurandi, Andrew Ozga, Irina M. Velsko, Christina Warinner

Submitted to FEMS Microbes (September 2021).

Also available as a preprint in *bioRxiv* (August 2021).

Synopsis: Different niches within the oral cavity vary in factors such as pH, oxygen levels, and salivary flow, all of which are likely to affect the distribution of microbial taxa, but it is unknown to which degree such patterns are present in archaeological dental calculus. In order to map out differences in the reconstructed oral microbiome across the dentition, we performed a dense sampling of dental calculus from four Chalcolithic (ca. 4500-5000 BP) Iberian individuals. We performed shotgun metagenomic sequencing on 87 dental calculus samples from various dental niches, as well as four mandible bone samples, in order to account for colonization of microbes from the burial ground.

We find that the main driver of variation in the dental calculus microbiome is generally the individual the sample originated from, indicating that inter-individual factors drive the distribution of taxa to a higher degree than intra-individual factors. A single dental calculus sample can therefore represent the entire oral cavity of an individual in future studies. We do, however, find minor patterns in the distribution of the microbial taxa within the oral cavity, such as a difference in aerotolerance of taxa between anterior and posterior teeth, which should be taken into account when interpreting results. We also find that occlusal calculus generally differs from calculus from other tooth surfaces, and should therefore not be directly compared to non-occlusal samples. Finally, we find DNA assigned to grapevine (*Vitis vinifera*) exhibiting ancient DNA damage in both dental calculus and mandibular bone samples, which indicates a postmortem inclusion, showing the necessity of including environmental controls or supplementary lines of evidence when studying eukaryotic taxa within archaeological dental calculus.

Taken together, these results provide researchers with an understanding of patterns in archaeological dental calculus within the oral cavity, aiding in study design and result interpretation.

Manuscript D

"Exploring the potential of archaeogenetic studies of dental calculus to shed light on past human migrations in the Pacific"

Zandra Fagernäs, Monica Tromp, Stuart Bedford, Hallie Buckley, Geoffrey Clark, John Dudgeon, James Flexner, Anatauarii Leal-Tamarii, Cecil M. Lewis Jr, Kathrin Nägele, Andrew Ozga, Adam B. Rohrlach, Cosimo Posth, Richard Shing, Matthew Spriggs, Edson Willie, Frédérique Valentin, Irina M. Velsko, Christina Warinner

In preparation for submission.

Synopsis: Archaeological dental calculus holds promise for the study of past human migrations, through studies of changes in the oral microbiome and evolution and adaptation of microbial taxa. In addition, DNA tends to be better preserved in dental calculus than in skeletal elemets, potentially allowing for improved recovery of genetic material in areas such as the tropics, where heat and humidity increase the rate of DNA degradation. We explore the potential of this application of archaeogenetic study of dental calculus using the Pacific islands as a case study. In total, we perform shotgun metagenomic sequencing of 101 dental calculus samples, from 12 different islands across the Pacific, spanning a time range of approximately 3000 years.

We find that the majority of the dental calculus samples are well-preserved, again confirming the exceptional preservation of biomolecules in this substrate. Sample preservation is, however, highly variable, and cannot be predicted by climatic factors of the island they originate from. The microbial community composition of the dental calculus samples does not exhibit a geographic or temporal pattern; instead the samples cluster by island of origin, indicating that local factors on each island are mainly driving microbial community composition. We find that phylogenetic analyses of common oral bacterial taxa show some degree of temporal and geographic patterns, but selection of study species among the oral microbiome is non-trivial. To summarize, we show that dental calculus has the potential to reflect past human migrations, and may be a valuable alternative study material in the absence of human DNA preservation in skeletal elements, but analysis methods need further refinement.

MANUSCRIPT A

Manuscript Nr. A

Title of manuscript: Dental Calculus

Authors: Zandra Fagernäs and Christina Warinner

Bibliographic information: In press for the *Handbook of Archaeological Science*, 2nd edition (eds. Mark Pollard, Ruth-Ann Armitage and Cheryl Makarewicz)

The candidate is

⊗ First author, o Co-first author, o Corresponding author, o Coauthor.

Status: In press

Proportion (in %) of authors in the publication (to be indicated from 20%):

Author	Concept	Data analysis	Experiment	Manuscript composition	Material provision
Zandra Fagernäs	50%	100%	-	70%	-
Christina Warinner	50%	-	-	30%	-

Note: A more thorough breakdown of author contributions can be found in Appendix A.

Dental Calculus

ZANDRA FAGERNÄS

Department of Archaeogenetics, Max Planck Institute for the Science of Human History, Jena

and

CHRISTINA WARINNER

Department of Anthropology, Harvard University, Cambridge, MA

Dental calculus, also known as tooth tartar, is a mineralized form of dental plaque, a bacterial biofilm that naturally forms on teeth. It is the only part of the body that fossilizes during life (Warinner *et al*. 2015a, 2015b), and it can preserve for thousands or even millions of years. Until recently, this inconspicuous mineralized biofilm was generally ignored or even discarded, but with the growth of microfossil analysis in palaeoethnobotany and the advent of powerful new technologies in genetic sequencing and mass spectrometry, dental calculus has emerged as an unexpectedly valuable material in studies of the human past (Warinner *et al*. 2015a, 2015b).

Today, dental calculus is the richest known source of ancient biomolecules in the archaeological record, containing more than an order of magnitude more DNA than bone (Mann *et al*. 2018), and because it entraps a wide variety of biomolecules and micro-remains, its study has many diverse applications in archaeology. Techniques that have been applied to the study of archaeological dental calculus include light microscopy, electron microscopy, Raman spectroscopy, high throughput genetic sequencing, and mass spectrometry, among others (Velsko and Warinner 2017). There is a wealth of knowledge about the past that can be gained by studying the micro-organisms, dietary debris, and human-derived molecules present within dental calculus. By combining multiple lines of evidence, information about the human host and their diet, lifestyle, disease, and even craft activities can be gained from this calcified matrix. However, as an emerging area of research, there are still many aspects of its biology and taphonomy that remain to be understood, and research on dental calculus is developing and expanding at a fast pace. Here we review what dental calculus is and how it is formed, the types

of information that can be gained by studying it, and current questions and challenges that remain to be addressed.

DENTAL CALCULUS AND ITS FORMATION

Dental calculus forms through the calcification of dental plaque, a bacterial biofilm that forms naturally on the surface of teeth. Teeth are made up of three types of mineralized tissue (enamel, dentine, cementum) and one soft tissue (pulp), and teeth are the only part of the skeleton that is normally visible during life. The teeth are held in place by the periodontal ligament, which attaches the cementum surrounding the dentine to the alveolar bone of the upper and lower jaws. In healthy individuals, only the enamel crown is exposed, while the upper part of the tooth root and surrounding structures are covered by gingival soft tissue. Although bacteria attach well to soft tissues, they cannot directly bind to the hydroxyapatite mineral that makes up the enamel crown of teeth. During life, the teeth are continually bathed in saliva, which coats the teeth with proteins that can bind to hydroxyapatite (Kolenbrander 1988, Ventura *et al*. 2017), forming the acquired enamel pellicle (AEP) – a thin, approximately 0.005 millimetre thick proteinaceous layer made up of mostly acidic and proline-rich proteins. The AEP protects the teeth from dietary acids, but it also enables bacterial colonization (Scannapieco 1994, Jin and Yip 2002, Jepsen *et al.* 2011).

Bacteria that can adhere directly to the AEP are known as early colonizers, and most of these bacteria belong to the genera *Actinomyces* and *Streptococcus* (Whittaker *et al*. 1996). Further colonization of the tooth surface proceeds in an ordered manner, with successive groups of bridging bacteria, such as *Fusobacterium*, binding onto prior colonizers using specific receptors (Bradshaw *et al*. 1998, Welch *et al*. 2016). Late and terminal colonizers, such as *Porphyromonas* and the archaeon *Methanobrevibacter*, are the last to establish themselves in the biofilm (Kolenbrander *et al*. 2002). The ordered microbial succession of the biofilm results from the metabolic relationships between specific bacterial species, with some species providing the nutrients or microenvironments necessary for subsequent colonizers to grow (Jin and Yip 2002, Kolenbrander *et al*. 2005, Dewhirst *et al*. 2010). Throughout this process, the bacteria of the

growing dental plaque secrete polysaccharides and other compounds, forming a biofilm matrix that provides physical support and protects them from the environment and the host immune system. This biofilm also passively entraps microdebris, dietary biomolecules, and host biomolecules from the oral environment (Radini *et al*. 2017).

Over time, dental plaque biofilms periodically and spontaneously mineralize, forming dental calculus. The causes of this mineralization are not yet fully understood, but may be related to increasingly inefficient nutrient diffusion as the biofilm grows, resulting in cell death and changes in ion balance in the biofilm (White 1997). After mineralization, the plaque biofilm reforms and the cycle repeats, resulting in a constantly growing and mineralizing biofilm that builds up in successive layers, resembling tree rings. This incremental, layered calculus does not remodel, but rather preserves an enduring physical record of the individual's lifetime, provided that the deposits are not removed by dental hygiene activities during life or by excavation or curatorial activities after death.

Although dental calculus typically forms on the enamel above the gingival margin (supragingival calculus), it can also form on the tooth root below the gingival margin (subgingival calculus; Figure 1) (Schroeder 1969). Prolonged inflammation can cause the gingiva to separate from the tooth, enlarging the gingival sulcus and forming a periodontal pocket in which subgingival plaque can grow (Schroeder 1969). Being below the gumline, subgingival calculus has little direct contact with saliva, but it is continuously bathed in gingival crevicular fluid (GCF), a fluid derived from blood plasma that lubricates the gingiva and transports immune-related cells and molecules into the oral environment. This constant exposure to GCF causes subgingival calculus to differ in composition compared to supragingival calculus, and the environment in which subgingival calculus forms is further supplied with less oxygen, which leads the microbial community to have a higher proportion of anaerobic microbes (Abusleme *et al.* 2013).

In living individuals, these two types of calculus can be distinguished both by their location on the tooth and by their morphology. Subgingival calculus often appears hard and dense, reflecting the limited space of the periodontal pocket in which it formed, whereas supragingival deposits tend to be larger and less dense. However, there is a great deal of variation, and both types can be

found in deposits of varying colour, shape and density, especially in cases of advanced periodontal disease. In archaeological specimens, distinguishing the two types of calculus can be difficult, especially if there has been substantial recession of the alveolar bone. In these cases, supragingival calculus may overgrow subgingival deposits as the alveolar bone recedes, resulting in mixed deposits of supra- and subgingival calculus.

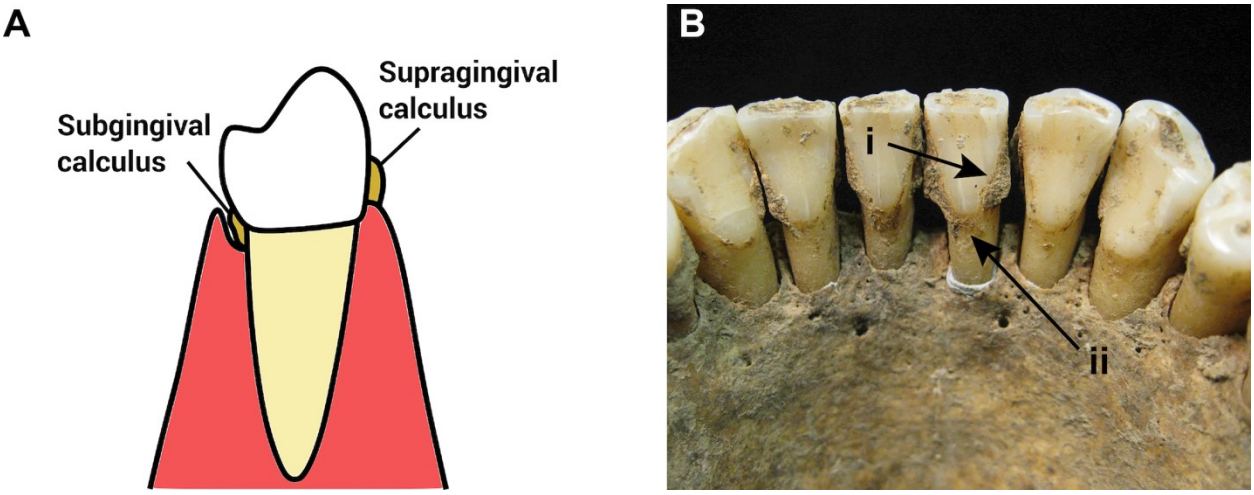


Figure 1 Location of dental calculus on teeth. (A) Diagram illustrating the locations of supra- and subgingival calculus on a living tooth. (B) Supragingival (i) and subgingival (ii) dental calculus deposits on the lower incisors (lingual view) of individual B61 from Dalheim, Germany, *c.*1100 CE. Large supragingival deposits often exhibit a lipping morphology where they had once overgrown the surface of the gingiva, as seen in (i). Alveolar bone recession is apparent for all visible teeth, an indication of generalized periodontitis, but subgingival calculus is only present on the two central incisors (ii). Photo credit: Christina Warinner.

DENTAL CALCULUS, IMMUNITY, AND PERIODONTAL DISEASE

The growth of dental plaque biofilms near or beneath the gingiva elicits an innate immune response from the host. Gingival epithelial cells produce antimicrobial peptides (AMPs) that are active against a broad range of microbes and slow their growth (Ji and Choi 2013). They also secrete compounds that attract neutrophils, a type of white blood cell involved in bacterial clearance. Neutrophils are an abundant cell type in the gingival sulcus and in GCF, and they are present even in clinically healthy gingival tissue – likely due to their near constant recruitment in

response to various oral bacteria. Neutrophils are able to neutralize bacteria through phagocytosis (ingestion) and by secreting potent proteases and other cell-killing molecules, as well as through NETosis, a kind of programmed cell death that is effective against biofilms. During heavy bacterial loads, when neutrophil activity is very high, they also cause inflammation and damage to host tissues, including both gingiva and bone (Ji and Choi 2013), which contributes to periodontal disease. Bacteria in dental plaque have varying susceptibilities to AMPs and phagocytosis by neutrophils. Late colonizing bacteria do not elicit a strong immune response, and are fairly resistant to AMPs and phagocytosis, thereby evading the innate immune response. Bridging taxa, on the other hand, are highly sensitive to AMPs and phagocytosis, and they also activate a stronger immune response, which leads them to be overall highly susceptible to the innate immune response (Ji and Choi 2013). In this way, the immune system slows down plaque progression by hindering bridging taxa from colonizing the biofilm, thereby also preventing the growth of late colonizing bacteria.

The growth of plaque biofilms plays an important role in the progression of periodontal diseases (gingivitis and periodontitis). Today, periodontal diseases are among the most prevalent diseases in industrialized societies, affecting up to 90% of the adult population worldwide (reviewed in Warinner 2016), with periodontitis affecting nearly half of the adult population in the United States (Eke *et al*. 2015). Periodontal diseases involve a chronic inflammation of the soft and hard tissues that support the teeth, and they can be divided into two types: gingivitis, which affects only the soft tissues and is reversible, and periodontitis, a more severe form that also affects hard tissues and causes the permanent destruction and loss of bone. Periodontitis is a leading cause of antemortem tooth loss and is also associated with a wide range of chronic inflammatory diseases, including cardiovascular disease and rheumatoid arthritis. Because periodontitis affects bone, it is readily observable in archaeological dentitions (Figure 1), but it has been little studied archaeologically due to the fact that there are currently no widely accepted osteological standards for its documentation or measurement in archaeological assemblages (Warinner 2016).

DENTAL CALCULUS COMPOSITION

As a mineralized biofilm, dental calculus consists of both organic (approximately 15-25%) and inorganic (approximately 75-85%) components (Schroeder 1969). In humans, the inorganic fraction of calculus consists almost entirely of calcium phosphates that have precipitated out of saliva and gingival crevicular fluid, and they occur primarily as hydroxyapatite, octacalcium phosphate, whitlockite, and brushite (Jin and Yip 2002, Schroeder 1969, Jepsen *et al*. 2011). Of these, brushite is mainly found during the earliest stages of calculus formation (Jin and Yip 2002), while hydroxyapatite is characteristic of more mature calculus. In other animal species with a more alkaline oral pH (e.g., dogs and pigs), the inorganic component of dental calculus is made up of predominantly calcium carbonates rather than calcium phosphates (Driessens and Verbeeck 1989).

The organic component of dental calculus is overwhelmingly made up of bacterial cells and their secreted matrix, but may also include traces of food remains (e.g., starch granules, pollen, proteins), host biomolecules (e.g., DNA, proteins), and assorted environmental debris (e.g., smoke particles, pigments, textile fibres). It has been estimated there are more than 200 million bacterial cells per milligram of plaque (Socransky *et al*. 1963), and bacterial DNA accounts for more than 99% of all DNA recovered from dental calculus (Warinner *et al*. 2014b, Ziesemer *et al*. 2019). To date, there is little evidence for the presence of human cells within dental calculus; rather, the low amounts of human DNA that have been found likely entered through inflammatory processes involving neutrophil-related immune activity (Mann *et al*. 2018). In contrast to DNA, however, the proteins within dental calculus have more diverse origins. Approximately 75% of the proteins in calculus originate from bacteria (Hendy *et al*. 2018a, Jersie-Christensen *et al*. 2018), while host salivary and immune proteins typically make up more than 20%, and a small and highly variable percentage originate from the diet (Hendy *et al*. 2018a). A large proportion of human proteins found in calculus are secretory proteins released by neutrophils or produced by the salivary glands (Warinner *et al*. 2014b). Lipids are also abundant in calculus and well preserved even in archaeological specimens (Velsko *et al*. 2017). These lipids likely originate from microbial cell walls and membranes, but additional sources cannot be excluded. A wide range of small molecules have also been identified in dental calculus, including dietary biomarkers (Velsko *et al*. 2017).

From the 1960s through the early 2000s, dental calculus was generally believed to be devoid of nucleic acids, including DNA (Jin and Yip 2002), and consequently little genetic work was performed on it. Subsequent research, however, demonstrated that dental calculus in fact contains very high amounts of DNA, and it is now recognized to be the richest known source of ancient DNA in the archaeological record (Warinner *et al*. 2015a). Compared to other archaeological materials, such as dentine or bone, the overall amount of DNA recovered from dental calculus is typically 10 to 1,000-fold higher (Warinner *et al*. 2014b, Ozga *et al*. 2016, Mann *et al*. 2018, Ziesemer *et al*. 2019). There are three factors that likely contribute to this difference. First, the starting cell density of dental calculus is much higher than that of dentine or bone. The cell density in subgingival dental plaque has been estimated at approximately 200 million bacterial cells per milligram of wet weight (Socransky *et al*. 1963), whereas bone is estimated to contain only ~600 cells per milligram (Bianconi *et al*. 2013). Even taking into account the fact that the genome size of human cells (6 billion base pairs, diploid) is more than 2000-fold larger than that of a bacterial cell (1-3 million base pairs), the high cell density of dental plaque still suggests that it should contain approximately two orders of magnitude more DNA than bone, which is consistent with empirical observations.

Second, decomposition progresses differently and more extensively in bone and dentine than in calculus, in large part due differences in their structure and organization. Living bone is porous and contains numerous channels through which blood vessels and nerves pass, connecting bone marrow and bone cells to the circulatory and nervous systems (Turner-Walker 2007). Dentine, although lacking blood vessels, nevertheless contains dentine tubules that connect the pulp chamber, which is filled with blood and nerves, to the tooth's surface. These channels and tubules, together with the pulp, provide passageways and nutrition for postmortem colonizing bacteria, and increase the rate of water penetration and chemical diagenesis (Turner-Walker 2007). Dental calculus, on the other hand, is typically solidly mineralized, without pores or channels, and it contains few readily accessible sources of nutrition for colonizing bacteria. Moreover, because dental calculus forms through the repeated death and rapid mineralization of dental plaque, it is largely biologically inert long before the death of the individual. Finally, the hydroxyapatite crystals in dental calculus are much larger than those produced in bone and dentine. DNA binds readily to hydroxyapatite, and this binding is thought to contribute to DNA

preservation in ancient samples (Warinner *et al*. 2015b). The larger hydroxyapatite crystals within dental calculus may provide greater opportunities for DNA binding, and thus preservation.

In addition to bacterial cells and host biomolecules, dental calculus is known to also entrap and preserve a wide range of debris and biomolecules from food and the environment. Such material becomes embedded in dental plaque when it is still soft and adhesive, and as the dental plaque biofilm calcifies, these biomolecules and fragments become encased within the mineral matrix. To date, a variety of dietary biomolecules have been identified in studies of modern dental plaque and calculus from individuals in the US and Europe, including chloroplast DNA sequences from manioc (Dewhirst *et al*. 2010), proteins specific to peanuts (Hendy *et al*. 2018a), and theobromine, a metabolite biomarker for chocolate and other caffeine-containing foods (Velsko *et al*. 2017). Small environmental fragments can also become entrapped in calculus through breathing, extramasticatory uses of teeth (e.g., such as holding or breaking materials), food preparation, and oral hygiene practices (Radini *et al*. 2017). For example, Juhola *et al.* (2019) identified bast fibres, animal hairs, and feathers in the dental calculus of Iron Age Scandinavians, suggesting exposure to these materials through textile, fur, and feather processing, and Hardy *et al*. (2016) identified pollen and microcharcoal from burning wood in Lower Paleolithic dental calculus, giving clues to the exposure of hominins to seasonal pollen and fire smoke. More examples of analyses of biomolecules and fragments trapped in dental calculus are described below.

HISTORY OF STUDY

Although references to archaeological dental calculus have occasionally appeared in the physical anthropology (Leigh 1925) and dentistry (Rowles 1961) literature since the 1920s, it was initially viewed as having low value in archaeological studies and was for many decades routinely cleaned off teeth and discarded prior to metric and non-metric trait studies of the dentition. Sustained interest in archaeological dental calculus can be traced to the 1970s and 1980s, when archaeologists and dental anthropologists began drawing attention to it in studies of oral health

and subsistence (Brothwell 1972, Evans 1973, Armitage 1975, Klepinger *et al*. 1977, Hillson 1979, Allison 1984, Cassidy 1984, Kennedy 1984, Dobney and Brothwell 1986). Most of this early work focused on describing the amount and distribution of calculus deposits. By the late 1980s, although researchers were increasingly recording the amount of calculus on archaeological teeth, the results were generally not comparable across studies, as there was no standardized system in place. In order to gain a better understanding of the etiology, frequency, and distribution of dental calculus, Dobney and Brothwell (1987) developed a standardized system for scoring dental calculus on archaeological specimens, and this system, together with others (Greene *et al*. 2005), is still in use today (Buikstra and Ubelaker 1994).

A turning point in dental calculus research occurred in 1975, when it was shown using microscopy that plant phytoliths, a type of dietary microfossil, could be recovered from the dental calculus of archaeological ungulates (Armitage 1975). Analyzing micro-remains in calculus quickly became widespread, focusing first on phytoliths (Middleton and Rovner 1994) and later also on starch granules (Cummings and Magennis 1997), as researchers began to explore ancient diets and environments using this new direct approach (see Warinner *et al*. 2015a and Radini *et al*. 2017 for reviews).

In parallel, explorations of archaeological dental calculus using SEM (scanning electron microscopy) revealed an abundance of well-preserved microbial cells (Dobney and Brothwell 1988), and subsequent antibody labelling (Linossier *et al*. 1996) confirmed that these bacteria were consistent with oral taxa and therefore likely endogenous. In situ preservation of DNA within archaeological calculus was first demonstrated by Preus *et al*. (2011) using gold-labeled antibody TEM (transmission electron microscopy), and successful DNA sequencing from archaeological dental calculus was first achieved by de la Fuente *et al*. (2013) using species-specific PCR primers targeting selected oral bacteria. Later that same year, Adler *et al*. (2013) performed the first metataxonomic study of ancient dental calculus, amplifying and sequencing a portion of the 16S rRNA gene to reconstruct the microbial community of an ancient oral microbiome for the first time. Shortly thereafter, Warinner *et al*. (2014b) published a shotgun metagenomic and metaproteomic study of archaeological dental calculus, allowing the first detailed taxonomic and functional investigation of the ancient oral microbiome, including an

inventory of ancient antibiotic resistance genes and the reconstruction of the first ancient oral microbial genome (of the bacterium *Tannerella forsythia*). This opened the door to numerous subsequent studies that have refined the genetic methodologies applied to ancient dental calculus (Ziesemer *et al*. 2015, Velsko *et al*. 2019, Fellows Yates *et al.* 2021b), characterized the oral microbiome of Neanderthals (Weyrich *et al*. 2017, Fellows Yates *et al.* 2021b), and explored the use of dental calculus as a source of DNA for human genome sequencing (Ozga *et al*. 2016, Ziesemer *et al*. 2019), among others.

At the same time that genetic methods were being developed, new approaches in mass spectrometry were also being pursued. Although early enthusiasm for carbon and nitrogen stable isotope analysis of dental calculus (Scott and Poulson 2012, Poulson *et al*. 2013) was later tempered by criticisms of its high sample size requirements and low data resolution (Eerkens *et al*. 2014, Salazar-García *et al*. 2014), early experiments with thermal desorption-gas chromatography-mass spectrometry (TD-GC-MS) and pyrolysis gas chromatography mass spectrometry (Py-GC-MS) yielded more promising results. Using these techniques, chemical evidence of smoke exposure and cooking was identified in Neanderthal calculus (Hardy *et al*. 2012). Although such small molecule research is still in an early phase, recent confirmation that a wide variety of microbial, host, and dietary metabolites is present in both modern and archaeological dental calculus (Hardy *et al*. 2012, Velsko *et al*. 2017) suggests that this is a promising area of research.

Alongside these developments, new ground was also being broken in protein mass spectrometry. Warinner *et al*. (2014b) were the first to apply a bottom-up proteomics approach to dental calculus, revealing a well-preserved proteome documenting bacterial activity and host immune response. Later that same year, milk proteins were identified in the calculus of the Greenland Norse, revealing changes in their dairying practices through time (Warinner *et al*. 2014a). Protein mass spectrometry has since been used to further explore ancient diets (Hendy *et al*. 2018a, Jeong *et al*. 2018, Scott *et al.* 2020) and oral health and disease (Jersie-Christensen *et al*. 2018, Fotakis *et al.* 2020), and recent studies have also aimed to combine different methodologies for studying archaeological dental calculus, such as DNA and protein analyses (Fagernäs *et al.* 2020, Fotakis *et al.* 2020), and DNA and plant micro-remains (Modi *et al.* 2020).

Dental calculus is ubiquitous in the archaeological record, found in all areas of the world and in populations from all time periods. The archaeological dental calculus that has been investigated to date spans five continents and—in the case of dental calculus from Miocene fossil apes—dates as far back as 12 million years ago (Hershkovitz *et al*. 1997, Fuss et al. 2018). Even focusing on palaeogenomic studies alone, the number of archaeological dental calculus samples studied to date show enormous geographic and temporal breadth spanning more than 100 000 years (Figure 2).

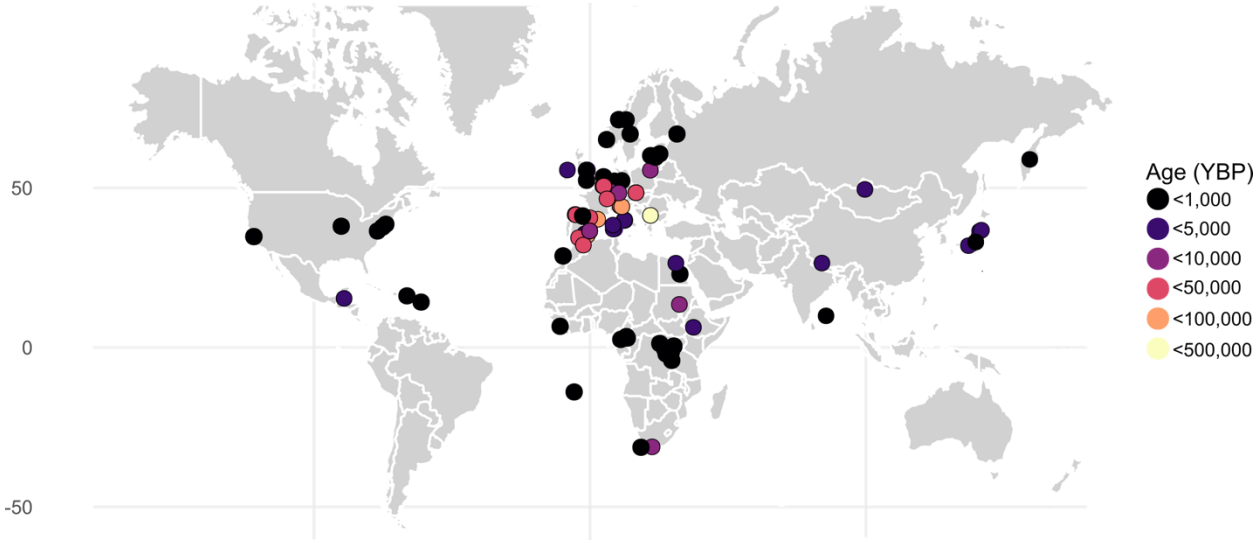


Figure 2 Meta-analysis of archaeological dental calculus studies, showing the geographic breadth, time depth, and research focus of palaeogenomics investigations from around the world. Data compiled from AncientMetagenomeDir (Fellows Yates *et al.* 2021a; downloaded on 2021-06-21).

Although dental calculus is ubiquitous among human populations, the amount of dental calculus found on the dentition of a given individual can vary greatly depending on many factors, including personal hygiene practices and health state, as well as salivary composition and flow (Mandel 1974, Epstein *et al*. 1980). At a population level, high carbohydrate diets have been associated with higher overall rates of dental calculus (Lieverse 1999), but the sample size used

to make this association was small, and the subject requires further study. In archaeological collections, dental calculus deposits on individual dentitions can vary from barely visible to hundreds of milligrams (Warinner *et al*. 2014b), but typical archaeological dentitions generally support deposits in the range of 5-25 mg. Fortunately, this amount is sufficient to perform many of the biomolecular and microfossil analyses described below.

PRESERVATION

An important aspect of dental calculus is its capacity to preserve and protect biomolecules over long periods of time. Initially, dental calculus was viewed as a fairly closed system. This was largely because elemental analysis of sectioned dental calculus showed little infiltration of soil material, and early metagenomic analysis found only low levels of environmental bacteria in the calcified matrix (Warinner *et al*. 2014b). However, as with all archaeological remains, dental calculus does undergo processes of decay and biomolecular degradation, as well as varying degrees of postmortem microbial colonization and environmental exchange. For example, it has been shown that small hydrophilic metabolites are disproportionately lost from dental calculus compared to larger, more hydrophobic metabolites (Velsko *et al*. 2017). This suggests that some movement of water through the calculus deposits does occur (Mann *et al*. 2018). Because of this, it is necessary to consider movement and exchange of hydrophilic metabolites, especially small organic acids, when interpreting the metabolite profiles of ancient dental calculus, as small molecules from plant roots and the burial environment can enter into the calculus postmortem.

Although the initial effects of decomposition are mitigated for dental calculus due to its densely mineralized nature at the time of death, nearly all archaeological dental calculus contains at least some exogenous DNA. Measuring this precisely can be difficult, but tools such as SourceTracker (Knights *et al*. 2011), SourcePredict (Borry 2019), and cuperdec (Fellows Yates *et al*. 2021b) can be useful for estimating overall sample preservation and inferring exogenous sources of DNA (Warinner *et al*. 2014b, Weyrich *et al*. 2017, Mann *et al*. 2018, Fellows Yates *et al.* 2021b). In addition, tools such as decontam (Davis *et al*. 2018) can be used to identify and remove specific DNA sequences that are likely to belong to contaminant taxa (Fellows Yates *et al.* 2021b),

PMDtools (Skoglund *et al*. 2014) can be used to remove DNA sequences that lack characteristic ancient DNA damage, and HOPS (Hübler *et al.* 2019) can be used to estimate the authenticity of ancient taxonomic identifications. However, care should still always be taken during sample excavation, handling, and storage to reduce contamination exposure.

Over time, as the proportion of original biomolecules decreases due to sample degradation, the relative proportion of exogenous material gradually increases (Figure 3). In addition to post-mortem colonization of bacteria from the burial environment, excavation, storage and handling can also introduce contamination. Factors that contribute to this contamination include handling samples without gloves, storing samples improperly, and working with samples in settings or laboratories that are not suited for archaeological materials. Conventional laboratories used for modern genetic work, for example, can introduce large amounts of contamination - from the laboratory environment (e.g., airborne PCR products), the worker, and reagents (Weyrich *et al*. 2019). It is thus essential to work within cleanroom facilities when conducting biomolecular analyses of ancient samples, because such cleanrooms maintain a level of cleanliness that is substantially higher than in a normal laboratory and include features such as air filtration systems to minimize airborne contamination, personal protective equipment to reduce contact contamination, and the use of high purity reagents to reduce reagent contamination. As an additional precaution, it is also advisable to collect environmental context samples (e.g., paired bone samples, archaeological sediments, museum storage swabs) when possible, as this can aid in the identification of postmortem contaminants.

Although laboratory work should only be performed in a dedicated cleanroom, dental calculus sampling is typically performed on site, such as in the field, at a storage facility, or in a museum. In order to ensure that as little contamination as possible is introduced when removing calculus from teeth, detailed recommendations for sampling have been proposed (e.g., Velsko and Warinner 2017, Hendy *et al*. 2018b, Fellows Yates *et al.* 2021c). These recommendations include sterilizing sampling equipment before and between sampling, wearing nitrile gloves when handling specimens (as a protein, latex rubber can contaminate protein studies), covering the working area with aluminium foil to provide a clean working surface, and collecting the calculus into laboratory grade microcentrifuge tubes. Although the latter is preferable for genetic,

proteomic, and micro-remains analysis, if organic residue analysis (ORA) is also planned, the samples can alternatively be collected into sterile foil pouches. Foil pouches can be difficult to work with and they easily tear, but they are necessary for ORA analysis because even trace plastic contamination can interfere with GC-MS (gas chromatography-mass spectrometry) measurements. By adhering to these recommendations, effective sampling can be achieved and post-excavation contamination can be kept to a minimum, ensuring the best possible results for downstream biomolecular analyses.

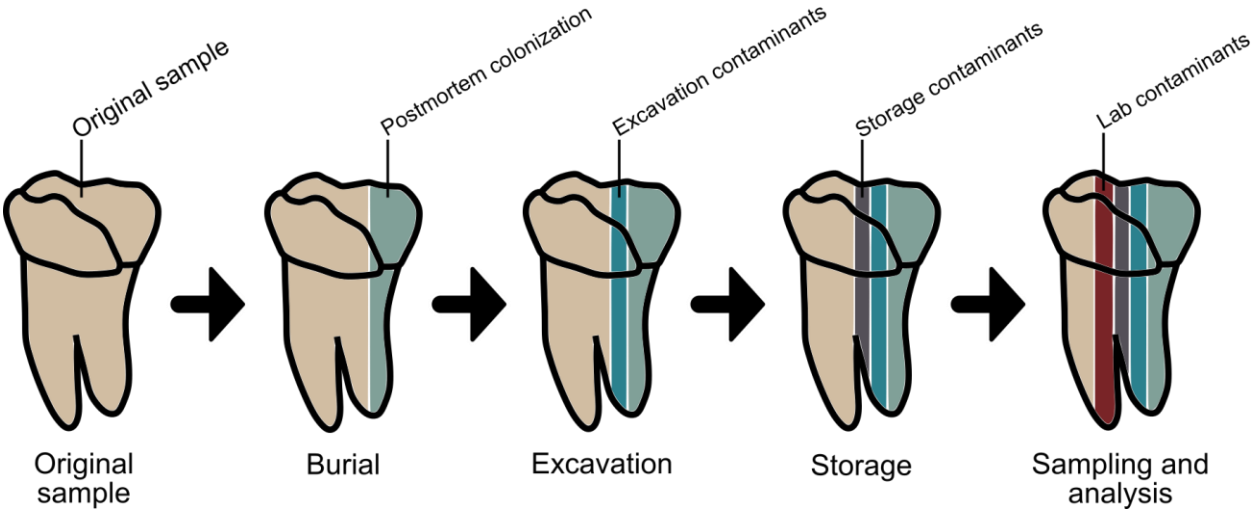


Figure 3 Theoretical model of the proportion of the original biomolecules that remain in a sample over time, as it is colonized by microbes and contaminated by various postmortem events and environments.

APPLICATIONS

Because multiple types of biomolecules and micro-remains are preserved over long time periods in dental calculus, studies of dental calculus have many different applications. Below, we discuss how studies of dental calculus have been used to investigate the oral microbiome, the host, diet, and craft activities in the past.

Oral microbiome

The vast majority of biomolecules in dental calculus stem from micro-organisms. Microbial DNA within archaeological dental calculus mainly derives from human-associated oral species, with a minor proportion originating from environmental organisms (Mann *et al*. 2018). The oral microbiome is potentially informative about an individual's health, diet, and lifestyle, as these factors can influence the microbiome.

DNA

In 2013, Adler and colleagues were the first to reconstruct the microbial communities present within ancient dental calculus, which they achieved by using Roche 454 technology to deeply sequence the V3 region of the 16S ribosomal RNA gene, a genetic marker that is commonly used to rapidly inventory the bacterial species that are present in a sample. The results suggested that the human oral microbiome had changed over time, following major changes in lifestyle, such as the transition from hunter-gatherers to farmers (Adler *et al*. 2013). However, it was later found that there are limitations to a 16S rRNA amplicon-based approach when studying ancient bacteria (Ziesemer *et al.* 2015). The variable regions of the 16S rRNA gene have often been targeted in metataxonomic studies because they can be used to classify sequences into taxonomic units without the need for extensive sequencing efforts. However, the V3 region, which is the variable region most suited for ancient microbiome studies, was shown to introduce biases in the reconstructed microbiomes (Ziesemer *et al*. 2015). The V3 region has extensive length polymorphisms, and due to the generally short DNA sequences in archaeological samples, taxa with longer target sequences are underrepresented. Thus, this approach was shown to be sensitive to sample age and preservation, and consequently return biased results. The method is no longer used for ancient samples, and today nearly all genetic work on ancient microbiomes instead uses an alternative approach known as shotgun metagenomic sequencing. Shotgun metagenomic sequencing is more suitable to studies of ancient microbiomes, as it is much less affected by biases in DNA degradation (Ziesemer *et al*. 2015, Mann *et al*. 2018).

The first shotgun metagenomic study of archaeological dental calculus was performed on samples from a medieval German cemetery (Warinner *et al*. 2014b), and subsequent studies have substantially broadened the geographic and temporal scope of analysed ancient oral microbiomes

(e.g., Mann *et al*. 2018) to include not only a wide variety of humans, but also Neanderthals (Weyrich *et al*. 2017, Fellows Yates *et al.* 2021b), primates (Ozga *et al*. 2019, Fellows Yates *et al.* 2021b), and other mammals (Brealey *et al.* 2020). To date, most studies have focused on describing the ancient oral microbiome, but there are many additional research questions that could be potentially answered by genetic studies of archaeological calculus, such as how the oral microbiome and specific microbial species, such as starch digesting oral streptococci, have evolved over time (Fellows Yates *et al.* 2021b). In addition, ancient oral microbiome studies also have the potential to reveal patterns of past human migration and interaction through phylogenetic analysis of their oral bacterial strains (Eisenhofer *et al.* 2020).

One surprising pattern that has emerged from genetic studies of archaeological dental calculus is that three species of bacteria that are today almost exclusively associated with periodontal disease, *Porphyromonas gingivalis*, *Tannerella forsythia*, and *Treponema denticola*, appear to have been both more prevalent and abundant in the past. However, this may not necessarily indicate a higher periodontal disease burden in ancient human populations. Because these taxa are late colonizers of the biofilm (Kolenbrander *et al*. 2002), this pattern may rather reflect the fact that plaque biofilms were able to grow to a higher degree of maturity in the past, when tooth brushing and other forms of oral hygiene were less common. In many studies, modern dental plaque has been used as a proxy for modern calculus when studying microbiome composition in ancient calculus, since studies of modern dental calculus are rare. However, Velsko *et al*. (2019) demonstrated that dental plaque is not a good proxy for dental calculus, due to the different maturation stages of the substrates. As dental plaque deposits grow, different species are able to colonize the biofilm in a specific order, creating a layered bacterial structure (Jin and Yip 2002, Dewhirst *et al*. 2010). The last to colonize and expand are anaerobes classically associated with periodontal disease. The large proportion of these late colonizing 'pathogenic' bacteria seen in ancient calculus may simply reflect the fact that ancient biofilms were able to grow and mature longer than what is typically seen in the dental plaque of individuals practicing regular tooth brushing, which disrupts the normal process of microbial succession.

Proteins and metabolites

In addition to genetic studies of the oral microbiome, proteins and metabolites within

archaeological dental calculus can also be highly informative about the microorganisms inhabiting the oral cavity, as well as host immunological response. For example, a study of medieval Danish dental calculus found an ancient proteome suggestive of oral health (Jersie-Christensen *et al*. 2018), something that would be much more difficult to deduce from metagenomic data alone. In the study, nearly 90% of the bacterial proteins could be assigned to the genus level or lower, enabling links to be made between the taxonomic composition of the calculus and its functional profile. In addition to proteins, metabolites are also powerful indicators of microbial activity (Takahashi *et al*. 2010), and metabolic pathway analysis can provide detailed information about health and disease processes in the oral cavity (Velsko *et al*. 2017).

Host

Although primarily a microbial biofilm, dental calculus contains trace amounts of host DNA, as well as host proteins and metabolites that can be informative about the physiology and health status of ancient individuals.

DNA

The proportion of endogenous ancient human DNA in archaeological dental calculus is typically very low, ranging from 0.005-0.5% (Mann *et al.* 2018, Ziesemer *et al*. 2019), but generally found consistently in samples from a wide range of preservation environments and time periods (Mann *et al*. 2018). When normalized for the very high amounts of total DNA recovered from dental calculus and the amount of input material used for DNA extraction from dental calculus (~2 mg) compared to dentine (~50 mg), the overall amount of human DNA in dental calculus can be similar to dentine from the same individual (Mann *et al*. 2018, Ziesemer *et al*. 2019). However, performing human genetic studies using dental calculus is generally more difficult than from dentine because less material is available for analysis, and because human DNA within dental calculus has been found to be more degraded than in surrounding skeletal tissues. For example, when comparing human DNA from paired dental calculus and dentine samples from the same individual, Mann *et al*. (2018) found that human DNA fragments were significantly shorter in calculus than in dentine, and human DNA within calculus was also significantly shorter than

bacterial DNA in the same calculus sample. This suggests that human DNA within dental calculus has undergone additional destructive processes, which are likely related to its origin and manner of incorporation into the biofilm. Human DNA within calculus is not cellular, but rather is incorporated either passively from saliva and the surrounding mucosa, or actively through processes of innate immune activity (Fuchs *et al*. 2007, Mann *et al*. 2018). The latter involves host programmed cell death and the release of oxidizing agents and proteases, as well as microbial defense using nucleases, which all likely contribute to the enhanced human DNA damage observed within calculus (Mann *et al*. 2018, Warinner *et al*. 2014b).

Although the low proportion and short fragment lengths of human DNA in dental calculus make studies of the human genome using this substrate more difficult, important human genetic information can still be recovered. For example, by using in solution capture technology to enrich for DNA belonging to the human mitochondrial genome, Ozga *et al*. (2016) were able to use dental calculus to reconstruct full mitochondrial genomes for 700-year-old Native American individuals from Illinois, USA. As this method focuses on ectopic microbial accretions (dental calculus) and not skeletal tissues, it provides an alternative method for obtaining genetic information that may prove useful for tribes, descendant communities, museums, and other stakeholders who wish to avoid destructive analysis of human remains. Although mitochondrial genome reconstruction from dental calculus has been highly successful, attempts to date to recover whole human genomes from dental calculus have not been as promising (Ziesemer *et al*. 2019). The problem, however, may relate with the specific approach that has been attempted – whole genome enrichment – which is known to suffer from issues of low on-target recovery and low complexity, even from dentine. If future methodological developments are able to overcome these problems, dental calculus could become an important substrate for reconstructing the genetic histories of populations who otherwise wish to avoid the destructive analysis of human remains.

Proteins and metabolites

In addition to DNA, proteins and metabolites from calculus can also yield valuable information about the human host. Human salivary and immune proteins make up as much as one fifth of the calculus proteome (Warinner *et al*. 2014b) and can be used to study the digestion of food and

human defense against microbes (Hendy *et al*. 2018a, Jersie-Christensen *et al*. 2018). For very well preserved samples, dental calculus metabolite profiles can also provide information about an individual's sex and age (Velsko *et al*. 2017).

Diet

Dental calculus can be used to directly study the individual diets of ancient people, enabling a higher precision in identifying specific dietary components than techniques that involve isotopic averaging over many years, such as carbon and nitrogen stable isotope analysis of collagen. During the formation of dental calculus, a diverse range of dietary debris and biomolecules can become entrapped within the growing bacterial biofilm and subsequent calcified matrix, and this dietary record preserves over long periods of time.

Microfossils

Microfossils, such as phytoliths, diatoms, pollen, and starch granules, may become incorporated into dental calculus (Radini *et al*. 2017). Phytoliths and diatoms both consist of biogenic silica, which leads them to be very resistant to salivary digestion and mastication, and they can preserve over exceptionally long time periods (Dudgeon and Tromp 2014). Phytoliths, also known as plant glass, are formed by the concretion of silica (mostly silicon dioxide) within plant tissues, and phytolith shapes differ between plant groups, giving clues about their taxonomic origins. Diatoms, on the other hand, are single-celled algae found in the environment, and they can be specific to certain types of water bodies and soil. In a study of phytoliths and diatoms in dental calculus from early 16th to late 19th century CE Rapa Nui (Easter Island), several plant species were identified, as well as the water source used by these individuals (Dudgeon and Tromp 2014). Phytoliths from palms – which no longer exist on the island but were once abundant – were identified from the calculus samples. The authors also identified regional differences in the amounts of diatoms in calculus, which may reflect differences in the use of permanent versus transient drinking sources (Dudgeon and Tromp 2014). In certain contexts, diatoms may also indicate local patterns of shellfish consumption since diatoms are especially abundant in bivalves, such as oysters, mussels and clams, which feed by filtering diatoms and other phytoplankton from the water. Thus, biogenic silica entrapped within dental calculus can provide

multiple types of information about the past, from food items and drinking sources to prehistoric environments and water utilization practices.

Starch is the main form of stored energy in plants, and starch is packaged into granules that are stored within specialized organelles in plant cells. Although starch granules are found in all plant tissues, they are especially abundant in seeds, fruits, roots, and specialized underground starch storage organs, such as tubers. Based on their size and morphology, starch granules may be identifiable to specific plant families or genera – and in some cases even species – and their taphonomic features can also indicate whether the starch was cooked prior to consumption (Henry *et al*. 2009, 2011). Different types of cooking cause different types of damage to starch, such that it is sometimes possible to discern specific cooking practices from starch grain morphology (Henry *et al*. 2009). For example, microfossils identified in the dental calculus of Neanderthals from Spy Cave in Belgium (~36 Kya) and from Shanidar Cave in Iraq (~46 Kya) have shown that Neanderthals consumed a variety of plants in their diet, and the recovered starch grains showed signs of cooking (Henry *et al*. 2011). In the Americas, the transformation of gelatinized starch into semi-crystalline particles known as starch spherulites is considered a diagnostic indicator of nixtamalization, a form of maize alkaline cooking that has been performed in Mesoamerica for millennia (Johnson and Marston 2020). Starch granules and their byproducts are thus highly informative about the diets and cooking practices of past populations.

Microfossils can survive within dental calculus for extraordinary lengths of time, making them especially valuable for studying the diets and behaviours of early hominins. For example, phytoliths recovered from the dental calculus of *Australopithecus sediba* (~2 Ma) has revealed that this hominin consumed fruit, leaves, and wood/bark from a diverse range of plants (Henry *et al*. 2012). In conjunction with dental microwear and stable isotope analyses, these results provide clues about the diet, environment, and behaviour of this early hominin. Similarly, a study of dental calculus from an early member of *Homo* at the site of Sima del Elefante in Spain, dated to 1.2-1.1 Ma, revealed the presence of starch granules from grasses, as well as other plants (Hardy *et al*. 2016). In addition, inedible plant fibres also embedded in the calculus suggested raw material processing, or perhaps even dental hygiene activities, and other environmental debris found within the calculus, including conifer pollen, insect fragments, and fungal spores, provided

further clues about the environment in which the individual lived (Hardy *et al*. 2016). Microfossils and environmental debris preserved in dental calculus are thus valuable sources of information about a wide range of activities and environmental exposures throughout human evolution, from diet and cooking practices to oral hygiene activities to airborne pollen exposure.

However, it is important to note that, as with other archaeological finds, microfossils obtained from calculus require authentication to ensure that they do not originate from contamination. For example, Power *et al*. (2015) showed that the microfossil assemblage identified in calculus from an Upper Palaeolithic individual from the El Mirón cave (Cantabria, Spain) differed substantially from microfossils identified in bone, sediment, and herbivore calculus from the same site. This suggested that the microfossils were original to the human calculus and not contaminants from the general burial environment. Precautions must also be taken to ensure that contamination does not occur in the laboratory. Previous studies have identified starch contaminants in common laboratory consumables (gloves, pipette tips, parafilm), chemicals (sodium polytungstate), and cleaning materials (tissue paper, paper towels, brooms, mops), as well as on laboratory benches, tables, and floors (Crowther *et al*. 2014). Nearly all of these contaminant starches originated from maize, wheat (and close relatives), and potato, which are widely used in the manufacture of paper, textiles, cosmetics, and plastics. Finds of these species should therefore be interpreted with caution. Rigorous cleaning procedures and special precautions are necessary to reduce and monitor postmortem and post-excavation contamination of samples in any ancient microfossil study.

DNA

Ancient DNA from dietary sources is generally identified in very low or trace amounts in dental calculus (Warinner *et al*. 2014b, Weyrich *et al*. 2017), and authentication of these finds can be challenging (Mann *et al.* 2020). Similarity in genetic sequences between species, missing species in databases, overrepresented species in databases, and contaminated published genomes can all confound species identification and authentication when very few DNA sequences of putative dietary origin are identified. Although DNA damage patterns are often used as a criterion for authenticity, confidently identifying this pattern typically requires hundreds to thousands of DNA sequences per dietary species (Warinner *et al*. 2017, Mann *et al.* 2020), which is rarely

found in dental calculus. Using DNA from archaeological dental calculus to study ancient diets thereby remains a challenging area where feasible methods for authentication are still lacking.

Proteins

In contrast to DNA, robust evidence for dietary protein has been identified in archaeological calculus and can be used to reliably infer specific components of ancient diets. For example, in a study of British dental calculus from a range of archaeological sites and time periods, dietary proteins were identified in approximately one third of individuals (Hendy *et al*. 2018a). Plant proteins and non-human animal derived proteins represented 0-2% of total identified proteins in these samples, and plant protein sequences were found from oat, pea and Brassicaceae, as well as from broader groups of plants. In another study of dental calculus from the Levant, plant proteins from staple crops (wheat, sesame), as well as spices (turmeric), fruit (banana), and oil (soybean), were identified together with phytoliths of wheat and date palm, providing a glimpse into the scope of long-distance trade during the Late Bronze and Iron Ages (Scott *et al.* 2020). Overall, microfossil analysis and archaeological context suggest that a wider range of plants were consumed than those identified by proteomics, which suggests that this method is more effective at identifying some foodstuffs compared to others. There are many possible explanations for this, including variation in protein preservation (Scott *et al.* 2020) and the fact that, especially for plants, current protein reference databases are strongly biased towards key domesticates and a few selected biological pathways (Hendy *et al*. 2018a).

Among animal proteins, milk proteins are among the best preserved dietary proteins found in dental calculus – in particular the whey protein beta-lactoglobulin (BLG). BLG was the first dietary protein identified in both modern (Warinner *et al.* 2014b) and ancient (Warinner *et al*. 2014a) dental calculus, and it has emerged as a powerful biomarker of past dairying, not only because it preserves well, but also because species-specific amino acid variants in this protein make it useful for distinguishing between the milks of different dairy livestock. To date, BLG has been identified in the dental calculus of numerous archaeological dairying societies throughout Eurasia (Warinner *et al.* 2014a, Hendy *et al*. 2018a, Jeong *et al*. 2018, Jersie-Christensen *et al.* 2018, Wilkin *et al.* 2020), and its identification has been used to resolve longstanding archaeological debates. For example, in a study of medieval Norse settlements on

Greenland, dental calculus BLG was found to decline in prevalence and abundance through time at the same time that stable isotope evidence indicated an increasing reliance on marine resources (Warinner *et al*. 2014a). This supported the hypothesis that environmental deterioration during the Little Ice Age had contributed to the collapse of the traditional dairy-based subsistence of the Norse settlers. As a proof of concept, the study also investigated known milk consumers from Europe, as well as a population from the island of St. Helena whose diet did not include milk, and showed that milk proteins were found in a quarter of Europeans and none of the individuals from St. Helena. Milk proteins have also helped to reveal the time depth of dairying in Asia. In two studies of prehistoric Mongolia (Jeong *et al*. 2018, Wilkin *et al*. 2020), milk proteins (BLG and alpha-S1-casein) have been used to trace the origins of ruminant dairying in the region to the Early Bronze age (*c*.3000 BCE), followed by the subsequent rise of horse milking during the Late Bronze (*c*.1200 BCE) and camel milking during the Mongol period (*c*.1200-1300 CE). Palaeoproteomics is thus a powerful tool with which to directly study ancient diets, both at the level of populations and individuals.

Metabolites

Although most metabolites in dental calculus originate from bacteria or host saliva and GCF, several metabolites of certain dietary origin have been identified (Velsko *et al*. 2017). For example, theobromine, a plant methylxanthine naturally found in plants that produce caffeine (e.g., cacao, coffee, tea, kola, and maté), was recently identified in modern dental calculus (Velsko *et al*. 2017). This suggests that small molecule studies of dental calculus could potentially be used to trace the consumption of foods with special properties, such as stimulants, in ancient populations.

Microscopy, genetic analysis, protein analysis, and small molecule metabolite studies of dental calculus all provide potential avenues for reconstructing ancient diets. Nevertheless, it is important to keep in mind that each provides a slightly different perspective and is subject to different strengths and weaknesses. For example, microfossils preserve over longer timescales than either DNA or proteins, but DNA and proteins, when preserved, often allow more precise taxonomic identifications to be made. It is also important to consider human behaviour when interpreting dietary evidence from dental calculus. Humans specifically target starch-rich and

protein-rich foods for consumption, and so starch-based and protein-based analyses, respectively, are often the most direct way to identify these foods. The apparent near absence of food-derived DNA within dental calculus should not be surprising given that DNA within foods is typically present in only trace amounts and may have been seriously damaged by cooking processes prior to consumption. In addition to these factors, not all foods or food parts appear to integrate into the oral biofilm equally. Among plant foods, for example, wheat starches are often found in abundance in the dental calculus of wheat-farming societies, but wheat DNA and proteins are uncommonly identified in dental calculus. This pattern may reflect the substantial size difference between the large starch-rich endosperm and the small germ in domesticated wheat seeds, as well as how the wheat is prepared, milled, and cooked. Likewise, among animal products, milk proteins are much more often identified than muscle proteins, even though societies practicing animal husbandry likely consumed both. Combining several different types of analyses may be necessary in order to reach a more complete view of the diet of past populations.

Craft activities

In addition to food, dental calculus has also been found to entrap a wide variety of non-edible fragments and environmental debris (Radini *et al*. 2017). Some of these debris have been associated with craft activities that may be specific to certain professions or methods of manufacturing. For example, GC-MS has been used to identify possible traces of oil shale or bitumen in Neanderthal dental calculus (Hardy *et al*. 2012). Bitumen is known to have been used as a hafting material by Neanderthals, and such finds suggest links between specific tools and makers. In another study, the dental calculus of a medieval religious woman was found to contain numerous microscopic blue particles. By combining SEM with energy dispersive X-ray spectroscopy (EDS) and Raman spectroscopy, the authors were able to identify the particles as originating from lapis lazuli. Lapis lazuli was the source of a rare and expensive blue pigment used in manuscript painting during the Middle Ages, and the authors concluded that the woman was likely a manuscript artist. Given that little historical information about female artists and scribes has survived from this time period, this finding provided important support to previous claims that women contributed to medieval book production (Radini *et al*. 2019). Dental calculus thus provides opportunities to infer information about the past behaviours, activities, and even

professions of specific individuals that might otherwise leave few traces in the archaeological record.

CONCLUSIONS

Over the past few decades, archaeologist perceptions of dental calculus have transformed from 'grime on teeth' to the source of a veritable treasure trove of intimate information about past lives. It has been an unusual journey for such a humble archaeological material, but clearly one that is not yet over. Current studies of dental calculus have only scratched the surface of what is possible, and as technologies continue to develop and new analysis methods become available, more questions will be answered, and new ones asked.

REFERENCES

Abusleme, L., Dupuy, A.K., Dutzan, N., Silva, N., Burleson, J.A., Strausbaugh, L.D., Gamonal, J. and Diaz, P.I. (2013). The subgingival microbiome in health and periodontitis and its relationship with community biomass and inflammation. *The International Society for Microbial Ecology (ISME) Journal*, **7**:1016–1025.

Adler, C.J., Dobney, K., Weyrich, L.S., Kaidonis, J., Walker, A.W. and 7 others (2013). Sequencing ancient calcified dental plaque shows changes in oral microbiota with dietary shifts of the Neolithic and Industrial revolutions. *Nature Genetics,* **45**:450–455.

Allison, M.J. (1984). Paleopathology in Peruvian and Chilean populations. In Armelagos, G.J. and Cohen, M.N. (eds) *Paleopathology at the Origins of Agriculture*, 515–529. Academic Press: New York.

Armitage, P.L.(1975). The extraction and identification of opal phytoliths from the teeth of ungulates. *Journal of Archaeological Science*, **2**:187–197.

Bianconi, E., Piovesan, A., Facchin, F., Beraudi, A., Casadei, R. and 8 others (2013). An estimation of the number of cells in the human body. *Annals of Human Biology*, **40**:463–471.

Borry, M. (2019). Sourcepredict: Prediction of metagenomic sample sources using dimension reduction followed by machine learning classification. *Journal of Open Source Software*, **4**:1540.

Bradshaw, D.J., Marsh, P.D., Watson, G.K. and Allison, C. (1998). Role of *Fusobacterium nucleatum* and coaggregation in anaerobe survival in planktonic and biofilm oral microbial communities during aeration. *Infection and Immunity,* **66**:4729–4732.

Brealey, J.C., Leitão, H.G., van der Valk, T., Xu, W., Bougiouri, K., Dalén, L. and Guschanski, K. (2020). Dental calculus as a tool to study the evolution of the mammalian oral microbiome. *Molecular Biology and Evolution,* **37**:3003-3022.

Brothwell, D.R. (1972). *Digging up Bones: The excavation, treatment and study of human skeletal remains.* British Museum: London.

Buikstra, J.E. and Ubelaker, D.H. (1994). *Standards for Data Collection from Human Skeletal Remains.* Arkansas Archeological Survey: Fayetteville, Ark.

Cassidy, C.M. (1984). Skeletal evidence for prehistoric subsistence adaptation in the Central Ohio River Valley. In Armelagos, G.J. and Cohen, M.N. (eds) *Paleopathology at the Origins of Agriculture*, 307–345. Academic Press: New York.

Crowther, A., Haslam, M., Oakden, N., Walde, D. and Mercader, J. (2014). Documenting contamination in ancient starch laboratories. *Journal of Archaeological Science*, **49**:90-104.

Cummings, L.S. and Magennis, A. (1997). A phytolith and starch record of food and grit in Mayan human tooth tartar. In Pinilla, A., Juan-Tresserras, J. and Machado, M.J. (eds) *The State-of-the-Art of Phytoliths in Soils and Plants*, 211-218. Monografias del Centro de Ciencias Medioambientales: Madrid.

Davis, N.M., Proctor, D.M., Holmes, S.P., Relman, D.A. and Callahan, B.J. (2018). Simple statistical identification and removal of contaminant sequences in marker-gene and metagenomics data. *Microbiome*, **6**:226.

da Silva Ventura, T.M., de Paula Silva Cassiano, L., de Souza e Silva, C.M., Taira, E.A., Leite, A. de L., Rios, D. and Buzalaf, M.A.R. (2017). The proteomic profile of the acquired enamel pellicle according to its location in the dental arches. *Archives of Oral Biology*, **79**:20–29.

de la Fuente, C., Flores, S. and Moraga, M. (2013). DNA from human ancient bacteria: A novel source of genetic evidence from archaeological dental calculus. *Archaeometry*, **55**:767–778.

Dewhirst, F.E., Chen, T., Izard, J., Paster, B.J., Tanner, A.C.R., Yu, W.-H., Lakshmanan, A. and Wade, W.G. (2010). The human oral microbiome. *Journal of Bacteriology,* **192**:5002–5017.

Dobney, K. and Brothwell, D. (1986). Dental calculus: its relevance to ancient diet and oral ecology. In Cruwys, E. and Foley, R.A. (eds) *Teeth and Anthropology,* 55–81. BAR 291, British Archaeological Reports: Oxford.

Dobney, K. and Brothwell, D. (1987). A method for evaluating the amount of dental calculus on teeth from archaeological sites. *Journal of Archaeological Science*, **14**:343–351.

Dobney, K. and Brothwell, D. (1988). A scanning electron microscope study of archaeological dental calculus. In Olsen, S.L. (ed.) *Scanning Electron Microscopy in Archaeology*, 372–385. British Archaeological Reports: Oxford.

Driessens, F.C.M. and Verbeeck, R.M.H. (1989). Possible pathways of mineralization of dental plaque. In Ten Cate, J.M. (ed.) *Recent Advances in the Study of Dental Calculus*, 7–17. Oxford University Press: Oxford.

Dudgeon, J.V. and Tromp, M. (2014). Diet, geography and drinking water in Polynesia: Microfossil research from archaeological human dental calculus, Rapa Nui (Easter Island). *International Journal of Osteoarchaeology*, **24**:634–648.

Eerkens, J.W., de Voogt, A., Dupras, T.L., Rose, S.C., Bartelink, E.J. and Francigny, V. (2014). Intra- and inter-individual variation in $\delta^{13}C$ and $\delta^{15}N$ in human dental calculus and comparison to bone collagen and apatite isotopes. *Journal of Archaeological Science*, **52**:64-71.

Eisenhofer, R., Kanzawa-Kiriyama, H., Shinoda, K. and Weyrich, L.S. (2020). Investigating the

demographic history of Japan using ancient oral microbiota. *Philosophical Transactions of the Royal Society B*, **375**:20190578.
Eke, P.I., Dye, B.A., Wei, L., Slade, G.D., Thornton-Evans, G.O., Borgnakke, W.S., Taylor, G.W., Page, R.C., Beck, J.D. and Genco, R.J. (2015). Update on prevalence of periodontitis in adults in the United States: NHANES 2009 to 2012. *Journal of Periodontology*, **86**:611–622.
Epstein, S.R., Mandel, I. and Scopp, I.W. (1980). Salivary composition and calculus formation in patients undergoing hemodialysis. *Journal of Periodontology,* **51**:336–338.
Evans, D.T. (1973). A preliminary evaluation of tooth tartar among the preconquest Maya of the Tayasal area, El Petén, Guatemala. *American Antiquity*, **38**:489–493.
Fagernäs, Z., García-Collado, M.I., Hendy, J., Hofman, C.A., Speller, C., Velsko, I. and Warinner, C. (2020). A unified protocol for simultaneous extraction of DNA and proteins from archaeological dental calculus. *Journal of Archaeological Science*,, **118**:105135.
Fellows Yates, J., Andrades Valtueña, A., Vägene, A.J., Cribdon, B., Velsko, I.M. and 12 others (2021a). Community-curated and standardised metadata of published ancient metagenomic samples with AncientMetagenomeDir. *Nature Scientific Data*, **8**:31.
Fellows Yates, J.A., Velsko, I.M., Aron, F., Posth, C., Hofman, C.A. and 44 others (2021b). The evolution and changing ecology of the African hominid oral microbiome. *Proceedings of the National Academy of Sciences of the USA,* **118**:e2021655118.
Fellows Yates, J.A., Aron, F., Neumann, G., Velsko, I.M., Skourtanioti, E. and 7 others (2021c). *A-Z of ancient DNA protocols for shotgun Illumina Next Generation sequencing, v.2,* June 15, 2021. DOI: 10.17504/protocols.io.bvt9n6r6
Fotakis, A.K., Denham, S.D., Mackie, M., Orbegozo, M.I., Mylopotamitaki, D. and 8 others (2020). Multi-omic detection of *Mycobacterium leprae* in archaeological human dental calculus. *Philisophical Transactions of the Royal Society B,* **375**:20190584.
Fuchs, T.A., Abed, U., Goosmann, C., Hurwitz, R., Schulze, I., Wahn, V., Weinrauch, Y., Brinkmann, V. and Zychlinsky, A. (2007). Novel cell death program leads to neutrophil extracellular traps. *Journal of Cell Biology,* **176**:231–241.
Fuss, J., Uhlig, G. and Böhme, M. (2018). Earliest evidence of caries lesion in hominids reveal sugar-rich diet for a Middle Miocene dryopithecine from Europe. *PloS ONE*, **13**(8):e0203307.
Greene, T.R., Kuba, C.L. and Irish, J.D. (2005). Quantifying calculus: a suggested new approach for recording an important indicator of diet and dental health. *Homo*, **56**:119–132.
Hardy, K., Buckley, S., Collins, M.J., Estalrrich, A., Brothwell, D. and 12 others (2012). Neanderthal medics? Evidence for food, cooking, and medicinal plants entrapped in dental calculus. *Naturwissenschaften*, **99**:617–626.
Hardy, K., Radini, A., Buckley, S., Sarig, R., Copeland, L., Gopher, A. and Barkai, R. (2016). Dental calculus reveals potential respiratory irritants and ingestion of essential plant-based nutrients at Lower Palaeolithic Qesem Cave Israel. *Quaternary International*, **398**:129–135.
Hendy, J., Warinner, C., Bouwman, A., Collins, M.J., Fiddyment, S. and 15 others (2018a). Proteomic evidence of dietary sources in ancient dental calculus. *Proceedings of the Royal Society B*, **285**:20180977.
Hendy, J., Welker, F., Demarchi, B., Speller, C., Warinner, C. and Collins, M.J. (2018b). A guide to ancient protein studies. *Nature Ecology & Evolution,* **2**:791–799.
Henry, A.G., Hudson, H.F. and Piperno, D.R. (2009). Changes in starch grain morphologies

from cooking. *Journal of Archaeological Science*, **36**:915–922.
Henry, A.G., Brooks, A.S. and Piperno, D.R. (2011). Microfossils in calculus demonstrate consumption of plants and cooked foods in Neanderthal diets (Shanidar III, Iraq; Spy I and II, Belgium). *Proceedings of the National Academy of Sciences of the USA*, **108**:486–491.
Henry, A.G., Ungar, P.S., Passey, B.H., Sponheimer, M., Rossouw, L., Bamford, M., Sandberg, P., de Ruiter, D.J. and Berger, L. (2012). The diet of *Australopithecus sediba*. *Nature*, **487**:90–93.
Hershkovitz, I., Kelly, J., Latimer, B., Rothschild, B.M., Simpson, S., Polak, J. and Rosenberg, M. (1997). Oral bacteria in Miocene Sivapithecus. *Journal of Human Evolution*, **33**:507–512.
Hillson, S.W. (1979). Diet and dental disease. *World Archaeology*, **11**:147–162.
Hübler, R., Key, F.M., Warinner, C., Bos, K.I., Krause, J. and Herbig, A. (2019). HOPS: automated detection and authentication of pathogen DNA in archaeological remains. *Genome Biology*, **20**:280.
Jeong, C., Wilkin, S., Amgalantugs, T., Bouwman, A.S., Taylor, W.T.T. and 21 others (2018). Bronze Age population dynamics and the rise of dairy pastoralism on the eastern Eurasian steppe. *Proceedings of the National Academy of Sciences of the USA,* **115**:E11248–E11255.
Jepsen, S., Deschner, J., Braun, A., Schwarz, F. and Eberhard, J. (2011). Calculus removal and the prevention of its formation. *Periodontology 2000*, **55**:167–188.
Jersie-Christensen, R.R., Lanigan, L.T., Lyon, D., Mackie, M., Belstrøm, D. and 7 others (2018). Quantitative metaproteomics of medieval dental calculus reveals individual oral health status. *Nature Communications,* **9**:4744.
Ji, S. and Choi, Y. (2013). Innate immune response to oral bacteria and the immune evasive characteristics of periodontal pathogens. *Journal of Periodontal & Implant Science,* **43**:3–11.
Jin, Y. and Yip, H.-K. (2002). Supragingival calculus: formation and control. *Critical Reviews in Oral Biology & Medicine,* **13**:426–441.
Johnson, E.S. and Marston, J.M. (2020). The experimental identification of nixtamalized maize through starch spherulites. *Journal of Archaeological Science*, **113**:105056.
Juhola, T., Henry, A.G., Kirkinen, T., Laakkonen, J. and Väliranta, M. (2019). Phytoliths, parasites, fibers, and feathers from dental calculus and sediment from Iron Age Luistari cemetery, Finland. *Quaternary Science Reviews,* **222**:105888.
Kennedy, K.A. (1984). Growth, nutrition and pathology in changing paleodemographic settings in South Asia. In Armelagos, G.J. and Cohen, M.N. (eds) *Paleopathology at the Origins of Agriculture*, 169–192. Academic Press: New York.
Klepinger, L.L., Kuhn, J.J. and Thomas, J. Jr (1977). Prehistoric dental calculus gives evidence for coca in early coastal Ecuador. *Nature*, **269**:506–507.
Knights, D., Kuczynski, J., Charlson, E.S., Zaneveld, J., Mozer, M.C., Collman, R.G., Bushman, F.D., Knight, R. and Kelley, S.T. (2011). Bayesian community-wide culture-independent microbial source tracking. *Nature Methods,* **8**:761–763.
Kolenbrander, P.E. (1988). Intergeneric coaggregation among human oral bacteria and ecology of dental plaque. *Annual Review of Microbiology,* **42**:627–656.
Kolenbrander, P.E., Andersen, R.N., Blehert, D.S., Egland, P.G., Foster, J.S. and Palmer, R.J. Jr. (2002). Communication among oral bacteria. *Microbiology and Molecular Biology*

Reviews, **66**:486–505.
Kolenbrander, P.E., Egland, P.G., Diaz, P.I. and Palmer, R.J. Jr. (2005). Genome-genome interactions: bacterial communities in initial dental plaque. *Trends in Microbiology,* **13**:11–15.
Leigh, R.W. (1925). Dental pathology of Indian tribes of varied environmental and food conditions. *American Journal of Physical Anthropology*, **8**:179–199.
Lieverse, A.R. (1999). Diet and the aetiology of dental calculus. *International Journal of Osteoarchaeology,* **9**:219–232.
Linossier, A., Gajardo, M. and Olavarria, J. (1996). Paleomicrobiological study in dental calculus: *Streptococcus mutans*. *Scanning Microscopy,* **10**:1005–1014.
Mackie, M., Hendy, J., Lowe, A.D., Sperduti, A., Holst, M., Collins, M.J. and Speller, C.F. (2017). Preservation of the metaproteome: variability of protein preservation in ancient dental calculus. *STAR: Science & Technology of Archaeological Research,* **3**:58-70.
Mandel, I.D. (1974). Biochemical aspects of calculus formation. II. Comparative studies of saliva in heavy and light calculus formers. *Journal of Periodontal Research,* **9**:211–221.
Mann, A.E., Sabin, S., Ziesemer, K., Vågene, Å.J., Schroeder, H. and 16 others (2018). Differential preservation of endogenous human and microbial DNA in dental calculus and dentin. *Scientific Reports*, **8**:9822.
Mann, A.E., Fellows Yates, J.A., Fagernäs, Z., Austin, R.M., Nelson, E.A. and Hofman, C.A. (2020). Do I have something in my teeth? The trouble with genetic analyses of diet from archaeological dental calculus. *Quaternary International,* https://doi.org/10.1016/j.quaint.2020.11.019.
Middleton, W.D. and Rovner, I. (1994). Extraction of opal phytoliths from herbivore dental calculus. *Journal of Archaeological Science*, **21**:469–473.
Modi, A., Pisaneschi, L., Zaro, V., Vai, S., Vergata, C., Casalone, E., Caramelli, D., Moggi-Cecchi, J., Mariotti Lippi, M. and Lari, M. (2020). Combined methodologies for gaining much information from ancient dental calculus: testing experimental strategies for simultaneously analysing DNA and food residues. *Archaeological and Anthropological Sciences,* **12**:10.
Ozga, A.T., Nieves-Colón, M.A., Honap, T.P., Sankaranarayanan, K., Hofman, C.A., Milner, G.R., Lewis, C.M., Jr, Stone, A.C. and Warinner, C. (2016). Successful enrichment and recovery of whole mitochondrial genomes from ancient human dental calculus. *American Journal of Physical Anthropology*, **160**:220–228.
Ozga, A.T., Gilby, I., Nockerts, R.S., Wilson, M.L., Pusey, A. and Stone, A.C. (2019). Oral microbiome diversity in chimpanzees from Gombe National Park. *Scientific Reports,* **9**:17354.
Poulson, S.R., Kuzminsky, S.C., Scott, G.R., Standen, V.G., Arriaza, B., Muñoz, I. and Dorio, L. (2013). Paleodiet in northern Chile through the Holocene: extremely heavy $\delta^{15}N$ values in dental calculus suggest a guano-derived signature? *Journal of Archaeological Science*, **40**:4576–4585.
Power, R.C., Salazar-García, D.C., Straus, L.G., González Morales, M.R. and Henry, A.G. (2015). Microremains from El Mirón Cave human dental calculus suggest a mixed plant–animal subsistence economy during the Magdalenian in Northern Iberia. *Journal of Archaeological Science*, **60**:39–46.
Preus, H.R., Marvik, O.J., Selvig, K.A. and Bennike, P. (2011). Ancient bacterial DNA (aDNA) in dental calculus from archaeological human remains. *Journal of Archaeological*

Science, **38**:1827–1831.
Radini, A., Nikita, E., Buckley, S., Copeland, L. and Hardy, K. (2017). Beyond food: The multiple pathways for inclusion of materials into ancient dental calculus. *American Journal of Physical Anthropology*, **162**:71–83.
Radini, A., Tromp, M., Beach, A., Tong, E., Speller, C., McCormick, M., Dudgeon, J.V., Collins, M.J., Rühli, F., Kröger, R. and Warinner, C. (2019). Medieval women's early involvement in manuscript production suggested by lapis lazuli identification in dental calculus. *Science Advances*, **5**:eaau7126.
Rowles, S. (1961). Further studies on the crystalline constituents of oral calculus. *Journal of Dental Research,* **40**:1284.
Salazar-García, D.C., Richards, M.P., Nehlich, O. and Henry, A.G. (2014). 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). *Journal of Archaeological Science,* **47**:70–77.
Scannapieco, F.A. (1994). Saliva-bacterium interactions in oral microbial ecology. *Critical Reviews in Oral Biology & Medicine,* **5**:203–248.
Schroeder, H.E. (1969). *Formation and Inhibition of Dental Calculus*. H. Huber Verlag: Bern.
Scott, A.S., Power, R.C., Altmann-Wendling, V., Artzy, M., Martin, M.A.S. and 9 others (2020). Evidence for long-distance trade in exotic foods and spices from South Asia to the Near East in the 2nd millennium BCE. *Proceedings of the National Academy of Sciences of the USA,* **118**:e2014956117.
Scott, G.R. and Poulson, S.R. (2012). Stable carbon and nitrogen isotopes of human dental calculus: a potentially new non-destructive proxy for paleodietary analysis. *Journal of Archaeological Science,* **39**:1388–1393.
Skoglund, P., Northoff, B.H., Shunkov, M.V., Derevianko, A.P., Pääbo, S., Krause, J. and Jakobsson, M. (2014). Separating endogenous ancient DNA from modern day contamination in a Siberian Neandertal. *Proceedings of the National Academy of Sciences of the USA*, **111**:2229–2234.
Socransky, S.S., Gibbons, R.J., Dale, A.C., Bortnick, L., Rosenthal, E. and Macdonald, J.B., (1963). The microbiota of the gingival crevice area of man. I. Total microscopic and viable counts and counts of specific organisms. *Archives of Oral Biology*, **8**:275–280.
Takahashi, N., Washio, J. and Mayanagi, G. (2010). Metabolomics of supragingival plaque and oral bacteria. *Journal of Dental Research,* **89**:1383–1388.
Turner-Walker, G. (2007). The chemical and microbial degradation of bones and teeth. In Pinhasi, R. and Mays, S. (eds) *Advances in Human Palaeopathology,* 3–29. Wiley: Chichester.
Velsko, I.M. and Warinner, C. (2017). Bioarchaeology of the human microbiome. *Bioarchaeology International*, **1**:86–99.
Velsko, I.M., Overmyer, K.A., Speller, C., Klaus, L., Collins, M.J. and 9 others (2017). The dental calculus metabolome in modern and historic samples. *Metabolomics*, **13**:134.
Velsko, I.M., Fellows Yates, J.A., Aron, F., Hagan, R.W., Frantz, L.A.F. and 6 others (2019). Microbial differences between dental plaque and historic dental calculus are related to oral biofilm maturation stage. *Microbiome*, **7**:102.
Warinner, C. (2016). Dental calculus and the evolution of the human oral microbiome. *Journal of the California Dental Association*, **44**:411–420.

Warinner, C., Hendy, J., Speller, C., Cappellini, E., Fischer, R. and 18 others (2014a). Direct evidence of milk consumption from ancient human dental calculus. *Scientific Reports*, **4**:7104.

Warinner, C., Rodrigues, J.F.M., Vyas, R., Trachsel, C., Shved, N. and 27 others (2014b). Pathogens and host immunity in the ancient human oral cavity. *Nature Genetics*, **46**:336–344.

Warinner, C., Speller, C. and Collins, M.J. (2015a). 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**:20130376.

Warinner, C., Speller, C., Collins, M.J. and Lewis, C.M. Jr. (2015b). Ancient human microbiomes. *Journal of Human Evolution,* **79**:125–136.

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

Welch, J.L.M., Rossetti, B.J. and Rieken, C.W. (2016). Biogeography of a human oral microbiome at the micron scale. *Proceedings of the National Academy of Sciences of the USA*, **113**:791–800.

Weyrich, L.S., Duchene, S., Soubrier, J., Arriola, L., Llamas, B. and 26 (2017). Neanderthal behaviour, diet, and disease inferred from ancient DNA in dental calculus. *Nature*, **544**:357–361.

Weyrich, L.S., Farrer, A.G., Eisenhofer, R., Arriola, L.A., Young, J., Selway, C.A., Handsley-Davis, M., Adler, C.J., Breen, J. and Cooper, A. (2019). Laboratory contamination over time during low-biomass sample analysis. *Molecular Ecology Resources,* **19**:982–996.

White, D.J. (1997). Dental calculus: recent insights into occurrence, formation, prevention, removal and oral health effects of supragingival and subgingival deposits. *European Journal of Oral Sciences,* **105**:508–522.

Whittaker, C.J., Klier, C.M. and Kolenbrander, P.E. (1996). Mechanisms of adhesion by oral bacteria. *Annual Review of Microbiology*, **50**:513–552.

Wilkin, S., Ventresca Miller, A., Taylor, W.T.T., Miller, B.K., Hagan, R.W. and 14 others (2020). Dairy pastoralism sustained Eastern Eurasian steppe populations for 5000 years. *Nature Ecology and Evolution,* **4**:346-355.

Ziesemer, K.A., Mann, A.E., Sankaranarayanan, K., Schroeder, H., Ozga, A.T. and 15 others (2015). Intrinsic challenges in ancient microbiome reconstruction using 16S rRNA gene amplification. *Scientific Reports*, **5**:16498.

Ziesemer, K.A., Ramos-Madrigal, J., Mann, A.E., Brandt, B.W., Sankaranarayanan, K. and 12 others (2019). The efficacy of whole human genome capture on ancient dental calculus and dentin. *American Journal of Physical Anthropology,* **168**:496–509.

MANUSCRIPT B

Manuscript Nr. B

Title of manuscript: A unified protocol for simultaneous extraction of DNA and proteins from archaeological dental calculus

Authors: Zandra Fagernäs, Maite I. García-Collado, Jessica Hendy, Courtney A. Hofman, Camilla Speller, Irina Velsko, Christina Warinner

Bibliographic information: Fagernäs, Zandra, Maite I. García-Collado, Jessica Hendy, Courtney A. Hofman, Camilla Speller, Irina Velsko, and Christina Warinner. 2020. "A Unified Protocol for Simultaneous Extraction of DNA and Proteins from Archaeological Dental Calculus." *Journal of Archaeological Science* 118 (June): 105135; doi: https://doi.org/10.1016/j.jas.2020.105135

The candidate is

⊗ First author, o Co-first author, o Corresponding author, o Coauthor.

Status: Published

Proportion (in %) of authors in the publication (to be indicated from 20%):

Author	Concept	Data analysis	Experiment	Manuscript composition	Material provision
Zandra Fagernäs	25%	100%	100%	70%	-
Maite I. García-Collado	-	-	-	-	50%
Jessica Hendy	25%	-	-	-	50%
Courtney A. Hofman	-	-	-	-	-
Camilla Speller	25%	-	-	-	-
Irina Velsko	-	-	-	-	-
Christina Warinner	25%	-	-	-	-

Note: A more thorough breakdown of author contributions can be found in Appendix A. Supplementary figures can be found in Appendix B. The supplementary data files, which contain numerous large tables, can be accessed in the online version of this manuscript (DOI: https://doi.org/10.1016/j.jas.2020.105135).

Journal of Archaeological Science 118 (2020) 105135

Contents lists available at ScienceDirect

Journal of Archaeological Science

journal homepage: http://www.elsevier.com/locate/jas

A unified protocol for simultaneous extraction of DNA and proteins from archaeological dental calculus

Zandra Fagernäs[a], Maite I. García-Collado[b], Jessica Hendy[a,c], Courtney A. Hofman[d,e], Camilla Speller[c,f], Irina Velsko[a], Christina Warinner[a,g,h,*]

[a] Department of Archaeogenetics, Max Planck Institute for the Science of Human History, Jena, 07745, Germany
[b] Department of Geography, Prehistory and Archaeology, University of the Basque Country, Leioa, 48940, Spain
[c] Department of Archaeology, University of York, York, YO10 5NG, United Kingdom
[d] Department of Anthropology, University of Oklahoma, Norman, 73019, United States
[e] Laboratories of Molecular Anthropology and Microbiome Research, University of Oklahoma, Norman, 73019, United States
[f] Department of Anthropology, University of British Columbia, Vancouver, V6T 1Z4, Canada
[g] Department of Anthropology, Harvard University, Cambridge, 02138, United States
[h] Faculty of Biological Sciences, Friedrich Schiller University, Jena, 07743, Germany

ARTICLE INFO

Keywords:
Ancient DNA
Palaeoproteomics
Oral microbiome
Dental plaque
Metagenomics

ABSTRACT

Archaeological materials are a finite resource, and efforts should be made to minimize destructive analyses. This can be achieved by using protocols combining extraction of several types of biomolecules or microparticles, which decreases the material needed for analyses while maximizing the information yield. Archaeological dental calculus is a source of several different types of biomolecules, as well as microfossils, and can tell us about the human host, microbiome, diet, and even occupational activities. Here, we present a unified protocol allowing for simultaneous extraction of DNA and proteins from a single sample of archaeological dental calculus. We evaluate the protocol on dental calculus from six individuals from a range of time periods and estimated preservation states, and compare it against previously published DNA-only and protein-only protocols. We find that most aspects of downstream analyses are unaltered by the unified protocol, although minor shifts in the recovered proteome can be detected, such as a slight loss of hydrophilic proteins. Total protein recovery depends on both the amount of starting material and choice of extraction protocol, whereas total DNA recovery is significantly reduced using the unified protocol (mean 43%). Nevertheless, total DNA recovery from dental calculus is generally very high, and we found no differences in DNA fragment characteristics or taxonomic profile between the protocols. In conclusion, the unified protocol allows for simultaneous extraction of two complementary lines of biomolecular evidence from archaeological dental calculus without compromising downstream results, thereby minimizing the need for destructive analysis of this finite resource.

1. Introduction

Biomolecular analysis is becoming increasingly feasible in archaeology as methods improve and costs decrease. However, archaeological materials are a finite resource, and there is a need to develop techniques that reduce the extent of destructive sampling, while concurrently maximizing the amount of information that can be obtained (Green and Speller, 2017). So-called non-destructive methods for extracting DNA (Bolnick et al., 2012; Rohland et al., 2004) and proteins (Fiddyment et al., 2015; Manfredi et al., 2017; van Doorn et al., 2011) from ancient materials have been developed, but many of these techniques are limited by downstream challenges, including lower or biased biomolecule recovery, higher rates of contamination, and a higher burden of sample degradation. Depending on the research question and the sample type, such techniques may not be suitable. Alternatively, it is possible to modify existing destructive methods to either simultaneously or sequentially extract multiple classes of biomolecules (e.g., DNA, proteins, lipids, metabolites) from the same specimen, thereby taking advantage of the effectiveness of destructive methods while improving efficiency and minimizing waste. Such an approach has previously been applied with success to combining methods for genetic analysis and radiocarbon dating of skeletal remains (Korlević et al., 2018) and to

* Corresponding author. Department of Archaeogenetics, Max Planck Institute for the Science of Human History, Jena, 07745, Germany.
E-mail address: warinner@shh.mpg.de (C. Warinner).

https://doi.org/10.1016/j.jas.2020.105135
Received 8 January 2020; Received in revised form 25 March 2020; Accepted 25 March 2020
Available online 24 April 2020

combining lipid analysis and radiocarbon dating (Berstan et al., 2008; Casanova et al., 2018).

Developing a unified protocol to recover multiple classes of biomolecules also presents several research advantages. Less material is required for analysis, which reduces sampling demands and mitigates variability that can arise from heterogeneous substrates. In addition, different classes of biomolecules can provide distinct but complementary lines of analysis, thereby strengthening the results of a study. For example, two studies of ancient dental calculus (Velsko et al., 2019; Warinner et al., 2014b) identified both DNA and proteins from the periodontal pathogen *Porphyromonas gingivalis*, thereby providing strong evidence of both its presence and activity. For highly degraded samples, targeting multiple classes of biomolecules also improves the chances of retrieving at least one successfully (Chen et al., 2019). DNA and proteins are known to have considerably differing maximum ages of survival, with proteins surviving for millions of years (Demarchi et al., 2016), whereas the oldest successfully recovered DNA, even from permafrost, is less than a million years old (Orlando et al., 2013). However, regardless of age, it can be difficult to accurately predict the analytical success of a given sample in advance, as the preservation of individual specimens can be highly dependent on local environmental factors.

One archaeological substrate for which a unified extraction protocol would be particularly useful is dental calculus. Archaeological dental calculus is a rich source of ancient biomolecules (e.g., DNA, proteins, metabolites) and microremains (e.g., plant microfossils, environmental debris) that originate from the host, microbes, food, and the environment (Radini et al., 2017; Velsko and Warinner, 2017). Forming through periodic calcification of dental plaque, dental calculus entraps and preserves such debris throughout an individual's lifetime, and it can persist over very long periods of time due to its densely mineralized nature. From a single sample of calculus, it is possible to gain information about the individual's genome (Ozga et al., 2016; Ziesemer et al., 2019), oral microbiome and health (Velsko et al., 2019, 2017; Warinner, 2016; Warinner et al., 2014b), diet (Hendy et al., 2018; Jeong et al., 2018; Warinner et al., 2014a), and even occupational activities (Radini et al., 2019). Nevertheless, the amount of dental calculus available per individual is typically low (on the order of tens of milligrams), which limits the number of biomolecule extractions, and therefore analyses, that can be performed.

Efficient protocols have been developed to isolate different types of biomolecules and microremains from ancient dental calculus, including DNA and proteins (Jeong et al., 2018; Mann et al., 2018), as well as phytoliths and starch granules (Henry and Piperno, 2008; Tromp et al., 2017). Currently, however, many of these protocols are incompatible. For example, calculus decalcification using HCl (for microfossil extraction) is incompatible with DNA analysis; strong heat denaturation (for protein extraction) is incompatible with DNA and starch granule analysis; use of proteinase K (for DNA extraction) is incompatible with protein analysis. DNA and protein extraction protocols are the most difficult to combine, as they typically involve multiple incompatible steps. Previous attempts have been made to modify and combine DNA and protein extraction protocols for archaeological teeth (Rusu et al., 2019) and dental calculus (Mackie et al., 2017); however, in each case the performance and efficiency of the combined protocol was not compared to similar, non-combined protocols. As such, the efficiency and potential biases of the combined protocols could not be systematically evaluated.

Here we present a unified protocol (UP) for the simultaneous extraction of DNA and proteins from archaeological dental calculus. This protocol is based on prior observations from genomic and proteomic experiments that, following decalcification, most DNA is recovered from the supernatant rather than the cell pellet, while most protein is recovered from the cell pellet rather than the supernatant. We evaluate the performance of the UP on a panel of dental calculus specimens originating from archaeological sites of different antiquities and estimated preservation states, and we apply the protocol to both high (10 mg) and low (2 mg) starting amounts of dental calculus. We then compare these results to those obtained using DNA-only (DO) and protein-only (PO) protocols in order to assess the efficiency of the UP and identify potential biases. We evaluate the UP on the basis of total DNA and protein recovery, DNA fragment length, DNA GC-content, amino acid properties, protein hydropathy and molecular weight, reconstructed microbiome, proteome composition, and contamination burden. Overall, we find that the UP introduces relatively few biases and is a more efficient use of starting material than performing separate DO and PO extractions. Only DNA recovery is strongly impacted, with the UP resulting in a 43% lower DNA yield on average compared to the DO protocol, which we propose is due to the absence of proteinase K during the initial separation of DNA following decalcification. Nevertheless, given that ancient dental calculus typically contains high amounts of DNA in excess of metagenomic sequencing requirements (Mann et al., 2018), this reduction in DNA recovery is unlikely to limit downstream investigations in most cases.

2. Methods

2.1. Sample material

Dental calculus samples were obtained from four archaeological sites representing different time periods and estimated preservation states (Table 1). Calculus samples from Driffield Terrace and Wighill had been previously analyzed and shown to have good protein preservation (Hendy et al., 2018; Mackie et al., 2017; Warinner et al., 2014a), while calculus samples from Rupert's Valley had been previously shown to have poor DNA preservation and a high contamination burden from environmental bacteria (Ziesemer et al., 2015). Calculus from San Martín de Dulantzi had not been previously analyzed using molecular techniques. In total, dental calculus from six individuals was analyzed in this study: one individual each from Driffield Terrace, Wighill and Rupert's Valley, and three individuals from San Martín de Dulantzi.

2.2. Laboratory procedures

All DNA and protein extractions were performed in dedicated

Table 1
Dental calculus samples analyzed in this study.

Sample ID	Archaeological ID	Site	Age	Preservation
Yorkshire, England				
DRT001	3DT21	Driffield Terrace	44-410 CE	Good
WIG001	WG1561	Wighill	1000–1550 CE	Good
Basque Country, Spain				
SMD017	SMD 1441-1-1440(044)	San Martín de Dulantzi	700–950 CE	Unknown
SMD046	SMD 2591-1-2590(159)	San Martín de Dulantzi	700–950 CE	Unknown
SMD051	SMD 2711-1-2710(171)	San Martín de Dulantzi	700–1150 CE	Unknown
St. Helena, South Atlantic Ocean				
RUV001	SK203	Rupert's Valley	1840–1872 CE	Poor

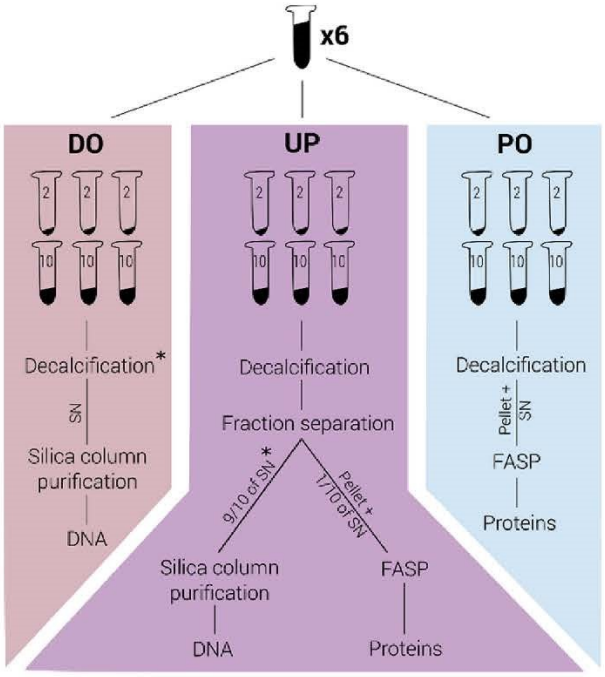


Fig. 1. Flowchart of extraction protocols and experiments. Amount of starting material is in milligrams. The asterisk denotes addition of proteinase K and 'SN' indicates supernatant.

archaeogenetic and palaeoproteomic facilities at the Max Planck Institute for the Science of Human History (Jena, Germany). DNA-only (DO) and protein-only (PO) extractions were performed alongside the unified (UP) extractions, and sample batches were randomized to avoid batch effects. For all extractions, two sample starting weights were used, 2 mg and 10 mg, in order to investigate the efficiency of the protocols on different amounts of starting material. Extractions for each sample, protocol, and starting weight were performed in triplicate (Fig. 1), resulting in a total of 72 DNA extracts and 72 protein extracts.

2.2.1. Sample collection and preparation

Dental calculus was collected from archaeological teeth using a sterile dental scaler and UV irradiated for 2 min to reduce potential surface contamination. Calculus from each individual was homogenized for 5 s at 20 Hz in an oscillating mill (MM200, Retsch) using 15 mm zirconium oxide beads. The resulting powder was subsampled for use in all subsequent extractions. Only nitrile gloves were used during sample processing, as latex gloves may introduce latex proteins into the samples.

2.2.2. DNA-only (DO) protocol

DNA extraction was performed following Dabney et al. (2013), with modifications adapted for dental calculus (Mann et al., 2018). Briefly, the powdered dental calculus was washed with 1 ml of 0.5 M EDTA to remove surface contaminants, and the supernatant was removed without incubation. The remaining dental calculus was decalcified in 1 ml of 0.5 M EDTA on a rotator at room temperature for 24 h, after which 50 μl of 10 mg/ml proteinase K (Sigma-Aldrich) was added. The samples were further decalcified for 48 h by rotation at room temperature (total decalcification time 72 h), followed by centrifugation at 18400 rcf for 1 min to pellet cell debris. The complete supernatant was then removed and mixed with 10 ml binding buffer (5 M guanidine hydrochloride, 0.12 M sodium acetate and 40% isopropanol). DNA purification was performed using the High Pure Viral Nucleic Acid kit (Roche Life Science) following the manufacturer's instructions. DNA was eluted from the column in two rounds of 50 μl of Qiagen EB buffer, to which Tween 20 had been added to a final concentration of 0.05%. DNA quantification was performed on 1 μl of the eluate with a Qubit HS assay (Thermo Fisher Scientific). DNA extractions were performed in batches of 3–12 samples, and one extraction blank was included per batch.

2.2.3. Protein-only (PO) protocol

Protein extraction and digestion was performed using a filter-aided sample preparation (FASP) protocol (Wiśniewski et al., 2009) modified for ancient dental calculus (Jeong et al., 2018). Briefly, the powdered dental calculus was decontaminated as described in section 2.2.2., and then decalcified in 1 ml of 0.5 M EDTA on a rotator at room temperature for 72 h. After decalcification, the samples were centrifuged at 18400 rcf for 1 min to pellet cellular material, and 200 μl of supernatant was transferred to a 30 kDa Microcon filter unit (Merck) containing 50 μl of 8 M urea. The mixture was centrifuged through the filter at 14000 rcf for 10 min. This procedure was repeated until all supernatant had been passed through the filter. The pellet was resuspended in 30 μl of lysis buffer (containing 4% w/v sodium dodecyl sulfate, 100 mM Tris hydrochloride and 0.1 M dithiothreitol) and incubated at 95 °C for 5 min, followed by centrifugation at 14000 rcf for 1 min to pellet cell debris. The supernatant was then mixed with 200 μl of 8 M urea, transferred to the corresponding Microcon filter unit used in the previous step, and centrifuged at 14000 rcf for 20 min. The filter was washed with 200 μl of 8 M urea in an additional centrifugation step. On the filter, proteins were alkylated by adding 100 μl of 0.5 M iodoacetamide solution and incubating for 5 min at room temperature in the dark. The filter units were then centrifuged at 14000 rcf for 12 min, and washed twice with 100 μl of 8 M urea and twice with 100 μl of 0.5 M NaCl. The proteins were digested overnight at 37 °C on the filter in a digestion buffer consisting of 117 μl of 0.05 M triethylammoniumbicarbonate and 3 μl of 0.4 μg/μl porcine trypsin (Pierce Trypsin Protease MS Grade, Thermo Fisher Scientific). Digested peptides were recovered from the filter by centrifugation at 14000 rcf, and acidified by adding 5% trifluoroacetic acid (TFA) to get a final concentration of 0.5% TFA. The acidified peptides were desalted using StageTips (C18 tips, Thermo Fisher Scientific) following the manufacturer's instructions, and dried to completion in a vacuum centrifuge (Martin Christ RVC 2-18). Protein quantification was performed by rehydrating the peptides in 20 μl of 3% acetonitrile/0.1% TFA and measuring the absorbance at 215 nm with a spectrophotometer (DeNovix DS-11 FX+). Afterwards, the samples were dried to completion in a vacuum centrifuge and stored at −80 °C until further analysis. Protein extraction and digestion was performed in batches of 12–24 samples, and one extraction blank was included per batch.

2.2.4. Unified protocol (UP)

Powdered dental calculus samples were decontaminated as described in section 2.2.2., and then decalcified in 1 ml of 0.5 M EDTA on a rotator at room temperature for 72 h. After decalcification, the samples were centrifuged at 18400 rcf for 1 min to pellet cellular material, and 900 μl of supernatant was transferred into a new tube containing 50 μl of 10 mg/ml proteinase K (Sigma-Aldrich) and incubated overnight at room temperature. To this, 10 ml binding buffer (5 M guanidine hydrochloride, 0.12 M sodium acetate and 40% isopropanol) was added, and DNA was purified and quantified using the method described in section 2.2.2. Proteins were extracted, digested and quantified from the remaining 100 μl supernatant and cell debris pellet using the modified FASP protocol as described in section 2.2.3. Unified extractions were performed in batches of 11–13 samples, and one extraction blank was included per batch.

2.2.5. DNA sequencing

One randomly selected DNA extract per individual per protocol (DO2, DO10, UP2 and UP10; 24 in total), as well as blanks from

extractions and library preparations (9 in total), were treated with partial uracil-DNA glycosylase treatment (Rohland et al., 2015) and prepared into double-stranded DNA libraries with dual indexing following published protocols (Kircher et al., 2012; Meyer and Kircher, 2010) using 20–25 μl extract each. The libraries were normalized and pooled in equimolar amounts and sequenced on an Illumina NextSeq with 75-bp paired-end chemistry to a depth of approximately 10 million reads for calculus samples and 2 million reads for blanks.

2.2.6. LC-MS/MS

Liquid chromatography tandem mass spectrometry (LC-MS/MS) was performed on one randomly selected protein digest per individual per protocol (PO2, PO10, UP2 and UP10; 24 in total), as well as extraction blanks (3 in total), at the Functional Genomics Centre Zürich (University of Zürich, Switzerland). The samples were analyzed on a Q-Exactive HF mass spectrometer (Thermo Fisher Scientific) coupled to an ACQUITY UPLC M-Class system (Waters AG), following the procedures and parameters described in Jeong et al. (2018).

2.3. Analysis workflow

Statistical analysis was performed in R v. 3.6.1 (R Core Team, 2018). Differences between protocols were tested with pairwise Wilcoxon tests and corrected for multiple testing using the Benjamini-Hochberg method (hereafter called pW-BH), unless otherwise noted. General R-packages used for data manipulation and creating graphical figures were tidyverse v. 1.2.1 (Wickham, 2017), janitor v. 1.2.0 (Firke, 2018), ggpubr v. 0.2.3 (Kassambara, 2018), gtools v. 3.8.1 (Warnes et al., 2018), ggsignif v. 0.6.0 (Ahlmann-Eltze, 2019), cowplot v. 1.1.0 (Wilke, 2017), rcartocolor v. 2.0.0 (Nowosad, 2018) and fuzzyjoin v. 0.1.4 (Robinson, 2018). Individual RUV001 was excluded from all statistical tests after it was found to exhibit a non-typical DNA and protein composition for calculus (indicating very poor preservation), however, results for this individual are shown in figures. R Markdown files for all analyses are available on GitHub (https://github.com/ZandraFagernas/unified_protocol).

2.3.1. Genetic analysis

2.3.1.1. DNA recovery. DNA yield was normalized by starting weight of dental calculus. The values were log-transformed, and linear mixed-effect models with the individual as the random effect were fitted to find the model that best predicts DNA recovery. All mixed effects models were fit using the R-package lme4 v. 1.1.21 (Bates et al., 2015), model selection tests via ANOVA were performed using the R-package lmerTest v. 3.1.0 (Kuznetsova et al., 2017) and Box-Cox transformations were identified using the R-package MASS (Venables and Ripley, 2002).

2.3.1.2. Data preprocessing. EAGER v.1.92.55 (Peltzer et al., 2016) was used to process the raw DNA reads and align sequences to the human reference genome (HG19). In this pipeline, adapter removal and read merging were performed by AdapterRemoval v. 2.2.0 (Schubert et al., 2016). On average, 87.3 ± 7.8% (mean ± standard deviation) of calculus reads and 36.8 ± 23.6% of blank reads merged. BWA v. 0.7.12 (Li and Durbin, 2009) was used for human genome alignment with default settings (-l 32, -n 0.01) without quality filtering. Unmapped reads were extracted with SAMtools v. 1.3 (Li et al., 2009) for downstream processing, and duplicates removed using DeDup v. 0.12.2 (Peltzer et al., 2016). Non-human reads were aligned to the NCBI nucleotide database (as of April 2016) using MALT v. 0.3.8 (Herbig et al., 2016) to assign taxonomy (settings –minPercentIdentity 85.0, –topPercent 1.0, –minSupportPercent 0.01). MEGAN v. 6.11.1 (Huson et al., 2016) was used to export genus- and species-level OTU-tables with summarized read counts from the MALT results.

2.3.1.3. Fragment characteristics. For calculations of GC-content and fragment length, adapter-trimmed, merged reads were aligned to *Actinomyces oris* strain T14V (NCBI assembly accession ASM155393v1) and *Leptotrichia buccalis* strain DSM 1135 (NCBI assembly accession ASM2390v1) using the EAGER pipeline described above, with the exception of mapping quality being set to 37. These oral bacterial species were chosen due to their high abundance in the samples, as well as their differential median GC-content (*A. oris* 68.3%, *L. buccalis* 29.6%).

2.3.1.4. Microbiome reconstruction. The R-package decontam v. 1.1.2 (Davis et al., 2018) was used to identify putative contaminants from genus- and species-level OTU tables. The prevalence method was used (with default threshold = 0.1) to identify OTUs with a higher prevalence in blanks than in samples, thereby likely being contaminants. A species-level OTU-table, with putative contaminants removed, was used to investigate the taxonomic profiles of the samples. The 20 most abundant species across all samples were calculated excluding individual RUV001, and the abundance of these species was compared across extraction protocols. The body site these species are associated with was extracted from the expanded Human Oral Microbiome Database (eHOMD; Escapa et al., 2018). A principal component analysis was conducted on a genus-level OTU-table, after multiplicative zero replacement using the R-package zCompositions v. 1.3.2.1 (Palarea-Albaladejo and Martín-Fernández, 2015) and CLR-transformation (Gloor et al., 2017). The two loadings with highest contribution to the separation along each PC in both the negative and positive directions were extracted and added to the PCA plots. A PERMANOVA (Anderson, 2001) was performed on the dataset without RUV001 using the 'adonis' function (on Euclidean distances with 999 permutations) from the R-package vegan v. 2.5–6 (Oksanen et al., 2019), in order to assess drivers of variation in microbiome composition.

2.3.1.5. Contamination. The species that were identified as putative contaminants (described in section 2.3.1.4) were further analyzed to evaluate if the proportion and composition of contaminants differed between extraction protocols.

2.3.2. Proteomic analysis

2.3.2.1. Protein recovery. Protein yield was normalized by starting weight of dental calculus. The values were log-transformed, and linear mixed-effect models were fitted in order to find the model that best predicts protein recovery, as described in section 2.3.1.1 .

2.3.2.2. Data preprocessing. MSConvert v. 3.0.11781 (ProteoWizard) was used to transform raw data files (.raw) to Mascot generic files (.mgf), using the 100 most intense peaks. The resulting files were searched using Mascot v. 2.6.0 (Matrix Science) against the Swiss-Prot database (as of January 2018) plus reversed decoys (in total 1.1 million entries). Fragment ion mass tolerance was set to 0.01 Da and parent ion tolerance to 10.0 ppm. Carbamidomethylation C was set as a fixed modification, and deamidation (N and Q) and oxidation (M and P) as variable modifications (Hendy et al., 2018; Jeong et al., 2018). Using Scaffold v. 4.8.9 (Proteome Software Inc.), the results were filtered to a 1% peptide false discovery rate (FDR), a 5% protein FDR, and a minimum support of two peptides. Decoy hits and common laboratory contaminants (collagen, keratin and serum albumin) were removed before subsequent analyses. The resulting dataset was exported from Scaffold at the levels of peptides, proteins, and protein clusters, to be used in downstream analyses.

2.3.2.3. Protein characteristics. Hydropathy values (a measure of hydrophobicity) were calculated for recovered proteins using the web application GRAVY Calculator (www.gravy-calculator.de/). Molecular weights of all identified proteins were obtained from the protein report file exported from Scaffold. Physicochemical properties of amino acids

were calculated using the R-package Peptides v. 2.4 (Osorio, 2015) from all unique identified peptides.

2.3.2.4. Proteome reconstruction. Scaffold was used to annotate protein clusters with gene ontology (GO) terms from NCBI (as of May 2019). Protein clusters were selected because this groups proteins by homology and functional similarity, thereby mitigating some of the uncertainty introduced by assigning proteins to species. The GO terms for cellular location were classified into three groups: 1) intracellular and organelle, 2) membranes, and 3) extracellular. Protein clusters with GO terms for more than one of these groups were classified as: 4) Various; clusters without GO terms for cellular location available were classified as: 5) Unknown. The top 20 most abundant protein clusters were calculated excluding RUV001. A table of the percentage of total spectra belonging to each protein cluster per individual was used as input for a PCA. Only protein clusters present in at least two of the samples were included, and the data was log2-transformed after pseudocount zero replacement (+1). The two loadings with highest contribution to separation along each PC in positive and negative directions were added to the PCA plots. Drivers of variation were assessed by performing a PERMANOVA (Anderson, 2001) using the 'adonis' function (on Euclidean distances with 999 permutations) from the R-package vegan v. 2.5–6, (Oksanen et al., 2019).

2.3.2.5. Contamination. Protein extraction blanks contained only four proteins in total: trypsin (TRYP_PIG), serum albumin (ALBU_BOVIN), collagen alpha-1 (I) chain (CO1A1), and collagen alpha-2 (I) chain (CO1A2). Porcine trypsin is a laboratory reagent used during protein digestion, but was not identified in any of the calculus samples. Bovine serum albumin is a common laboratory contaminant (www.thegpm.org/crap/), and therefore serum albumin was classified as a contaminant regardless of assigned species, although some of it is likely endogenous. While some collagen recovered from dental calculus is also likely endogenous (released from the gingiva and periodontium bone during inflammation), its frequent recovery from extraction blanks led us to classifying it as a contaminant for the purposes of this study. Keratins, although not identified in our blanks, are common laboratory contaminants (www.thegpm.org/crap/) and were classified as contaminants. The proportion of spectra in the calculus samples belonging to these proteins was compared across extraction protocols.

3. Results

3.1. Genetic analysis

3.1.1. DNA recovery

To investigate how the DO and UP protocols differ in performance for DNA extraction, we compared normalized DNA yield for both protocols (Data S1). Only choice of extraction method, not starting material mass, was found to be a significant predictor of DNA recovery, with UP having lower DNA recovery ($p < 0.001$). On average, DO10 yielded 1.8 ± 0.2 (mean ± standard deviation) fold more DNA per mg calculus than UP10, and DO2 yielded 1.7 ± 0.6 fold more than UP2 (Fig. 2). The largest decrease in yield was seen in WIG001, the sample with best estimated preservation. The poorly preserved individual RUV001 showed the opposite trend, with UP10 yielding 1.2 fold more DNA than DO10, and UP2 yielding 1.4 fold more DNA than DO2. However, RUV001 had very low DNA recovery overall, and the absolute difference in DNA recovery between the protocols was negligible (only 0.7 ng/mg on average).

3.1.2. DNA fragment characteristics

Given that there are differences in DNA recovery based on protocol, we next tested whether there are differences in the characteristics of the DNA fragments recovered by each protocol. Previous studies have shown that DNA extracted from ancient dental calculus may yield biases in microbial taxonomic profiles due to selective loss of short, AT-rich sequences (Mann et al., 2018). We calculated the mean fragment length and GC-content per sequence library from alignments to *Actinomyces oris* and *Leptotrichia buccalis* (Fig. 3). The only significant shifts observed for GC-content and fragment length were found in *A. oris*, where GC-content is significantly different between DO2 and DO10, as well as between UP2 and DO10 (pW-BH, $p < 0.05$ in both cases). In addition, a considerable shift in median fragment length can be seen among reads from our selected bacterial species, where *L. buccalis* shows median fragment lengths of 43–51 bp, whereas *A. oris* has median fragment lengths of 46–75 bp (with RUV001 excluded). The average GC-content of *L. buccalis* reads, 31.5%–34.4%, is also higher than the genome average of 29.6% (dashed grey line in Fig. 3A), which all agrees with the previously-reported loss of short, AT-rich sequences (Mann et al., 2018).

3.1.3. Microbiome reconstruction

Ancient dental calculus has been found to be dominated by oral

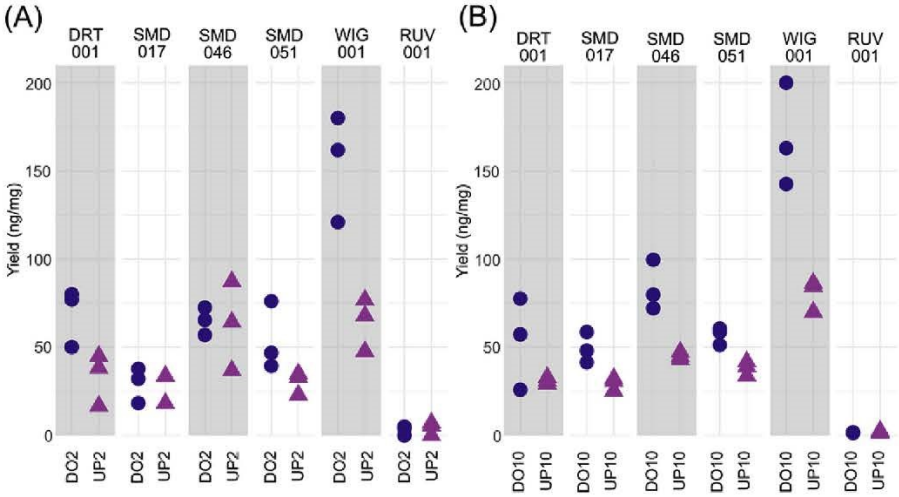


Fig. 2. DNA yield differs significantly by extraction protocol but not starting mass of dental calculus. Normalized DNA yield for triplicate extracts of (A) DO2 and UP2, and (B) DO10 and UP10, ordered by individual from oldest to most recent.

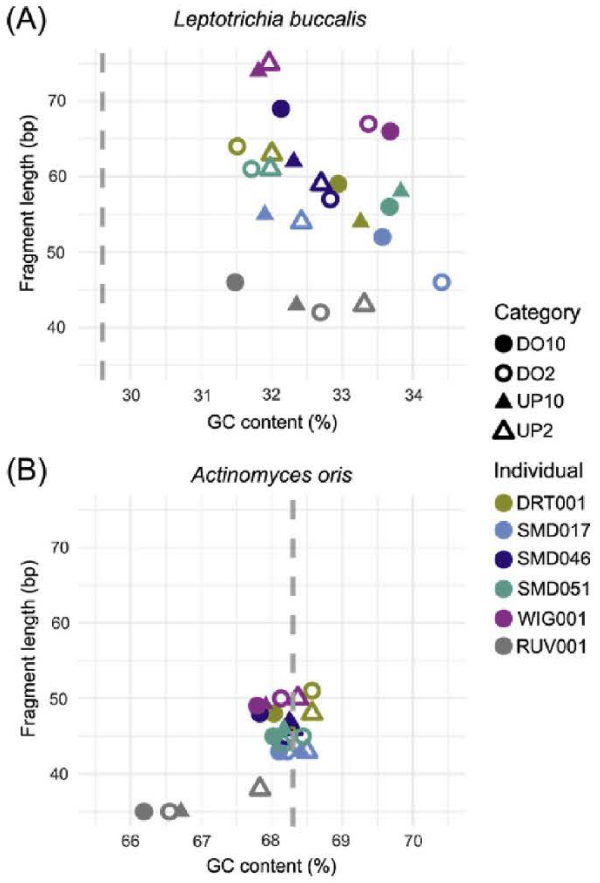


Fig. 3. Median fragment length by GC-content for (A) *L. buccalis* and (B) *A. oris* show no consistent changes by choice of extraction protocol. Dashed grey line shows expected median GC-content of genome.

bacteria from the phyla Firmicutes, Actinobacteria, Proteobacteria, and Bacteroidetes (Warinner et al., 2014b), which accords with the known composition of dental plaque biofilms today (Dewhirst et al., 2010). The DNA sequences obtained from dental calculus in this study are overwhelmingly bacterial in origin (Data S2), and these four phyla account for 83.5% of the identified bacterial species (excluding RUV001). Dental calculus from the poorly preserved individual RUV001 contains considerably more identified species (641 ± 19 species) than the other individuals (181 ± 25 species), which is consistent with RUV001 being a highly degraded sample infiltrated by diverse environmental taxa (Fig. 4A). There are no significant differences in the number of species identified between the different extraction protocols or starting weights, nor in the proportions of the four most abundant phyla (pW-BH, $p > 0.05$ in all cases). For the 20 most abundant species, the majority of which were oral species, no significant taxonomic differences were identified between the extraction protocols (pW-BH, $p > 0.05$ in all cases); rather taxonomic differences are primarily driven by individual (Fig. 4B).

A principal component analysis on genus level read counts shows no systematic shifts by extraction protocol or starting weight (Fig. 4C). The samples cluster by individual (PERMANOVA, $p < 0.001$), and the genera separating the samples are mostly environmental. This indicates that the main drivers of separation are bacteria from the burial environment. When the poorly preserved individual (RUV001) is added to the PCA (Fig. S1), it is separated from the other individuals along PC1, which also separates oral from environmental genera as the major loadings, indicating that this individual has a distinct microbial composition compared to the well-preserved samples.

3.1.4. Contamination

Contamination is a major issue for studies of ancient biomolecules, and laboratory procedures can influence contamination levels (Key et al., 2017; Salter et al., 2014; Warinner et al., 2017). We therefore assessed whether the choice of extraction protocol affects the level of estimated contamination observed for the samples. Using the R-package decontam (Davis et al., 2018), a total of 154 species were identified as putative contaminants (Data S3). Of these, the majority belong to Bacteria (n = 95) and Animalia (n = 33). Excluding RUV001, the proportion of putative contaminants averaged 0.05 ± 0.03%, ranging from 0.01% to 0.13% of total reads per extract (Fig. S2), and did not significantly differ between extraction protocols or starting weights (pW-BH, $p > 0.05$). Contamination in the poorly preserved individual, RUV001, was >100-fold higher (9.39 ± 0.36%) than in the other samples, a finding consistent with previous analysis of calculus from this site (Ziesemer et al., 2015).

3.2. Proteomic analysis

3.2.1. Protein recovery

To evaluate UP performance, we compared total protein recovery between the PO and UP protocols (Fig. 5; Data S1). The best fitting model to predict protein recovery was found to be an interaction between starting weight and protocol ($p < 0.001$). When normalized for input weight, UP2 extractions yielded 2.0 ± 0.3 fold more protein than UP10 extractions, and PO2 extractions yielded 3.2 ± 0.4 fold more protein than PO10 extractions (although in both cases absolute recovery was higher for 10 mg extractions). This pattern suggests that the FASP protocol used in this study may be limiting protein recovery efficiency at higher amounts of starting material. For the 2 mg extractions PO2 yielded 1.3 ± 0.1 fold more proteins than UP2, whereas for the 10 mg extractions the opposite pattern was observed, where UP10 yielded 1.3 ± 0.1 fold more proteins than PO10. The increase in normalized protein recovery through UP10 over PO10 may also be caused by the limitation of the amount of input material in FASP, as UP10 has a lower amount of input material (90% of the supernatant is excluded from the protein extraction). However, FASP is known to have relatively low consistency between technical replicates (Sielaff et al., 2017), and this variation may have influenced our results. In addition, protein concentration measurement may be less accurate at lower protein inputs. Further work is needed to determine whether FASP is indeed more efficient with lower input amounts of dental calculus, and to what degree other factors may be contributing to our observed pattern.

3.2.2. Protein characteristics

In the UP, only 10% of the supernatant is reserved for protein extraction. Hydrophilic proteins, which we expect to be mainly present in the supernatant, could thereby be lost to the DNA extraction fraction. Ideally, a unified protocol should not have a skewed representation of hydrophobic and/or hydrophilic proteins compared to a protein-only extraction method. To test for this potential bias, we compared the hydropathy values of proteins and physicochemical properties of each amino acid of peptides recovered using the UP and PO protocols (Data S4). The mean protein hydropathy values did not significantly differ between the two extraction protocols (pW-BH, $p > 0.05$; Fig. 6A). However, significant differences in amino acid properties were found with respect to basic amino acids (Fig. 6B), which are hydrophilic. The proportion of basic amino acids was significantly different between UP10 and PO10, UP10 and PO2, UP2 and PO10, and UP2 and PO2 (pW-BH, $p < 0.05$ in all cases), with the PO protocol consistently recovering a higher proportion of basic amino acids. Thus, the UP may result in an underrepresentation of hydrophilic proteins, particularly those containing basic residues. Further, the proportion of acidic amino acids (Fig. S3) was significantly higher in UP2 than PO10 (pW-BH, $p < 0.05$).

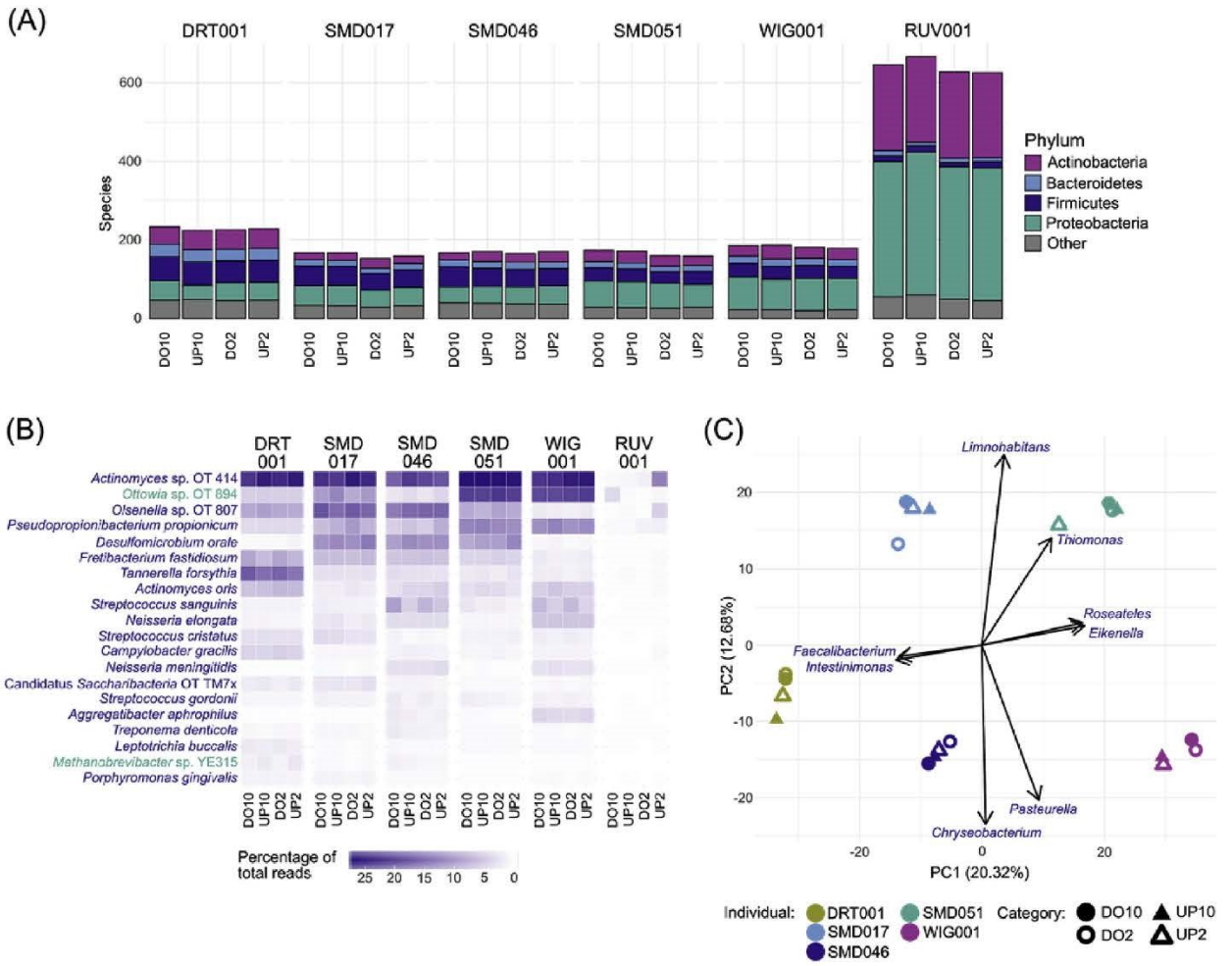


Fig. 4. Microbiome profiles are consistent between extraction protocols. (A) Number of species identified in each sample, grouped by phylum. Phyla containing <5% of total species are grouped under "Other". (B) Heatmap of percentage of total assigned reads for the 20 most abundant species, with species names colored according to if they are oral (blue) or unassigned (green) in eHOMD. OT oral taxon. (C) PCA on genus level taxa, without individual RUV001. The top genera contributing to separation for each principal component as major loadings are indicated by arrows. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

Finally, we tested whether the UP exhibited biases based on protein size. We found that the mean molecular weight of recovered proteins was not significantly affected by choice of extraction protocol or starting weight (pW-BH, $p > 0.05$; Fig. S4).

3.2.3. *Proteome reconstruction*

The number of identified protein clusters (Data S5) did not significantly differ by extraction protocol or starting weight (pW-BH, $p > 0.05$; Fig. 7A). The highest number of identified protein clusters (67.5 ± 9.3) was observed for WIG001, while the lowest (17.8 ± 4.3) was observed for RUV001, which is consistent with the differing preservation of these two samples. Among the identified protein clusters, most are found in various cellular locations, followed by intracellular/organelle and extracellular proteins clusters; only a small number of membrane protein clusters were identified. The choice of extraction protocol did not significantly affect the proportions of these groups (pW-BH, $p > 0.05$ in all cases). Of the 20 most abundant protein clusters, most are associated with host defenses against microbes or bacterial cellular processes (Fig. 7B). This is consistent with previous findings for modern and well-preserved ancient dental calculus (Jersie-Christensen et al., 2018; Velsko et al., 2019; Warinner et al., 2014b). Although some protein clusters seem to differ in abundance by extraction method by visual inspection, statistical significance is not reached with this design, due to the need to correct p-values for multiple testing and for the abundances being dependent.

In a principal component analysis of identified protein clusters, the samples were found to cluster by individual and extraction method (PERMANOVA, $p < 0.001$ for both), but not starting weight (Fig. 7C). The major protein cluster separating extraction protocols along PC1 is Elongation factor Tu (EFTU_NEIMA), which is involved in protein biosynthesis. It is a hydrophobic protein (hydropathy index −0.136) and would thereby be found at a higher proportion in the UP, where hydrophilic proteins are lost to the aqueous fraction. This pattern is also present in the abundance of this protein cluster (Fig. 7B). Adding RUV001 to the PCA did not considerably alter the relationships between the samples (Fig. S5).

3.2.4. *Contamination*

Of the total identified spectra per sample, 0.39–6.11% were assigned to collagen, keratin and serum albumin proteins, which were classified

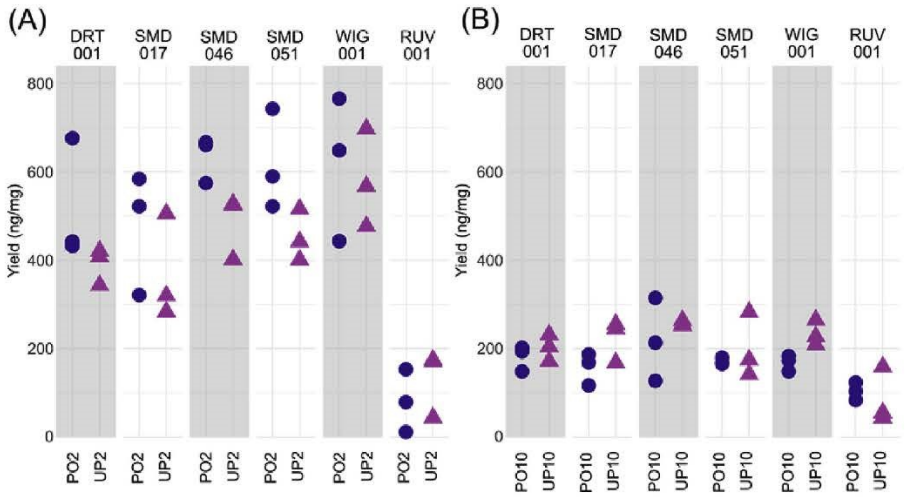


Fig. 5. Protein yield differs by starting material mass and extraction method. Protein yield normalized by sample mass for triplicate extracts of (A) PO2 and UP2, and (B) PO10 and UP10, ordered by individual from oldest to most recent.

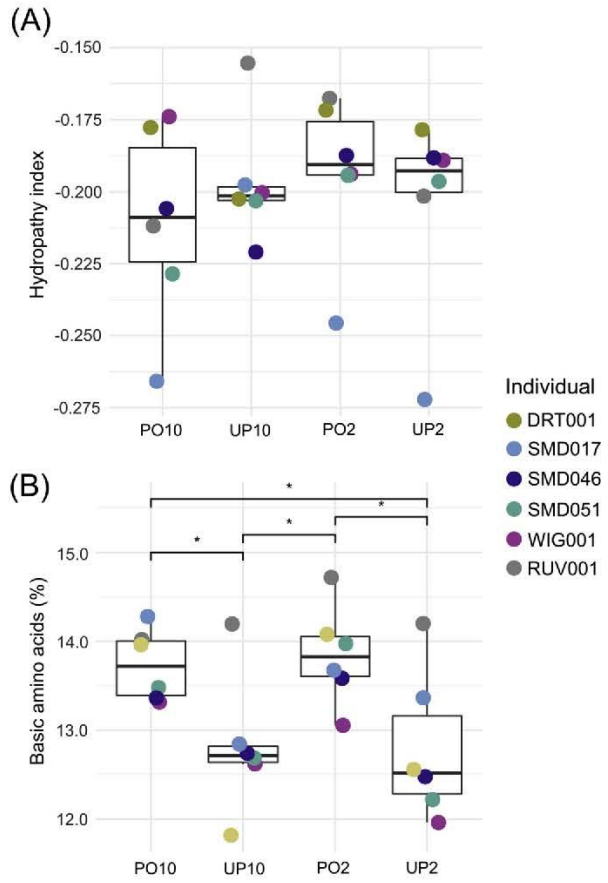


Fig. 6. Protein hydropathy is unaffected by extraction protocol, but basic amino acids may be lost in UP. (A) Hydropathy index of identified proteins. (B) Proportion of basic amino acids among recovered peptides, with significant pairwise tests indicated by brackets (*$p < 0.05$; **$p < 0.01$; ***$p < 0.001$).

as probable contaminants (Fig. S6). There was no correlation between contaminant spectra and extraction protocol or starting weight, neither overall nor in the proportions of the different contaminants (pW-BH, $p > 0.05$ in all cases).

4. Discussion and conclusions

Protocols that allow for the simultaneous extraction of different biomolecules, such as DNA and proteins, can reduce sampling demands for archaeological remains, thus supporting their long-term curation and preservation. For a unified protocol to be an acceptable alternative to simply performing two separate extractions, the unified protocol should recover at least 50% of the biomolecules that each individual protocol recovers separately. Otherwise, splitting the sample in half and performing two separate extractions will have a higher overall yield. Further, the protocol should introduce no major biases, and contamination levels should not be altered. Careful testing of protocols is therefore necessary to evaluate if all the above criteria are met.

In this study, we developed a unified protocol for the simultaneous extraction of DNA and proteins, and compared its performance to separate DNA and protein extraction protocols. During the unified protocol, archaeological dental calculus is decalcified in EDTA, and different fractions of this solution are used to recover DNA and proteins, based on the expected solubility of these biomolecules. Because ancient DNA is highly fragmented, and the sugar-phosphate backbone of DNA is negatively charged and hydrophilic, we expect the majority of DNA molecules to be present in the aqueous fraction. Proteins, on the other hand, are biochemically complex and may be very large (e.g. collagen), membrane-bound, or otherwise have reduced solubility in their native (non-denatured) state. Consequently, we expect them to mostly pellet out of solution during centrifugation of the decalcified sample. Some proteins are, however, highly hydrophilic, and therefore we included a portion (10%) of the aqueous fraction in our protein extraction. Following this initial partitioning of DNA into the aqueous supernatant and proteins into the pellet (plus a small portion of the supernatant), we proceeded to isolate each biomolecule using optimized methods.

Overall, we found that a significant amount of DNA (mean 43%) is lost in the unified protocol, presumably to the pellet, and the effect is greatest for well-preserved samples. One possible explanation for this pattern is that proteinase K is not added to the sample until after the fractions have been separated in the unified protocol, and a substantial portion of DNA may be bound to insoluble proteins or trapped within

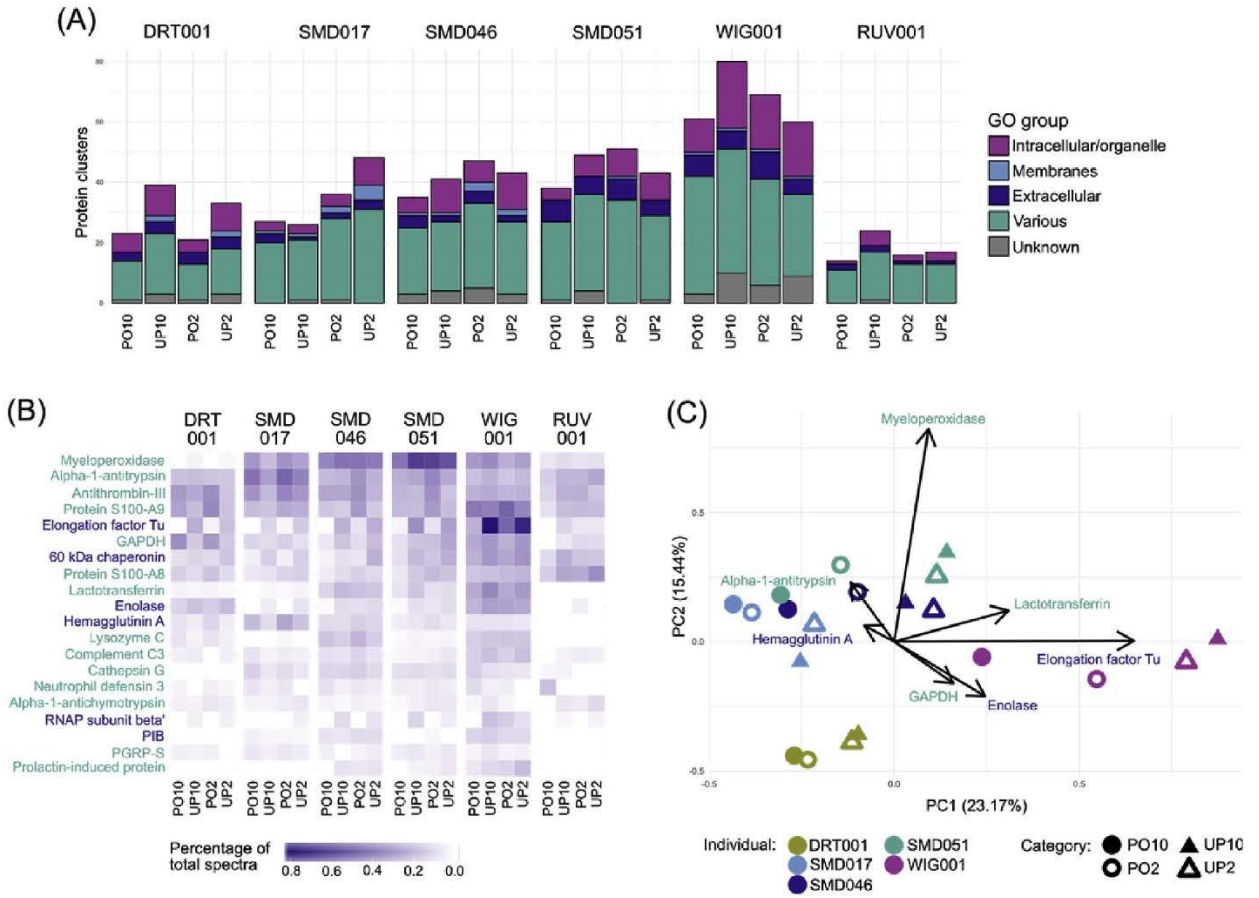


Fig. 7. Choice of extraction protocol does not induce major biases in proteome reconstruction. (A) Number of identified protein clusters per extraction protocol, colored by GO cellular location. (B) Percentage of spectra assigned to the top 20 most abundant protein clusters per sample. (C) PCA of protein clusters without RUV001, where the top protein clusters contributing to separation along each principal component as major loadings are indicated by arrows. Protein cluster names in (B) and (C) are colored according to: Blue prokaryotes, green eukaryotes. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

cellular structures that are pelleted during centrifugation. In the DNA-only protocol, the entire sample is exposed to proteinase K, which breaks up proteins and cellular membranes, thereby facilitating greater DNA release. Unfortunately, this is not a problem that is easily overcome. Proteinase K cannot be added to the entire sample in the unified protocol, as this enzyme produces peptides too small for successful analysis using LC-MS/MS. On the other hand, using an alternative enzyme compatible with LC-MS/MS, such as trypsin, is not feasible due to the risk of its autodigestion when left active for an extended period of time, such as during calculus decalcification; such autodigestion can result in off-target cleavages performed by the enzyme, which reduces downstream protein identifications. Trypsin digestion during decalcification would also preclude the use of FASP for protein extraction, as this protocol relies on size exclusion to remove EDTA and other lysis components that would interfere with downstream analysis. For protein extractions, acids can be used in place of EDTA for decalcification, which may assist with pellet digestion, but even weak acids can cause DNA to undergo hydrolysis and depurination, thus making acids unsuitable for genetic analysis.

Although the unified protocol does lead to reduced DNA recovery, it is still more efficient than performing two separate biomolecular extractions. In addition, the DNA losses are mitigated by the fact that the starting amount of DNA within archaeological dental calculus is typically very high (however see Austin et al. (2019) for exceptions). For example, in a study of paired dental calculus and dentine samples, Mann et al. (2018) reported a median DNA recovery from calculus of 72.1 ng/mg compared to 4.8 ng/mg for dentin. However, given the significantly reduced DNA recovery through the unified protocol, it is advisable to keep some material aside for a potential future DNA-only extraction, especially in cases where preservation status is unknown. Beyond reducing DNA recovery, we found that the unified protocol does not affect downstream genetic analyses or change the amount of identified contaminants. Although we sequenced only one of the three extraction replicates, thereby reducing our statistical power, the results are consistent over multiple individuals, and thus appear generalizable.

Protein recovery using the unified protocol was significantly different from the protein-only protocol, but recovery was only decreased by 20.6% on average for the 2 mg extractions, confirming that the majority of proteins are present in the initial cell debris pellet. For the 10 mg extractions, on the other hand, protein recovery was surprisingly increased (mean 25.3%) through the unified protocol. Equally unexpectedly, protein recovery (normalized by weight of input material) was found to be higher when using 2 mg of starting material than when using 10 mg, a trend that was not reflected in DNA extractions. This suggests that the version of the FASP protocol used for protein purification may be limiting protein yield at higher amounts of input material. This also explains the increase in recovery through the unified protocol when using 10 mg starting material, as the amount of input

material is lower in the unified protocol (where only 10% of the supernatant is used for protein extraction) than in the protein-only protocol. However, it is not clear from our results what the limiting factor in the FASP protocol could be, and overall, we observed a high variance in protein recovery. FASP has a relatively high variability between technical replicates (Sielaff et al., 2017), which may cause the variance in our results. Further, spectrophotometric measurements of peptide bonds at 215 nm absorbance can have a high error rate if other organic compounds have not been sufficiently removed, which is possible here given the complex nature of dental calculus and the variety of chemicals used in FASP. Aside from protein recovery, there were no observable differences in downstream analyses between the two starting material weights.

In addition to the difference in protein recovery between the protocols, a significantly lower proportion of basic amino acids among downstream identified peptides was observed through the unified protocol, suggesting that the unified protocol may have a slight bias against hydrophilic proteins. The bias is likely minor, as no difference in overall hydropathy of proteins was observed. However, if a specific, hydrophilic protein is the target of a research study, further testing is recommended before implementing the unified protocol.

Although most dental calculus samples in this study exhibited good biomolecular preservation, the 19th century Rupert's Valley dental calculus was known to be poorly preserved, and it performed differently than the other samples in most analyses. It was previously reported to have poor DNA preservation and high contamination by environmental bacteria (Ziesemer et al., 2015), which we likewise observed in our genetic analyses. However, our protein analyses did not show high levels of environmental bacteria. Instead, fewer protein clusters were identified than in other samples, likely due to a lack of sufficient representation of environmental bacteria in the protein database. It is therefore important to note that choice of database may affect assessment of sample preservation.

Moving forward, it may be possible to combine the unified protocol with palaeoethnobotanical and other microparticle analyses. Recently, it was shown that microparticle recovery from dental calculus following EDTA decalcification is both possible and more successful compared to conventional decalcification using HCl (Tromp et al., 2017). During the unified protocol, a visible pellet of insoluble cell debris often remains after DNA and protein extraction, and recovery of microparticles may be possible, especially for phytoliths, which are not damaged by the heating step during protein extraction. Extending the unified protocol to also incorporate microremains will further broaden the range of information that can be obtained from a single sample of dental calculus, while further decreasing destructive demands on this valuable archaeological resource.

Accession numbers

Genetic data have been deposited in the ENA under the accession PRJEB35483. Raw and processed LC-MS/MS data files have been deposited in the ProteomeXchange Consortium via the PRIDE partner repository under the identifier PXD015817.

Author contributions

Z.F., C.S., J.H., and C.W. designed the study. C.S. and J.H. conceived of the unified protocol. Z.F. performed the experiments. Z.F. analyzed the data. M.I.G.C., C.S., J.H., and C.A.H. provided materials and resources. Z.F. wrote the manuscript with input from C.W., I.V. and the other authors.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This work was supported by the US National Science Foundation (BCS-1516633, BCS-1528698 to C.W., and BCS-1643318 to C.W. and C. H.), the Max Planck Society, and the Government of the Basque Country (IT1193-19). The authors thank Franziska Aron, Raphaela Stahl, and Richard Hagan for laboratory assistance, and James Fellows Yates, A. Ben Rohrlach, Alexander Hübner, and Ashley Scott for assistance with data analysis. The authors also thank the personnel at the Functional Genomics Centre Zürich for their assistance with data generation. We thank York Archaeological Trust for Excavation and Research Ltd (Christine McDonnell), York Osteoarchaeology Ltd (Malin Holst, Katie Keefe), Andrew Pearson, and Iterbide S.C. (Javier Niso, Miguel Loza) for providing access to skeletal collections.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jas.2020.105135.

References

Ahlmann-Eltze, C., 2019. ggsignif: significance Brackets for "ggplot2.". https://CRAN.R-project.org/package=ggsignif.
Anderson, M.J., 2001. A new method for non-parametric multivariate analysis of variance: non-parametric manova for ecology. Austral Ecol. 26, 32–46.
Austin, R.M., Sholts, S.B., Williams, L., Kistler, L., Hofman, C.A., 2019. Opinion: to curate the molecular past, museums need a carefully considered set of best practices. Proc. Natl. Acad. Sci. U. S. A 116, 1471–1474.
Bates, D., Mächler, M., Bolker, B., Walker, S., 2015. Fitting linear mixed-effects models using lme4. J. Stat. Software 67.
Berstan, R., Stott, A.W., Minnitt, S., Bronk Ramsey, C., Hedges, R.E.M., Evershed, R.P., 2008. Direct dating of pottery from its organic residues: new precision using compound-specific carbon isotopes. Antiquity 82, 702–713.
Bolnick, D.A., Bonine, H.M., Mata-Míguez, J., Kemp, B.M., Snow, M.H., LeBlanc, S.A., 2012. Nondestructive sampling of human skeletal remains yields ancient nuclear and mitochondrial DNA. Am. J. Phys. Anthropol. 147, 293–300.
Casanova, E., Knowles, T.D.J., Williams, C., Crump, M.P., Evershed, R.P., 2018. Practical considerations in high-precision compound-specific radiocarbon analyses: eliminating the effects of solvent and sample cross-contamination on accuracy and precision. Anal. Chem. 90, 11025–11032.
Chen, F., Welker, F., Shen, C.-C., Bailey, S.E., Bergmann, I., Davis, S., Xia, H., Wang, H., Fischer, R., Freidline, S.E., Yu, T.-L., Skinner, M.M., Stelzer, S., Dong, G., Fu, Q., Dong, G., Wang, J., Zhang, D., Hublin, J.-J., 2019. A late middle pleistocene denisovan mandible from the Tibetan plateau. Nature 569, 409–412.
Dabney, J., Knapp, M., Glocke, I., Gansauge, M.-T., Weihmann, A., Nickel, B., Valdiosera, C., García, N., Pääbo, S., Arsuaga, J.-L., Meyer, M., 2013. Complete mitochondrial genome sequence of a Middle Pleistocene cave bear reconstructed from ultrashort DNA fragments. Proc. Natl. Acad. Sci. U. S. A 110, 15758–15763.
Davis, N.M., Proctor, D.M., Holmes, S.P., Relman, D.A., Callahan, B.J., 2018. Simple statistical identification and removal of contaminant sequences in marker-gene and metagenomics data. Microbiome 6, 226.
Demarchi, B., Hall, S., Roncal-Herrero, T., Freeman, C.L., Woolley, J., Crisp, M.K., Wilson, J., Fotakis, A., Fischer, R., Kessler, B.M., Rakownikow Jersie-Christensen, R., Olsen, J.V., Haile, J., Thomas, J., Marean, C.W., Parkington, J., Presslee, S., Lee-Thorp, J., Ditchfield, P., Hamilton, J.F., Ward, M.W., Wang, C.M., Shaw, M.D., Harrison, T., Domínguez-Rodrigo, M., MacPhee, R.D.E., Kwekason, A., Ecker, M., Kolska Horwitz, L., Chazan, M., Kröger, R., Thomas-Oates, J., Harding, J.H., Cappellini, E., Penkman, K., Collins, M.J., 2016. Protein sequences bound to mineral surfaces persist into deep time. Elife 5.
Dewhirst, F.E., Chen, T., Izard, J., Paster, B.J., Tanner, A.C.R., Yu, W.-H., Lakshmanan, A., Wade, W.G., 2010. The human oral microbiome. J. Bacteriol. 192, 5002–5017.
Escapa, I.F., Chen, T., Huang, Y., Gajare, P., Dewhirst, F.E., Lemon, K.P., 2018. New insights into human nostril microbiome from the expanded human oral microbiome database (eHOMD): a resource for the microbiome of the human aerodigestive tract. mSystems 3.
Fiddyment, S., Holsinger, B., Ruzzier, C., Devine, A., Binois, A., Albarella, U., Fischer, R., Nichols, E., Curtis, A., Cheese, E., Teasdale, M.D., Checkley-Scott, C., Milner, S.J., Rudy, K.M., Johnson, E.J., Vnouček, J., Garrison, M., McGrory, S., Bradley, D.G., Collins, M.J., 2015. Animal origin of 13th-century uterine vellum revealed using noninvasive peptide fingerprinting. Proc. Natl. Acad. Sci. U. S. A 112, 15066–15071.
Firke, S., 2018. Janitor: simple tools for examining and cleaning dirty data. https://CRAN.R-project.org/package=janitor.
Gloor, G.B., Macklaim, J.M., Pawlowsky-Glahn, V., Egozcue, J.J., 2017. Microbiome datasets are compositional: and this is not optional. Front. Microbiol. 8, 2224.

Green, E.J., Speller, C.F., 2017. Novel substrates as sources of ancient DNA: prospects and hurdles. Genes 8.
Hendy, J., Warinner, C., Bouwman, A., Collins, M.J., Fiddyment, S., Fischer, R., Hagan, R., Hofman, C.A., Holst, M., Chaves, E., Klaus, L., Larson, G., Mackie, M., McGrath, K., Mundorff, A.Z., Radini, A., Rao, H., Trachsel, C., Velsko, I.M., Speller, C.F., 2018. Proteomic evidence of dietary sources in ancient dental calculus. Proc. Biol. Sci. 285.
Henry, A.G., Piperno, D.R., 2008. Using plant microfossils from dental calculus to recover human diet: a case study from Tell al-Raqa'i, Syria. J. Archaeol. Sci. 35, 1943–1950.
Herbig, A., Maixner, F., Bos, K.I., Zink, A., Krause, J., Huson, D.H., 2016. MALT: Fast Alignment and Analysis of Metagenomic DNA Sequence Data Applied to the Tyrolean Iceman (bioRxiv).
Huson, D.H., Beier, S., Flade, I., Górska, A., El-Hadidi, M., Mitra, S., Ruscheweyh, H.-J., Tappu, R., 2016. MEGAN community edition - interactive exploration and analysis of large-scale microbiome sequencing data. PLoS Comput. Biol. 12, e1004957.
Jeong, C., Wilkin, S., Amgalantugs, T., Bouwman, A.S., Taylor, W.T.T., Hagan, R.W., Bromage, S., Tsolmon, S., Trachsel, C., Grossmann, J., Littleton, J., Makarewicz, C. A., Krigbaum, J., Burri, M., Scott, A., Davaasambuu, G., Wright, J., Irmer, F., Myagmar, E., Boivin, N., Robbeets, M., Rühli, F.J., Krause, J., Frohlich, B., Hendy, J., Warinner, C., 2018. Bronze Age population dynamics and the rise of dairy pastoralism on the eastern Eurasian steppe. Proc. Natl. Acad. Sci. U. S. A 115, E11248–E11255.
Jersie-Christensen, R.R., Lanigan, L.T., Lyon, D., Mackie, M., Belstrøm, D., Kelstrup, C.D., Fotakis, A.K., Willerslev, E., Lynnerup, N., Jensen, L.J., Cappellini, E., Olsen, J.V., 2018. Quantitative metaproteomics of medieval dental calculus reveals individual oral health status. Nat. Commun. 9, 4744.
Kassambara, A., 2018. Ggpubr: "ggplot2" based publication ready plots. https://CRAN.R-project.org/package=ggpubr.
Key, F.M., Posth, C., Krause, J., Herbig, A., Bos, K.I., 2017. Mining metagenomic data sets for ancient DNA: recommended protocols for authentication. Trends Genet. 33, 508–520.
Kircher, M., Sawyer, S., Meyer, M., 2012. Double indexing overcomes inaccuracies in multiplex sequencing on the Illumina platform. Nucleic Acids Res. 40, e3.
Korlević, P., Talamo, S., Meyer, M., 2018. A combined method for DNA analysis and radiocarbon dating from a single sample. Sci. Rep. 8, 4127.
Kuznetsova, A., Brockhoff, P.B., Christensen, R.H.B., 2017. lmerTest package: tests in linear mixed effects models. J. Stat. Software 82.
Li, H., Durbin, R., 2009. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics 25, 1754–1760.
Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R., 1000 Genome Project Data Processing Subgroup, 2009. The sequence alignment/map format and SAMtools. Bioinformatics 25, 2078–2079.
Mackie, M., Radini, A., Speller, C.F., 2017. The sustainability of dental calculus for archaeological research. In: Favreau, J., Patalano, R. (Eds.), Shallow Pasts, Endless Horizons: Sustainability & Archaeology:Proceedings of the 48th Annual Chacmool Archaeology Conference, p. 8.
Manfredi, M., Barberis, E., Gosetti, F., Conte, E., Gatti, G., Mattu, C., Robotti, E., Zilberstein, G., Koman, I., Zilberstein, S., Marengo, E., Righetti, P.G., 2017. Method for noninvasive analysis of proteins and small molecules from ancient objects. Anal. Chem. 89, 3310–3317.
Mann, A.E., Sabin, S., Ziesemer, K., Vågene, Å.J., Schroeder, H., Ozga, A.T., Sankaranarayanan, K., Hofman, C.A., Fellows Yates, J.A., Salazar-García, D.C., Frohlich, B., Aldenderfer, M., Hoogland, M., Read, C., Milner, G.R., Stone, A.C., Lewis Jr., C.M., Krause, J., Hofman, C., Bos, K.I., Warinner, C., 2018. Differential preservation of endogenous human and microbial DNA in dental calculus and dentin. Sci. Rep. 8, 9822.
Meyer, M., Kircher, M., 2010. Illumina sequencing library preparation for highly multiplexed target capture and sequencing. Cold Spring Harb. Protoc. 2010 db. prot5448.
Nowosad, J., 2018. CARTOColors' palettes. https://nowosad.github.io/rcartocolor.
Oksanen, J., Blanchet, F.G., Friendly, M., Kindt, R., Legendre, P., McGlinn, D., Minchin, P.R., O'Hara, R.B., Simpson, G.L., Solymos, P., Stevens, M.H.H., Szoecs, E., Wagner, H., 2019. Vegan: community ecology package. https://CRAN.R-project.org/package=vegan.
Orlando, L., Ginolhac, A., Zhang, G., Froese, D., Albrechtsen, A., Stiller, M., Schubert, M., Cappellini, E., Petersen, B., Moltke, I., Johnson, P.L.F., Fumagalli, M., Vilstrup, J.T., Raghavan, M., Korneliussen, T., Malaspinas, A.-S., Vogt, J., Szklarczyk, D., Kelstrup, C.D., Vinther, J., Dolocan, A., Stenderup, J., Velazquez, A.M.V., Cahill, J., Rasmussen, M., Wang, X., Min, J., Zazula, G.D., Seguin-Orlando, A., Mortensen, C., Magnussen, K., Thompson, J.F., Weinstock, J., Gregersen, K., Røed, K.H., Eisenmann, V., Rubin, C.J., Miller, D.C., Antczak, D.F., Bertelsen, M.F., Brunak, S., Al-Rasheid, K.A.S., Ryder, O., Andersson, L., Mundy, J., Krogh, A., Gilbert, M.T.P., Kjær, K., Sicheritz-Ponten, T., Jensen, L.J., Olsen, J.V., Hofreiter, M., Nielsen, R., Shapiro, B., Wang, J., Willerslev, E., 2013. Recalibrating Equus evolution using the genome sequence of an early Middle Pleistocene horse. Nature 499, 74–78.
Osorio, D., 2015. Peptides: a package for data mining of antimicrobial peptides. R J 7, 4–14.
Ozga, A.T., Nieves-Colón, M.A., Honap, T.P., Sankaranarayanan, K., Hofman, C.A., Milner, G.R., Lewis Jr., C.M., Stone, A.C., Warinner, C., 2016. Successful enrichment and recovery of whole mitochondrial genomes from ancient human dental calculus. Am. J. Phys. Anthropol. 160, 220–228.
Palarea-Albaladejo, J., Martín-Fernández, J.A., 2015. zCompositions — R package for multivariate imputation of left-censored data under a compositional approach. Chemometr. Intell. Lab. Syst. 143, 85–96.
Peltzer, A., Jäger, G., Herbig, A., Seitz, A., Kniep, C., Krause, J., Nieselt, K., 2016. EAGER: efficient ancient genome reconstruction. Genome Biol. 17, 60.
Radini, A., Nikita, E., Buckley, S., Copeland, L., Hardy, K., 2017. Beyond food: the multiple pathways for inclusion of materials into ancient dental calculus. Am. J. Phys. Anthropol. 162 (Suppl. 63), 71–83.
Radini, A., Tromp, M., Beach, A., Tong, E., Speller, C., McCormick, M., Dudgeon, J.V., Collins, M.J., Rühli, F., Kröger, R., Warinner, C., 2019. Medieval women's early involvement in manuscript production suggested by lapis lazuli identification in dental calculus. Sci. Adv. 5 eaau7126.
R Core Team, 2018. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/.
Robinson, D., 2018. Fuzzyjoin: Join Tables Together on Inexact Matching. https://CRAN.R-project.org/package=fuzzyjoin.
Rohland, N., Harney, E., Mallick, S., Nordenfelt, S., Reich, D., 2015. Partial uracil–DNA–glycosylase treatment for screening of ancient DNA. Phil. Trans. R. Soc. B 370 (20130624).
Rohland, N., Siedel, H., Hofreiter, M., 2004. Nondestructive DNA extraction method for mitochondrial DNA analyses of museum specimens. Biotechniques 36, 814–816, 818–821.
Rusu, I., Paica, I., Vulpoi, A., Radu, C., Mircea, C., Dobrinescu, C., Bodolică, V., Kelemen, B., 2019. Dual DNA-protein extraction from human archeological remains. Archaeol. Anthropol. Sci. 11, 3299–3307.
Salter, S.J., Cox, M.J., Turek, E.M., Calus, S.T., Cookson, W.O., Moffatt, M.F., Turner, P., Parkhill, J., Loman, N.J., Walker, A.W., 2014. Reagent and laboratory contamination can critically impact sequence-based microbiome analyses. BMC Biol. 12, 87.
Schubert, M., Lindgreen, S., Orlando, L., 2016. AdapterRemoval v2: rapid adapter trimming, identification, and read merging. BMC Res. Notes 9, 88.
Sielaff, M., Kuharev, J., Bohn, T., Hahlbrock, J., Bopp, T., Tenzer, S., Distler, U., 2017. Evaluation of FASP, SP3, and iST protocols for proteomic sample preparation in the low microgram range. J. Proteome Res. 16, 4060–4072.
Tromp, M., Buckley, H., Geber, J., Matisoo-Smith, E., 2017. EDTA decalcification of dental calculus as an alternate means of microparticle extraction from archaeological samples. J. Archaeol. Sci. Rep. 14, 461–466.
van Doorn, N.L., Hollund, H., Collins, M.J., 2011. A novel and non-destructive approach for ZooMS analysis: ammonium bicarbonate buffer extraction. Archaeol. Anthropol. Sci. 3, 281.
Velsko, I.M., Fellows Yates, J.A., Aron, F., Hagan, R.W., Frantz, L.A.F., Loe, L., Martinez, J.B.R., Chaves, E., Gosden, C., Larson, G., Warinner, C., 2019. Microbial differences between dental plaque and historic dental calculus are related to oral biofilm maturation stage. Microbiome 7, 102.
Velsko, I.M., Overmyer, K.A., Speller, C., Klaus, L., Collins, M.J., Loe, L., Frantz, L.A.F., Sankaranarayanan, K., Lewis Jr., C.M., Martinez, J.B.R., Chaves, E., Coon, J.J., Larson, G., Warinner, C., 2017. The dental calculus metabolome in modern and historic samples. Metabolomics 13, 134.
Velsko, I.M., Warinner, C., 2017. Bioarchaeology of the human microbiome. Bioarchaeol. Int. 1, 86–99.
Venables, W.N., Ripley, B.D., 2002. Modern Applied Statistics with S. Springer Publishing Company, Incorporated.
Warinner, C., 2016. Dental calculus and the evolution of the human oral microbiome. J. Calif. Dent. Assoc. 44, 411–420.
Warinner, C., Hendy, J., Speller, C., Cappellini, E., Fischer, R., Trachsel, C., Arneborg, J., Lynnerup, N., Craig, O.E., Swallow, D.M., Fotakis, A., Christensen, R.J., Olsen, J.V., Liebert, A., Montalva, N., Fiddyment, S., Charlton, S., Mackie, M., Canci, A., Bouwman, A., Rühli, F., Gilbert, M.T.P., Collins, M.J., 2014a. Direct evidence of milk consumption from ancient human dental calculus. Sci. Rep. 4, 7104.
Warinner, C., Herbig, A., Mann, A., Fellows Yates, J.A., Weiß, C.L., Burbano, H.A., Orlando, L., Krause, J., 2017. A robust framework for microbial archaeology. Annu. Rev. Genom. Hum. Genet. 18, 321–356.
Warinner, C., Rodrigues, J.F.M., Vyas, R., Trachsel, C., Shved, N., Grossmann, J., Radini, A., Hancock, Y., Tito, R.Y., Fiddyment, S., Speller, C., Hendy, J., Charlton, S., Luder, H.U., Salazar-García, D.C., Eppler, E., Seiler, R., Hansen, L.H., Castruita, J.A. S., Barkow-Oesterreicher, S., Teoh, K.Y., Kelstrup, C.D., Olsen, J.V., Nanni, P., Kawai, T., Willerslev, E., von Mering, C., Lewis Jr., C.M., Collins, M.J., Gilbert, M.T. P., Rühli, F., Cappellini, E., 2014b. Pathogens and host immunity in the ancient human oral cavity. Nat. Genet. 46, 336–344.
Warnes, G.R., Bolker, B., Lumley, T., 2018. Gtools: various R programming tools. https://CRAN.R-project.org/package=gtools.
Wickham, H., 2017. Tidyverse: easily install and load the "tidyverse". https://CRAN.R-project.org/package=tidyverse.
Wilke, C.O., 2017. Cowplot: streamlined plot theme and plot annotations for "ggplot2". https://CRAN.R-project.org/package=cowplot.
Wiśniewski, J.R., Zougman, A., Nagaraj, N., Mann, M., 2009. Universal sample preparation method for proteome analysis. Nat. Methods 6, 359–362.
Ziesemer, K.A., Mann, A.E., Sankaranarayanan, K., Schroeder, H., Ozga, A.T., Brandt, B. W., Zaura, E., Waters-Rist, A., Hoogland, M., Salazar-García, D.C., Aldenderfer, M., Speller, C., Hendy, J., Weston, D.A., MacDonald, S.J., Thomas, G.H., Collins, M.J., Lewis, C.M., Hofman, C., Warinner, C., 2015. Intrinsic challenges in ancient microbiome reconstruction using 16S rRNA gene amplification. Sci. Rep. 5, 16498.
Ziesemer, K.A., Ramos-Madrigal, J., Mann, A.E., Brandt, B.W., Sankaranarayanan, K., Ozga, A.T., Hoogland, M., Hofman, C.A., Salazar-García, D.C., Frohlich, B., Milner, G.R., Stone, A.C., Aldenderfer, M., Lewis Jr., C.M., Hofman, C.L., Warinner, C., Schroeder, H., 2019. The efficacy of whole human genome capture on ancient dental calculus and dentin. Am. J. Phys. Anthropol. 168, 496–509.

MANUSCRIPT C

Manuscript Nr. C

Title of manuscript: Understanding the microbial biogeography of ancient human dentitions to guide study design and interpretation

Authors: Zandra Fagernäs, Domingo C. Salazar-García, Azucena Avilés Fernández, María Haber Uriarte, Amanda Henry, Joaquín Lomba Maurandi, Andrew Ozga, Irina M Velsko, Christina Warinner

Bibliographic information: Submitted to FEMS Microbes (September 2021). Also available as a pre-print at: https://doi.org/10.1101/2021.08.16.456492

The candidate is

⊗ First author, o Co-first author, o Corresponding author, o Coauthor.

Status: Submitted

Proportion (in %) of authors in the publication (to be indicated from 20%):

Author	Concept	Data analysis	Experiment	Manuscript composition	Material provision
Zandra Fagernäs	33%	100%	95%	70%	-
Domingo C. Salazar-García	-	-	-	-	25%
Azucena Avilés Fernández	-	-	-	-	25%
María Haber Uriarte	-	-	-	-	25%
Amanda Henry	-	-	-	-	-
Joaquín Lomba Maurandi	-	-	-	-	25%
Andrew Ozga	-	-	-	-	-
Irina M Velsko	33%	-	-	-	-
Christina Warinner	33%	-	-	-	-

Note: A more thorough breakdown of author contributions can be found in Appendix A. Supplementary figures can be found in Appendix C. The supplementary data, which contains numerous large tables, can be accessed online in the pre-print version of this manuscript (DOI: https://doi.org/10.1101/2021.08.16.456492).

Understanding the microbial biogeography of ancient human dentitions to guide study design and interpretation

Zandra Fagernäs[1], Domingo C. Salazar-García[2,3,4], Azucena Avilés Fernández[5], María Haber[5], Amanda Henry[6], Joaquín Lomba Maurandi[5], Andrew Ozga[7], Irina M Velsko[1], Christina Warinner[1,8,9]

1. Department of Archaeogenetics, Max Planck Institute for the Science of Human History, Jena, Germany.
2. Departament de Prehistòria, Arqueologia i Història Antiga, Universitat de València, València, Spain.
3. Grupo de Investigación en Prehistoria IT-1223-19 (UPV-EHU)/IKERBASQUE-Basque Foundation for Science, Vitoria, Spain.
4. Department of Geological Sciences, University of Cape Town, Cape Town, South Africa.
5. Departamento de Prehistoria, Arqueología, Historia Antigua, Historia Medieval y Ciencias y Técnicas Historiográficas, Universidad de Murcia, Murcia, Spain.
6. Faculty of Archaeology, Leiden University, Leiden, the Netherlands.
7. Halmos College of Arts and Sciences, Nova Southeastern University, Fort Lauderdale, FL, USA.
8. Faculty of Biological Sciences, Friedrich Schiller University, Jena, Germany.
9. Department of Anthropology, Harvard University, Cambridge, MA, USA

Keywords

Ancient DNA / Microbiome / Dental calculus / Metagenomics / Biogeography / Archaeology

Abstract

The oral cavity is a heterogeneous environment, varying in factors such as pH, oxygen levels, and salivary flow. These factors affect the microbial community composition and distribution of species in dental plaque, but it is not known how well these patterns are reflected in archaeological dental calculus. In most archaeological studies, a single sample of dental calculus is studied per individual and is assumed to represent the entire oral cavity. However, it is not known if this sampling strategy introduces biases into studies of the ancient oral microbiome. Here, we present the results of a shotgun metagenomic study of a dense sampling of dental calculus from four Chalcolithic individuals from the southeast Iberian peninsula (ca. 4500-5000 BP). Inter-individual differences in microbial composition are found to be much larger than intra-individual differences, indicating that a single sample can indeed represent an individual in most cases. However, there are minor spatial patterns in species distribution within the oral cavity that should be taken into account when designing a study or interpreting results. Finally, we show that plant DNA identified in the samples may be of postmortem origin, showing the importance of including environmental controls or several lines of biomolecular evidence.

Introduction

Dental calculus forms when the dental plaque biofilm mineralizes during life (Jin & Yip, 2002), a periodic occurrence that encapsulates microbes, host biomolecules, food residues, and particles from the environment (Velsko & Warinner, 2017). After the death of an individual, biomolecules within dental calculus can be preserved for tens of thousands of years (Fellows Yates et al., 2021), largely protected from environmental processes within the mineral matrix. Studies of archaeological dental calculus have rapidly increased in number in recent years, in part due to an elevated interest in the evolution of the oral microbiome and a growing understanding of the plethora of ancient biomolecules and information that can be recovered from this semi-fossilized microbial biofilm. However, there are still many unknown factors regarding the formation and preservation of archaeological dental calculus, and further method development is therefore necessary.

In carrying out comparative studies of ancient dental calculus, researchers aim to set up a sampling strategy that mitigates biases caused by intra-individual variability of the studied individuals. However, as archaeological dental calculus is often found in small quantities, especially in individuals dating far back in time, and is not always present on the same teeth across individuals, it may not always be possible to adhere to such a sampling scheme. Pre- and post-mortem tooth loss can further complicate sampling designs, as does working with calculus samples that were dislocated from the teeth during handling or storage, such that the precise tooth of origin is unknown. Due to such sampling constraints, some studies have pooled and homogenized calculus from several teeth for analysis (Warinner et al., 2014), which may partly mitigate spatial biases, but this approach requires the presence and collection of larger amounts of calculus, which is a finite archaeological substrate. In light of these challenges, most ancient oral microbiome studies implicitly assume that a single sample can be representative of the entire dentition, regardless of the tooth niche from which the calculus sample is obtained, and analyze only a single dental calculus deposit per individual. The oral cavity, however, is not a uniform environment, and thus microbial communities may vary across the dentition, potentially leading to bias when comparing across individuals from whom different teeth were sampled.

Differences in the microbial composition of different oral tissues, such as buccal mucosa, keratinized gingiva, saliva, and teeth, have been reported in present-day humans (Aas et al., 2005; Ding & Schloss, 2014; Eren et al., 2014; Mark Welch et al., 2016; Proctor et al., 2018; Utter et al., 2020). Further, differences in dental plaque microbial communities have been previously reported between mandibular and maxillary teeth (Haffajee et al., 2009; Simon-Soro & Tomás, 2013), between tooth position (e.g. anterior vs. posterior teeth) (Haffajee et al., 2009; Proctor et al., 2018), between tooth surfaces (e.g. buccal vs. lingual) (Proctor et al., 2018;

Simon-Soro & Tomás, 2013), and between supragingival and subgingival plaque (Eren et al., 2014; Simon-Soro & Tomás, 2013). Local variations in oral physiological conditions, such as salivary flow rate, salivary composition, oxygen availability, and mechanical abrasion during mastication, may contribute to these subtle spatial microbial differences in dental plaque. However, while such spatial differences have been detected in the microbial composition of dental plaque, it is not known whether these patterns are also reflected in dental calculus. Dental calculus represents a fully mature stage of oral biofilm development that is often disrupted in living individuals practicing oral hygiene, leading to a distinct microbial profile between dental plaque and dental calculus (Kazarina et al., 2021; Velsko et al., 2019). Overall, dental calculus typically contains higher proportions of late colonizer taxa that thrive in the anaerobic environment created as the biofilm matures, and thus its composition may be less spatially variable than developing plaque biofilms, which are more dynamic and subject to periodically disruptive forces such as toothbrushing (Velsko et al., 2019).

However, evaluating intra-individual microbial variation in dental calculus across the dental arcade, and thus determining the degree to which a single sample can represent an individual, is challenging. Dense sampling of calculus is often hindered by missing teeth or a lack of calculus deposits distributed across the entire dental arcade. Consequently, previous studies have attempted to identify microbial spatial patterns across the dentition by instead sampling diverse individual teeth from a large number of individuals (Farrer et al., 2018), but this introduces a number of uncontrolled variables, such as individual differences, different biological and absolute ages of samples, different postmortem conditions, and differing degrees of preservation and degradation, which may introduce biases or otherwise alter the observable spatial patterns. Further, this approach does not allow for comparisons of how much of the variation in the dental calculus microbiome stems from intra- vs. inter-individual differences.

To determine the degree to which tooth selection matters in dental calculus sampling for comparative ancient microbiome studies, we conducted a systematic analysis of microbial spatial variation in four nearly complete human dentitions with low to heavy dental calculus deposits from the Iberian Chalcolithic site of Camino del Molino (ca. 4500-5000 BP). With dense sampling across tooth types (incisor, canine, premolar, molar) and tooth surfaces (buccal, labial, interproximal, occlusal), we performed shotgun metagenomic analysis of 87 dental calculus samples. We find that the main source of variation in the oral microbiome is the sampled individual, and therefore one randomly selected sample can, for most purposes, be used to represent an individual in population-level comparative studies. However, minor intra-individual patterns in community composition, functional potential, and species abundances are detectable with respect to tooth position (anterior vs. posterior), dental calculus deposit size, and tooth surface, although with low effect sizes. Only occlusal calculus, which is uncommon and may

indicate injury or physiological dysfunction, considerably differed in composition. We found that ancient human DNA is randomly distributed across the dentition, and no spatial patterns were observed with respect to postmortem environmental contamination. Finally, we found that ancient grapevine (*Vitis vinifera*) DNA was present in the dental calculus we analyzed; however, it was also present in mandibular bone, suggesting a postmortem origin. This study contributes to an awareness of spatial variation in dental calculus microbial community composition that aims to aid researchers in developing robust study designs and valid interpretations for ancient oral microbiome studies.

Materials and Methods

Samples

Dental calculus was collected from four Chalcolithic (4500-5000 BP) individuals from the southeastern Iberian archaeological site of Camino del Molino near the city of Caravaca de la Cruz in Murcia, Spain, excavated during a salvage excavation in 2008 (Díaz Navarro et al., 2019; Haber-Uriarte et al., 2011; Lomba Maurandi, López Martínez, & Ramos Martínez, 2009; Lomba Maurandi et al., 2009). The Camino del Molino communal burial is a natural pit with a 7 meter diameter circular base and a depth of 4 meters (of which only the lower 2 meters were used for burial), which was likely covered and sealed by a perishable structure (Lomba Maurandi et al., 2009). The upper layers of the site were destroyed in the early 20th century as a result of agricultural terracing, but the damage did not extend to the burial deposits. Approximately 1,300 human individuals representing a broad demographic profile were buried at the site between 2800-2400 BCE (Haber-Uriarte et al., 2011). The site was chosen for this study because prior dental calculus research at the site had shown excellent oral microbiome preservation (Mann et al., 2018; Ziesemer et al., 2015), and microfossil studies of the dental calculus had been conducted (Power et al., 2014), and because the large number of individuals excavated from the site made it possible to select suitable individuals with nearly complete dentitions and sufficient dental calculus for this study. The four selected individuals were adults and had dental calculus present on most teeth (Figure 1, S1), allowing near comprehensive sampling. Dental notation below follows the FDI World Dental Federation standard (Peck & Peck, 1993); molar enamel wear is reported as a Brothwell score from 1 (none) to 7 (obliteration of crown and wear of roots) (Brothwell, 1972), and dental calculus deposits are graded from 1 (slight) to 4 (gross) according to Dobney and Brothwell (1987).

Individual CM55. Individual CM55 (35-39 year old female) had a complete mandible and a partial, fragmented maxilla, with a total of 22 teeth (Figure 1). Alveolar bone loss and reactive

bone formation was observed throughout the mandibular periodontium, suggesting generalized periodontitis. Gross carious lesions were present in teeth 17, 35, 37, 45, and 47. Molar enamel wear was low (Brothwell stage 2). Dental calculus deposits were grade 1-2 in size, except on left premolars and molars, where they reached grade 4. The excessive calculus accumulation on the left posterior teeth, including on the occlusal surfaces, suggests that this individual had experienced pain on the left side of the mouth and avoided mastication on this side. Although no skeletal trauma was apparent, CM55 had experienced antemortem tooth loss of teeth 36 and 38, and a large carious lesion was present in 37. Significant alveolar recession and reactive bone formation was also evident around 24, but damage to the left maxilla prohibited further inspection of the bone supporting the upper molar teeth.

Individual CM59. Individual CM59 (25-35 year old male) had an intact mandible and a partial, fragmented maxilla, with a total of 25 teeth (Figure 1). Molar enamel wear was minimal (Brothwell stage 1-2), and no gross carious lesions were observed. Dental calculus deposits were grade 1-2 in size. Alveolar recession was slight across the periodontium, and in general the individual exhibited good dental health.

Individual CM82. Individual CM82 (35-45 year old female) had a complete mandible and a partial, fragmented maxilla, with a total of 23 teeth (Figure 1). Heavy enamel wear (Brothwell stage 4) was observed on the molar teeth. Dental calculus deposits were grade 1-2 in size. A large bone abscess was present adjacent to the healed alveolar bone where teeth 37 and 38 had been lost antemortem. Alveolar recession was pronounced around the molars, and healing was incomplete for four molars that had been lost antemortem. Gross carious lesions were present in teeth 16, 18, 45 and 46.

Individual CM165. Individual CM165 (25-30 year old likely female) had a near complete mandible and maxilla, with a total of 29 teeth (Figure 1). Although an adult, the individual retained deciduous tooth 52 and the corresponding adult tooth 12 was absent, suggesting agenesis. CM165 also had a partially impacted tooth 38. Gross carious lesions were present in teeth 37, 38, and 48. Postmortem bone breakage made the alveolar margin difficult to assess, but where observable recession was not pronounced. Molar enamel wear was low (Brothwell stage 2-3), and dental calculus deposits were grade 1 in size. Overall, the individual exhibited relatively good dental health.

Dental calculus collection was performed in an ancient DNA cleanroom environment at the University of Oklahoma (individual CM55) and the Max Planck Institute for Human History (individuals CM59, CM82 and CM165) under sterile conditions following Warinner, Velsko and Fellow Yates (2019), and supragingival dental calculus was separately collected from four different surfaces on each tooth: lingual, buccal, interproximal and occlusal. For each individual,

a bone sample (approximately 50 mg) was also collected from the mandibular ramus to be used as a control for microbes characteristic of the local burial environment. As bone is mainly assumed to be free of microbes during life, in the absence of disease, the microbes identified from archaeological bone stem from the burial environment, and represent taxa that have colonized the remains, including dental calculus, postmortem. A subset of dental calculus samples was selected from each individual for metagenomic analysis (Figure 1). This subsampling was performed with the aim of achieving a balanced representation of dental sites and surfaces across individuals, as well as a consistent sample mass for analysis. For all individuals, dental sites or surfaces with < 1mg of dental calculus were generally excluded from analysis. For individuals CM59, CM82, and CM165, half of the dentition was sampled (right or left, depending on completeness and calculus abundance), but samples from the paired left/right side were also included as needed to balance out the sampling scheme with respect to tooth site and sample mass; this was particularly necessary for CM82. For individual CM55, dental calculus across the entire available dentition was sampled. Although dental calculus was mostly present only on the tooth buccal and lingual surfaces, the massive calculus deposits on the left molars of CM55 enabled the analysis of occlusal calculus for this individual. In addition, eight interproximal sites in CM59 and CM165 yielded sufficient calculus for analysis and were also sampled. In total, 87 calculus samples were selected from the four individuals for metagenomic analysis (Supplementary data 1).

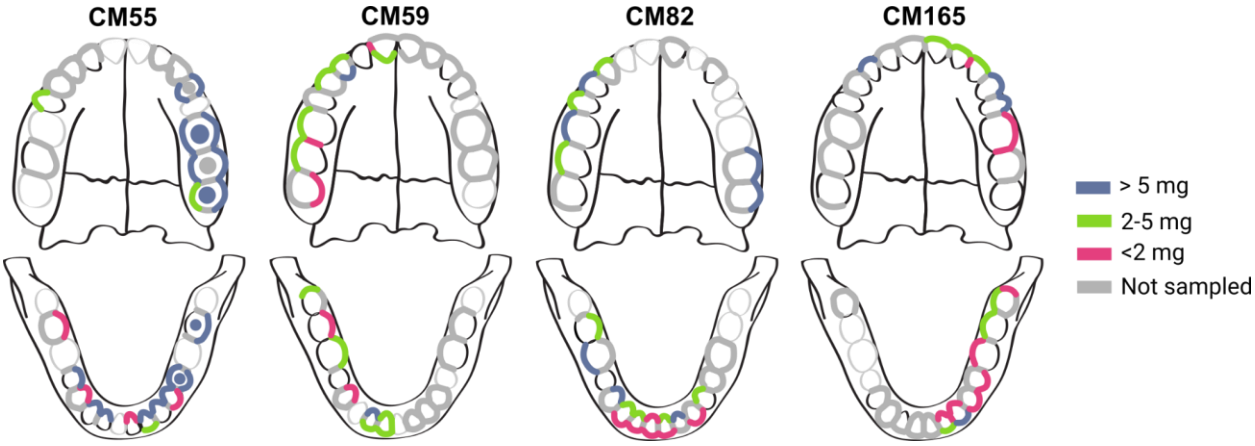


Figure 1. Study sampling design. Dental calculus deposits investigated in this study are highlighted and correspond to the sampled tooth surface (buccal, lingual, interproximal, occlusal). The color of the highlighting indicates the initial mass of the dental calculus deposit on the teeth that were analyzed: <2 mg (pink); 2-5.0 mg (green); >5.0 mg (blue). Teeth that were present are indicated in black outline; teeth that were absent are indicated in light gray outline. Dental calculus that was present but excluded from analysis (due to sampling design or insufficient starting mass) is marked in dark gray.

Laboratory methods

Surface contamination was reduced by UV irradiation (30 s on both sides), followed by a washing step in 1 mL of 0.5 M EDTA (without incubation). DNA was extracted from the calculus and bone samples using a modified version of (Dabney et al., 2013) adapted for dental calculus (Aron et al., 2020; Mann et al., 2018) and allowing for potential future protein extraction from the same samples (Fagernäs et al., 2020). Briefly, the samples were decalcified in 1 mL 0.5 M EDTA for three days, after which the cell debris pellet and 100 µl of the supernatant was frozen at -20 °C and set aside for future analyses (Fagernäs et al., 2020). To the remaining 900 µL supernatant, proteinase K (Sigma-Aldrich) was added, and the samples were incubated at room temperature overnight. The supernatant was then mixed with binding buffer (5 M guanidine hydrochloride, 0.12 M sodium acetate, 40% isopropanol) and DNA was purified using a High Pure Viral Nucleic Acid kit (Roche Life Science) according to the manufacturer's instructions. DNA was eluted in Qiagen EB buffer, to which Tween 20 had been added to a final concentration 0.05%. DNA was quantified using a Qubit HS assay (Thermo Fisher Scientific). Extraction blanks (one per batch) were processed alongside the samples. The full extraction protocol is available at (Aron et al., 2020).

Extracted DNA was processed with a partial uracil-DNA-glycosylase treatment (Aron et al., 2020; Rohland et al., 2015) and was prepared into double-stranded libraries with dual indexing (Kircher et al., 2012; Meyer & Kircher, 2010; Stahl et al., 2019). Library blanks were processed alongside the samples, one per batch. The DNA libraries were shotgun sequenced on an Illumina NextSeq with 75-bp paired-end chemistry. Dental calculus samples were sequenced to a depth of 10.1 ± 1.4 M reads (average ± standard deviation), bone samples to 6.6 ± 2.0 M reads, and blanks to 1.7 ± 0.7 M reads.

Data analysis

Preprocessing

The EAGER v.1.92.56 (Peltzer et al., 2016) pipeline was used for preprocessing of the raw data. Adapter removal and merging of reads were performed using AdapterRemoval v. 2.3.1 (Schubert et al., 2016). The reads were mapped to the human reference genome HG19 using BWA aln v. 0.7.12 (Li & Durbin, 2009) with default settings (-l 32, -n 0.01), and unmapped reads were extracted with SAMtools v. 1.3 (Li et al., 2009) for downstream microbiome analyses. The unmapped reads were aligned to a custom RefSeq database (Fellows Yates et al., 2021) using MALT v. 0.4.0 (Herbig et al., 2016) (settings -id 85.0 -top 1 -supp 0.01). This database contains all bacterial and archaeal assemblies at scaffold/chromosome/complete levels (as of November 2018), with max 10 randomly selected genomes per species (prioritizing more complete

genomes), as well as the human HG19 reference genome. A preliminary screening for eukaryotic DNA was also performed as described above, using the NCBI full nt database (as of October 2017), but the custom RefSeq database was chosen for further analyses, as it has been shown to yield a higher percentage aligned sequences for dental calculus (Fellows Yates et al., 2021). OTU tables with summarized read counts at genus level were exported through MEGAN v. 6.17.0 (Huson et al., 2016) (Supplementary data 2 and 3). The R-package decontam v. 1.6.0 (Davis et al., 2018) was used to identify putative laboratory and environmental contaminants from OTU tables, using the prevalence method with two sets of controls (cutoff 0.8 for each): mandibular bone from the sampled individuals in this study and previously published bone samples from Bronze Age Mongolia (Fellows Yates et al., 2021; Jeong et al., 2018), and laboratory extraction and library preparation blanks .

Preservation assessment

A genus-level OTU table was used as input for SourceTracker v.1.0.1 (Knights et al., 2011). Included were also comparative samples from published shotgun microbiome studies, including 10 non-industrialized gut samples (Obregon-Tito et al., 2015; Rampelli et al., 2015), 11 industrialized gut samples (Gevers et al., 2012; Sankaranarayanan et al., 2015), 10 skin samples (Oh et al., 2016), 11 subgingival and 10 supragingival plaque samples (Gevers et al., 2012), 10 archaeological bone samples (Fellows Yates et al., 2021), 10 modern dental calculus samples (Fellows Yates et al., 2021) and 10 archaeological sediment samples (Slon et al., 2017). During the SourceTracker analysis, the samples were rarefied to 10,000 reads, with a training data rarefaction of 5,000. A principal component analysis was conducted on summarized genus level read counts of all samples, blanks and sources (including an additional 9 modern dental calculus samples). Multiplicative zero replacement was conducted using the R-package zCompositions v. 1.3.4 (Palarea-Albaladejo & Martín-Fernández, 2015) and the data was CLR-transformed (Gloor et al., 2017). The non-human DNA sequences were also mapped to the *Tannerella forsythia* representative genome (strain 9212) using EAGER v. 1.92.38 as described above. The output from DamageProfiler v. 0.3.10 (Neukamm et al., 2021) was used to visualize damage curves for the samples, and fragment length was extracted from the output table from EAGER.

Community composition

Analyses of community composition were conducted on the MALT taxon tables, where putative contaminants had been removed, following recommendations for compositional data (Gloor et al., 2017). Significant differences in community composition of samples in selected metadata groups were tested using a PERMANOVA with the R-package vegan v. 2.5.6 (Oksanen et al., 2019), using euclidean distance and 9999 permutations, and individuals as strata when needed.

A PCA was conducted as described above. Alpha diversity was analyzed using a species-level OTU table, and Shannon Index and Inverse Simpson Index were computed using the R package microbiome v. 1.8.0 (Lahti & Shetty, 2012).

Differential abundance

Differential abundance of species was calculated using Songbird v1.0.1 (Morton et al., 2019) (--formula "Jawbone+ToothSurface+ToothPosition+DepositMass_scaled+Individual", --epochs 10000 and --differential-prior 0.5). Tensorboard v. 1.14.0 was used for model checking. Input was a species-level OTU table, where putative contaminants were removed. Further, taxa present in fewer than three samples per individual were removed, and thereafter taxa absent in one or more of the individuals. This stringent filtering was applied in order to avoid any potential remaining contaminants or mismapping to influence the results. Two separate analyses were conducted, one without occlusal samples, and one including occlusal samples.

Functional analysis

The functional profiles of the microbial communities were extracted from the non-human DNA sequences using HUMAnN v. 2.8.0 (Franzosa et al., 2018), using the CocoPhlAn nucleotide database and the UniRef90 protein database. The output was normalized to copies per million, and translated into KEGG orthologies. Gene families were analyzed, without taking into account species assignments, and putative contaminants were removed from the dataset using decontam as described above (threshold 0.5 for both blanks and bones). A PCA was conducted, and drivers of variation identified using PERMANOVA, all as described above for community composition.

Human reads

In order to investigate the amount of host human DNA in the samples, while controlling for contaminating human DNA, the raw reads were mapped to the human HG19 genome as described above, with the exception of filtering for mapping quality (-q 37). Duplicates were removed using DeDup v. 0.12.2 (Peltzer et al., 2016), and the reads were filtered for a PMD (post-mortem damage) score of 3 using PMDtools v.0.6 (Skoglund et al., 2014), thereby only retaining damaged ancient reads. This is likely an underrepresentation of the number of ancient reads, as not all DNA fragments will have damage. However, assuming a consistent rate of postmortem damage accumulation over the dental arcade, the bias will be even across all sampling sites, and the patterns of damaged reads can be assumed to also represent patterns of total endogenous human reads. It was noted that occlusal samples generally have a higher percentage damage than other samples, and were therefore excluded from this analysis, as

they break the assumption of equal damage. Deposit mass was accounted for in the analysis, as a positive correlation was found between deposit mass and DNA damage.

Plant DNA

During preliminary eukaryotic screening of the dental calculus samples, it was observed that the samples contain a considerable amount of DNA mapping to grapevine (*Vitis vinifera*). To further explore this pattern, the complete experimental dataset of dental calculus, mandibular bone controls, and blanks were mapped to the grapevine representative genome (GCA_000003745.2 12X) using EAGER as described above, with mapping quality set to 37. Damage profiles, specifically cytosine to thymine (C to T) transitions typical for ancient DNA, were created using DamageProfiler v.0.3.10 (Neukamm et al., 2021).

General statistics

Unless otherwise stated, data was processed in R v. 3.6.1 (R Core Team, 2019), using packages tidyverse 1.3.0 (Wickham et al., 2019), ggpubr v.0.3.0 (Kassambara, 2018), readxl v.1.3.1 (Wickham & Bryan, 2019), janitor v.2.0.1 (Firke, 2018), and ggeffects v.0.14.3 (Lüdecke, 2018). In order to investigate patterns across the dentition, linear mixed-effects models (LME) were fitted to the variables in question using lme4 v.1.1.23 (Bates et al., 2015), with the individual as the random effect when required. Model selection was performed with ANOVA using lmerTest v.3.1.2 (Kuznetsova et al., 2017) and Box-Cox transformations identified using MASS v.7.3.51.4 (Venables & Ripley, 2002). Explanatory variables in all tests are: jawbone (mandible/maxilla), tooth surface (lingual/buccal/interproximal/occlusal), tooth position (anterior/posterior), and mass of the original calculus deposit (scaled and centered continuous variable). Incisors and canines are treated as anterior teeth; premolars and molars were treated as posterior teeth. Unless otherwise noted, occlusal calculus, which was only obtained from a single individual, was excluded from most analyses because these samples were found to break the assumption of homogeneous distribution of variance (euclidean distances, ANOVA, p=0.001). 2D illustration of DNA yield, human DNA, and environmental contamination across the dental arcade was performed with ‘ili (Protsyuk et al., 2018), and can be accessed at https://tinyurl.com/eyjcs674. All R Markdown files have been deposited at https://github.com/ZandraFagernas/dental_arcade.

Results

Preservation and authentication

Total DNA yield from a sample, normalized by the mass of the dental calculus sample used for DNA extraction, may vary depending on preservation and organic matter content of the sample, and may bias downstream taxonomic profiles (Fagernäs et al., 2020). Occlusal samples were excluded from this analysis, as it was noted during sampling that their consistency was different from all other samples. Using linear mixed effects modeling, we tested whether tooth surface, tooth position, jawbone or deposit mass influenced the mass-normalized DNA yield from a sample. We found that none of these factors outperformed the null model (LME, individual as random effect), and therefore normalized DNA yield cannot be predicted by these variables (Figure 2A).

Prior to oral microbiome analysis, the archaeological dental calculus in this study was first evaluated for preservation and authenticity of the ancient oral microbiome. This is important because poor dental calculus preservation and contamination with environmental microbes can bias or interfere with downstream analyses. A PCA on genus level read counts shows that all the archaeological dental calculus samples cluster together with modern dental plaque samples, and are clearly separated from archaeological bone, gut and sediment samples (Figure 2D). To further assess preservation of the dental calculus samples, the contribution of different source environments to the composition of the samples was estimated using SourceTracker (Knights et al., 2011). All samples were estimated to have a majority contribution from oral microbiome sources, indicating good preservation of the oral microbiome (Figure S2). Some samples were estimated to have a minor contribution from the skin microbiome. Minor estimated contributions from the gut microbiome and sediment are also present, but are expected because gut and oral taxa are similar and can be difficult to distinguish using short read data, and because archaeological samples typically contain some soil contamination, even after washing. After taking these factors into consideration, all dental calculus samples were determined to be sufficiently well preserved for inclusion in downstream analyses.

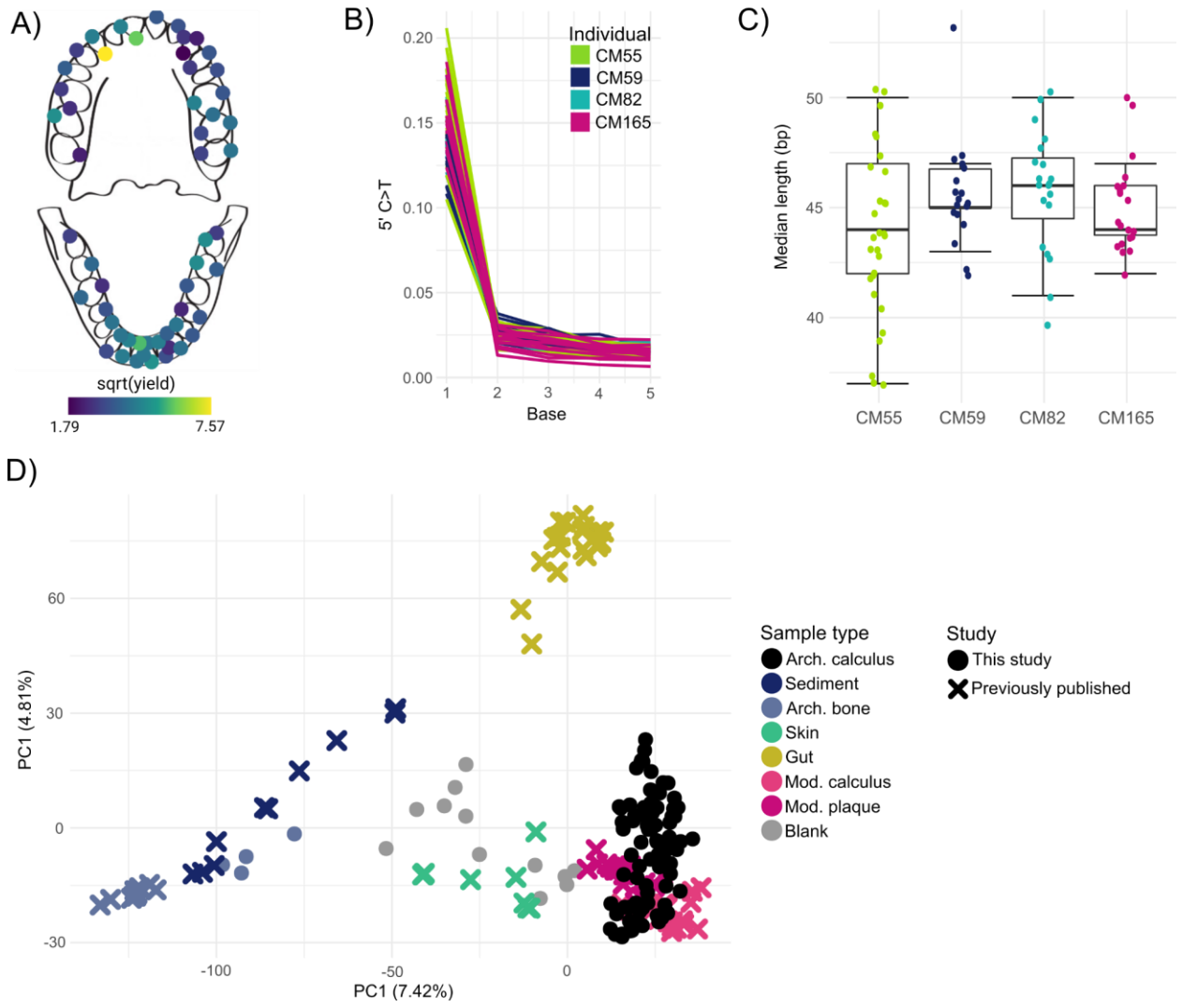


Figure 2. Preservation assessment of dental calculus samples. (A) normalized DNA yield (in ng DNA per mg calculus) across the dental arcade averaged across individuals. (B) C to T transitions at the 5' end of DNA fragments aligning to *Tannerella forsythia*, consistent with ancient DNA. Note that the sharp drop from the first to the second base is due to treatment with uracil-DNA-glycosylase. (C) DNA aligning to *T. forsythia* has short median fragment lengths, consistent with ancient DNA. (D) PCA on genus level read counts of samples from this study, before removing putative contaminants; dental calculus from this study forms a cluster overlapping with modern plaque and calculus, indicating good oral microbiome preservation.

We next assessed DNA damage patterns in the dental calculus as an indicator of authenticity. DNA from archaeological samples accumulates specific forms of damage over time, which can be seen as C to T transitions at the ends of DNA fragments and a high degree of DNA fragmentation (Dabney et al., 2013). We generated a damage plot for fragments mapping to the

prevalent oral bacterium *Tannerella forsythia* (Figure 2B), and all four individuals exhibit damage patterns typical for ancient DNA that has undergone partial UDG-treatment (Rohland et al., 2015). The fragment length distributions of reads mapping to *T. forsythia* show that most samples have a median length <50 bp, as is expected for ancient samples (Figure 2C). Thus, taken together, the microbial DNA present within the dental calculus of the four Chalcolithic individuals in this study is consistent with an ancient and endogenous oral microbiome.

Community composition

To determine whether local environmental and spatial variables along the dental arcade influence microbial community composition, we analyzed patterns of variation with the dental calculus samples. We found that the main driver of variation in a genus-level PCA was the individual from whom the sample originated (PERMANOVA, $p<0.001$, $R^2=0.14$; Figure 3A), indicating that the main differences in community composition are found between individuals. When controlling for the variation introduced by the individual, the mass of the calculus deposit and tooth position (anterior *vs*. posterior) were also found to be significant drivers of variation in community composition (PERMANOVA, individual as strata, $p=0.040$ and $R^2=0.024$ for mass; $p=0.021$ and $R^2=0.032$ for tooth position; Figure 3B). However, although the differences are statistically significant in this study, when doing population-level comparisons they are unlikely to cause biases, as the R^2 values are very low.

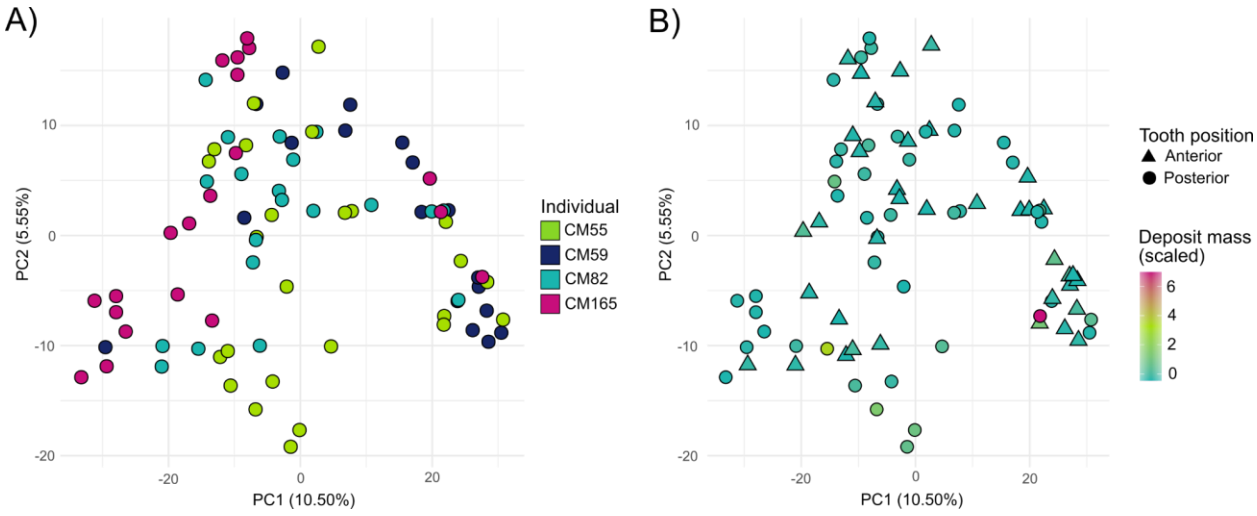


Figure 3. PCA on genus-level read counts. (A) All dental calculus samples plotted together, and coloured by individual. (B) Same data as A, but with samples coloured by initial deposit mass (scaled variable) and shapes representing tooth position.

We next examined alpha diversity within the dataset. Alpha diversity is a measure of how rich in species the community in a certain sample is, which may be of importance when selecting samples for a community composition study. Using the inverse Simpson Index, the mass of the original calculus deposit was found to be a significant predictor of diversity in the dental calculus samples (LME, individual as random effect, $p=0.025$), with diversity slightly increasing with deposit mass (Figure S3). In contrast, the null model fits the Shannon Index best, indicating that alpha diversity does not vary across the oral cavity for any of the tested variables. The Simpson index takes into account evenness, and is less influenced by rare species than the Shannon Index, indicating that the generally large number of rare species in archaeological dental calculus may erode any spatial patterns in alpha diversity.

Differential taxonomic abundance

Due to different local environmental conditions in different areas of the oral cavity, small differences in microbial composition have been reported across the dentition in present-day dental plaque (Haffajee et al., 2009; Proctor et al., 2018; Simon-Soro & Tomás, 2013). It is, however, not known if such patterns can be detected in archaeological samples, after both biofilm maturation during life and postmortem degradation over time. Here, we find that there are slight taxonomic differences with respect to tooth surface and initial deposit mass. First, we observe differences in taxa between anterior (incisors and canines) and posterior (premolar and molars) teeth, where the taxa that are more abundant in the anterior teeth are more often aerobic or facultatively anaerobic, while the taxa that are most associated with posterior teeth are anaerobic (Figure 4A). Second, interproximal spaces seem to be enriched in species belonging to the genera *Methanobrevibacter* and *Olsenella* (Figure 4B), which are both acid tolerant anaerobes. Finally, the species *Actinobaculum* sp. oral taxon 183 and *Fusobacterium* sp. oral taxon 203 are found at a higher abundance in low mass dental calculus deposits, as compared to high mass deposits (Figure 4C). However, little is known about the physiology or role in the dental plaque biofilm of these taxa. Fusobacteria are generally secondary colonizers in the dental plaque biofilm, and bind to several other bacterial taxa (Kolenbrander, 1988). A 2D model showing the spatial distributions of the taxa in Figure 4A-C across the dentition can be found at https://tinyurl.com/eyjcs674.

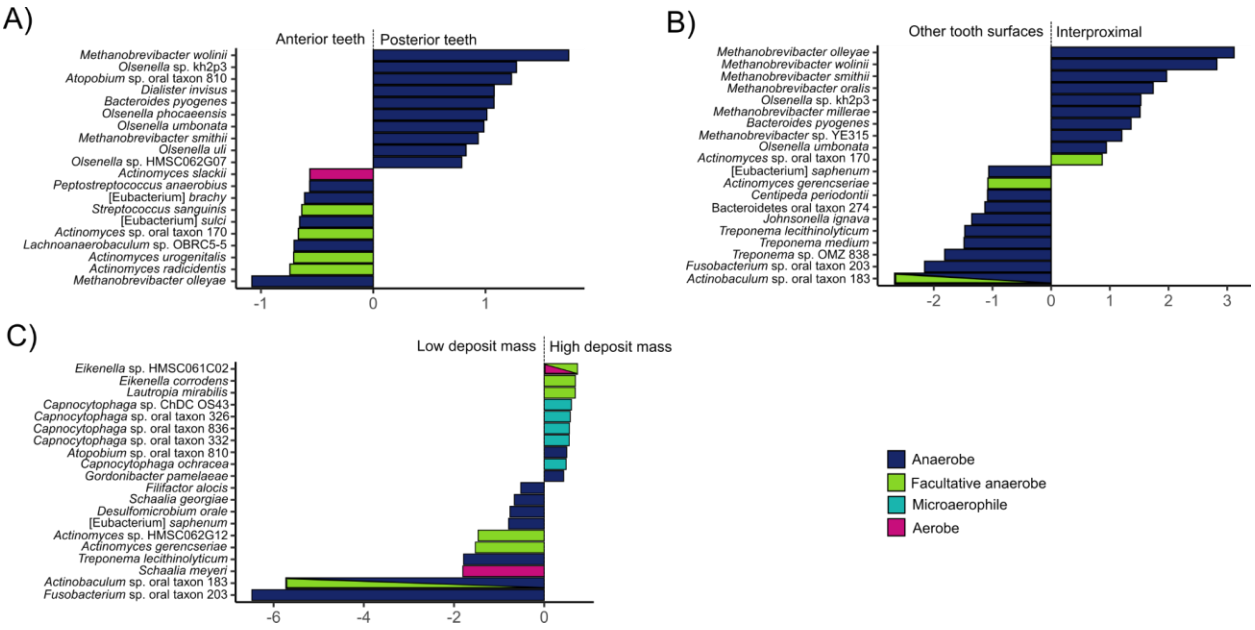


Figure 4. Differential abundance of species across the oral cavity. A) Species associated with posterior (premolars and molars) vs. anterior (incisors and canines) teeth. B) Species associated with interproximal spaces vs. all other tooth sites. C) Species associated with high vs. low initial deposit mass. Only the top ten taxa most associated with each factor are shown.

Functional profile

In addition to their taxonomic composition, microbial communities may also differ in their gene content, and therefore functional potential. It has been seen that although microbial community composition may be similar between individuals, they can differ in the functional potential of the microbiome (Fellows Yates et al., 2021). To evaluate whether there are potential functional differences across the dental arcade, we analyzed the genes present in the dental calculus metagenomes. In total, 2,791 gene families were identified in the dataset, after removing putative contaminants that were identified from blanks and bone samples. The individual was found to be the strongest driver of variation (PERMANOVA, p=0.001 and R^2=0.12; Figure 5), and after accounting for this, tooth surface (p=0.020, R^2=0.049), tooth position (p=0.039, R^2=0.029), and deposit mass (p=0.013, R^2=0.036) were found to significantly drive functional variation (PERMANOVA, individual as strata). However, these factors explain only a very minor part of the variation, as can be seen by the low R^2 values. It should also be noted that the tooth surface variable breaks the assumption of homogeneity of variance for this analysis, which may affect the results of the PERMANOVA.

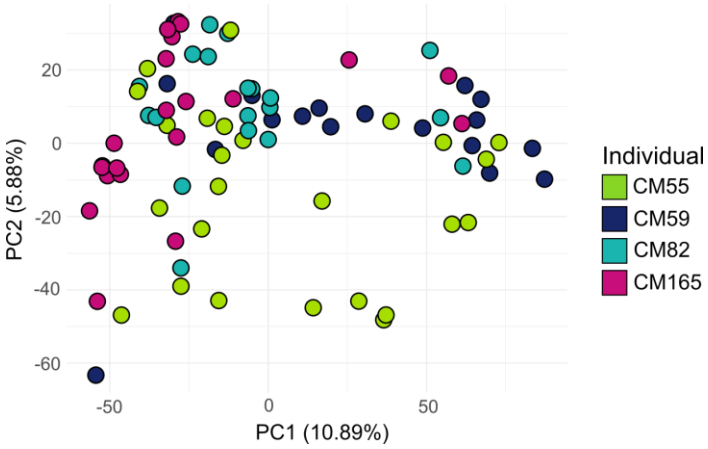


Figure 5. Functional profile of dental calculus samples. PCA of gene families, normalized to copies per million, with colour indicating individual.

Human genetic content

Although dental calculus generally contains a very low proportion of human DNA (Mann et al., 2018), different enrichment approaches have been used to increase the human DNA fraction enough to study the human genome (Ozga et al., 2016; Ziesemer et al., 2019). Human DNA from dental calculus is mainly derived from a single individual, the host (Ozga et al., 2016). Human DNA may in theory be differentially incorporated into dental calculus across the dental arcade, depending on salivary flow, inflammation, or disease, among other factors. We investigated the presence and relative abundance of ancient human DNA in our samples to assess potential spatial patterning of human host DNA in calculus. To focus our analysis on host ancient DNA, we restricted our analysis to only DNA fragments with C to T DNA damage. As a slight positive correlation was found between deposit mass and damage (Figure S4), deposit mass was accounted for in this analysis. We found that the best fitting model for predicting the proportion of human reads in the dental calculus samples is a null model, indicating that the distribution of human DNA in dental calculus does not significantly vary according to tooth surface, tooth position, or jawbone (LME, deposit mass as random effect; Figure 6A).

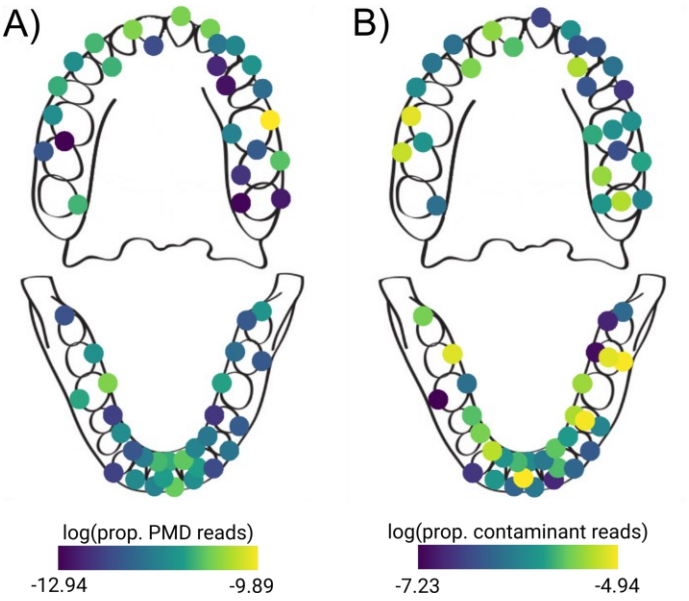


Figure 6. Distribution of ancient human reads and environmental contaminant reads across the dental arcade. A) Proportion of human reads with post-mortem damage (PMD) of the total number of sequenced reads, averaged across individuals for each sampling location. B) Proportion of reads that stem from putative environmental contaminant taxa out of all reads, averaged across all four individuals for each sampling location.

Postmortem environmental colonization

Whether contamination by infiltration of environmental microbes from the burial context is introduced in a non-random way across the oral cavity is not known. Because different properties of calculus across the dentition could make certain regions more susceptible to external colonization, we investigated the distribution of environmental contaminant reads in our samples. On species level, a total of 215 taxa (out of 556 taxa) were identified as putative environmental contaminants in the entire dataset, using the bone samples from the mandibles as a proxy for colonizing microbes from the burial ground. This analysis was performed at the species level, as it is possible for taxa in the same genus to grow in different habitats. We tested whether the distribution of these species across the dentition was influenced by tooth surface, tooth position, jawbone or deposit mass using linear mixed effects modeling; however, we found that none of these factors outperformed the null model. Therefore, it appears that contamination is randomly distributed across the dental arcade (LME, individual as random effect; Figure 6B).

Occlusal calculus

The occlusal dental calculus analyzed in this study differed from the calculus from other tooth surfaces in several ways, and was therefore excluded from most analyses. During sampling, occlusal calculus was found to have a different consistency from the other calculus, being less dense and having less structural integrity. Occlusal calculus was found to have a higher amount of DNA damage than other calculus. For reads mapping to *Tannerella forsythia*, a model including tooth surface and deposit mass best predicted damage at the 1st base at the 5' end of the fragment (LME, individual as random effect, p=0.018), with occlusal samples having higher levels of damage than other samples (Figure S4). Further, occlusal calculus samples broke the assumption of homogeneity of dispersion for the community composition, which may be due to the fact that they were only collected from a single individual, and from only posterior teeth on the same side of the mouth. Overall, we found that despite forming on posterior teeth, occlusal calculus samples are somewhat enriched in aerotolerant species, possibly due to their more exposed location on the tooth, compared to the lingual and labial surfaces of the posterior teeth that directly abut the tongue and buccal mucosa, respectively (Figure 8).

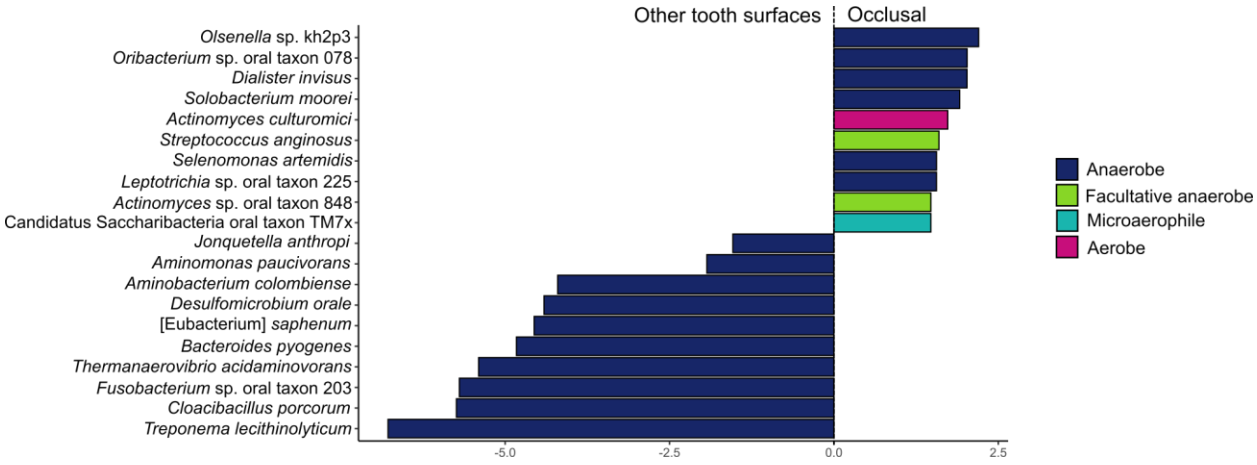


Figure 8. Differential abundance of taxa in occlusal samples compared to others. Only the top ten taxa are shown, and the bars are coloured by aerotolerance of the taxa.

Plant DNA

Ancient dental calculus is a potentially valuable source of information about ancient diets, as it is possible to directly study diet-related biomolecules and microfossils incorporated in the calculus during an individual's lifetime. Researchers have previously attempted to identify dietary sources using DNA from dental calculus (Warinner et al., 2014; Weyrich et al., 2017), an approach that also has many difficulties due to the exceptionally low number of dietary DNA

sequences typically found in dental calculus (Mann et al., 2020). The dental calculus samples in this study contained trace amounts of plant DNA fragments (between 170-1578 reads per individual, or 0.002-0.011% of total reads) mapping to the grapevine (*Vitis vinifera*) genome, which is currently and historically widely cultivated in the region. However, it was noticed that similar numbers of grapevine reads (205-2119 reads, or 0.003-0.034%) were also recovered in the mandibular bone control samples (Figure 7A). Both sets of reads were found to have C to T damage typical of ancient DNA (7-9% for bones and 3-14% for calculus; Figure 7B), but at lower levels than observed for the oral bacterium *T. forsythia* (11-21%; Figure 2B). The presence of grapevine reads in both dental calculus and bone, together with the lower amount of damage, suggests a likely postmortem origin of the grapevine DNA. However, a dietary origin of the grapevine DNA cannot completely be excluded, as a wild variety has been documented in the region since the Palaeolithic (Aura et al., 2005; Iriarte-Chiapusso et al., 2017).

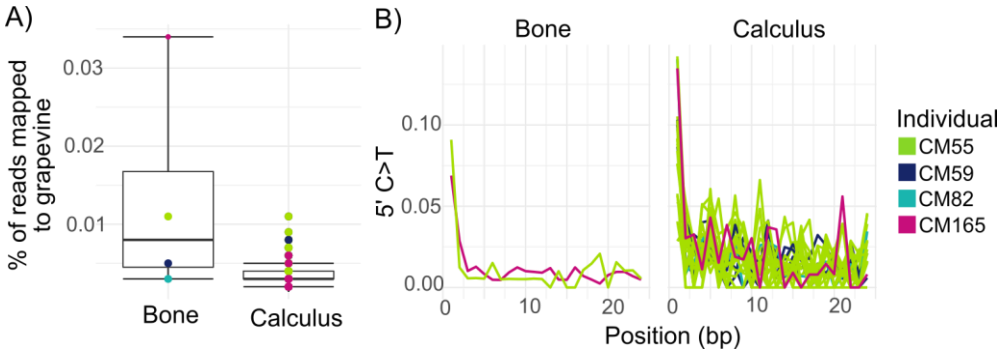


Figure 7. Presence of grapevine DNA in bones and dental calculus samples. A) The percentage of reads that aligned to the grapevine genome per sample. B) C to T miscoding lesions at the 5' end of the read, for each sample with >500 reads aligning to grapevine.

Discussion

A potentially uneven distribution of microbes in microbiomes can cause biases in downstream analyses if spatial variation is not taken into account during sampling design and data interpretation. Archaeological dental calculus provides a valuable window into the evolution of the oral microbiome, but to date it has not been clear to what degree microbial taxa are spatially patterned across the dentition and, thus, to what degree sampling strategy might impact comparative studies of dental calculus microbial communities. The results of present-day dental plaque studies cannot be directly applied to dental calculus because the two substrates reflect different levels of biofilm maturity and have slightly different composition (Velsko et al., 2019), and in previous studies of spatial variation in archaeological dental calculus, which sampled diverse individual teeth from a large number of individuals (Farrer et al., 2018), potentially

confounding factors such as individual, temporal, environmental, and taphonomic differences were not controlled for. Here, we have presented a systematic study of intra-individual variation in archaeological dental calculus by focusing on intensive, comprehensive sampling of the dentitions of four contemporaneous individuals from the same burial context.

Overall, we find that although there are small differences in the spatial distribution of anaerobic and aerotolerant taxa, as well as minor associations between taxonomic composition and initial calculus deposit size, these factors account for very little of the overall microbial and functional variation within dental calculus. Spatial patterns in the oral microbiome that have been identified in studies of modern dental plaque (Haffajee et al., 2009; Proctor et al., 2018; Simon-Soro & Tomás, 2013) are not obvious in this study. Such patterns may have been present during life but were subsequently lost over time due to taphonomy, or these patterns may simply not be present in calculus. Although taphonomic processes, such as C to T damage accumulation and DNA fragmentation, as well as postmortem colonization of the body by environmental taxa, may obscure oral microbiome spatial patterns, we did not find these factors to correlate with the microbial patterns we observed. A study of modern dental calculus that investigates species spatial patterning will be needed to determine if the patterns observed in dental plaque are maintained as the biofilm matures and calcifies into dental calculus.

Although this study investigated a small number of individuals from a single archaeological site, the purpose of this study design was to limit the number of potentially confounding factors, such as different sample ages, different burial conditions, and different storage and handling practices after excavation. Microbial spatial patterning may differ in other populations or at other archaeological sites, and this warrants further investigation. However, as the species profiles of human dental calculus appear to be more consistent across time, space, and health status than dental plaque (Fellows Yates et al., 2021; Velsko et al., 2019), it is possible that any variation will be very minor.

Although we observed few spatial patterns in archaeological dental calculus, we find that occlusal calculus may represent a special exception. Dental calculus rarely accumulates on the occlusal surfaces of teeth, in part due to the abrasive forces of mastication, and large deposits of occlusal calculus are generally indicative of physiological injury or dysfunction. Here, only one individual had occlusal calculus, but this calculus had a distinct texture, higher DNA damage, and different levels of taxonomic dispersion compared to other dental calculus in the study, even from the same individual. Although further research on a larger number of individuals is necessary, occlusal calculus is likely not representative of oral microbiome communities, and therefore should be avoided in comparative studies of microbial variation across individuals.

Beyond microbes, dental calculus is also valuable because it entraps dietary and other

environmental debris during life, and thus can provide clues about the foods and activities of past societies (Hardy et al., 2009; Leonard et al., 2015; Power et al., 2015; Radini et al., 2017). Although dietary proteins have been shown to preserve within dental calculus (Hendy et al., 2018; Scott et al., 2021; Wilkin et al., 2020), the metagenomic recovery of dietary DNA from calculus has yielded more equivocal results (Mann et al., 2020). The recovery and authentication of eukaryotic DNA in metagenomic datasets is not trivial due to complicating factors such as the very low number of non-host eukaryotic DNA fragments typically found in dental calculus and the problem of microbial contamination in eukaryotic reference genomes, which can lead to false positives (Mann et al., 2020). Here, we show that an additional complicating factor may be contamination from the environment, and specifically from nearby agricultural fields, or other types of postmortem inclusion in remains. It is therefore advisable to include environmental controls, such as bone or sediment samples, in metagenomic studies of diet. Another authentication aid may be the use of complementary dietary identification methods, such as microfossil analysis or palaeoproteomics. Through proteomic analyses, for example, it is possible to deduce the part of an organism from which the biomolecules originate, such as seed proteins from plant seeds, or milk proteins from dairy products. Combining methods may thus aid researchers in establishing the plausibility of a given organism being incorporated into dental calculus as a food as opposed to environmental contamination.

To conclude, we find that in most applications a single sample of archaeological dental calculus can be used to represent an individual in comparative studies of the ancient oral microbiome. The use of a single sample instead of multiple samples, either pooled or studied separately, reduces the destructive demands on this finite archaeological material. However, as there are minor spatial patterns present, care should be taken to record the sampling location within the oral cavity for each dental calculus sample, whenever possible. This makes it possible to later reevaluate findings if systematic biases are suspected.

Accession numbers

Genetic data have been deposited in the ENA under the accession PRJEB46022.

Acknowledgements

This work was supported by the Werner Siemens Stiftung (“Paleobiotechnology” to C.W.), the European Research Council (STG–677576 “HARVEST” to A.G.H.), the Generalitat Valenciana (GV/2015/060 and CIDEGENT/2019/061 to DCSG), the Spanish government (PID2019-

111207GA-I00 and EUR2020-112213 to DCSG), and the Max Planck Society. The authors thank James Fellows Yates for assistance with data analysis.

References

Aas, J. A., Paster, B. J., Stokes, L. N., Olsen, I., & Dewhirst, F. E. (2005). Defining the normal bacterial flora of the oral cavity. *Journal of Clinical Microbiology*, *43*(11), 5721–5732.

Aron, F., Hofman, C., Fagernäs, Z., Velsko, I., Skourtanioti, E., Brandt, G., & Warinner, C. (2020). Ancient DNA extraction from dental calculus v1. *protocols.io*. ZappyLab, Inc.

Aron, F., Neumann, G. U., & Brandt, G. (2020). Half-UDG treated double-stranded ancient DNA library preparation for Illumina sequencing v1. *protocols.io*. ZappyLab, Inc.

Aura, J. E., Carrión, Y., Estrelles, E., & Jordà, G. P. (2005). Plant economy of hunter-gatherer groups at the end of the last Ice Age: plant macroremains from the cave of Santa Maira (Alacant, Spain) ca. 12000–9000 b.p. *Vegetation History and Archaeobotany*, *14*(4), 542–550.

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Softwarem, 67*(1), 1-48.

Brothwell, D. R. (1972). *Digging up bones: The excavation, treatment and study of human skeletal remains*. British Museum.

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

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

Davis, N. M., Proctor, D. M., Holmes, S. P., Relman, D. A., & Callahan, B. J. (2018). Simple statistical identification and removal of contaminant sequences in marker-gene and metagenomics data. *Microbiome*, *6*(1), 226.

Díaz Navarro, S., Haber Uriarte, M., Avilés Fernández, A., & Lomba Maurandi, J. (2019). Aproximación paleodemográfica a una comunidad calcolítica mediterránea. El enterramiento colectivo de Camino del Molino (Caravaca de la Cruz, Murcia). *XXI Congreso Internacional de La Sociedad Española de Antropología Física (Granada, 24-26 de Julio de 2019)*.

Ding, T., & Schloss, P. D. (2014). Dynamics and associations of microbial community types across the human body. *Nature*, *509*(7500), 357–360.

Dobney, K., & Brothwell, D. (1987). A method for evaluating the amount of dental calculus on teeth from archaeological sites. *Journal of Archaeological Science*, *14*(4), 343–351.

Eren, A. M., Borisy, G. G., Huse, S. M., & Mark Welch, J. L. (2014). Oligotyping analysis of the human oral microbiome. *Proceedings of the National Academy of Sciences of the United States of America*, *111*(28), E2875–E2884.

Fagernäs, Z., García-Collado, M. I., Hendy, J., Hofman, C. A., Speller, C., Velsko, I., & Warinner, C. (2020). A unified protocol for simultaneous extraction of DNA and proteins from archaeological dental calculus. *Journal of Archaeological Science*, *118*, 105135.

Farrer, A. G., Bekvalac, J., Redfern, R., Gully, N., Dobney, K., Cooper, A., & Weyrich, L. S. (2018). Biological and cultural drivers of oral microbiota in Medieval and Post-Medieval London, UK. *bioRxiv* 343889.

Fellows Yates, J. A., Velsko, I. M., Aron, F., Posth, C., Hofman, C. A., Austin, R. M., Parker, C. E., Mann, A. E., Nägele, K., Arthur, K. W., Arthur, J. W., Bauer, C. C., Crevecoeur, I., Cupillard, C., Curtis, M. C., Dalén, L., Díaz-Zorita Bonilla, M., Díez Fernández-Lomana, J. C., Drucker, D. G., Escribano Escrivá, E., Francken, M., Gibbon, V. E., González Morales, M. R., Grande Mateu, A., Harvati, K., Henry, A. G., Humphrey, L., Menéndez, M., Mihailović, D., Peresani, M., Rodríguez Moroder, S., Roksandic, M., Rougier, H., Sázelová, S., Stock, J. T., Straus, L. G., Svoboda, J., Teßmann, B., Walker, M. J., Power, R. C., Lewis, C. M., Sankaranarayanan, K., Guschanski, K., Wrangham, R. W., Dewhirst, F. E., Salazar-García, D. C., Krause, J., Herbig, A., & Warinner, C. (2021). The evolution and changing ecology of the African hominid oral microbiome. *Proceedings of the National Academy of Sciences of the United States of America*, *118*(20), e2021655118.

Firke, S. (2018). *Janitor: Simple tools for examining and cleaning dirty data*. R package version.

Franzosa, E. A., McIver, L. J., Rahnavard, G., Thompson, L. R., Schirmer, M., Weingart, G., Lipson, K. S., Knight, R., Caporaso, J. G., Segata, N., & Huttenhower, C. (2018). Species-level functional profiling of metagenomes and metatranscriptomes. *Nature Methods*, *15*(11), 962–968.

Gevers, D., Knight, R., Petrosino, J. F., Huang, K., McGuire, A. L., Birren, B. W., Nelson, K. E., White, O., Methé, B. A., & Huttenhower, C. (2012). The Human Microbiome Project: a community resource for the healthy human microbiome. *PLoS Biology*, *10*(8), e1001377.

Gloor, G. B., Macklaim, J. M., Pawlowsky-Glahn, V., & Egozcue, J. J. (2017). Microbiome Datasets Are Compositional: And This Is Not Optional. *Frontiers in Microbiology*, *8*, 2224.

Haber-Uriarte, M., Avilés-Fernández, A., & Lomba-Maurandi. (2011). Estudio antropológico preliminar de los restos humanos calcolíticos del enterramiento múltiple de Camino del Molino (Carava de la Cruz, Murcia). In D. Turbón, L. Fañanás, C. Rissech, A. Rosa (Ed.), *Biodiversidad Humana y Evolución* (pp. 236–242). Barcelona.

Haffajee, A. D., Teles, R. P., Patel, M. R., Song, X., Yaskell, T., & Socransky, S. S. (2009). Factors affecting human supragingival biofilm composition. II. Tooth position. *Journal of Periodontal Research*, *44*(4), 520–528.

Hardy, K., Blakeney, T., Copeland, L., Kirkham, J., Wrangham, R., & Collins, M. (2009). Starch granules, dental calculus and new perspectives on ancient diet. *Journal of Archaeological Science*, *36*(2), 248–255.

Hendy, J., Warinner, C., Bouwman, A., Collins, M. J., Fiddyment, S., Fischer, R., Hagan, R., Hofman, C. A., Holst, M., Chaves, E., Klaus, L., Larson, G., Mackie, M., McGrath, K., Mundorff, A. Z., Radini, A., Rao, H., Trachsel, C., Velsko, I. M., & Speller, C. F. (2018). Proteomic evidence of dietary sources in ancient dental calculus. *Proceedings of the Royal Society B: Biological Sciences*, *285*(1883).

Herbig, A., Maixner, F., Bos, K. I., Zink, A., Krause, J., & Huson, D. H. (2016). MALT: Fast alignment and analysis of metagenomic DNA sequence data applied to the Tyrolean Iceman. *bioRxiv* 050559.

Huson, D. H., Beier, S., Flade, I., Górska, A., El-Hadidi, M., Mitra, S., Ruscheweyh, H.-J., & Tappu, R. (2016). MEGAN Community Edition - Interactive Exploration and Analysis of Large-Scale Microbiome Sequencing Data. *PLoS Computational Biology*, *12*(6), e1004957.

Iriarte-Chiapusso, M. J., Ocete-Pérez, C. A., Hernández-Beloqui, B., & Ocete-Rubio, R. (2017). *Vitis vinifera* in the Iberian Peninsula: A review. *Plant Biosystems - An International Journal Dealing with All Aspects of Plant Biology*, *151*(2), 245–257.

Jeong, C., Wilkin, S., Amgalantugs, T., Bouwman, A. S., Taylor, W. T. T., Hagan, R. W., Bromage, S., Tsolmon, S., Trachsel, C., Grossmann, J., Littleton, J., Makarewicz, C. A., Krigbaum, J., Burri, M., Scott, A., Davaasambuu, G., Wright, J., Irmer, F., Myagmar, E., Boivin, N., Robbeets, M., Rühli, F. J., Krause, J., Frohlich, B., Hendy, J., & Warinner, C. (2018). Bronze Age population dynamics and the rise of dairy pastoralism on the eastern Eurasian steppe. *Proceedings of the National Academy of Sciences of the United States of America*, *115*(48), E11248–E11255.

Jin, Y., & Yip, H. K. (2002). SUPRAGINGIVAL CALCULUS: FORMATION AND CONTROL. *Critical Reviews in Oral Biology and Medicine, 13*(5), 426–441.

Kassambara, A. (2018). *ggpubr: "ggplot2" Based Publication Ready Plots*.

Kazarina, A., Petersone-Gordina, E., Kimsis, J., Kuzmicka, J., Zayakin, P., Griškjans, Ž., Gerhards, G., & Ranka, R. (2021). The Postmedieval Latvian Oral Microbiome in the Context of Modern Dental Calculus and Modern Dental Plaque Microbial Profiles. *Genes*, *12*(2).

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

Knights, D., Kuczynski, J., Charlson, E. S., Zaneveld, J., Mozer, M. C., Collman, R. G., Bushman, F. D., Knight, R., & Kelley, S. T. (2011). Bayesian community-wide culture-independent microbial source tracking. *Nature Methods*, *8*(9), 761–763.

Kolenbrander, P. E. (1988). Intergeneric coaggregation among human oral bacteria and ecology of dental plaque. *Annual Review of Microbiology*, *42*, 627–656.

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest Package: Tests in Linear Mixed Effects Models. *Journal of Statistical Software*, *82*(13), 1-26.

Lahti, L., & Shetty, S. (2012). *microbiome R package*.

Leonard, C., Vashro, L., O'Connell, J. F., & Henry, A. G. (2015). Plant microremains in dental calculus as a record of plant consumption: A test with Twe forager-horticulturalists. *Journal of Archaeological Science: Reports*, *2*, 449–457.

Li, H., & Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* , *25*(14), 1754–1760.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R., & 1000 Genome Project Data Processing Subgroup. (2009). The Sequence Alignment/Map format and SAMtools. *Bioinformatics* , *25*(16), 2078–2079.

Lomba Maurandi, J., López Martínez, M., & Ramos Martínez, F. (2009). Un excepcional sepulcro del Calcolítico: Camino del Molino (Caravaca de la Cruz). *XX Jornadas de Patrimonio Cultural de la Región de Murcia*, 205–219.

Lomba Maurandi, J., López Martínez, M., Ramos Martínez, F., & Avilés Fernández, A. (2009). El enterramiento múltiple, calcolítico, de Camino del Molino(Caravaca, Murcia). Metodología y primeros resultados de un yacimiento excepcional. *TRABAJOS DE PREHISTORIA*, *66*(2).

Lüdecke, D. (2018). ggeffects: Tidy Data Frames of Marginal Effects from Regression Models. *Journal of Open Source Software, 3*(26), 772.

Mann, A. E., Fellows Yates, J. A., Fagernäs, Z., Austin, R. M., Nelson, E. A., & Hofman, C. A. (2020). Do I have something in my teeth? The trouble with genetic analyses of diet from archaeological dental calculus. *Quaternary International, in press.*

Mann, A. E., Sabin, S., Ziesemer, K., Vågene, Å. J., Schroeder, H., Ozga, A. T., Sankaranarayanan, K., Hofman, C. A., Fellows Yates, J. A., Salazar-García, D. C., Frohlich, B., Aldenderfer, M., Hoogland, M., Read, C., Milner, G. R., Stone, A. C., Lewis, C. M., Jr, Krause, J., Hofman, C., Bos, K. I., & Warinner, C. (2018). Differential preservation of endogenous human and microbial DNA in dental calculus and dentin. *Scientific Reports*, *8*(1), 9822.

Mark Welch, J. L., Rossetti, B. J., Rieken, C. W., Dewhirst, F. E., & Borisy, G. G. (2016). Biogeography of a human oral microbiome at the micron scale. *Proceedings of the National Academy of Sciences of the United States of America*, *113*(6), E791–E800.

Meyer, M., & Kircher, M. (2010). Illumina sequencing library preparation for highly multiplexed target capture and sequencing. *Cold Spring Harbor Protocols*, *2010*(6), db.prot5448.

Morton, J. T., Marotz, C., Washburne, A., Silverman, J., Zaramela, L. S., Edlund, A., Zengler, K., & Knight, R. (2019). Establishing microbial composition measurement standards with reference frames. *Nature Communications*, *10*(1), 2719.

Neukamm, J., Peltzer, A., & Nieselt, K. (2021). DamageProfiler: Fast damage pattern calculation for ancient DNA. *Bioinformatics*, btab190.

Obregon-Tito, A. J., Tito, R. Y., Metcalf, J., Sankaranarayanan, K., Clemente, J. C., Ursell, L. K., Zech Xu, Z., Van Treuren, W., Knight, R., Gaffney, P. M., Spicer, P., Lawson, P., Marin-Reyes, L., Trujillo-Villarroel, O., Foster, M., Guija-Poma, E., Troncoso-Corzo, L., Warinner, C., Ozga, A. T., & Lewis, C. M. (2015). Subsistence strategies in traditional societies distinguish gut microbiomes. *Nature Communications*, *6*, 6505.

Oh, J., Byrd, A. L., Park, M., NISC Comparative Sequencing Program, Kong, H. H., & Segre, J. A. (2016). Temporal Stability of the Human Skin Microbiome. *Cell*, *165*(4), 854–866.

Oksanen, J., Blanchet, F. G., Friendly, M., Kindt, R., Legendre, P., McGlinn, D., Minchin, P. R., O'Hara, R. B., Simpson, G. L., Solymos, P., Stevens, M. H. H., Szoecs, E., & Wagner, H. (2019). *vegan: Community Ecology Package*.

Ozga, A. T., Nieves-Colón, M. A., Honap, T. P., Sankaranarayanan, K., Hofman, C. A., Milner, G. R., Lewis, C. M., Jr, Stone, A. C., & Warinner, C. (2016). Successful enrichment and recovery of whole mitochondrial genomes from ancient human dental calculus. *American Journal of Physical Anthropology*, *160*(2), 220–228.

Palarea-Albaladejo, J., & Martín-Fernández, J. A. (2015). zCompositions — R package for multivariate imputation of left-censored data under a compositional approach. *Chemometrics and Intelligent Laboratory Systems*, *143*, 85–96.

Peck, S., & Peck, L. (1993). A time for change of tooth numbering systems. *Journal of Dental Education*, *57*(8), 643–647.

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

Power, R. C., Salazar-García, D. C., Wittig, R. M., Freiberg, M., & Henry, A. G. (2015). Dental calculus evidence of Taï Forest Chimpanzee plant consumption and life history transitions. *Scientific Reports*, *5*, 15161.

Power, R. C., Salazar-García, D. C., Wittig, R. M., & Henry, A. G. (2014). Assessing use and suitability of scanning electron microscopy in the analysis of micro remains in dental calculus. *Journal of Archaeological Science*, *49*, 160–169.

Proctor, D. M., Fukuyama, J. A., Loomer, P. M., Armitage, G. C., Lee, S. A., Davis, N. M., Ryder, M. I., Holmes, S. P., & Relman, D. A. (2018). A spatial gradient of bacterial diversity in the human oral cavity shaped by salivary flow. *Nature Communications*, *9*(1), 681.

Protsyuk, I., Melnik, A. V., Nothias, L.-F., Rappez, L., Phapale, P., Aksenov, A. A., Bouslimani, A., Ryazanov, S., Dorrestein, P. C., & Alexandrov, T. (2018). 3D molecular cartography using LC-MS facilitated by Optimus and 'ili software. *Nature Protocols*, *13*(1), 134–154.

Radini, A., Nikita, E., Buckley, S., Copeland, L., & Hardy, K. (2017). Beyond food: The multiple pathways for inclusion of materials into ancient dental calculus. *American Journal of Physical Anthropology*, *162 Suppl 63*, 71–83.

Rampelli, S., Schnorr, S. L., Consolandi, C., Turroni, S., Severgnini, M., Peano, C., Brigidi, P., Crittenden, A. N., Henry, A. G., & Candela, M. (2015). Metagenome Sequencing of the Hadza Hunter-Gatherer Gut Microbiota. *Current Biology*, *25*(13), 1682–1693.

R Core Team. (2019). R: A language and environment for statistical computing. *R Foundation for Statistical Computing, Vienna, Austria*.

Rohland, N., Harney, E., Mallick, S., Nordenfelt, S., & Reich, D. (2015). Partial uracil-DNA-glycosylase treatment for screening of ancient DNA. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *370*(1660), 20130624.

Sankaranarayanan, K., Ozga, A. T., Warinner, C., Tito, R. Y., Obregon-Tito, A. J., Xu, J., Gaffney, P. M., Jervis, L. L., Cox, D., Stephens, L., Foster, M., Tallbull, G., Spicer, P., & Lewis, C. M. (2015). Gut Microbiome Diversity among Cheyenne and Arapaho Individuals from Western Oklahoma. *Current Biology*, *25*(24), 3161–3169.

Schubert, M., Lindgreen, S., & Orlando, L. (2016). AdapterRemoval v2: rapid adapter trimming, identification, and read merging. *BMC Research Notes*, *9*, 88.

Scott, A., Power, R. C., Altmann-Wendling, V., Artzy, M., Martin, M. A. S., Eisenmann, S., Hagan, R., Salazar-García, D. C., Salmon, Y., Yegorov, D., Milevski, I., Finkelstein, I., Stockhammer, P. W., & Warinner, C. (2021). Exotic foods reveal contact between South Asia and the Near East during the second millennium BCE. *Proceedings of the National Academy of Sciences of the United States of America*, *118*(2).

Simon-Soro, A., & Tomás, I. (2013). Microbial geography of the oral cavity. *Journal of Dental Research, 92*(7):616-21.

Skoglund, P., Northoff, B. H., Shunkov, M. V., Derevianko, A. P., Pääbo, S., Krause, J., & Jakobsson, M. (2014). Separating endogenous ancient DNA from modern day contamination in a Siberian Neandertal. *Proceedings of the National Academy of Sciences of the United States of America*, *111*(6), 2229–2234.

Slon, V., Hopfe, C., Weiß, C. L., Mafessoni, F., de la Rasilla, M., Lalueza-Fox, C., Rosas, A., Soressi, M., Knul, M. V., Miller, R., Stewart, J. R., Derevianko, A. P., Jacobs, Z., Li, B.,

Roberts, R. G., Shunkov, M. V., de Lumley, H., Perrenoud, C., Gušić, I., Kućan, Ž., Rudan, P., Aximu-Petri, A., Essel, E., Nagel, S., Nickel, B., Schmidt, A., Prüfer, K., Kelso, J., Burbano, H. A., Pääbo, S., & Meyer, M. (2017). Neandertal and Denisovan DNA from Pleistocene sediments. *Science*, *356*(6338), 605–608.

Stahl, R., Warinner, C., Velsko, I., Orfanou, E., Aron, F., & Brandt, G. (2019). Illumina double-stranded DNA dual indexing for ancient DNA v1. *protocols.io*. ZappyLab, Inc.

Utter, D. R., Borisy, G. G., Murat Eren, A., Cavanaugh, C. M., & Mark Welch, J. L. (2020). Metapangenomics of the oral microbiome provides insights into habitat adaptation and cultivar diversity. *Genome Biology, 21*(293).

Velsko, I. M., Fellows Yates, J. A., Aron, F., Hagan, R. W., Frantz, L. A. F., Loe, L., Martinez, J. B. R., Chaves, E., Gosden, C., Larson, G., & Warinner, C. (2019). Microbial differences between dental plaque and historic dental calculus are related to oral biofilm maturation stage. *Microbiome*, *7*(1), 102.

Velsko, I. M., & Warinner, C. G. (2017). Bioarchaeology of the human microbiome. *Bioarchaeology International*, *1*(1-2), 86–99.

Venables, W. N., & Ripley, B. D. (2002). *Modern Applied Statistics with S*. Springer.

Warinner, C., Rodrigues, J. F. M., Vyas, R., Trachsel, C., Shved, N., Grossmann, J., Radini, A., Hancock, Y., Tito, R. Y., Fiddyment, S., Speller, C., Hendy, J., Charlton, S., Luder, H. U., Salazar-García, D. C., Eppler, E., Seiler, R., Hansen, L. H., Castruita, J. A. S., Barkow-Oesterreicher, S., Teoh, K. Y., Kelstrup, C. D., Olsen, J. V., Nanni, P., Kawai, T., Willerslev, E., von Mering, C., Lewis Jr, C. M., Collins, M. J., Gilbert, M. T. P., Rühli, F., & Cappellini, E. (2014). Pathogens and host immunity in the ancient human oral cavity. *Nature Genetics*, *46*(4), 336–344.

Warinner, C., Velsko, I., & Fellows Yates, J. A. (2019). Dental calculus field-sampling protocol (Warinner version) v1. *protocols.io*. ZappyLab, Inc.

Weyrich, L. S., Duchene, S., Soubrier, J., Arriola, L., Llamas, B., Breen, J., Morris, A. G., Alt, K. W., Caramelli, D., Dresely, V., Farrell, M., Farrer, A. G., Francken, M., Gully, N., Haak, W., Hardy, K., Harvati, K., Held, P., Holmes, E. C., Kaidonis, J., Lalueza-Fox, C., de la Rasilla, M., Rosas, A., Semal, P., Soltysiak, A., Townsend, G., Usai, D., Wahl, J., Huson, D. H., Dobney, K., & Cooper, A. (2017). Neanderthal behaviour, diet, and disease inferred from ancient DNA in dental calculus. *Nature*, *544*(7650), 357–361.

Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L. D., François, R., Grolemund, G., Hayes, A., Henry, L., Hester, J., Kuhn, M., Pedersen, T. L., Miller, E., Bache, S. M., Müller, K., Ooms, J., Robinson, D., Seidel, D. P., Spinu, V., Takahashi, K., Vaughan, D., Wilke, C., Woo, K., & Yutani, H. (2019). Welcome to the tidyverse. In *Journal of Open Source Software, 4*(43), 1686.

Wickham, H., & Bryan, J. (2019). *readxl: Read Excel Files*

Wilkin, S., Ventresca Miller, A., Taylor, W. T. T., Miller, B. K., Hagan, R. W., Bleasdale, M., Scott, A., Gankhuyg, S., Ramsøe, A., Trachsel, C., Nanni, P., Grossmann, J., Orlando, L., Horton, M., Stockhammer, P., Myagmar, E., Boivin, N., Warinner, C., & Hendy, J. (2020). Dairy pastoralism sustained Eastern Eurasian steppe populations for 5000 years. *Nature Ecology and Evolution*, *4*(3), 346–355.

Ziesemer, K. A., Mann, A. E., Sankaranarayanan, K., Schroeder, H., Ozga, A. T., Brandt, B. W., Zaura, E., Waters-Rist, A., Hoogland, M., Salazar-García, D. C., Aldenderfer, M., Speller, C.,

Hendy, J., Weston, D. A., MacDonald, S. J., Thomas, G. H., Collins, M. J., Lewis, C. M., Hofman, C., & Warinner, C. (2015). Intrinsic challenges in ancient microbiome reconstruction using 16S rRNA gene amplification. *Scientific Reports*, *5*, 16498.

Ziesemer, K. A., Ramos-Madrigal, J., Mann, A. E., Brandt, B. W., Sankaranarayanan, K., Ozga, A. T., Hoogland, M., Hofman, C. A., Salazar-García, D. C., Frohlich, B., Milner, G. R., Stone, A. C., Aldenderfer, M., Lewis, C. M., Jr, Hofman, C. L., Warinner, C., & Schroeder, H. (2019). The efficacy of whole human genome capture on ancient dental calculus and dentin. *American Journal of Physical Anthropology*, *168*(3), 496–509.

MANUSCRIPT D

Manuscript Nr. D

Title of manuscript: Exploring the potential of archaeogenetic studies of dental calculus to shed light on past human migrations in the Pacific

Authors: Zandra Fagernäs, Monica Tromp, Stuart Bedford, Hallie Buckley, Geoffrey Clark, John Dudgeon, James Flexner, Anatauarii Leal-Tamarii, Cecil M. Lewis Jr, Kathrin Nägele, Andrew Ozga, Adam B. Rohrlach, Cosimo Posth, Richard Shing, Matthew Spriggs, Edson Willie, Frédérique Valentin, Irina M. Velsko, Christina Warinner

Bibliographic information: In preparation for submission.

The candidate is

⊗ First author, o Co-first author, o Corresponding author, o Coauthor.

Status: The manuscript presented here will be published together with a currently ongoing project containing a multifaceted analysis of Pacific dental calculus.

Proportion (in %) of authors in the publication (to be indicated from 20%):

Author	Concept	Data analysis	Experiment	Manuscript composition	Material provision
Zandra Fagernäs	20%	90%	57%	70%	-
Monica Tromp	20%	-	29%	-	-
Stuart Bedford	-	-	-	-	-
Hallie Buckley	-	-	-	-	-
Geoffrey Clark	-	-	-	-	-
John Dudgeon	-	-	-	-	-
James Flexner	-	-	-	-	-
Anatauarii Leal-Tamarii	-	-	-	-	-
Cecil M. Lewis, Jr	-	-	-	-	-
Kathrin Nägele	20%	-	-	-	-

Andrew Ozga	-	-	-	-	-
Adam B. Rohrlach	-	-	-	-	-
Cosimo Posth	-	-	-	-	-
Richard Shing	-	-	-	-	-
Matthew Spriggs	-	-	-	-	-
Edson Willie	-	-	-	-	-
Frédérique Valentin	-	-	-	-	-
Irina M Velsko	20%	-	-	-	-
Christina Warinner	20%	-	-	-	-

Note: A more thorough breakdown of author contributions can be found in Appendix A. Supplementary tables and figures can be found in Appendix D.

Exploring the potential of archaeogenetic studies of dental calculus to shed light on past human migrations in the Pacific

Zandra Fagernäs[1,2], Monica Tromp[3,4,5], Stuart Bedford[6,7], Hallie Buckley[5], Geoffrey Clark[6], John Dudgeon[8], James Flexner[9], Anatauarii Leal-Tamarii[10], Cecil M. Lewis, Jr[11], Kathrin Nägele[1,2], Andrew Ozga[12], Adam B. Rohrlach[2,13], Cosimo Posth[1,14], Richard Shing[15], Matthew Spriggs[15,16], Edson Willie[15], Frédérique Valentin[17], Irina M. Velsko[1,2], Christina Warinner[1,2,18,19]

1. Department of Archaeogenetics, Max Planck Institute for the Science of Human History, Jena, Germany
2. Department of Archaeogenetics, Max Planck Institute for Evolutionary Anthropology, Leipzig, Germany
3. Department of Archaeology, Max Planck Institute for the Science of Human History, Jena, Germany
4. Southern Pacific Archaeological Research, Archaeology Programme, University of Otago, Dunedin, New Zealand
5. Department of Anatomy, School of Biomedical Sciences, University of Otago, Dunedin, New Zealand
6. Department of Archaeology and Natural History, College of Asia and the Pacific, The Australian National University, Canberra, Australia
7. Department of Linguistic and Cultural Evolution, Max Planck Institute for the Science of Human History, Jena, Germany
8. Department of Anthropology, Idaho State University, Pocatello, ID, USA
9. School of Philosophical and Historical Inquiry, University of Sydney, Sydney, Australia
10. Direction de la Culture et du Patrimoine, Puna'auia, French Polynesia
11. Department of Anthropology, University of Oklahoma, Norman, OK, USA
12. Halmos College of Arts and Sciences, Nova Southeastern University, Fort Lauderdale, FL, USA
13. ARC Centre of Excellence for Mathematical and Statistical Frontiers, School of Mathematical Sciences, The University of Adelaide, Adelaide, Australia.
14. Department of Archaeo- and Palaeogenetics, Eberhard Karls Universität Tübingen, Tübingen, Germany
15. Vanuatu Cultural Centre, Port-Vila, Vanuatu
16. School of Archaeology and Anthropology, College of Arts & Social Sciences, The Australian National University, Canberra, Australia
17. CNRS UMR 7041, MSH Mondes, Nanterre, France
18. Faculty of Biological Sciences, Friedrich Schiller University, Jena, Germany.
19. Department of Anthropology, Harvard University, Cambridge, MA, USA.

Abstract

The human microbiome offers an alternative to population genetics in the study of ancient human migrations, especially in cases where human DNA preservation is poor or skeletal material is not available for destructive analyses. As human populations migrate, and lifestyles and cultural practices change, their accompanying microbes also evolve and respond to the new living environment. The Pacific islands have experienced multiple waves of human migrations, some of them taking place over very short time periods, making it an ideal case study for exploring the potential of using the human microbiome to study migration. Here, we present results from a metagenomic study of archaeological dental calculus from 101 individuals, originating from 12 islands in the Pacific and spanning a time range of approximately 3000 years. We find that the majority of the dental calculus samples are preserved well enough to be used in further analyses, although preservation is highly variable. Variation in the microbial community composition was found to not be driven by time period or geography, but rather by local factors specific to the island of origin. In contrast, bacterial phylogeny does reflect time period and geography to some extent, but methodology and the selection of study species pose challenges. We conclude that archaeological dental calculus has the potential to yield information about past human migrations, and provides a new and alternative source of data that can complement studies of the human genome.

1. Introduction

Unravelling the complex history of how humanity spread across the world, settled down, mixed with other populations, and migrated further to reach new areas, requires interdisciplinary studies drawing on diverse lines of evidence. Archaeology, biological anthropology, and archaeogenetics are often combined in order to gain a multifaceted view of our past. New methodologies have the promise to extend our understanding of our past even more, which is particularly useful in areas with especially complex migration histories and cultural changes over time. One such region is the Pacific islands.

The Pacific islands have experienced multiple waves of migration and human settlement throughout the Pleistocene and Holocene. First, the land mass of Sahul, comprising what is today mainland Australia, Tasmania, the Aru Islands and New Guinea, was settled as early as 65000 BP (before present), at a time when sea levels were lower (Clarkson et al., 2017). The next major migration took place substantially later, around 5500 BP, as seafaring Austronesian-speaking people from the area of present-day Taiwan began migrating through Island Southeast Asia and into Oceania (Gray et al., 2009). At around 3300 BP, the Lapita cultural complex developed in the Bismarck Archipelago, as a part of the Austronesian expansion, and it was the descendants of this population that first reached the islands of Remote Oceania, moving beyond the main Solomons chain (Pugach et al., 2018; Sheppard, 2011). Finally, an expansion of people with Papuan-related ancestry, also originating from the Bismarck Archipelago, took place sometime during the next millennium, reaching the Reef-Santa Cruz group and Vanuatu and also impacting the genetic structure of Austronesian populations in Polynesia (Lipson et al., 2018, 2020; Posth et al., 2018; Skoglund et al., 2016). Hereafter, there was constant population movement and contact throughout the east Pacific, with people migrating and settling new areas. The last islands of the Pacific to be reached by humans, Aotearoa (New Zealand), Hawai'i and Rapa Nui, were colonized by about 1000-700 BP (Athens et al., 2014; Wilmshurst et al., 2011).

Archaeogenetic studies of ancient human migrations are conventionally conducted by analyzing human DNA from skeletal elements, such as teeth and the petrous portion of the temporal bone. These particular skeletal elements have been shown to have consistently good DNA preservation compared to other parts of the skeleton (Parker et al., 2020). Human population relationships can be studied directly by comparing the genetic variation of individuals from different geographic areas and time periods. This can, however, be challenging if the studied populations are closely related and migrations take place over short periods of time. Another factor complicating studies of ancient human migrations in the Pacific islands is DNA preservation. High temperatures and humidity generally increase the rate of DNA decay

(Dabney et al., 2013; Smith et al., 2003), and in tropical areas of the world, such as the Pacific islands, human DNA preservation can be highly variable (Posth et al., 2018). In addition, many Indigenous people do not wish to perform destructive analyses on the skeletal remains of their ancestors (Aranui, 2020; Fox, 2021).

Archaeological dental calculus has been proposed as an alternative study material to overcome some of these problems affecting skeletal remains (Eisenhofer et al., 2019; Tromp et al., 2016), but this approach has not yet been widely explored. Dental calculus forms when the bacterial biofilm known as dental plaque periodically calcifies on the surfaces of teeth. The dental plaque microbes become encased within the mineral matrix, which preserves their biomolecules, and this enables studies of oral microbial communities stretching thousands of years back in time (Fellows Yates et al., 2021). Because microbes have shorter generation times than humans, they offer the possibility to study the migrations of closely related populations over shorter timescales. Further, DNA within dental calculus is generally better preserved than DNA in skeletal tissues from the same individual (Mann et al., 2018), making it a promising substrate to study in areas where skeletal DNA preservation is generally poor, such as the tropics. Although dental calculus microbial community composition has been found to be relatively stable over long time periods, it is possible to trace the phylogenetic histories of individual strains or species of bacteria (Fellows Yates et al., 2021; Mégraud et al., 2016), which can aid in tracing the migration histories and cultural behaviors of their human hosts (Eisenhofer et al., 2020).

In addition to preserving DNA from the oral microbiome, dental calculus also preserves several other types of information about the lives of past individuals. For example, the presence of non-commensal microbes can be informative about disease, as pathogens have the potential to become encased in the dental calculus mineral matrix (Fotakis et al., 2020), and environmental debris can enter the oral cavity and become embedded in dental plaque, preserving information about the living environment and activities of past individuals (Hardy et al., 2016; Radini et al., 2019). Food particles may also become encased in dental calculus, but are generally best studied using proteomics or microscopy approaches (Dudgeon & Tromp, 2014; Geber et al., 2019; Warinner et al., 2014), as very low amounts of dietary DNA are typically recovered from dental calculus, making genetic authentication difficult (Mann et al., 2020). Taken together, dental calculus has the potential to provide a holistic view of the lives of past individuals, all from a single sample that is no more invasive than a routine dental cleaning. In addition, as dental calculus is a bacterial biofilm and not a human tissue, it may in some cases be more readily available for destructive analyses (Ozga et al., 2016; Tromp et al., 2016).

Here, we explore the possibility of using archaeological dental calculus to study past human migrations, using the Pacific islands as a case study. Shotgun metagenomic sequencing was

performed for a total of 101 dental calculus samples from 12 Pacific islands, spanning a time range of nearly 3000 years. We show that preservation is variable, but that the majority of the samples have a well-preserved oral microbiome, providing further support to the exceptional preservation of dental calculus. We find that variation in the microbial community composition does not have a clear geographic or temporal structure, but rather appears to be influenced by local factors specific to each island. Phylogenetic analyses of individual oral bacterial taxa exhibit temporal trends, but the ability to detect this signal is dependent on species selection. Finally, we show that DNA extraction and library preparation methods influence microbial community reconstruction, and this should be taken into account when comparing data across studies. Overall, we find that metagenomic analysis of archaeological dental calculus has the potential to reveal information about past human migration patterns. When combined with human DNA analysis and other approaches, such as palaeodietary studies using palaeoproteomics and microremains, it promises to greatly enrich our understanding of the dynamic biological and cultural processes that accompanied migrations in the Pacific islands.

2. Material and methods

Laboratory methods

A total of 101 archaeological dental calculus samples were processed in this study (Table 1). The dental calculus samples were processed in three groups, hereafter referred to as Jena, Oklahoma and Otago, named after the respective processing lab. Each sample group was processed using a slightly different extraction and library preparation protocol, as detailed below. A major finding of this study is that the processing methods have an effect on downstream community composition results, which has not been previously demonstrated for dental calculus metagenomics. These results are described below. Note that samples from Efate were analysed in two groups - **Efate** (samples processed in Jena, <1000 BP) and **Efate 3000 BP** (samples processed in Otago, 3000-2400 BP). Temporal information for the dental calculus samples in this study was obtained either through direct radiocarbon dating of the individual or by cultural association of the burial.

Table 1. Archaeological dental calculus samples included in this study. 'n' indicates the number of individuals analyzed from each site. The 'Mean age BP' is the range of estimated mean time periods of the individuals from the site (where a range is not given, all individuals were estimated to the same time BP).

Island	Area/island group	n	Mean age BP	Processing lab
Efate	Vanuatu	5	990-152	Jena
Efate	Vanuatu	16	3000-2400	Otago
Flores	Lesser Sunda Islands	3	3000	Oklahoma
Futuna	Vanuatu	3	1100	Jena
Raiatea	Society Islands	2	135	Jena
Rapa Nui	Rapa Nui	18	350	Oklahoma
Rurutu	Austral Islands	1	795	Oklahoma
Taumako	Duff Islands	17	485	Jena
Tongatapu	Tonga	6	2510-2495	Jena
Uripiv	Vanuatu	6	2500-2100	Otago
Vao	Vanuatu	5	2100	Otago
Viti Levu	Fiji	6+11	1500	Oklahoma+Jena
Watom	Bismarck Archipelago	2	2500	Otago

Jena: All processing took place in a dedicated cleanroom facility at the Max Planck Institute for the Science of Human History (MPI-SHH, Jena, Germany). Total DNA was extracted from 0.5-7 mg of dental calculus per individual, using a silica column-based extraction protocol optimized for the recovery of short DNA fragments, adapted for dental calculus (Aron et al., 2020; Dabney et al., 2013; Mann et al., 2018). The extracts were prepared into double-stranded libraries with partial uracil-DNA-glycosylase (UDG) treatment (Aron et al., 2020; Rohland et al., 2015) and dual indexing (Kircher et al., 2012; Meyer & Kircher, 2010; Stahl et al., 2019). The libraries were sequenced to a depth of 10.5 ± 2.3 million reads (mean ± standard deviation) on an Illumina NextSeq with 75-bp paired-end sequencing chemistry. Blanks were processed alongside the samples for both extraction and library preparation. Two samples, SIG040 and SIG046, were extracted and sequenced twice, as it was suspected that burial sand may have been unintentionally included during the first processing round.

Oklahoma: Total DNA was extracted from 0.8-12.8 mg dental calculus per individual following Ozga *et al.* (2016) (a method very similar to that performed in Jena) at the ancient DNA facility of the University of Oklahoma Laboratories of Molecular Anthropology and Microbiome Research (LMAMR, Norman, OK, USA). Blanks were processed alongside the samples. The extracts were thereafter shipped to MPI-SHH, where they were prepared into libraries, as described above, alongside the Jena samples. The samples were sequenced to a depth of 10.6

± 1.5 million reads on the same NextSeq flow cells as the Jena samples with 75-bp paired-end sequencing chemistry.

Otago: Total DNA was extracted from approximately 1-17 mg of dental calculus per individual using a phenol-chloroform aDNA extraction protocol (Tromp, 2016). Briefly, dental calculus samples were washed with ultrapure water and allowed to dry in a laminar flow hood overnight. A second wash was performed using 1 ml of 0.5 M EDTA, with a 30 min incubation time. The supernatant was removed, and the samples were thereafter demineralized in 1 ml of 0.5 M EDTA for up to 72 hours, until fully demineralized. The supernatant was added to a tube with 750 µl of phenol:chloroform:isoamyl (25:24:1), vortexed, and left on a rotator for 10 min. After centrifuging, the aqueous phase was transferred to 750 µl of phenol:chloroform:isoamyl alcohol (25:24:1). The incubation step was repeated, after which the aqueous phase was transferred to 750 µl of chloroform:isoamyl alcohol (24:1). After vortexing and mixing by inversion, the mixture was centrifuged and the aqueous phase transferred to 13 ml of 6 M GuSCN and 200 µl of silica suspension, and left on a nutator for 30 min. After centrifugation, the supernatant was removed and the silica was resuspended in 1 ml of GuSCN binding buffer, and the supernatant discarded after centrifugation (three times in total). The silica pellet was air dried for 15 min, and DNA eluted twice in 60 µl TE (heated to 65-75°C). Blanks were processed alongside samples through extraction and library preparation. Double-stranded libraries were prepared by blunt-end repairing the DNA strands, and thereafter ligating and filling in adapters. The libraries were amplified using KAPA HiFi enzyme, and no UDG-treatment was performed. The libraries were sequenced using an Illumina MiSeq 75-bp paired-end sequencing chemistry to 8.3 ± 7.2 million reads at the Otago Genomics Facility (Otago, New Zealand).

General data processing

Data processing and analyses were conducted in R v.4.1.0 (R Core Team, 2021), unless otherwise stated. General packages used were *tidyverse* v.1.3.1 (Wickham et al., 2019), *readxl* v.1.3.1 (Wickham & Bryan, 2019), *ggpubr* v.0.4.0 (Kassambara, 2018) and *janitor* v.2.1.0 (Firke, 2021). The colour palette for the study is from the R package *microshades* v.0.0.0.9000 (Dahl et al., 2021).

Preprocessing

DNA sequencing data was preprocessed using the nf-core/eager v.2.3.3 pipeline (Fellows Yates et al., 2021). Default options were used unless otherwise stated. Poly-G stretches were removed from the raw data, as they are a common by-product of the two-colour chemistry sequencing strategy used by Illumina's NextSeq. Human DNA was removed from the dataset by

mapping to the human reference genome GRCh38, and only unmapped reads were retained for downstream microbiome analyses. To produce an OTU table of the microbes present in the samples, the dataset was aligned using MALT v.0.4.1 (Herbig et al., 2016; Vågene et al., 2018) to a custom database containing all bacterial and archaeal assemblies (scaffold/chromosome/ complete levels, up to 10 randomly selected genomes per species) from RefSeq and the human HG19 reference genome (Fellows Yates et al., 2021) via the nf-core/eager pipeline. In addition, the dataset was aligned to the NCBI nucleotide database (as of October 2017), to screen for eukaryotic DNA. MEGAN v.6.17.0 (Huson et al., 2016) was used to export OTU tables from the resulting MALT-produced rma6 files, using summarized read counts at both the genus and species level.

A comparative dataset of published microbiome studies was also processed using the same procedures, consisting of 10 non-industrialized gut samples (Obregon-Tito et al., 2015; Rampelli et al., 2015), 11 industrialized gut samples (Gevers et al., 2012; Sankaranarayanan et al., 2015), 10 skin samples (Oh et al., 2016), 11 subgingival and 10 supragingival plaque samples (Gevers et al., 2012), 10 archaeological bone samples (Fellows Yates et al., 2021), 10 modern dental calculus samples (Fellows Yates et al., 2021) and 10 archaeological sediment samples (Slon et al., 2017). In addition, 10 archaeological bone samples from Taumako and 10 from Viti Levu were included as local environmental controls; this data had been produced during genetic screening of human remains for human population genetic studies at MPI-SHH (unpublished data). Because bones are free of microbes during life, microbes detected in these samples provide a good proxy for local post-mortem colonization (Warinner et al., 2014).

Putative environmental and laboratory contaminants in the dental calculus samples were identified using the R package *decontam* v.1.6.0 (Davis et al., 2018), with the prevalence method. The samples were separated into groups based on the processing lab, and blanks and archaeological bones were used as proxies for contamination sources (cut-off 0.25 for all groups, for both blanks and bones).

Preservation

Preservation was assessed using three methods. First, preservation was investigated and visualized using SourceTracker v.1.01 (Knights et al., 2011). A genus-level OTU table was used as input, and the reference metagenomic datasets described above were used as sources. During the SourceTracker analysis, rarefaction was performed to 10,000 reads, with a training data rarefaction of 5,000 reads. Next, genus-level read counts of all dental calculus samples and sources (including an additional 9 modern dental calculus samples from the same study as listed above) were compared using principal component analysis (PCA). The R-package

zCompositions v.1.3.4 (Palarea-Albaladejo & Martín-Fernández, 2015) was used for multiplicative zero replacement, and the data was thereafter CLR-transformed (centered log ratio) (Gloor et al., 2017). Finally, the R-package *cuperdec* (Fellows Yates, 2020; Fellows Yates et al., 2021) was used to identify well-preserved samples (burn-in method, cut-off 55%, database from Fellows Yates et al. (2021)), and 65 (out of 101) samples were carried on to further analyses based on this analysis.

In order to evaluate patterns in preservation, a dataset consisting of annual average temperature and annual total rainfall was compiled for each island in this study (Zepner et al., 2021). Missing data for some islands was obtained from alternative sources (*Climates to Travel - World Climate Guide* and *Timeanddate.com*, retrieved 2021-06-21). Annual average evapotranspiration was compiled from the *ArcGIS Average Annual Evapotranspiration Map*, (2020), which is based on data from the MOD16 Global Evapotranspiration Product. The proportion of taxa estimated to originate from the oral microbiome (i.e., dental plaque and dental calculus) by SourceTracker was used as a proxy for preservation of the archaeological dental calculus samples. The effects of environmental variables and/or sample age on preservation were investigated using beta regression with a complementary log-log link function to account for observed heteroscedasticity. Using ANOVA, it was found that the model was not significantly improved by adding the random effects of laboratory and/or island ($p»1$). Step AIC (using both directions) was used for model selection. To reliably estimate parameters for the model, statistically influential data points were removed, and the model-fitting process was repeated until a stable dataset was reached. Model fit was measured using the R^2 value as suggested by Ferrari and Cribari-Neto (2004) for beta regression models.

Community composition

A PCA was conducted on the genus-level OTU table, as described above. During initial visualization, it was noted that the dental calculus samples clustered based on the laboratory in which they were processed, suggesting that low abundance laboratory-specific contaminants might be driving the main axes of variation. Therefore, stringent filtering was applied before analyses of community composition. First, only taxa identified in samples from all three labs were included. Visual inspection of the removed taxa confirmed that they were primarily known laboratory contaminants. Second, the function comBat from the R package *sva* v.3.41.0 (Leek et al., 2019) was applied to the remaining OTU table, after multiplicative zero replacement and CLR-transformation. It should be noted that, unlike for patterns of preservation, the ‘Island’ and sample ‘Age’ variables are confounded with the ‘Lab’ variable, meaning that some island- and age-related patterns may be removed through this approach. Drivers of variation in the community composition were tested with a PERMANOVA from the R-package *vegan* v.2.5.7

(Oksanen et al., 2019), with euclidean distances and 9999 permutations. It should be noted that the 'Island' groups break the assumption of homogeneity of variance (betadispersal of euclidean distance, ANOVA, $p<0.001$), which may affect the results of the PERMANOVA.

Phylogenetic analyses

The nf-core/eager pipeline was used, as described above, to map the non-human reads of the well-preserved samples to the abundant and prevalent oral bacteria *Tannerella forsythia* (strain 92A2, assembly GCA_00238215.1) and Anaerolineaceae bacterium oral taxon 439 (assembly GCA_001717545.1). Through nf-core/eager, duplicates were removed using Picard MarkDuplicates v.2.22.9, and prior to mapping, damage was clipped off of the reads (two bases for libraries with partial UDG treatment, and seven bases for non-treated libraries). Genotyping was performed with GATK UnifiedGenotyper, allowing for heterozygous calls and using all sites, with the SNP likelihood model. A minimum base coverage of 5 was required. The SNPs were further filtered in order to construct the phylogenies with only homozygous SNPs (defined as the major nucleotide having a frequency greater than 0.9), using MultiVCFanalyzer v.0.0.87 (Bos et al., 2014). Only samples with at least 1000 SNPs and a mean genome-wide coverage of at least 2X (for *T. forsythia*) or 5X (Anaerolineaceae bacterium oral taxon 439) were included. The coverage requirement was increased for Anaerolineaceae bacterium oral taxon 439, because its percentage of heterozygous SNPs was higher. The trees were rooted using the reconstructed genomes of corresponding taxa obtained from a well-preserved dental calculus sample from El Collado, Spain, 9475-9300 BP (Fellows Yates et al., 2021). The phylogenetic trees were constructed and visualized using R packages *ape* v.5.5 (Paradis & Schliep, 2019), *ade4* v.1.7.17 (Dray & Dufour, 2007), *adegenet* v.2.1.3 (Jombart, 2008) and *ggtree* v.1.99.1 (Yu et al., 2017).

Eukaryotic DNA

In addition to analyzing microbial DNA, we also investigated putative eukaryotic DNA within the samples. Identification of eukaryotic taxa within metagenomic datasets is challenging and requires multiple steps of validation (Mann et al., 2020). Within the well-preserved dental calculus samples, we identified DNA from five non-human eukaryotic species of interest: cattle (*Bos taurus*), dog (*Canis lupus familiaris*), broad fish tapeworm (*Dibothriocephalus latus*), bamboo (*Fargesia denudata*), and wheat (*Triticum aestivum*) (Table S1). Other eukaryotic DNA present in the dataset belonged to commonly recognized contaminants, were assigned to genomes with known contamination (Mann et al., 2020), or represented highly unlikely taxa; these assignments were excluded from subsequent analysis. nf-core/eager was used to align the DNA sequences to reference genomes (Table S2) for the five species of interest (with

mapping quality set to 37), in order to evaluate the authenticity of the finds. Damage patterns were investigated for individuals with at least 200 reads mapping to the specific taxa (Mann et al., 2020). For taxa where 200 reads were not reached for any individuals, the damage patterns of the 10 individuals with the highest number of reads were investigated. For bamboo, only the chloroplast genome was available. For wheat, the mapping was restricted to the mitochondrial genome in this initial step, as the full wheat genome is very large (15.4 Gb).

3. Results

Preservation

Preservation is a persistent challenge in archaeogenetic studies of the tropics, as both temperature and humidity contribute to the decay of DNA (Dabney et al., 2013; Smith et al., 2003). Poor endogenous DNA preservation, as well as postmortem overgrowth of environmental microbes, can affect downstream analyses, making assessment of sample preservation a critical first step. Using SourceTracker analysis (Knights et al., 2011), we estimated the proportion of microbial taxa originating from endogenous and contaminant sources for each of the dental calculus samples in this study (Figure 1A). At least 50% of microbial content of 71 of the 101 archaeological dental calculus samples is estimated to originate from the oral microbiome, indicating good preservation of the dental calculus for these samples, with minimal contamination from exogenous sources such as soil and skin. However, the remaining samples exhibit variable microbial contributions from sediment and archaeological bone, indicating postmortem colonization, and in some cases complete overgrowth, by microbes from the burial environment. A few samples have a high contribution of skin microbes, indicating that they may have, at some point, been handled without appropriate protective equipment (e.g., gloves). Some islands have overall very good preservation, such as Efate, Flores, Raiatea, Rurutu, Taumako and Tongatapu, whereas for other islands, such as Efate 3000BP, Uripiv and Viti Levu, preservation is highly variable. The effect of sample age and environmental variables on the proportion of taxa that could be assigned to oral sources (dental plaque and dental calculus) was investigated using beta regression. Only sample age was found to be a significant predictor of preservation ($p=0.046$, $R^2=0.083$). However, as the R^2 value is very low, this indicates that other factors more greatly affect preservation, but these factors remain unknown. Nevertheless, local environmental factors at the burial site, such as soil type, humidity, and pH, are plausible candidates.

To assess sample preservation further, we performed a beta-diversity analysis between the archaeological dental calculus samples and the same source metagenomic datasets used in

SourceTracker, and visualized the similarity between samples using a PCA (Figure 1B). In the PCA plot, the archaeological dental calculus samples fall on a cline in between modern oral microbiome samples and archaeological bone/sediment samples, showing a gradient of well-preserved to poorly preserved samples.

After visually inspecting preservation using SourceTracker and a PCA, samples that were well-preserved enough to be carried on to further analyses were selected using the R-package *cuperdec* (Fellows Yates, 2020) (cut-off 55% with the burn-in method; Figure S1). In total, 65 dental calculus samples were sufficiently well-preserved and were carried on to further analyses (Figure 1C). Preservation does not show a clear geographic or temporal trend, although younger sites in general tend to have a higher proportion of well-preserved samples.

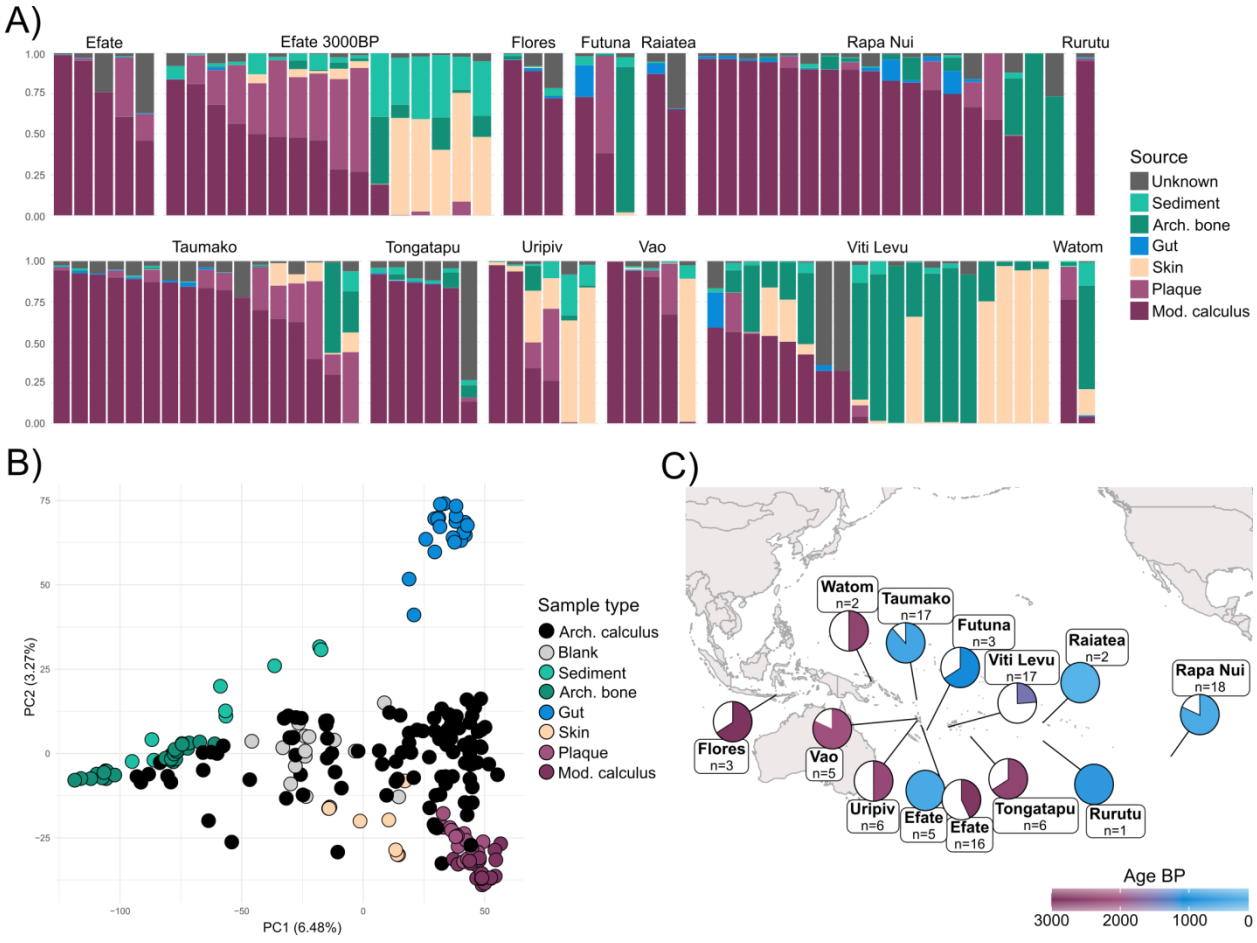


Figure 1. Preservation assessment of dental calculus samples. (**A**) SourceTracker analysis of genus-level OTU tables, using metagenomic studies as sources. Each bar represents a sample, coloured by the proportion of each contributing source. (**B**) PCA of comparative metagenomic samples, archaeological dental calculus samples, and laboratory blanks. (**C**) Map of sample sites, with the colored part of the pie chart indicating the fraction of samples that are

considered to be well-preserved, and hence carried on to further analyses. The pies are colored according to the mean age BP of samples from the island. The total number of screened samples from the island is denoted by n.

Community composition

Focusing only on the dental calculus samples with good preservation, we next addressed the question of whether the dental calculus microbial communities cluster based on geography or time. To this end, we performed a beta-diversity analysis and visualized the samples using PCA. During initial assessment, it was noted that the first two principal components were strongly driven by the laboratory in which the sample was processed (Figure 2A). Factor analysis confirmed that the taxa driving the separation of the Jena and Oklahoma samples in PC2 consist of common environmental taxa found in soils and sediments (Jena: *Mesorhizobium*, *Desulfovibrio*; Oklahoma: *Butyricicoccus*, *Anaerotruncus*), suggesting that differences in ambient environmental microbes distinguish the background contamination of these labs.

In contrast, the separation of Otago in PC1 is primarily driven by differences in the abundances of oral taxa (Otago: *Corynebacterium*, *Ottowia*; Jena/Oklahoma: *Lactobacillus*, *Anaerococcus*). We found that these oral taxa have very different GC-contents: *Lactobacillus* and *Anaerococcus* species, which characterize the Jena and Oklahoma samples, generally have low GC-contents around 30-35%, whereas *Ottowia* and *Corynebacterium* species, which characterize the Otago samples, have high GC-contents around 50-70%. This suggests that the different extraction methods may be biased towards different DNA properties, and specifically the Otago method employing a phenol/chloroform separation may be inefficient at recovering taxa with low GC-contents. Additionally, the Otago samples are generally older, and AT-rich fragments are known to be lost at a higher rate in ancient samples (Mann et al., 2018), which may contribute to a generally higher GC-content in the Otago samples. Finally, the samples from Otago have a higher amount of DNA damage than the other samples (for well-preserved samples, 21.6-50.9% on the 1st base at the 5'end, versus 2.4-24.6% for the Jena/Oklahoma samples), based on mapping to the oral species *Tannerella forsythia*. However, the Otago samples were not UDG-treated to remove ancient DNA damage, which explains the higher level of C to T transitions. Nevertheless, damage has been shown to not significantly influence species assignments in ancient metagenomes (Eisenhofer & Weyrich, 2019; Velsko et al., 2018), and so differences in UDG treatment and damage levels are unlikely to be the cause of the differentiation observed in PC1.

To account for the observed bias between laboratories, we only retained taxa identified in samples from all three laboratories, which primarily removed known laboratory contaminants. Thereafter, we used the comBat function from the R package *sva* (Leek et al., 2019) to remove remaining biases caused by processing laboratory (Figure 2B).

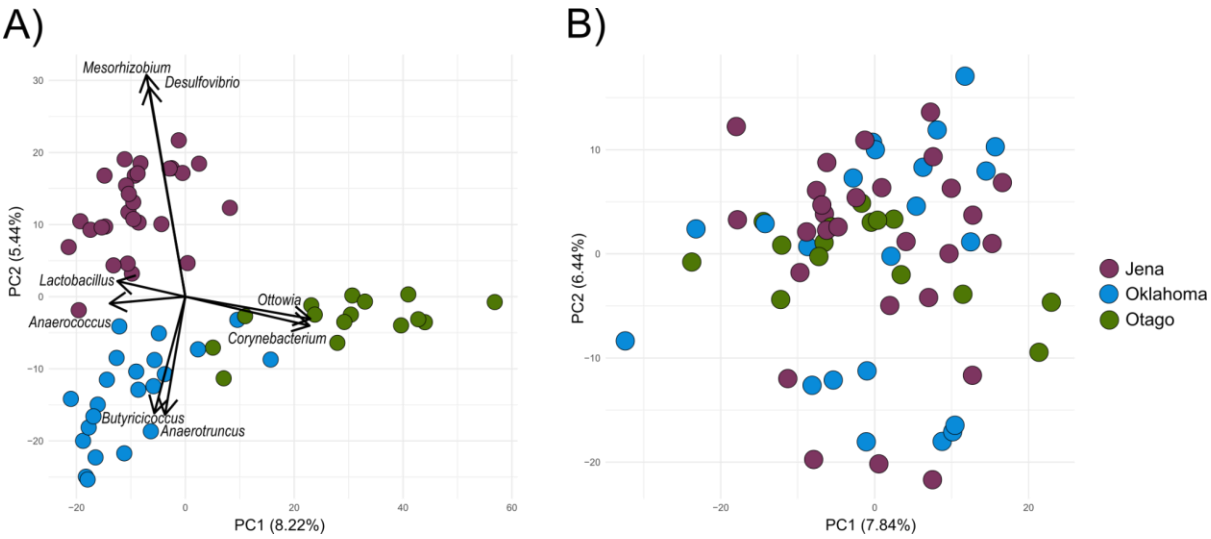


Figure 2. Sample processing laboratory significantly affects community composition. A) PCA, coloured by the processing lab, of genus-level read counts prior to removing bias caused by processing lab. **B)** PCA, coloured by processing lab, of genus-level read counts, after controlling for batch effects based on the processing lab.

Next, we assessed whether the dental calculus microbial community contained a geographic or temporal signal through PCA. We found that the island from which the individual originated is a significant driver of variation in community composition (PERMANOVA, $p<0.001$, $R^2=0.25$; Figure 3), while sample age is not. However, geographical clustering of different islands was not evident, suggesting that local conditions, such as cultural or preservational differences, have a bigger impact on the reconstructed oral microbiome composition than sample age or host migratory patterns. Moreover, a high degree of consistency in microbial community composition was also evident when studying the distribution of genera across the samples (Figure S2), and no clear patterns were observed by time or between islands. This stability in the dental calculus community composition has also been previously shown, including over even longer periods of time and across continents (Fellows Yates et al., 2021).

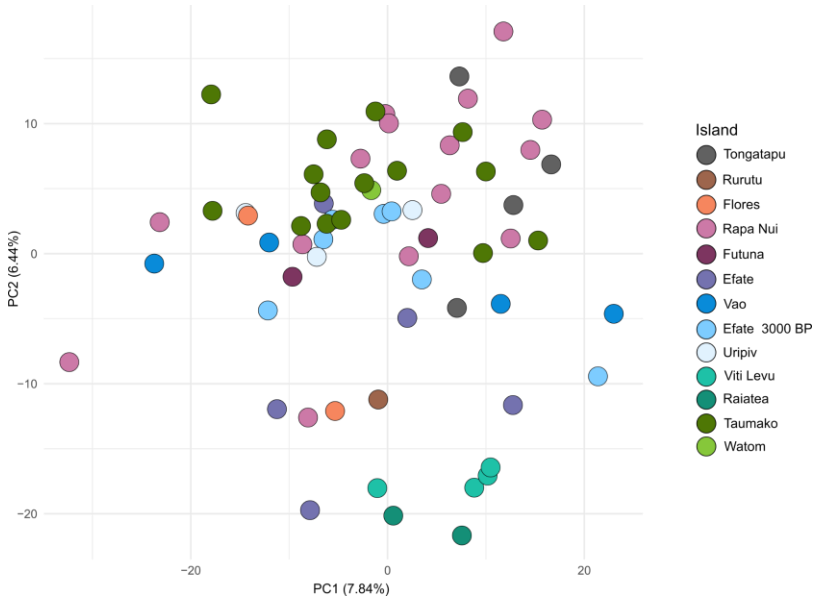


Figure 3. Island is a significant driver of community composition. PCA of genus-level read counts, coloured by the island of origin.

Phylogenetic analyses

Although microbial community composition appears to be fairly resistant to change over time, individual microbial taxa may still exhibit phylogeographic patterns, as has been previously noted, for example, with the gastric pathogen *Helicobacter pylori* (Mégraud et al., 2016). Therefore, examining genomic variation in individual species may provide insights into microbial evolution and adaptation during human migrations which cannot be observed in community-wide metrics. Phylogenetic trees were constructed for *Tannerella forsythia* and Anaerolineaceae bacterium oral taxon 439, as they are both abundant and prevalent in this dataset, ensuring sufficient coverage for genomic reconstruction, and both have previously been studied phylogenetically in archaeological dental calculus (Bravo-Lopez et al., 2020; Eisenhofer et al., 2020; Fellows Yates et al., 2021; Ottoni et al., 2021).

In both phylogenetic trees, samples from the same islands generally cluster together (Figure 4), indicating that the bacterial strains among individuals living on the same island resemble each other. *Tannerella forsythia* does not, however, exhibit a clear time-related pattern in the tree topography (Figure 4A). A temporal pattern is more apparent for Anaerolineaceae bacterium oral taxon 439, as older samples generally fall more basal in the tree (Figure 4B). The percentage of heterozygous SNPs is generally quite low for *T. forsythia* (<20% for all samples), indicating that there is only one dominant species/strain per sample aligning to the reference

genome (Warinner et al., 2017). For Anaerolineaceae bacterium oral taxon 439, however, the percentage of sites with heterozygous SNPs is much higher, indicating that several strains or species in each sample may be aligning to this reference genome. It is likely that the Anaerolineaceae bacterium oral taxon 439 phylogeny is not of a single taxon, but rather a collection of closely related species or strains. Within a cluster of samples, there is also a tendency for samples with higher levels of heterozygosity to fall basal to other samples (e.g., the cluster of samples from Taumako and Rurutu).

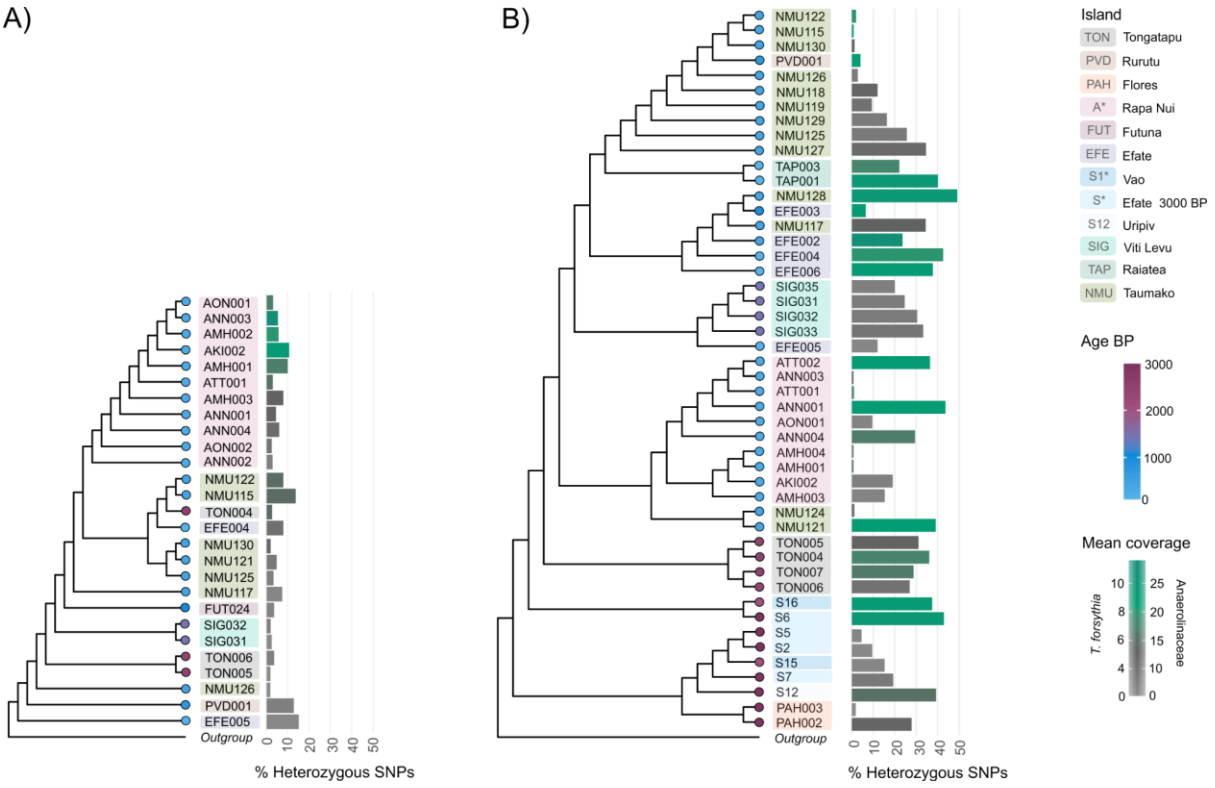


Figure 4. Phylogenetic trees show that bacterial genomes from the same island resemble each other. A) A neighbor-joining tree of *Tannerella forsythia* from samples with >2X genomic coverage, using only homozygous SNPs. **B)** A neighbor-joining tree of Anaerolineaceae bacterium oral taxon 439, including only samples with >5X genomic coverage of the taxon and using only homozygous SNPs. For both trees, the age of the sample (in years BP) is shown as colored circles on tree tips, island of origin as a colored box behind sample IDs, the percentage of heterozygous SNPs as a bar, and the mean coverage of the genome as the color of the bar.

Eukaryotic DNA

In addition to tracing the microbial changes in dental calculus across the Pacific islands, we sought to examine the likelihood of recovering food-derived DNA that may offer insight into

dietary patterns across these sites. Dietary DNA may become incorporated into dental calculus during the life of an individual; however, eukaryotic DNA is often only recovered in very small amounts from archaeological dental calculus, and authenticating these finds can provide a major challenge (Mann et al., 2020). It should also be noted that some potential dietary items in the Pacific, such as kava (*Piper methysticum*) or the giant swamp taro (*Cyrtosperma merkusii*), may not be present in reference databases as they do not have sequenced nuclear or organelle genomes. Some dietary items may therefore be missed, or DNA fragments may be aligning to closely related species.

Potentially authentic hits, based on DNA damage patterns, were detected for cattle and dog, but only in a limited number of samples. For the other species investigated here (broad fish tapeworm, bamboo and wheat), the DNA sequences do not exhibit a typical ancient DNA damage pattern, and their authenticity can thereby not be confirmed. In some cases, this may be due to the low number of DNA sequences that were recovered, or the sequences may be modern contaminants or aligning to an inaccurate reference genome (Mann et al., 2020).

Cattle with typical ancient DNA damage were detected in a sample from Flores (209 reads). It should, however, be noted that these DNA fragments most likely stem from another bovid related to cattle, which is found in the Pacific islands but does not have a sequenced genome, such as the banteng (*Bos javanicus)*. Potentially authentic dog DNA was detected in one sample from Flores (237 reads) and two samples from Tongatapu (361 and 202 reads, respectively). Dogs were brought to Oceania by the Lapita people, and were present in Near Oceania during the time of the Flores and Tongatapu individuals (Gonzalez et al., 2013). However, it has been suggested that dogs were not very prominent in early Lapita settlements, and they likely did not arrive in Remote Oceania until after the Lapita period (Greig et al., 2015), bringing the plausibility of this find, especially in the Tongatapu samples, into question. It should be noted that both cattle and dog DNA are commonly detected contaminants in genetic laboratories, and these finds need to be complemented with other methods, such as proteomics, before their authenticity can be confirmed.

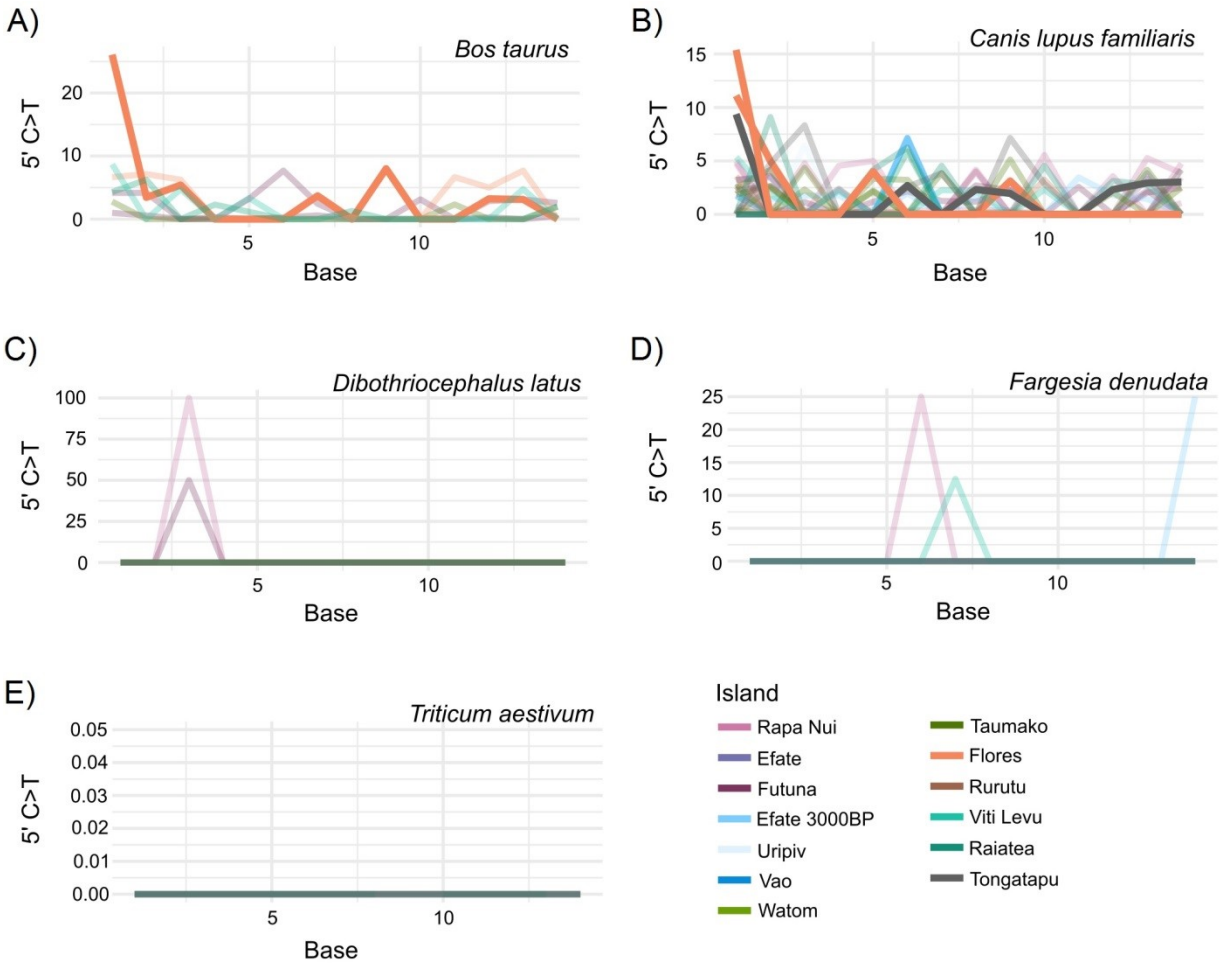


Figure 5. Damage patterns of eukaryotic DNA in archaeological dental calculus samples. A) Cattle (*Bos taurus*), **B)** Dog (*Canis lupus familiaris*), **C)** Broad fish tapeworm (*Dibothriocephalus latus*), **D)** Bamboo (*Fargesia denudata*), and **E)** Wheat (*Triticum aestivum*). A and B contain all samples with >200 reads, whereas C, D and E contain the ten samples with the highest number of reads each. Damage patterns which appear authentic are highlighted, and non-authentic damage patterns are faded. Lines are colored based on which island the individual stems from.

4. Discussion

The prospect of studying human migration through archaeological dental calculus is very intriguing, as it would allow for a holistic study of the life of an individual through just one sample - from migration and diet, to health and disease, and even occupational activities. In the Pacific, dental calculus may prove to be an especially valuable study material, given the rapid migrations between islands and significant cultural changes over a short time. However, no

study has to date attempted to investigate past human migrations through the oral microbiome, although the prospect has been discussed (Eisenhofer et al., 2019; Tromp et al., 2016).

In this study, we find that the majority of archaeological dental calculus samples from islands across the Pacific preserve an authentic oral microbiome signal. Considering the high temperatures and humidity in the area, and variable success in human DNA extraction from skeletal elements (Posth et al., 2018), this provides further support to the exceptional preservation of biomolecules in dental calculus, also seen in other studies (Mann et al., 2018). No climatic variables were found to predict dental calculus preservation, which suggests that smaller-scale local factors may drive preservation, such as soil biogeochemistry and the microclimate of the burial site. Further studies are required to investigate which local factors may contribute to preservation, but as of now, sample preservation cannot be predicted prior to destructive analyses.

The impact of migrations and cultural changes on oral microbial communities can be studied by comparing the relationships between samples from different archaeological sites and time periods at both a community-wide level, and at a genomic level. Here, sample clustering is mainly driven by island of origin for both community profile- and individual genome-level analyses, and far fewer patterns are seen when related to larger geographic areas or time period of the samples. This may be caused by cultural or preservational effects on the reconstructed microbiome. As individuals from the same island (and, as is the case for many islands in this study, from the same archaeological site) would likely have lived in close proximity, consumed the same types of food, lived in comparable conditions, or had similar cultural practices, this may cause their oral microbiomes to resemble each other more closely than individuals from nearby islands. Further, as preservation differs between sites and islands, decay of the genetic information may progress in a similar manner in samples from the same island, leading to differences between samples from different island populations that were not present during the life of an individual. With the experimental design of this study we were not able to disentangle whether the observed differences are from one or the other, but it is a critical point for investigating in future ancient dental calculus studies. It should also be noted that dating still needs to be ascertained for some samples included in this study.

While broad geographic and temporal patterns are weak in this data set, minor patterns can be seen in the phylogenetic analyses of bacterial taxa. Intriguingly, the branching patterns partly reflect migration patterns seen in human population genetic analyses. For example, Anaerolineaceae bacterium oral taxon 439 in samples from Wallacea (Flores, 3000 BP) fall basal to all other samples, apparently reflecting the direction of migration from Island Southeast Asia to Near to Remote Oceania (Gray et al., 2009; Pugach et al., 2018; Sheppard, 2011). In

phylogenetic trees for both *Tannerella forsythia* and Anaerolineaceae bacterium oral taxon 439, the samples from Rapa Nui, the most remote island in this study, fall in a single cluster that is not basal to any other samples, potentially reflecting that this island was the last to be colonized of the islands represented in this study (Wilmshurst et al., 2011). Phylogenetic analyses of bacterial taxa in dental calculus thereby show potential to reflect human migrations, although there are several issues to overcome in order to refine these analyses.

A major issue faced by researchers using metagenomic data to reconstruct past microbial genomes is the presence of multiple closely-related species and strains within a microbial community. When studying past migrations through human population genetics, there is no question of the appropriate reference species genome, which yields high certainty in aligning reads and calling SNPs. While contamination may still be a cause for concern, the majority of the studied DNA will be from a single individual. Attempting to study migration patterns through the microbiome, on the other hand, comes with a higher degree of inherent uncertainty with attributing microbial DNA to particular species and strains. Through issues such as contaminated ('dirty') reference genomes, which include sequences not derived from the species of interest, or incomplete databases, it is possible that sequencing reads from multiple species are inappropriately aligned to a reference genome (Mann et al., 2020; Steinegger & Salzberg, 2020). In such cases, it is no longer the genetic information of a single species that is studied, and this introduces 'false' genetic variation (stemming from a mixture of species, not genetic variation within a species). In addition, microbes have complex ways of exchanging genetic information, such as horizontal gene transfer, which allows exchange of DNA between distantly related species. This poses a challenge for phylogenetic studies, as genetic information is not only transmitted vertically, and a considerable amount of genetic variation can be introduced in a single generation. These factors complicate phylogeographic studies of microbiomes, introducing background noise and potentially eroding patterns caused by migration of the host.

One way to overcome such issues is enrichment of a target genome, which increases the proportion of the sequenced DNA that stems from a specific organism and is a popular technique in archaeogenetic pathogen research (Spyrou et al., 2019). This is, however, a quite costly method, and only feasible in cases where the target organism has multiple high-quality genomes available, which is not the case for many oral taxa. It also comes with its own limitations, such as representing only the known variation in the captured taxon, which may miss genetic variation that existed in past populations. Another method is *de novo* assembly from deeply sequenced metagenomic datasets (Wibowo et al., 2021). This method avoids the issue of requiring a reference genome, and allows for discovery of novel species and novel genetic variation, but does require good preservation and very high sequencing depths. Finally, to

reduce the background noise of the genetic information, analyses may be restricted to the phylogeographic study of a smaller set of genes, such as housekeeping genes (Lorén et al., 2014). As the study of migratory patterns through ancient host-associated microbiomes is still in its infancy, method development will be fundamental in order to explore the full potential of this field.

5. Outlook

Archaeological dental calculus shows promise in the study of past human migrations. However, further method development and refinement will be necessary before the full potential of the field can be realized. The results presented here are part of a larger multidisciplinary research project studying human migration in the Pacific islands using archaeological dental calculus, and the analyses will be further improved and expanded prior to publication of the project. In order to refine the results further, and bring them into the bigger picture, reference samples of published ancient and modern oral microbiomes will be added. This will assist with assessing where the variation seen in the microbiome composition, and the variation within species, stems from. The phylogenetic methods will be used to investigate specific gene classes, such as virulence or housekeeping genes, in order to remove some of the potential noise in the genetic variation. Dating of the samples will be refined, to ensure accurate time-related results. Finally, the metagenomic results will be combined with microfossil analyses of these samples, in order to gain a more holistic view of the migrations and cultural changes in the area, together with pre-existing linguistic and archaeological studies.

6. Acknowledgements

This work was supported by the Werner Siemens Stiftung (“Paleobiotechnology” to C.W.), the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany’s Excellence Strategy EXC 2051, Project-ID 390713860 (C.W.), and the Max Planck Society. It was also funded by a University of Otago Doctoral Scholarship, a Royal Society of New Zealand Skinner Fund grant and an Otago Centre for Electron Microscopy Student Research Award awarded to M.T. We also thank M. Abong of the Vanuatu Cultural Centre (VCC) for assistance during calculus sampling. Sampling of the Teouma dental calculus was done through a research agreement with the Vanuatu National Cultural Council. The Teouma Archaeological Project is a joint initiative of the Vanuatu National Museum and the Australian National University (ANU), directed by M.S. and S.B. and at different times R. Regenvanu and M. Abong, both former Directors of the VCC. Funding of the project was provided by the Australian Research Council

(grant no. DP 0556874), the National Geographic Society (grant no. SRC 8038–06), the Pacific Biological Foundation, the Department of Archaeology and Natural History and School of Archaeology and Anthropology at the ANU, the Snowy Mountains Engineering Corporation Foundation and B. Powell. The laboratory research and travel for excavation of the skeletal remains were funded by The Royal Society of New Zealand Marsden Fund (grant nos. UOO0407 and 09-UOO-106) and a University of Otago Research Grant awarded to H.B. The support of the leaseholder M. R. Monvoisin and family is acknowledged, as is the support and assistance of the traditional landowners and population of Eratap Village. Detailed acknowledgement by the authors of the Vanuatu studies (S.B., H.B., J.F., R.S., M.S. and F.V.) are given in the published site reports for each location.

3. References

Aranui, A. (2020). Restitution or a loss to science? Understanding the Importance of Returning Māori Ancestral Remains. *Museum & Society*, *18*(1), 19–29.

ArcGIS Average Annual Evapotranspiration map. (2020). https://www.arcgis.com/home/webmap/viewer.html?layers=ad3f8cc18fc74e6894ee220acd15020a

Aron, F., Hofman, C., Fagernäs, Z., Velsko, I., Skourtanioti, E., Brandt, G., & Warinner, C. (2020). Ancient DNA extraction from dental calculus v1. *protocols.io*. ZappyLab, Inc.

Aron, F., Neumann, G. U., & Brandt, G. (2020). Half-UDG treated double-stranded ancient DNA library preparation for Illumina sequencing v1. *protocols.io*. ZappyLab, Inc.

Athens, J. S., Reith, T. M., & Dye, T. S. (2014). A palaeoenvironmental and archaeological model-based age estimate for the colonization of Hawai'i. *American Antiquity*, *79*, 144–155.

Bos, K. I., Harkins, K. M., Herbig, A., Coscolla, M., Weber, N., Comas, I., Forrest, S. A., Bryant, J. M., Harris, S. R., Schuenemann, V. J., Campbell, T. J., Majander, K., Wilbur, A. K., Guichon, R. A., Wolfe Steadman, D. L., Cook, D. C., Niemann, S., Behr, M. A., Zumarraga, M., Bastida, R., Huson, D., Nieselt, K., Young, D., Parkhill, J., Buikstra, J. E., Gagneux, S., Stone, A. C., & Krause, J. (2014). Pre-Columbian mycobacterial genomes reveal seals as a source of New World human tuberculosis. *Nature*, *514*(7523), 494–497.

Bravo-Lopez, M., Villa-Islas, V., Rocha Arriaga, C., Villaseñor-Altamirano, A. B., Guzmán-Solís, A., Sandoval-Velasco, M., Wesp, J. K., Alcantara, K., López-Corral, A., Gómez-Valdés, J., Mejía, E., Herrera, A., Meraz-Moreno, A., Moreno-Cabrera, M. de la L., Moreno-Estrada, A., Nieves-Colón, M. A., Olvera, J., Pérez-Pérez, J., Iversen, K. H., Rasmussen, S., Sandoval, K., Zepeda, G., & Ávila-Arcos, M. C. (2020). Paleogenomic insights into the red complex bacteria *Tannerella forsythia* in Pre-Hispanic and Colonial individuals from Mexico. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *375*(1812), 20190580.

Clarkson, C., Jacobs, Z., Marwick, B., Fullagar, R., Wallis, L., Smith, M., Roberts, R. G., Hayes, E., Lowe, K., Carah, X., Florin, S. A., McNeil, J., Cox, D., Arnold, L. J., Hua, Q., Huntley, J., Brand, H. E. A., Manne, T., Fairbairn, A., Shulmeister, J., Lyle, L., Salinas, M., Page, M., Connell, K., Park, G., Norman, K., Murphy, T., & Pardoe, C. (2017). Human occupation of

northern Australia by 65,000 years ago. *Nature*, *547*(7663), 306–310.
Climates to Travel - world climate guide. Retrieved June 21, 2021, from https://www.climatestotravel.com/
Dabney, J., Knapp, M., Glocke, I., Gansauge, M.-T., Weihmann, A., Nickel, B., Valdiosera, C., García, N., Pääbo, S., Arsuaga, J.-L., & Meyer, M. (2013). Complete mitochondrial genome sequence of a Middle Pleistocene cave bear reconstructed from ultrashort DNA fragments. *Proceedings of the National Academy of Sciences of the United States of America*, *110*(39), 15758–15763.
Dabney, J., Meyer, M., & Pääbo, S. (2013). Ancient DNA damage. *Cold Spring Harbor Perspectives in Biology*, *5*(7).
Dahl, E., Karstens, L., & Neer, E. (2021). *microshades: A custom color palette for improving data visualization*.
Davis, N. M., Proctor, D. M., Holmes, S. P., Relman, D. A., & Callahan, B. J. (2018). Simple statistical identification and removal of contaminant sequences in marker-gene and metagenomics data. *Microbiome*, *6*(1), 226.
Dray, S., & Dufour, A.-B. (2007). The ade4 Package: Implementing the Duality Diagram for Ecologists. *Journal of Statistical Software, Articles*, *22*(4), 1–20.
Dudgeon, J. V., & Tromp, M. (2014). Diet, Geography and Drinking Water in Polynesia: Microfossil Research from Archaeological Human Dental Calculus, Rapa Nui (Easter Island). *International Journal of Osteoarchaeology*, *24*(5), 634–648.
Eisenhofer, R., Anderson, A., Dobney, K., Cooper, A., & Weyrich, L. S. (2019). Ancient Microbial DNA in Dental Calculus: A New method for Studying Rapid Human Migration Events. *The Journal of Island and Coastal Archaeology*, *14*(2), 149–162.
Eisenhofer, R., Kanzawa-Kiriyama, H., Shinoda, K.-I., & Weyrich, L. S. (2020). Investigating the demographic history of Japan using ancient oral microbiota. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *375*(1812), 20190578.
Eisenhofer, R., & Weyrich, L. S. (2019). Assessing alignment-based taxonomic classification of ancient microbial DNA. *PeerJ*, *7*, e6594.
Fellows Yates, J. A. (2020). *cuperdec: cumulative percent decay curves*.
Fellows Yates, J. A., Lamnidis, T. C., Borry, M., Valtueña, A. A., Fagernäs, Z., Clayton, S., Garcia, M. U., Neukamm, J., & Peltzer, A. (2021). Reproducible, portable, and efficient ancient genome reconstruction with nf-core/eager. *PeerJ*, *9*, e10947.
Fellows Yates, J. A., Velsko, I. M., Aron, F., Posth, C., Hofman, C. A., Austin, R. M., Parker, C. E., Mann, A. E., Nägele, K., Arthur, K. W., Arthur, J. W., Bauer, C. C., Crevecoeur, I., Cupillard, C., Curtis, M. C., Dalén, L., Díaz-Zorita Bonilla, M., Díez Fernández-Lomana, J. C., Drucker, D. G., Escribano Escrivá, E., Francken, M., Gibbon, V. E., González Morales, M. R., Grande Mateu, A., Harvati, K., Henry, A. G., Humphrey, L., Menéndez, M., Mihailović, D., Peresani, M., Rodríguez Moroder, S., Roksandic, M., Rougier, H., Sázelová, S., Stock, J. T., Straus, L. G., Svoboda, J., Teßmann, B., Walker, M. J., Power, R. C., Lewis, C. M., Sankaranarayanan, K., Guschanski, K., Wrangham, R. W., Dewhirst, F. E., Salazar-García, D. C., Krause, J., Herbig, A., & Warinner, C. (2021). The evolution and changing ecology of the African hominid oral microbiome. *Proceedings of the National Academy of Sciences of the United States of America*, *118*(20), e2021655118.
Ferrari, S., & Cribari-Neto, F. (2004). Beta Regression for Modelling Rates and Proportions.

Journal of Applied Statistics, *31*(7), 799–815.
Firke, S. (2021). *janitor: Simple Tools for Examining and Cleaning Dirty Data*.
Fotakis, A. K., Denham, S. D., Mackie, M., Orbegozo, M. I., Mylopotamitaki, D., Gopalakrishnan, S., Sicheritz-Pontén, T., Olsen, J. V., Cappellini, E., Zhang, G., Christophersen, A., Gilbert, M. T. P., & Vågene, Å. J. (2020). Multi-omic detection of *Mycobacterium leprae* in archaeological human dental calculus. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *375*(1812), 20190584.
Fox, K. (2021). Lei Niho Palaoa and Digital Tools to Safeguard Against the Illicit Use of Ancient DNA. *Hawaiian Archaeology*, *15*.
Geber, J., Tromp, M., Scott, A., Bouwman, A., Nanni, P., Grossmann, J., Hendy, J., & Warinner, C. (2019). Relief food subsistence revealed by microparticle and proteomic analyses of dental calculus from victims of the Great Irish Famine. *Proceedings of the National Academy of Sciences of the United States of America*, *116*(39), 19380–19385.
Gevers, D., Knight, R., Petrosino, J. F., Huang, K., McGuire, A. L., Birren, B. W., Nelson, K. E., White, O., Methé, B. A., & Huttenhower, C. (2012). The Human Microbiome Project: a community resource for the healthy human microbiome. *PLoS Biology*, *10*(8), e1001377.
Gloor, G. B., Macklaim, J. M., Pawlowsky-Glahn, V., & Egozcue, J. J. (2017). Microbiome Datasets Are Compositional: And This Is Not Optional. *Frontiers in Microbiology*, *8*, 2224.
Gonzalez, A., Clark, G., O'Connor, S., & Matisoo-Smith, L. (2013). A 3000 Year Old Dog Burial in Timor-Leste. *Australian Archaeology*, *76*(1), 13–20.
Gray, R. D., Drummond, A. J., & Greenhill, S. J. (2009). Language phylogenies reveal expansion pulses and pauses in Pacific settlement. *Science*, *323*(5913), 479–483.
Greig, K., Walter, R., & Matisoo-Smith, L. (2015). Dogs and people in Southeast Asia and the Pacific. In M. Oxenham & H. R. Buckley (Eds.), *The Routledge Handbook of Bioarchaeology in Southeast Asia and the Pacific Islands*. Taylor and Francis.
Hardy, K., Radini, A., Buckley, S., Sarig, R., Copeland, L., Gopher, A., & Barkai, R. (2016). Dental calculus reveals potential respiratory irritants and ingestion of essential plant-based nutrients at Lower Palaeolithic Qesem Cave Israel. *Quaternary International: The Journal of the International Union for Quaternary Research*, *398*, 129–135.
Herbig, A., Maixner, F., Bos, K. I., Zink, A., Krause, J., & Huson, D. H. (2016). MALT: Fast alignment and analysis of metagenomic DNA sequence data applied to the Tyrolean Iceman. *bioRxiv* 050559.
Huson, D. H., Beier, S., Flade, I., Górska, A., El-Hadidi, M., Mitra, S., Ruscheweyh, H.-J., & Tappu, R. (2016). MEGAN Community Edition - Interactive Exploration and Analysis of Large-Scale Microbiome Sequencing Data. *PLoS Computational Biology*, *12*(6), e1004957.
Jombart, T. (2008). adegenet: a R package for the multivariate analysis of genetic markers. *Bioinformatics* , *24*(11), 1403–1405.
Kassambara, A. (2018). *ggpubr: "ggplot2" Based Publication Ready Plots*.
Kircher, M., Sawyer, S., & Meyer, M. (2012). Double indexing overcomes inaccuracies in multiplex sequencing on the Illumina platform. *Nucleic Acids Research*, *40*(1), e3.
Knights, D., Kuczynski, J., Charlson, E. S., Zaneveld, J., Mozer, M. C., Collman, R. G., Bushman, F. D., Knight, R., & Kelley, S. T. (2011). Bayesian community-wide culture-independent microbial source tracking. *Nature Methods*, *8*(9), 761–763.
Leek, J. T., Johnson, W. E., Parker, H. S., Fertig, E. J., Jaffe, A. E., Storey, J. D., Zhang, Y., &

Torres, L. C. (2019). *sva: Surrogate Variable Analysis*.
Lipson, M., Skoglund, P., Spriggs, M., Valentin, F., Bedford, S., Shing, R., Buckley, H., Phillip, I., Ward, G. K., Mallick, S., Rohland, N., Broomandkhoshbacht, N., Cheronet, O., Ferry, M., Harper, T. K., Michel, M., Oppenheimer, J., Sirak, K., Stewardson, K., Auckland, K., Hill, A. V. S., Maitland, K., Oppenheimer, S. J., Parks, T., Robson, K., Williams, T. N., Kennett, D. J., Mentzer, A. J., Pinhasi, R., & Reich, D. (2018). Population Turnover in Remote Oceania Shortly after Initial Settlement. *Current Biology, 28*(7), 1157–1165.e7.
Lipson, M., Spriggs, M., Valentin, F., Bedford, S., Shing, R., Zinger, W., Buckley, H., Petchey, F., Matanik, R., Cheronet, O., Rohland, N., Pinhasi, R., & Reich, D. (2020). Three Phases of Ancient Migration Shaped the Ancestry of Human Populations in Vanuatu. *Current Biology*, *30*(24), 4846–4856.e6.
Lorén, J. G., Farfán, M., & Fusté, M. C. (2014). Molecular phylogenetics and temporal diversification in the genus *Aeromonas* based on the sequences of five housekeeping genes. *PloS One*, *9*(2), e88805.
Mann, A. E., Fellows Yates, J. A., Fagernäs, Z., Austin, R. M., Nelson, E. A., & Hofman, C. A. (2020). Do I have something in my teeth? The trouble with genetic analyses of diet from archaeological dental calculus. *Quaternary International*, in press.
Mann, A. E., Sabin, S., Ziesemer, K., Vågene, Å. J., Schroeder, H., Ozga, A. T., Sankaranarayanan, K., Hofman, C. A., Fellows Yates, J. A., Salazar-García, D. C., Frohlich, B., Aldenderfer, M., Hoogland, M., Read, C., Milner, G. R., Stone, A. C., Lewis, C. M., Jr, Krause, J., Hofman, C., Bos, K. I., & Warinner, C. (2018). Differential preservation of endogenous human and microbial DNA in dental calculus and dentin. *Scientific Reports*, *8*(1), 9822.
Mégraud, F., Lehours, P., & Vale, F. F. (2016). The history of *Helicobacter pylori*: from phylogeography to paleomicrobiology. *Clinical Microbiology and Infection*, *22*(11), 922–927.
Meyer, M., & Kircher, M. (2010). Illumina sequencing library preparation for highly multiplexed target capture and sequencing. *Cold Spring Harbor Protocols*, *2010*(6), db.prot5448.
Obregon-Tito, A. J., Tito, R. Y., Metcalf, J., Sankaranarayanan, K., Clemente, J. C., Ursell, L. K., Zech Xu, Z., Van Treuren, W., Knight, R., Gaffney, P. M., Spicer, P., Lawson, P., Marin-Reyes, L., Trujillo-Villarroel, O., Foster, M., Guija-Poma, E., Troncoso-Corzo, L., Warinner, C., Ozga, A. T., & Lewis, C. M. (2015). Subsistence strategies in traditional societies distinguish gut microbiomes. *Nature Communications*, *6*, 6505.
Oh, J., Byrd, A. L., Park, M., NISC Comparative Sequencing Program, Kong, H. H., & Segre, J. A. (2016). Temporal Stability of the Human Skin Microbiome. *Cell*, *165*(4), 854–866.
Oksanen, J., Blanchet, F. G., Friendly, M., Kindt, R., Legendre, P., McGlinn, D., Minchin, P. R., O'Hara, R. B., Simpson, G. L., Solymos, P., Stevens, M. H. H., Szoecs, E., & Wagner, H. (2019). *vegan: Community Ecology Package*.
Ottoni, C., Borić, D., Cheronet, O., Sparacello, V., Dori, I., Coppa, A., Antonović, D., Vujević, D., Douglas Price, T., Pinhasi, R., & Cristiani, E. (2021). Tracking the transition to agriculture in Southern Europe through ancient DNA analysis of dental calculus. *Proceedings of the National Academy of Sciences of the United States of America*, *118*(32).
Ozga, A. T., Nieves-Colón, M. A., Honap, T. P., Sankaranarayanan, K., Hofman, C. A., Milner, G. R., Lewis, C. M., Jr, Stone, A. C., & Warinner, C. (2016). Successful enrichment and recovery of whole mitochondrial genomes from ancient human dental calculus. *American*

Journal of Physical Anthropology, *160*(2), 220–228.
Palarea-Albaladejo, J., & Martín-Fernández, J. A. (2015). zCompositions — R package for multivariate imputation of left-censored data under a compositional approach. *Chemometrics and Intelligent Laboratory Systems*, *143*, 85–96.
Paradis, E., & Schliep, K. (2019). ape 5.0: an environment for modern phylogenetics and evolutionary analyses in R. *Bioinformatics 35*, 526–528.
Parker, C., Rohrlach, A. B., Friederich, S., Nagel, S., Meyer, M., Krause, J., Bos, K. I., & Haak, W. (2020). A systematic investigation of human DNA preservation in medieval skeletons. *Scientific Reports*, *10*(1), 18225.
Posth, C., Nägele, K., Colleran, H., Valentin, F., Bedford, S., Kami, K. W., Shing, R., Buckley, H., Kinaston, R., Walworth, M., Clark, G. R., Reepmeyer, C., Flexner, J., Maric, T., Moser, J., Gresky, J., Kiko, L., Robson, K. J., Auckland, K., Oppenheimer, S. J., Hill, A. V. S., Mentzer, A. J., Zech, J., Petchey, F., Roberts, P., Jeong, C., Gray, R. D., Krause, J., & Powell, A. (2018). Language continuity despite population replacement in Remote Oceania. *Nature Ecology & Evolution*, *2*(4), 731–740.
Pugach, I., Duggan, A. T., Merriwether, D. A., Friedlaender, F. R., Friedlaender, J. S., & Stoneking, M. (2018). The Gateway from Near into Remote Oceania: New Insights from Genome-Wide Data. *Molecular Biology and Evolution*, *35*(4), 871–886.
Radini, A., Tromp, M., Beach, A., Tong, E., Speller, C., McCormick, M., Dudgeon, J. V., Collins, M. J., Rühli, F., Kröger, R., & Warinner, C. (2019). Medieval women's early involvement in manuscript production suggested by lapis lazuli identification in dental calculus. *Science Advances*, *5*(1), eaau7126.
Rampelli, S., Schnorr, S. L., Consolandi, C., Turroni, S., Severgnini, M., Peano, C., Brigidi, P., Crittenden, A. N., Henry, A. G., & Candela, M. (2015). Metagenome Sequencing of the Hadza Hunter-Gatherer Gut Microbiota. *Current Biology*, *25*(13), 1682–1693.
R Core Team. (2021). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing.
Rohland, N., Harney, E., Mallick, S., Nordenfelt, S., & Reich, D. (2015). Partial uracil-DNA-glycosylase treatment for screening of ancient DNA. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *370*(1660), 20130624.
Sankaranarayanan, K., Ozga, A. T., Warinner, C., Tito, R. Y., Obregon-Tito, A. J., Xu, J., Gaffney, P. M., Jervis, L. L., Cox, D., Stephens, L., Foster, M., Tallbull, G., Spicer, P., & Lewis, C. M. (2015). Gut Microbiome Diversity among Cheyenne and Arapaho Individuals from Western Oklahoma. *Current Biology*, *25*(24), 3161–3169.
Sheppard, P. J. (2011). Lapita Colonization across the Near/Remote Oceania Boundary. *Current Anthropology*, *52*(6), 799–840.
Skoglund, P., Posth, C., Sirak, K., Spriggs, M., Valentin, F., Bedford, S., Clark, G. R., Reepmeyer, C., Petchey, F., Fernandes, D., Fu, Q., Harney, E., Lipson, M., Mallick, S., Novak, M., Rohland, N., Stewardson, K., Abdullah, S., Cox, M. P., Friedlaender, F. R., Friedlaender, J. S., Kivisild, T., Koki, G., Kusuma, P., Merriwether, D. A., Ricaut, F-X., Wee, J. T. S., Patterson, N., Krause, J., Pinhasi, R., & Reich, D. (2016). Genomic insights into the peopling of the Southwest Pacific. *Nature*, *538*(7626), 510–513.
Slon, V., Hopfe, C., Weiß, C. L., Mafessoni, F., de la Rasilla, M., Lalueza-Fox, C., Rosas, A., Soressi, M., Knul, M. V., Miller, R., Stewart, J. R., Derevianko, A. P., Jacobs, Z., Li, B.,

Roberts, R. G., Shunkov, M. V., de Lumley, H., Perrenoud, C., Gušić, I., Kućan, Ž., Rudan, P., Aximu-Petri, A., Essel, E., Nagel, S., Nickel, B., Schmidt, A., Prüfer, K., Kelso, J., Burbano, H. A., Pääbo, S., & Meyer, M. (2017). Neandertal and Denisovan DNA from Pleistocene sediments. *Science*, *356*(6338), 605–608.

Smith, C. I., Chamberlain, A. T., Riley, M. S., Stringer, C., & Collins, M. J. (2003). The thermal history of human fossils and the likelihood of successful DNA amplification. *Journal of Human Evolution*, *45*(3), 203–217.

Spyrou, M. A., Bos, K. I., Herbig, A., & Krause, J. (2019). Ancient pathogen genomics as an emerging tool for infectious disease research. *Nature Reviews. Genetics*, *20*(6), 323–340.

Stahl, R., Warinner, C., Velsko, I., Orfanou, E., Aron, F., & Brandt, G. (2019). Illumina double-stranded DNA dual indexing for ancient DNA v1. *protocols.io*. ZappyLab, Inc.

Steinegger, M., & Salzberg, S. L. (2020). Terminating contamination: large-scale search identifies more than 2,000,000 contaminated entries in GenBank. *Genome Biology*, *21*(1), 115.

timeanddate.com. Retrieved June 21, 2021, from https://www.timeanddate.com/

Tromp, M. (2016). *Lapita plants, people and pigs* (Thesis, Doctor of Philosophy). University of Otago.

Tromp, M., Dudgeon, J. V., Buckley, H. R., & Matisoo-Smith, E. A. (2016). Dental calculus and plant diet in Oceania. In M. Oxenham & H. Buckley (Eds.), *The Routledge Handbook of Bioarchaeology in Southeast Asia and the Pacific Islands* (pp. 599–622). Routledge.

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

Velsko, I. M., Frantz, L. A. F., Herbig, A., Larson, G., & Warinner, C. (2018). Selection of Appropriate Metagenome Taxonomic Classifiers for Ancient Microbiome Research. *mSystems*, *3*(4).

Warinner, C., Hendy, J., Speller, C., Cappellini, E., Fischer, R., Trachsel, C., Arneborg, J., Lynnerup, N., Craig, O. E., Swallow, D. M., Fotakis, A., Christensen, R. J., Olsen, J. V., Liebert, A., Montalva, N., Fiddyment, S., Charlton, S., Mackie, M., Canci, A., Bouwman, A., Rühli, F., Gilbert, M. T. P., & Collins, M. J. (2014). Direct evidence of milk consumption from ancient human dental calculus. *Scientific Reports*, *4*, 7104.

Warinner, C., Herbig, A., Mann, A., Fellows Yates, J. A., Weiß, C. L., Burbano, H. A., Orlando, L., & Krause, J. (2017). A Robust Framework for Microbial Archaeology. *Annual Review of Genomics and Human Genetics*, *18*, 321–356.

Warinner, C., Rodrigues, J. F. M., Vyas, R., Trachsel, C., Shved, N., Grossmann, J., Radini, A., Hancock, Y., Tito, R. Y., Fiddyment, S., Speller, C., Hendy, J., Charlton, S., Luder, H. U., Salazar-García, D. C., Eppler, E., Seiler, R., Hansen, L. H., Castruita, J. A. S., Barkow-Oesterreicher, S., Teoh, K. Y., Kelstrup, C. D., Olsen, J. V., Nanni, P., Kawai, T., Willerslev, E., von Mering, C., Lewis Jr, C. M., Collins, M. J., Gilbert, M. T. P., Rühli, F., & Cappellini, E. (2014). Pathogens and host immunity in the ancient human oral cavity. *Nature Genetics*, *46*(4), 336–344.

Wibowo, M. C., Yang, Z., Borry, M., Hübner, A., Huang, K. D., Tierney, B. T., Zimmerman, S., Barajas-Olmos, F., Contreras-Cubas, C., García-Ortiz, H., Martínez-Hernández, A., Luber, J. M., Kirstahler, P., Blohm, T., Smiley, F. E., Arnold, R., Ballal, S. A., Pamp, S. J., Russ, J.,

Maixner, F., Rota-Stabelli, O., Segata, N., Reinhard, K., Orozco, L., Warinner, C., Snow, M., LeBlanc, S., & Kostic, A. D. (2021). Reconstruction of ancient microbial genomes from the human gut. *Nature*, *594*(7862), 234–239.

Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L. D., François, R., Grolemund, G., Hayes, A., Henry, L., Hester, J., Kuhn, M., Pedersen, T. L., Miller, E., Bache, S. M., Müller, K., Ooms, J., Robinson, D., Seidel, D. P., Spinu, V., Takahashi, K., Vaughan, D., Wilke, C., Woo, K., & Yutani, H. (2019). Welcome to the tidyverse. In *Journal of Open Source Software, 4*(43), 1686.

Wickham, H., & Bryan, J. (2019). *readxl: Read Excel Files*.

Wilmshurst, J. M., Hunt, T. L., Lipo, C. P., & Anderson, A. J. (2011). High-precision radiocarbon dating shows recent and rapid initial human colonization of East Polynesia. *Proceedings of the National Academy of Sciences of the United States of America*, *108*(5), 1815–1820.

Yu, G., Smith, D. K., Zhu, H., Guan, Y., & Lam, T. T.-Y. (2017). Ggtree : An r package for visualization and annotation of phylogenetic trees with their covariates and other associated data. *Methods in Ecology and Evolution / British Ecological Society*, *8*(1), 28–36.

Zepner, L., Karrasch, P., Wiemann, F., & Bernard, L. (2021). ClimateCharts.net – an interactive climate analysis web platform. *International Journal of Digital Earth*, *14*(3), 338–356.

DISCUSSION

The manuscripts in this thesis demonstrate developments in the methods, standards, and approaches used for the study of DNA and proteins in archaeological dental calculus, and how they can open up new doors in the research of this information-filled archaeological material. However, there is still much progress to be done, as biomolecular dental calculus research expands and starts shifting focus from the initial descriptive studies, to more intricate research into different parts of the human past, and researchers gain a better understanding of the development and preservation of this microbial substrate.

The stability of the dental calculus microbiome

To date, many archaeogenetic studies of dental calculus put their main focus on describing and comparing the entire microbial community composition in samples from different time periods or regions of the world. A similar trend is seen in palaeoproteomic studies of dental calculus as well, where the focus has been on describing the composition of the proteome and origin of various proteins, except in cases where the aim has been dietary studies. While these studies are essential for understanding the composition and preservation of dental calculus, it is becoming increasingly apparent that the microbial community as a whole is quite resistant to taxonomic shifts stemming from altered environments, such as changes in human diet or other cultural practices (Eisenhofer *et al.*, 2020; Fellows Yates *et al.*, 2021; Ottoni *et al.*, 2021), as is also demonstrated in Manuscript D. The lack of marked differences in the community composition in different niches within the oral cavity of an individual, as shown in Manuscript C, further demonstrates that environmental differences do not have a strong effect on community composition.

Why is it that the microbial community composition in dental calculus does not seem to reflect changes in the environment? It would be logical for the community to be altered, e.g. as the nutrient supply changes when the human host starts consuming a new type of diet, or changes their behaviour or cultural practices. Two factors may be contributing to this lack of detected changes: i) The nature of how dental plaque forms and functions, and ii) the fact that researchers are reconstructing microbiomes from ancient DNA.

Dental plaque forms in layers of microbes, in a very specific order - first primary colonizers attach to proteins on the enamel surface, then bridging taxa attach to the primary colonizers, and finally secondary colonizers attach to the bridging taxa (Kolenbrander *et al.*, 2002; see

further details is Manuscript A). Each new layer of colonizing taxa depends on specific taxa displaying specific carbohydrate ligands or proteins, to be able to adhere to the biofilm. In addition to this specific layered structure, there is also communication, coaggregation and cooperation between dental plaque taxa (Kolenbrander *et al.*, 2002, 2005), as taxa can be found in specific groups that are dependent on each other (Mark Welch *et al.*, 2016; Palmer *et al.*, 2017; Kim *et al.*, 2020; Mark Welch, Ramírez-Puebla and Borisy, 2020). For example, a relationship between the early colonizer genera *Streptococcus* and *Veillonella* has been found, where the presence of oral *Veillonella* species increases the biofilm formation of a number of *Streptococcus* species (Mashima and Nakazawa, 2014). This intricate and complex structure of dental plaque and interaction between taxa may contribute to the stability of the community over time, as there is little room for major changes. This is demonstrated by a core community of bacteria generally being identified for different areas in the oral cavity (Mark Welch, Dewhirst and Borisy, 2019; Mark Welch, Ramírez-Puebla and Borisy, 2020), as well as by the presence of a core microbial community in dental calculus, which has persisted for at least 8 million years in African hominids (Fellows Yates *et al.*, 2021). This stability may lead to the lack of a marked spatial pattern within the oral cavity seen in Manuscript C. However, as such spatial patterns are seen in studies of modern dental plaque (Haffajee *et al.*, 2009; Simon-Soro and Tomás, 2013; Eren *et al.*, 2014; Proctor *et al.*, 2018), it is possible that particular aspects of ancient metagenomic studies affect our ability to detect patterns.

The challenges that arise from the study of ancient metagenomic datasets may contribute to changes not being detected in the community composition. When the DNA extracted from archaeological dental calculus is shotgun sequenced, only a small proportion of the total DNA fragments present in the sample are sequenced, due to the high DNA yield and complexity of this substrate (Mann *et al.*, 2018). As is demonstrated in Manuscript B, there is no observable difference in the reconstructed community composition between 2 mg and 10 mg of starting material, indicating that even when using only a fifth of the currently-recommended amount of starting material, the complexity is high enough that the overall community composition is not altered. This means that low abundance species may not be detected, as they account for a much smaller proportion of DNA fragments than high-abundance species. If the alterations in the community compositions are only affecting low-abundance species, it is possible that these will not be observed through shotgun sequencing.

Further, identifying taxa in metagenomic studies of archaeological dental calculus is currently largely dependent on reference databases, to which the sequenced DNA fragments are aligned to achieve species identifications. Seeing as a large part of oral taxa cannot be cultured using traditional microbiological methods (Thompson *et al.*, 2015), and therefore cannot be sequenced from a pure culture, yielding a good-quality reference genome, there are taxa in

dental calculus that are not present in databases. Diversity that was present in ancient times, but has subsequently been lost from our oral microbiomes, is also available in reference databases. The species we are able to identify in studies of dental calculus from diverse geographic regions and time periods are therefore limited to the genomes present in a database, which are unlikely to encompass all the taxa that are present in the sample. For example, Fellows Yates *et al.* (2021) reported that using two different databases, on average 36% and 51% of reads from modern human dental calculus were aligned, respectively, with the percentages being even lower for Neanderthals (26% and 36% on average). This means that a large proportion of the data in these samples cannot be analysed and a substantial amount of variation may be unreported, potentially contributing to the low differentiation between the studied populations. The same issue is present in proteomic studies, where databases can be even smaller and more biased towards model organisms, which may lead to false conclusions. For example, in Manuscript B, we find that a very poorly preserved sample does not show a high level of environmental bacterial proteins, although a large amount of environmental bacterial DNA is detected. This is likely due to the lack of environmental bacteria in proteomic databases, and if only proteomic studies had been performed on this sample, they may have led to the conclusion that this is a well-preserved sample representative of an ancient oral metaproteome. Being aware of database-induced biases is therefore essential.

In Manuscript C, we find that the major differences in microbial community composition are found between individuals, not within the oral cavity. The four studied individuals were all from the same archaeological site and of similar ages, which indicates that the differences in microbial community composition may have been driven by lifestyle-related differences between these individuals. If this is the case, it will be necessary to include several individuals per population when studying differences between populations, in order to be able to statistically account for individual-level variation. Although we conclude that a single sample of dental calculus can represent the entire oral cavity of an individual in most cases, it cannot necessarily represent an entire group of individuals. Further, in Manuscript D we show that the microbial community composition in dental calculus differs between Pacific islands, but not in a manner associated with human migration patterns or age of the sample, indicating that local conditions have a significant effect on the reconstructed microbiome. This may be caused by minor differences in diet and lifestyle between the islands, or differences in preservation caused by environmental conditions of the burial site, such as soil type and pH. If the latter is true, potential differences detected in reconstructed microbiomes between populations may be caused by differing conditions in the burial site, which needs to be taken into account when interpreting results.

Contamination from the burial ground, or after excavation, may also obscure any differences between microbiome samples. On the other hand, different contaminants present in groups of samples, such as seen in Manuscript D, may also artificially increase the separation between populations. Although there are methods for identifying contaminant taxa available (e.g., Davis *et al.*, 2018), they need further optimization to reach their full potential. Currently, they rely on identifying taxa through a higher presence in control samples as compared to the studied samples, which means that common contaminants can be easier to identify than rare contaminants. Additionally, damage patterns in the DNA sequences can be investigated to distinguish modern contaminants from ancient microbes. This is, however, currently an unfeasible approach for the large number of taxa often identified in microbiome samples, and cannot distinguish between contamination from the burial soil shortly after death, and endogenous microbiome taxa, as they will both exhibit typical ancient DNA damage patterns. Optimizing methods for detection and removal of contaminants will be essential, in order to ensure that they are not affecting comparisons in microbial community composition.

To conclude, the lack of a change in the microbial community composition over time may either be due to the stability of the dental plaque biofilm itself, or our lack of ability to identify these changes through studying ancient metagenomes. Manuscripts B, C and D in this thesis all contribute to understanding this pattern, but there are still pieces missing in the puzzle. In order to elucidate which one of these factors (or perhaps both) are driving this pattern, more studies of modern dental calculus from various human populations across the world are necessary, as well as increased efforts to sequence taxa that are missing from databases.

Seeing the trees in the forest

Seeing as few changes can be detected in the dental calculus microbiome community composition, it is likely that future studies need to shift focus to individual taxa within the community, and their evolution and function, to elucidate changes that have occurred over time in our oral microbiome. Genomic analyses have provided tremendous insights into past microbes in archaeogenetic pathogen research, such as gene presence/absence, phylogeography, recombination, and dating of strain emergence for various pathogens (Spyrou *et al.*, 2019). These types of analyses can also be employed for ancient oral microbes from a microbiome sample, provided that appropriate methods are applied. Two major challenges when working with single species within microbiomes are i) selection of target species, and ii) achieving sufficient genomic coverage.

In archaeogenetic pathogen analysis, the target species is pre-determined in cases where the cause of death of an individual is known from the archaeological context, or a species of interest has been identified through data screening. In ancient microbiome research, however, there are hundreds of species present, and which ones will be informative, depends on the research question. Selection of target species may be straightforward, if the aim is to study e.g. periodontal disease or the evolution of virulence of certain species, as these have already been researched in modern microbiological studies. However, if the goal is to study human migration through the microbiome, researchers need to identify taxa that have a phylogeographic signal. One requirement for the presence of a phylogeographic signal is that the microbiome/microbial species is inherited from parent to child, as transmission between non-related individuals or from the environment may obscure the signal, through introducing variation from unknown sources.

In general, the oral microbiome is seeded during birth and from skin and oral contact with the mother during feeding and caregiving, but some microbes are also transmitted through other routes, such as contact with other individuals (Kaan, Kahharova and Zaura, 2021). Although the oral microbiota of children are more similar to that of their mothers, than to other adults, their composition and diversity do still significantly differ (Jo *et al.*, 2021). This means that a part of the oral microbiome will have a parent-child transmission route, and thereby have the potential to show a phylogeographic signal, whereas the rest of the oral microbiome will not. In Manuscript D, two oral bacterial taxa were analyzed for signs of human migration patterns: Anaerolineaceae bacterium oral taxon 439 and *Tannerella forsythia*. Both of these taxa were chosen due to previous successful phylogenetic studies from ancient dental calculus (Bravo-Lopez *et al.*, 2020; Eisenhofer *et al.*, 2020; Ottoni *et al.*, 2021), and their high prevalence and abundance in this dataset. Very little is, however, known about Anaerolineaceae bacterium oral taxon 439, and its role and ecology within the human oral cavity, so its transmission mode cannot be ascertained. In fact, the high heterozygosity observed for many individuals in Manuscript D brings into question if perhaps several closely related taxa or strains are aligning to this reference genome, in which case the phylogenetic tree is not reflecting the evolutionary history of a single taxon. *Tannerella forsythia* is, on the other hand, more thoroughly characterised, and studies have indicated that it may be transmitted from parent to child (Okada *et al.*, 2004; Tamura *et al.*, 2006), although transmission from pets may also be possible (Yamasaki *et al.*, 2012). Deducing transmission modes of microbes and microbiomes in the past is, however, challenging, as cultural practices have changed over time, and by studying modern day transmission, some transmission modes of the past may be left unaccounted for. More thorough studies of both Anaerolineaceae bacterium oral taxon 439 and *Tannerella forsythia* are necessary, especially in modern oral microbiome samples, in order to evaluate their suitability for phylogeographic studies.

In addition, horizontal gene transfer in microbes can erode any phylogeographic signal that would otherwise be present. Horizontal gene transfer is common between bacteria living in similar oral niches, as well as from transient bacteria that are not part of the local community (Roberts and Kreth, 2014). This gene transfer may change the structure and function of biofilms over time, as the acquired genes may help bacteria e.g. adhere to a biofilm, resist antibiotics, or better utilize nutrients in the oral environment (Roberts and Kreth, 2014). Such gene transfer events may have introduced noise into the phylogenetic analyses in Manuscript D, where reads from other species may have aligned to horizontally transferred genes contained in the reference genome. These would be genes transferred into the studied species from other, often closely-related species, which have a different evolutionary history. Although identifying such gene transfer events is standard in modern microbial genetics, it is challenging in ancient microbiome research, where sequencing of a pure culture is not possible, and damage and fragmentation of the genetic information makes both alignment of DNA fragments to reference genomes and *de novo* assembly challenging. Generally, very high coverage of the target taxon makes identification of such events easier, as the genome can be more securely reconstructed.

In order to achieve high enough coverage of a genome for phylogenetic analysis, ancient pathogen research often utilizes target enrichment methods, as sequencing to a high enough depth without enrichment is seldom feasible due to cost (Spyrou *et al.*, 2019). This approach may also be useful during phylogenetic and genomic studies of single taxa in ancient microbiome research. However, the lack of high-quality reference genomes for many oral taxa, mentioned earlier, poses a challenge for such an approach. The design of target enrichment methods is challenging as only a limited amount of the potential genetic variation in the selected species will be represented, and the enrichment method may fail to recover a substantial portion of the pan genome. Additionally, the presence of closely-related taxa in the oral cavity (Beall *et al.*, 2014), as well as different strains of the same species with varying gene content (Roberts and Kreth, 2014), will further complicate enrichment method design for taxa from the oral microbiome. Hopefully, additional sequencing of modern oral taxa will be achieved, and many more complete genomes deposited in public databases during the coming years, as the cost of whole-genome sequencing is decreasing, and new methods are being developed for culturing, such as novel culture media, co-culture methods, and *in vitro* biofilm models (Thompson *et al.*, 2015) .

In addition to shifting focus to single taxa, functional analyses of the dental calculus microbiome may provide insights into evolution of the community, as this type of analysis informs about the processes in the microbial community. Functional analyses of archaeological dental calculus have already provided insights into e.g. shifting diets between populations (Fellows Yates *et al.*,

2021; Ottoni *et al.*, 2021), antimicrobial resistance (Warinner, *et al.*, 2014; Ottoni *et al.*, 2021), virulence (Warinner *et al.*, 2014), keystone species (Jacobson *et al.*, 2020), and separation in microbial functional diversity between host populations (Jacobson *et al.*, 2020; Fellows Yates *et al.*, 2021). The functional profile of the oral microbiome differs between individuals, and to a smaller degree between oral niches within an individual (Manuscript C), indicating that it may reflect even minor variations in the environmental conditions, and phenotypic plasticity within species allow for adaptation to these specific conditions. The within-individual patterns that were detected in Manuscript C were very minor, and unlikely to affect the outcome of studies, but the difference in functional profile between individuals was more marked. Seeing as the four studied individuals were from the same burial site, from the same time period, and of very similar ages, this points to lifestyle differences between individuals driving functional potential, as is also the case for the microbial community composition. If oral microbiome functional potential in human populations are compared, it will therefore be necessary to include several individuals per population, in order to account for such individual differences. As researchers keep learning more about the functional diversity of the oral microbiome in the past, this may lead to further discoveries about the evolution and adaptation of the oral microbiome over time.

To conclude, knowledge of the microbial community composition of ancient and modern dental calculus is essential for understanding this substrate. However, research into the microbiome does not need to end there, but can instead be extended into single species within the microbiome, and functional analyses of the community. This type of multifaceted research into archaeological dental calculus will lead to a holistic view of the microbiome, its function, its adaptation, and its evolution.

Limitations of archaeological dental calculus

Archaeological dental calculus contains a plethora of information about the past, as described in Manuscript A. We can learn about the host (Ozga *et al.*, 2016; Ziesemer *et al.*, 2019), the microbiome (Warinner *et al.*, 2014; Weyrich *et al.*, 2017; Fellows Yates *et al.*, 2021), pathogens (Warinner *et al.*, 2014; Fotakis *et al.*, 2020), diet (Warinner *et al.*, 2014; Jeong *et al.*, 2018; Geber *et al.*, 2019), and even occupational activities of past humans (Radini *et al.*, 2019). It may seem like the potential for knowledge yield from dental calculus is limitless. This is not the case though; due to the nature of this calcified biofilm, consisting of a mixture of microbes, human molecules, and dietary particles accumulated over a long period of time, there are some limitations to which methods in archaeological science can be applied.

Radiocarbon dating is an essential tool in archaeological sciences, allowing for direct dating of objects in archaeological sites, using carbon from organic substrates. In cases where destructive analyses of the human skeletal material is not permitted, as discussed in Manuscript D, dating dental calculus deposits could present an alternative approach. Radiocarbon dating is, however, generally not performed on dental calculus deposits. One reason may be that the material requirements for radiocarbon dating are quite high for most substrates, ranging between 50 and 600 mg (*Oxford Radiocarbon Accelerator Unit*, 2021), which exceeds the amount of dental calculus present on most archaeological individuals. Further, collagen is generally preferred for radiocarbon dating of human remains, as it is possible to check for modern-day contamination using the amino acid composition or carbon:nitrogen ratio (Grün, 2006). Dental calculus generally contains low amounts of human collagen, as the vast majority of the proteins are of microbial origin (Jersie-Christensen *et al.*, 2018). Therefore, although radiocarbon dating of dental calculus may in theory be possible, bones and teeth are a much more suitable and recommendable alternative. Radiocarbon dating of phytoliths, which are plant silica particles often recovered from dental calculus, may be an option (Piperno, 2016), provided that the accuracy of the method improves.

In archaeology, the overall diet of an individual or a population is commonly studied through stable isotopes in skeletal elements. The ratio of stable isotopes of elements such as carbon and nitrogen, extracted from bone and enamel, can inform about which broad category of food an individual was consuming (Schwarcz and Schoeninger, 1991). This method has also been applied to archaeological dental calculus. Scott and Poulson (2012) found stable isotope values in archaeological dental calculus to match expected values from literature, but were not able to provide an explanation as to how the studied elements are incorporated into dental calculus and in which way they reflect diet. In a more thorough test, Salazar-García *et al.* (2014) compared the stable isotope values of dental calculus and bone, and found that although the dental calculus values did fall within the expected range for the population, there was no correlation with the values in bone material of the same individual. The authors attribute this to the mixture of different substrates that form dental calculus (microbes, human host, diet, etc.), that would all have different sources of stable isotopes, as well as a potentially uneven distribution of different sources within the oral cavity. Further, the timing during an individual's life when the dental calculus deposit is formed, and the amount of time it has been accumulating, are unknown. Although Eerkens *et al.* (2014) showed a correlation across archaeological sites between dental calculus and bone isotope values, they also see greater variation in the values in calculus than in bone. As is shown in Manuscript C, there is variability in the microbial community composition between individuals within the same site, which may contribute to this variation in stable isotope values in dental calculus. Further, we show that the community composition somewhat differs by deposit mass, meaning that the time the deposit has been accumulating may affect the

stable isotope values. To summarize, very little is known about what is actually reflected in the stable isotope ratios in dental calculus, and how this is related to diet, and without further studies with regard to these fundamental questions, stable isotope research on archaeological dental calculus cannot be assumed to lead to accurate results.

Research into ancient pathogens through biomolecular methods has become increasingly common during recent years, as methods have developed to allow for the thorough study of a single bacterial species (Spyrou *et al.*, 2019). Although most studies focus on the pulp cavities of teeth, or skeletal lesions, dental calculus has also proved to have the potential to contain pathogens that are not part of the oral microbiome (Fotakis *et al.*, 2020). This is, however, a rare find, due to the way microbes generally are incorporated into dental calculus - by colonization from the saliva, adhering to the tooth surface and each other (Jin and Yip, 2002; Kolenbrander *et al.*, 2005). Many non-oral pathogens are, on the other hand, mainly transported around the human body via the bloodstream. In order for these pathogens to be present in dental calculus, direct contact between blood and dental calculus would be required, which mainly occurs during severe periodontitis, or presence of the pathogen in other oral soft tissues (Fotakis *et al.*, 2020). Molecules from the air have been shown to be incorporated into dental calculus (Hardy *et al.*, 2016), indicating that it is theoretically possible also for respiratory pathogens become encased in dental calculus, as they pass through the oral cavity. However, both ways for pathogen incorporation described above would only lead to a small amount of pathogens being present, as compared to the taxa in the oral microbiome, and recovering them would require either prior knowledge of their presence, targeted capture methods, or pure luck.

Finally, although the biomolecular information in dental calculus mainly consists of microbes, human molecules do also become incorporated into the matrix. This provides an opportunity to study the human genome without the need for destructive analyses of human skeletal elements. DNA enrichment technologies have been developed to increase the proportion of the human genome in the sequenced data from dental calculus, both for the entire human genome (Ziesemer *et al.*, 2019) and the mitochondrial genome (Ozga *et al.*, 2016). The initial percentage of human DNA in dental calculus is, however, very low, generally less than 0.5% of all sequenced DNA fragments (Mann *et al.*, 2018; Ziesemer *et al.*, 2019), making even target enrichment approaches challenging and costly. In addition, human DNA in dental calculus is more fragmented than the microbial DNA, possibly due to it mainly being incorporated from neutrophils during immune reactions (Mann *et al.*, 2018). Seeing as there are very low amounts of human DNA in archaeological dental calculus samples, and this DNA is highly fragmented, dental calculus is a suboptimal substrate for human genetic studies, and bones and teeth are generally preferable. However, in cases where DNA preservation is poor in skeletal elements, or destructive analyses of the skeletal elements are not permitted, dental calculus may offer an

alternative. As there are no spatial patterns within the oral cavity of human DNA incorporation into dental calculus (Manuscript C), human DNA can be analyzed from any dental calculus deposit, without the need to aim for a specific sampling design.

The importance of study design

The pace of research is high as a new field develops, leaving little time for planning and designing studies, as novel results are constantly published and the risk of scooping is high. Having a proper study design may, however, hinder biased or even false results from being published, or lower the risk of study outcomes being unpublishable. Biases are unavoidable in all studies; what is important is being aware of them and designing the studies to avoid biases altering results. In modern microbiome studies, different biases have been mapped out at all stages of a study (Nearing, Comeau and Langille, 2021), and ancient microbiome analyses have their own additional biases, which researchers need to consider. Prior to each study, potential biases should be carefully outlined, so that they can be accounted for through study design or statistical analyses. Manuscripts B, C and D in this thesis all contribute to knowledge of study design and result interpretation, at different stages of the research process, and aid future studies to be designed in a manner that maximizes information output and minimizes the risk of result-altering biases.

In Manuscript B, we show that both DNA and proteins can be extracted from one single sample through one laboratory protocol, allowing researchers to minimize the amount of destructive analyses. However, there are two (minor) biases that this unified protocol introduces - lower DNA yield and an underrepresentation of peptides with basic amino acids. The lower DNA yield will in most cases not affect study outcome, as dental calculus is in general very rich in DNA (Mann *et al.*, 2018), far in excess of what is commonly sequenced, and the loss was found to be most pronounced in well-preserved samples. As no differences could be detected in downstream analyses between 2 mg and 10 mg of dental calculus being analyzed, with either the unified or a DNA-only protocol, this further supports that a lower yield will not affect most study outcomes. However, in cases with very poor DNA preservation, it may be necessary to increase the amount of input material, in order to mitigate the lower yield through the unified protocol. If this is not possible, due to a small amount of available sample, it may be advisable to alter the study design to only extract one type of biomolecule, in order to avoid destruction of the sample without achieving any results. The second bias, the underrepresentation of basic amino acids which may potentially lead to a lower recovery of hydrophilic proteins, may be accounted for in statistical analyses, if the proteome composition is compared between samples extracted through different protocols. If a hydrophilic protein is the target of a study, it is

recommendable to conduct further experiments to ensure that this protein can be successfully recovered. To summarize, if biases from a laboratory protocol are known from the start of a study, and the aim is clear, study design and statistical analyses can be adapted to account for biases. This is, however, only possible if a new laboratory protocol has been thoroughly and systematically compared to existing standard protocols, as is the case in Manuscript B.

When designing a study, samples need to be selected so that they are not introducing biases themselves. As shown in Manuscript C, the data itself may be inherently biased, if sampling is conducted from different oral niches. Although the patterns that were found were minor, they could still lead to inappropriate interpretations. For example, if differences in sampling location are not considered during analysis and result interpretation, the gradient in oxygen tolerance we detected within the oral cavity may lead to erroneous conclusions. Although it may not be possible to design a study in a manner that avoids introducing any spatial biases, especially in cases where very little dental calculus is available or sampling is restricted, being aware of these biases means they can be accounted for.

In Manuscript D, we see another example of study design introducing biases into the data. In this case, the dental calculus samples were processed in three different labs, which caused significant differences in the reconstructed microbiome composition. This shows how important it is to, whenever possible, randomize sample processing from beginning to end, while also trying to keep conditions and protocols as similar as possible. Batch effects are a commonly recognized issue in molecular biology, where unexpected differences between processing batches can arise although experiments have been conducted in the same manner (Bálint *et al.*, 2018). These differences may arise from different laboratory personnel processing the samples, different batches of reagents being used, or even something as seemingly trivial as different temperatures in the laboratory. For ancient biomolecule research, contaminants from the laboratory can be a major issue, as they bias or swamp out the original community profile. Contaminants in archaeogenetic studies have been found to vary between laboratories, seasons, and personnel (Weyrich *et al.*, 2019). If samples from different archaeological sites are processed in separate batches, there is therefore a risk of introducing different contaminants into the batches, as well as having a different level of efficacy in the processing, leading to a separation between sites during analyses that is not due to a biological reason. Randomization is thereby highly recommendable, through sampling, extraction, and library preparation, as well as keeping factors such as the environment, reagents, and personnel as consistent as possible across batches in the same study (Nearing, Comeau and Langille, 2021).

Future directions

The study of biomolecules in archaeological dental calculus has come a long way since the first studies were published in the early 2010s. With the manuscripts in this thesis, methods are refined and pushed even further, opening up new opportunities for this research topic. Yet, there is still much to be done.

The manuscripts presented in this thesis provide answers to questions that archaeological scientists have long been asking, but they also open up new questions and possibilities. For example, in Manuscript B we demonstrate that it is possible to successfully extract both DNA and proteins from a single sample of dental calculus. This leads to asking, can we use this method also for other archaeological materials, such as e.g. human bone? Proteins have been shown to preserve in human skeletal material for a longer time than DNA (Chen *et al.*, 2019; Welker *et al.*, 2020), and are informative of hominin phylogeny (Welker *et al.*, 2020). However, the genome is bound to contain a higher amount of information about human evolution than the proteome, as codon redundancy leads to a higher number of nucleotide polymorphisms than amino acid polymorphisms, and only a small fraction of the genome is expressed in proteins. A logical course of action would therefore be to simultaneously extract both DNA and proteins from very rare ancient samples, so that proteins can be analysed in cases where DNA is not preserved. Seeing as the ancient dental calculus DNA extraction protocols tend to be very similar to ancient skeletal material DNA extraction protocols (Velsko, Skourtanioti and Brandt, 2019; Aron *et al.*, 2020), this unified protocol should be applicable to skeletal material. However, this assumption needs to be tested, as there might be unexpected differences in e.g. the proportion of the proteome that is recovered, or the DNA yield from a material that contains less DNA to begin with.

Furthermore, in Manuscript C we demonstrate that there are only minor differences in the metagenomic content across the dental arcade. This does, however, lead to the question if the same holds true for proteins. A majority of the palaeoproteomic studies of archaeological dental calculus have so far focused on diet, as dietary proteins are well preserved within the dental calculus matrix and can directly inform us about the dietary habits of a single individual (Warinner *et al.*, 2014; Hendy *et al.*, 2018; Geber *et al.*, 2019; Scott *et al.*, 2021). There may, however, be a spatial bias in the incorporation of dietary fragments into dental calculus, as teeth that are more commonly used for mastication may contain more proteins from food that needs to be chewed. Performing a similar experiment as in Manuscript C, but studying proteins instead of DNA, may illuminate which oral sites are more likely to yield dietary proteins, as well as if there are spatial patterns in the microbial and human proteins in dental calculus.

Archaeological dental calculus has long been seen as a material of low value, and its analysis has been suggested to be ‘non-destructive’, as opposed to similar analyses performed on bones and teeth. However, as we continue to discover an increasing amount of information that can be gained from this substrate, this view will hopefully change. Dental calculus is a finite archaeological material, and destructive analyses may hinder any future analyses of the same material. Care should therefore be taken to preserve this archaeological substrate for future generations - whether it be by using combined extraction protocols, refining protocols so that less starting material can be used, or simply by carefully defining the study question and design, minimizing the risk of failure prior to destructive analyses.

SUMMARY

Studies of ancient microbiomes can shed light on several aspects of human history and microbial evolution, as well as aid in understanding modern day diseases. During recent years, studies of archaeological dental calculus have gained much interest, as this calcified microbial biofilm preserves biomolecules over long periods of time, and can hold information about health, disease, diet, the human host, and occasionally even the living environment and occupational activities of an individual. However, continued advances in the research of dental calculus are essential, in order to make sure that this finite archaeological material is utilized in the best way possible. This thesis presents several such developments, beginning with a thorough review of formation and occurrence of archaeological dental calculus, as well as its past and current study in Manuscript A, in order to set a basis for understanding the research into this topic. Thereafter, state-of-the-art archaeogenetic and palaeoproteomic methods are employed to advance the study of archaeological dental calculus at different stages in the research process, from project design and laboratory processing, to analysis and result interpretation.

In Manuscript B, I present and thoroughly evaluate a protocol for simultaneously extracting both DNA and proteins from archaeological dental calculus. I find that this unified protocol is a successful approach, introducing only minor biases into downstream analyses, which can be accounted for in project design. This protocol allows researchers to maximize the amount of information that is extracted from each sample, while minimizing the amount of destructive analyses that are performed. To investigate potential patterns in the microbial community across the dental arcade, I conduct a metagenomic study of a dense sampling of archaeological dental calculus from different oral niches in Manuscript C. This study shows that although there are minor patterns present, the between-individual variation is generally more significant, and thereby a single sample can represent the entire oral cavity of an individual. However, seeing as some patterns are detected, care should be taken to document sampling location within the oral cavity, in order to be able to account for such biases during analysis and result interpretation. Finally, in Manuscript D, I apply metagenomic analysis of dental calculus to a new approach - studying ancient human migration. By using the Pacific region as a case study, I show that the microbial community composition as a whole does not reflect human migration, but its structure is rather driven by local factors on each studied island. Phylogenetic studies of single taxa do, on the other hand, show time- and geography related patterns. Metagenomic studies of archaeological dental calculus therefore have potential for studies of human migration, in cases where genetic analyses of the human genome are not feasible.

Taken together, these four manuscripts push forward the development of the field of biomolecular research of archaeological dental calculus. They provide researchers with valuable

information for study design and result interpretation, as well as introduce new types of analyses into the field.

ZUSAMMENFASSUNG

Studien über alte Mikrobiome können verschiedene Aspekte der menschlichen Geschichte und der mikrobiellen Evolution beleuchten und zum Verständnis moderner Krankheiten betragen. In den letzten Jahren haben Studien von archäologischem Zahnstein viel Interesse geweckt, da dieser verkalkte, mikrobielle Biofilm Biomoleküle über lange Zeiträume konserviert und Informationen über Gesundheit, Krankheit, Ernährung, den menschlichen Wirt und gelegentlich sogar über das Lebensumfeld und die Tätigkeiten einer Person enthalten kann. Allerdings sind kontinuierliche Fortschritte bei der Erforschung von Zahnstein unerlässlich, um sicherzustellen, dass dieses endliche archäologische Material bestmöglich genutzt wird. In dieser Dissertation werden mehrere solcher Entwicklungen vorgestellt, beginnend mit einem gründlichen Überblick über die Entstehung und das Vorkommen von archäologischem Zahnstein, sowie dessen vergangene und aktuelle Erforschung in Manuskript A, um eine Grundlage für das Verständnis der Forschung zu diesem Thema zu schaffen. Danach werden modernste archäogenetische und paläoproteomische Methoden angewandt, um die Untersuchung von archäologischem Zahnstein in verschiedenen Stadien des Forschungsprozesses voranzutreiben, von der Projektplanung über die Verarbeitung im Labor bis hin zur Analyse und Interpretation der Ergebnisse.

In Manuskript B stelle ich ein Protokoll zur gleichzeitigen Extraktion von DNA und Proteinen aus archäologischem Zahnstein vor und bewerte es eingehend. Ich stelle fest, dass dieses einheitliche Protokoll ein erfolgreicher Ansatz ist, der nur geringfügige Verzerrungen in nachgelagerte Analysen verursacht, die bei der Projektplanung berücksichtigt werden können. Dieses Protokoll ermöglicht es den Forschern, die Menge an Informationen, die aus jeder Probe gewonnen wird, zu maximieren und gleichzeitig die Menge an zerstörerischen Analysen, die durchgeführt werden, zu minimieren. Um potenzielle Muster in der mikrobiellen Gemeinschaft im gesamten Zahnbogen zu untersuchen, führe ich eine metagenomische Studie mit einer dichten Stichprobe von archäologischem Zahnstein aus verschiedenen oralen Nischen in Manuskript C durch. Diese Studie zeigt, dass es zwar geringfügige Muster gibt, die Variation zwischen den einzelnen Individuen jedoch im Allgemeinen signifikanter ist, so dass eine einzige Probe die gesamte Mundhöhle eines Individuums darstellen kann. Da jedoch einige Muster festgestellt werden, sollte darauf geachtet werden, dass der Ort der Probenahme in der Mundhöhle dokumentiert wird, um derartige Verzerrungen bei der Analyse und der Interpretation der Ergebnisse berücksichtigen zu können. In Manuskript D schließlich verwende ich die metagenomische Analyse von Zahnstein für einen neuen Ansatz: die Untersuchung alter menschlicher Migration. Anhand der Pazifikregion als Fallstudie zeige ich, dass die Zusammensetzung der mikrobiellen Gemeinschaft als Ganzes nicht die menschliche Migration widerspiegelt, sondern vielmehr durch lokale Faktoren auf jeder untersuchten Insel bestimmt

wird. Phylogenetische Studien einzelner Taxa zeigen hingegen zeit- und geografisch bedingte Muster. Metagenomische Studien an archäologischem Zahnstein haben daher das Potenzial, die menschliche Migration zu untersuchen, wenn genetische Analysen des menschlichen Genoms nicht möglich sind.

Zusammengenommen bringen diese vier Manuskripte die Entwicklung des Feldes der biomolekularen Forschung von archäologischem Zahnstein voran. Sie liefern Forschern wertvolle Informationen für Studiendesign und Ergebnisinterpretation und führen neue Arten von Analysen in das Feld ein.

REFERENCES

Adler, C. J., Dobney, K., Weyrich, L. S., Kaidonis, J., Walker, A. W., Haak, W., Bradshaw, C. J. A., Townsend, G., Sołtysiak, A., Alt, K. W., Parkhill, J., & Cooper, A. (2013). Sequencing ancient calcified dental plaque shows changes in oral microbiota with dietary shifts of the Neolithic and Industrial revolutions. *Nature Genetics*, *45*(4), 450–455.

Aouizerat, T., Gutman, I., Paz, Y., Maeir, A. M., Gadot, Y., Gelman, D., Szitenberg, A., Drori, E., Pinkus, A., Schoemann, M., Kaplan, R., Ben-Gedalya, T., Coppenhagen-Glazer, S., Reich, E., Saragovi, A., Lipschits, O., Klutstein, M., & Hazan, R. (2019). Isolation and Characterization of Live Yeast Cells from Ancient Vessels as a Tool in Bio-Archaeology. *mBio*, *10*(2).

Aron, F., Hofman, C., Fagernäs, Z., Velsko, I., Skourtanioti, E., Brandt, G., & Warinner, C. (2020). Ancient DNA extraction from dental calculus v1. *protocols.io*. ZappyLab, Inc.

Austin, J. J., Smith, A. B., & Thomas, R. H. (1997). Palaeontology in a molecular world: the search for authentic ancient DNA. *TREE*, *12*(8), 303–306.

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

Bálint, M., Márton, O., Schatz, M., Düring, R.-A., & Grossart, H.-P. (2018). Proper experimental design requires randomization/balancing of molecular ecology experiments. *Ecology and Evolution*, *8*(3), 1786–1793.

Barlow, G. M., Yu, A., & Mathur, R. (2015). Role of the Gut Microbiome in Obesity and Diabetes Mellitus. *Nutrition in Clinical Practice, 30*(6), 787–797.

Beall, C. J., Campbell, A. G., Dayeh, D. M., Griffen, A. L., Podar, M., & Leys, E. J. (2014). Single cell genomics of uncultured, health-associated *Tannerella* BU063 (Oral Taxon 286) and comparison to the closely related pathogen *Tannerella forsythia*. *PloS One*, *9*(2), e89398.

Boyd, W. C., & Boyd, L. G. (1937). Blood Grouping Tests on 300 Mummies. *The Journal of Immunology*, *32*(4), 307–319.

Bravo-Lopez, M., Villa-Islas, V., Rocha Arriaga, C., Villaseñor-Altamirano, A. B., Guzmán-Solís, A., Sandoval-Velasco, M., Wesp, J. K., Alcantara, K., López-Corral, A., Gómez-Valdés, J., Mejía, E., Herrera, A., Meraz-Moreno, A., Moreno-Cabrera, M. de la L., Moreno-Estrada, A., Nieves-Colón, M. A., Olvera, J., Pérez-Pérez, J., Iversen, K. H., Rasmussen, S., Sandoval, K., Zepeda, G., & Ávila-Arcos, M. C. (2020). Paleogenomic insights into the red complex bacteria *Tannerella forsythia* in Pre-Hispanic and Colonial individuals from Mexico. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *375*(1812), 20190580.

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

Cappellini, E., Jensen, L. J., Szklarczyk, D., Ginolhac, A., da Fonseca, R. A. R., Stafford, T. W., Holen, S. R., Collins, M. J., Orlando, L., Willerslev, E., Gilbert, M. T. P., & Olsen, J. V.

(2012). Proteomic analysis of a pleistocene mammoth femur reveals more than one hundred ancient bone proteins. *Journal of Proteome Research*, *11*(2), 917–926.

Carmody, R. N., Sarkar, A., & Reese, A. T. (2021). Gut microbiota through an evolutionary lens. *Science*, *372*(6541), 462–463.

Chen, F., Welker, F., Shen, C.-C., Bailey, S. E., Bergmann, I., Davis, S., Xia, H., Wang, H., Fischer, R., Freidline, S. E., Yu, T.-L., Skinner, M. M., Stelzer, S., Dong, G., Fu, Q., Dong, G., Wang, J., Zhang, D., & Hublin, J.-J. (2019). A late Middle Pleistocene Denisovan mandible from the Tibetan Plateau. *Nature*, *569*(7756), 409–412.

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

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

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

Davis, N. M., Proctor, D. M., Holmes, S. P., Relman, D. A., & Callahan, B. J. (2018). Simple statistical identification and removal of contaminant sequences in marker-gene and metagenomics data. *Microbiome*, *6*(1), 226.

DeSalle, R., Gatesy, J., Wheeler, W., & Grimaldi, D. (1992). DNA sequences from a fossil termite in Oligo-Miocene amber and their phylogenetic implications. *Science*, *257*(5078), 1933–1936.

Devault, A. M., Golding, G. B., Waglechner, N., Enk, J. M., Kuch, M., Tien, J. H., Shi, M., Fisman, D. N., Dhody, A. N., Forrest, S., Bos, K. I., Earn, D. J. D., Holmes, E. C., & Poinar, H. N. (2014). Second-pandemic strain of *Vibrio cholerae* from the Philadelphia cholera outbreak of 1849. *The New England Journal of Medicine*, *370*(4), 334–340.

Dickson, J. H., Oeggl, K., & Stanton, D. (2017). “Forest Moss”: no part of the European Neanderthal diet. *Antiquity*, *91*(359), e3.

Downs, E. F., & Lowenstein, J. M. (1995). Identification of archaeological blood proteins: A cautionary note. *Journal of Archaeological Science*, *22*(1), 11–16.

Eerkens, J. W., de Voogt, A., Dupras, T. L., Rose, S. C., Bartelink, E. J., & Francigny, V. (2014). Intra- and inter-individual variation in δ13C and δ15N in human dental calculus and comparison to bone collagen and apatite isotopes. *Journal of Archaeological Science*, *52,* 64-71.

Eisenhofer, R., Anderson, A., Dobney, K., Cooper, A., & Weyrich, L. S. (2019). Ancient Microbial DNA in Dental Calculus: A New method for Studying Rapid Human Migration Events. *The Journal of Island and Coastal Archaeology*, *14*(2), 149–162.

Eisenhofer, R., Kanzawa-Kiriyama, H., Shinoda, K.-I., & Weyrich, L. S. (2020). Investigating the demographic history of Japan using ancient oral microbiota. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *375*(1812), 20190578.

Eisenhofer, R., & Weyrich, L. S. (2019). Assessing alignment-based taxonomic classification of ancient microbial DNA. *PeerJ*, *7*, e6594.

Eren, A. M., Borisy, G. G., Huse, S. M., & Mark Welch, J. L. (2014). Oligotyping analysis of the human oral microbiome. *Proceedings of the National Academy of Sciences of the United States of America*, *111*(28), E2875–E2884.

Evershed, R. P. (1993). Biomolecular archaeology and lipids. *World Archaeology*, *25*(1), 74–93.

Fellows Yates, J. A., Velsko, I. M., Aron, F., Posth, C., Hofman, C. A., Austin, R. M., Parker, C. E., Mann, A. E., Nägele, K., Arthur, K. W., Arthur, J. W., Bauer, C. C., Crevecoeur, I., Cupillard, C., Curtis, M. C., Dalén, L., Díaz-Zorita Bonilla, M., Díez Fernández-Lomana, J. C., Drucker, D. G., Escribano Escrivá, E., Francken, M., Gibbon, V. E., González Morales, M. R., Grande Mateu, A., Harvati, K., Henry, A. G., Humphrey, L., Menéndez, M., Mihailović, D., Peresani, M., Rodríguez Moroder, S., Roksandic, M., Rougier, H., Sázelová, S., Stock, J. T., Straus, L. G., Svoboda, J., Teßmann, B., Walker, M. J., Power, R. C., Lewis, C. M., Sankaranarayanan, K., Guschanski, K., Wrangham, R. W., Dewhirst, F. E., Salazar-García, D. C., Krause, J., Herbig, A., & Warinner, C. (2021). The evolution and changing ecology of the African hominid oral microbiome. *Proceedings of the National Academy of Sciences of the United States of America*, *118*(20), e2021655118.

Flandroy, L., Poutahidis, T., Berg, G., Clarke, G., Dao, M.-C., Decaestecker, E., Furman, E., Haahtela, T., Massart, S., Plovier, H., Sanz, Y., & Rook, G. (2018). The impact of human activities and lifestyles on the interlinked microbiota and health of humans and of ecosystems. *The Science of the Total Environment*, *627*, 1018–1038.

Fordyce, S. L., Ávila-Arcos, M. C., Rasmussen, M., Cappellini, E., Romero-Navarro, J. A., Wales, N., Alquezar-Planas, D. E., Penfield, S., Brown, T. A., Vielle-Calzada, J.-P., Montiel, R., Jørgensen, T., Odegaard, N., Jacobs, M., Arriaza, B., Higham, T. F. G., Ramsey, C. B., Willerslev, E., & Gilbert, M. T. P. (2013). Deep sequencing of RNA from ancient maize kernels. *PloS One*, *8*(1), e50961.

Fotakis, A. K., Denham, S. D., Mackie, M., Orbegozo, M. I., Mylopotamitaki, D., Gopalakrishnan, S., Sicheritz-Pontén, T., Olsen, J. V., Cappellini, E., Zhang, G., Christophersen, A., Gilbert, M. T. P., & Vågene, Å. J. (2020). Multi-omic detection of Mycobacterium leprae in archaeological human dental calculus. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *375*(1812), 20190584.

Fox, K., & Hawks, J. (2019). Use ancient remains more wisely. *Nature*, *572*(7771), 581–583.

Geber, J., Tromp, M., Scott, A., Bouwman, A., Nanni, P., Grossmann, J., Hendy, J., & Warinner, C. (2019). Relief food subsistence revealed by microparticle and proteomic analyses of dental calculus from victims of the Great Irish Famine. *Proceedings of the National Academy of Sciences of the United States of America*, *116*(39), 19380–19385.

Gillings, M. R., Paulsen, I. T., & Tetu, S. G. (2015). Ecology and Evolution of the Human Microbiota: Fire, Farming and Antibiotics. *Genes*, *6*(3), 841–857.

Green, R. E., Krause, J., Ptak, S. E., Briggs, A. W., Ronan, M. T., Simons, J. F., Du, L., Egholm, M., Rothberg, J. M., Paunovic, M., & Pääbo, S. (2006). Analysis of one million base pairs of Neanderthal DNA. *Nature*, *444*(7117), 330–336.

Grün, R. (2006). Direct dating of human fossils. *American Journal of Physical Anthropology*, *Suppl 43*, 2–48.

Haffajee, A. D., Teles, R. P., Patel, M. R., Song, X., Yaskell, T., & Socransky, S. S. (2009). Factors affecting human supragingival biofilm composition. II. Tooth position. *Journal of Periodontal Research*, *44*(4), 520–528.

Hardy, K., Buckley, S., Collins, M. J., Estalrrich, A., Brothwell, D., Copeland, L., García-Tabernero, A., García-Vargas, S., de la Rasilla, M., Lalueza-Fox, C., Huguet, R., Bastir, M., Santamaría, D., Madella, M., Wilson, J., Cortés, A. F., & Rosas, A. (2012). Neanderthal medics? Evidence for food, cooking, and medicinal plants entrapped in dental calculus. *Die Naturwissenschaften*, *99*(8), 617–626.

Hardy, K., Radini, A., Buckley, S., Sarig, R., Copeland, L., Gopher, A., & Barkai, R. (2016). Dental calculus reveals potential respiratory irritants and ingestion of essential plant-based nutrients at Lower Palaeolithic Qesem Cave Israel. *Quaternary International: The Journal of the International Union for Quaternary Research*, *398*, 129–135.

Harper, K. N., & Armelagos, G. J. (2013). Genomics, the origins of agriculture, and our changing microbe-scape: time to revisit some old tales and tell some new ones. *American Journal of Physical Anthropology*, *152 Suppl 57*, 135–152.

Hendy, J. (2021). Ancient protein analysis in archaeology. *Science Advances*, *7*(3), eabb9314.

Hendy, J., Colonese, A. C., Franz, I., Fernandes, R., Fischer, R., Orton, D., Lucquin, A., Spindler, L., Anvari, J., Stroud, E., Biehl, P. F., Speller, C., Boivin, N., Mackie, M., Jersie-Christensen, R. R., Olsen, J. V., Collins, M. J., Craig, O. E., & Rosenstock, E. (2018). Ancient proteins from ceramic vessels at Çatalhöyük West reveal the hidden cuisine of early farmers. *Nature Communications*, *9*(1), 4064.

Hendy, J., Warinner, C., Bouwman, A., Collins, M. J., Fiddyment, S., Fischer, R., Hagan, R., Hofman, C. A., Holst, M., Chaves, E., Klaus, L., Larson, G., Mackie, M., McGrath, K., Mundorff, A. Z., Radini, A., Rao, H., Trachsel, C., Velsko, I. M., & Speller, C. F. (2018). Proteomic evidence of dietary sources in ancient dental calculus. *Proceedings of the Royal Society B: Biological Sciences*, *285*(1883).

Hendy, J., Welker, F., Demarchi, B., Speller, C., Warinner, C., & Collins, M. J. (2018). A guide to ancient protein studies. *Nature Ecology & Evolution*, *2*(5), 791–799.

Hershkovitz, I., Kelly, J., Latimer, B., Rothschild, B. M., Simpson, S., Polak, J., & Rosenberg, M. (1997). Oral bacteria in miocene *Sivapithecus*. *Journal of Human Evolution*, *33*(4), 507–512.

Higuchi, R., Bowman, B., Freiberger, M., Ryder, O. A., & Wilson, A. C. (1984). DNA sequences from the quagga, an extinct member of the horse family. *Nature*, *312*(5991), 282–284.

Hübler, R., Key, F. M., Warinner, C., Bos, K. I., Krause, J., & Herbig, A. (2019). HOPS: automated detection and authentication of pathogen DNA in archaeological remains. *Genome Biology*, *20*, 280.

Jacobson, D. K., Honap, T. P., Monroe, C., Lund, J., Houk, B. A., Novotny, A. C., Robin, C., Marini, E., & Lewis, C. M., Jr. (2020). Functional diversity of microbial ecologies estimated from ancient human coprolites and dental calculus. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *375*(1812), 20190586.

Jensen, T. Z. T., Niemann, J., Iversen, K. H., Fotakis, A. K., Gopalakrishnan, S., Vågene, Å. J., Pedersen, M. W., Sinding, M.-H. S., Ellegaard, M. R., Allentoft, M. E., Lanigan, L. T., Taurozzi, A. J., Nielsen, S. H., Dee, M. W., Mortensen, M. N., Christensen, M. C., Sørensen, S. A., Collins, M. J., Gilbert, M. T. P., Sikora, M., Rasmussen, S., & Schroeder, H. (2019). A 5700 year-old human genome and oral microbiome from chewed birch pitch. *Nature Communications*, *10*(1), 5520.

Jeong, C., Wilkin, S., Amgalantugs, T., Bouwman, A. S., Taylor, W. T. T., Hagan, R. W., Bromage, S., Tsolmon, S., Trachsel, C., Grossmann, J., Littleton, J., Makarewicz, C. A.,

Krigbaum, J., Burri, M., Scott, A., Davaasambuu, G., Wright, J., Irmer, F., Myagmar, E., Boivin, N., Robbeets, M., Rühli, F. J., Krause, J., Frohlich, B., Hendy, J., & Warinner, C. (2018). Bronze Age population dynamics and the rise of dairy pastoralism on the eastern Eurasian steppe. *Proceedings of the National Academy of Sciences of the United States of America*, *115*(48), E11248–E11255.

Jersie-Christensen, R. R., Lanigan, L. T., Lyon, D., Mackie, M., Belstrøm, D., Kelstrup, C. D., Fotakis, A. K., Willerslev, E., Lynnerup, N., Jensen, L. J., Cappellini, E., & Olsen, J. V. (2018). Quantitative metaproteomics of medieval dental calculus reveals individual oral health status. *Nature Communications*, *9*(1), 4744.

Jin, Y., & Yip, H.-K. (2002). Supragingival calculus: formation and control. *Critical Reviews in Oral Biology and Medicine*, *13*(5), 426–441.

Jones, E. D., & Bösl, E. (2021). Ancient human DNA: A history of hype (then and now). *Journal of Social Archaeology*, *21*(2), 236–255.

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

Jo, R., Yama, K., Aita, Y., Tsutsumi, K., Ishihara, C., Maruyama, M., Takeda, K., Nishinaga, E., Shibasaki, K.-I., & Morishima, S. (2021). Comparison of oral microbiome profiles in 18-month-old infants and their parents. *Scientific Reports*, *11*(1), 861.

Kaan, A. M. M., Kahharova, D., & Zaura, E. (2021). Acquisition and establishment of the oral microbiota. *Periodontology 2000*, *86*(1), 123–141.

Kim, D., Barraza, J. P., Arthur, R. A., Hara, A., Lewis, K., Liu, Y., Scisci, E. L., Hajishengallis, E., Whiteley, M., & Koo, H. (2020). Spatial mapping of polymicrobial communities reveals a precise biogeography associated with human dental caries. *Proceedings of the National Academy of Sciences of the United States of America*, *117*(22), 12375–12386.

Kistler, L., Ware, R., Smith, O., Collins, M., & Allaby, R. G. (2017). A new model for ancient DNA decay based on paleogenomic meta-analysis. *Nucleic Acids Research*, *45*(11), 6310–6320.

Knights, D., Kuczynski, J., Charlson, E. S., Zaneveld, J., Mozer, M. C., Collman, R. G., Bushman, F. D., Knight, R., & Kelley, S. T. (2011). Bayesian community-wide culture-independent microbial source tracking. *Nature Methods*, *8*(9), 761–763.

Kolenbrander, P. E., Andersen, R. N., Blehert, D. S., Egland, P. G., Foster, J. S., & Palmer, R. J., Jr. (2002). Communication among oral bacteria. *Microbiology and Molecular Biology Reviews*, *66*(3), 486–505.

Kolenbrander, P. E., Egland, P. G., Diaz, P. I., & Palmer, R. J., Jr. (2005). Genome-genome interactions: bacterial communities in initial dental plaque. *Trends in Microbiology*, *13*(1), 11–15.

Kontopoulos, I., Penkman, K., Mullin, V. E., Winkelbach, L., Unterländer, M., Scheu, A., Kreutzer, S., Hansen, H. B., Margaryan, A., Teasdale, M. D., Gehlen, B., Street, M., Lynnerup, N., Liritzis, I., Sampson, A., Papageorgopoulou, C., Allentoft, M. E., Burger, J., Bradley, D. G., & Collins, M. J. (2020). Screening archaeological bone for palaeogenetic and palaeoproteomic studies. *PloS One*, *15*(6), e0235146.

Korlević, P., Talamo, S., & Meyer, M. (2018). A combined method for DNA analysis and radiocarbon dating from a single sample. *Scientific Reports*, *8*(1), 4127.

LeBlanc, J. G., Milani, C., de Giori, G. S., Sesma, F., van Sinderen, D., & Ventura, M. (2013). Bacteria as vitamin suppliers to their host: a gut microbiota perspective. *Current Opinion in Biotechnology*, *24*(2), 160–168.

Lederberg, J., & McCray, A. T. (2001). 'Ome Sweet 'Omics - A Genealogical Treasury of Words. *The Scientist*, *15*(7), 8.

Lugli, G. A., Milani, C., Mancabelli, L., Turroni, F., Ferrario, C., Duranti, S., van Sinderen, D., & Ventura, M. (2017). Ancient bacteria of the Ötzi's microbiome: a genomic tale from the Copper Age. *Microbiome*, *5*(1), 5.

Mann, A. E., Fellows Yates, J. A., Fagernäs, Z., Austin, R. M., Nelson, E. A., & Hofman, C. A. (2020). Do I have something in my teeth? The trouble with genetic analyses of diet from archaeological dental calculus. *Quaternary International,* in press.

Mann, A. E., Sabin, S., Ziesemer, K., Vågene, Å. J., Schroeder, H., Ozga, A. T., Sankaranarayanan, K., Hofman, C. A., Fellows Yates, J. A., Salazar-García, D. C., Frohlich, B., Aldenderfer, M., Hoogland, M., Read, C., Milner, G. R., Stone, A. C., Lewis, C. M., Jr, Krause, J., Hofman, C., Bos, K. I., & Warinner, C. (2018). Differential preservation of endogenous human and microbial DNA in dental calculus and dentin. *Scientific Reports*, *8*(1), 9822.

Mark Welch, J. L., Dewhirst, F. E., & Borisy, G. G. (2019). Biogeography of the Oral Microbiome: The Site-Specialist Hypothesis. *Annual Review of Microbiology*, *73*, 335–358.

Mark Welch, J. L., Ramírez-Puebla, S. T., & Borisy, G. G. (2020). Oral Microbiome Geography: Micron-Scale Habitat and Niche. *Cell Host & Microbe*, *28*(2), 160–168.

Mark Welch, J. L., Rossetti, B. J., Rieken, C. W., Dewhirst, F. E., & Borisy, G. G. (2016). Biogeography of a human oral microbiome at the micron scale. *Proceedings of the National Academy of Sciences of the United States of America*, *113*(6), E791–E800.

Mashima, I., & Nakazawa, F. (2014). The influence of oral *Veillonella* species on biofilms formed by *Streptococcus* species. *Anaerobe*, *28*, 54–61.

Mégraud, F., Lehours, P., & Vale, F. F. (2016). The history of *Helicobacter pylori*: from phylogeography to paleomicrobiology. *Clinical Microbiology and Infection, 22*(11), 922–927.

Modi, A., Pisaneschi, L., Zaro, V., Vai, S., Vergata, C., Casalone, E., Caramelli, D., Moggi-Cecchi, J., Mariotti Lippi, M., & Lari, M. (2020). Combined methodologies for gaining much information from ancient dental calculus: testing experimental strategies for simultaneously analysing DNA and food residues. *Archaeological and Anthropological Sciences*, *12*(1), 10.

Nearing, J. T., Comeau, A. M., & Langille, M. G. I. (2021). Identifying biases and their potential solutions in human microbiome studies. *Microbiome*, *9*(1), 113.

NIH HMP Working Group, Peterson, J., Garges, S., Giovanni, M., McInnes, P., Wang, L., Schloss, J. A., Bonazzi, V., McEwen, J. E., Wetterstrand, K. A., Deal, C., Baker, C. C., Di Francesco, V., Howcroft, T. K., Karp, R. W., Lunsford, R. D., Wellington, C. R., Belachew, T., Wright, M., Giblin, C., David, H., Mills, M., Salomon, R., Mullins, C., Akolkar, B., Begg, L., Davis, C., Grandison, L., Humble, M., Khalsa, J., Little, A. R., Peavy, H., Pontzer, C., Portnoy, M., Sayre, M. H., Starke-Reed, P., Zakhari, S., Read, J., Watson, B., & Guyer, M. (2009). The NIH Human Microbiome Project. *Genome Research*, *19*(12), 2317–2323.

Okada, M., Hayashi, F., Soda, Y., Zhong, X., Miura, K., & Kozai, K. (2004). Intra-familial distribution of nine putative periodontopathogens in dental plaque samples analyzed by PCR. *Journal of Oral Science*, *46*(3), 149–156.

Ostrom, P. H., Schall, M., Gandhi, H., Shen, T.-L., Hauschka, P. V., Strahler, J. R., & Gage, D. A. (2000). New strategies for characterizing ancient proteins using matrix-assisted laser desorption ionization mass spectrometry. *Geochimica et Cosmochimica Acta*, *64*(6), 1043–1050.

Ottoni, C., Borić, D., Cheronet, O., Sparacello, V., Dori, I., Coppa, A., Antonović, D., Vujević, D., Douglas Price, T., Pinhasi, R., & Cristiani, E. (2021). Tracking the transition to agriculture in Southern Europe through ancient DNA analysis of dental calculus. *Proceedings of the National Academy of Sciences of the United States of America*, *118*(32).

Oxford Radiocarbon Accelerator Unit. (2021). https://c14.arch.ox.ac.uk/

Ozga, A. T., Nieves-Colón, M. A., Honap, T. P., Sankaranarayanan, K., Hofman, C. A., Milner, G. R., Lewis, C. M., Jr, Stone, A. C., & Warinner, C. (2016). Successful enrichment and recovery of whole mitochondrial genomes from ancient human dental calculus. *American Journal of Physical Anthropology*, *160*(2), 220–228.

Palmer, R. J., Jr, Shah, N., Valm, A., Paster, B., Dewhirst, F., Inui, T., & Cisar, J. O. (2017). Interbacterial Adhesion Networks within Early Oral Biofilms of Single Human Hosts. *Applied and Environmental Microbiology*, *83*(11), e00407- 17.

Piperno, D. R. (2016). Phytolith radiocarbon dating in archaeological and paleoecological research: a case study of phytoliths from modern Neotropical plants and a review of the previous dating evidence. *Journal of Archaeological Science*, *68*, 54–61.

Proctor, D. M., Fukuyama, J. A., Loomer, P. M., Armitage, G. C., Lee, S. A., Davis, N. M., Ryder, M. I., Holmes, S. P., & Relman, D. A. (2018). A spatial gradient of bacterial diversity in the human oral cavity shaped by salivary flow. *Nature Communications*, *9*(1), 681.

Prüfer, K., de Filippo, C., Grote, S., Mafessoni, F., Korlević, P., Hajdinjak, M., Vernot, B., Skov, L., Hsieh, P., Peyrégne, S., Reher, D., Hopfe, C., Nagel, S., Maricic, T., Fu, Q., Theunert, C., Rogers, R., Skoglund, P., Chintalapati, M., Dannemann, M., Nelson, B. J., Key, F. M., Rudan, P., Kućan, Ž., Gušić, I., Golovanova, L. V., Doronichev, V. B., Patterson, N., Reich, D., Eichler, E. E., Slatkin, M., Schierup, M. H., Andrés, A. M., Kelso, J., Meyer, M., & Pääbo, S. (2017). A high-coverage Neandertal genome from Vindija Cave in Croatia. *Science*, *358*(6363), 655–658.

Radini, A., Tromp, M., Beach, A., Tong, E., Speller, C., McCormick, M., Dudgeon, J. V., Collins, M. J., Rühli, F., Kröger, R., & Warinner, C. (2019). Medieval women's early involvement in manuscript production suggested by lapis lazuli identification in dental calculus. *Science Advances*, *5*(1), eaau7126.

Ramsøe, A., Crispin, M., Mackie, M., McGrath, K., Fischer, R., Demarchi, B., Collins, M. J., Hendy, J., & Speller, C. (2021). Assessing the degradation of ancient milk proteins through site-specific deamidation patterns. *Scientific Reports*, *11*(1), 7795.

Ramsøe, A., van Heekeren, V., Ponce, P., Fischer, R., Barnes, I., Speller, C., & Collins, M. J. (2020). DeamiDATE 1.0: Site-specific deamidation as a tool to assess authenticity of members of ancient proteomes. *Journal of Archaeological Science*, *115*, 105080.

Roberts, A. P., & Kreth, J. (2014). The impact of horizontal gene transfer on the adaptive ability of the human oral microbiome. *Frontiers in Cellular and Infection Microbiology*, *4*, 124.

Salazar-García, D. C., Richards, M. P., Nehlich, O., & Henry, A. G. (2014). 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). *Journal of Archaeological Science*, *47*, 70–77.
Schroeter, E. R., & Cleland, T. P. (2016). Glutamine deamidation: an indicator of antiquity, or preservational quality? *Rapid Communications in Mass Spectrometry: RCM*, *30*(2), 251–255.
Schwarcz, H. P., & Schoeninger, M. J. (1991). Stable isotope analyses in human nutritional ecology. *American Journal of Physical Anthropology*, *34*(S13), 283–321.
Scott, A., Power, R. C., Altmann-Wendling, V., Artzy, M., Martin, M. A. S., Eisenmann, S., Hagan, R., Salazar-García, D. C., Salmon, Y., Yegorov, D., Milevski, I., Finkelstein, I., Stockhammer, P. W., & Warinner, C. (2021). Exotic foods reveal contact between South Asia and the Near East during the second millennium BCE. *Proceedings of the National Academy of Sciences of the United States of America*, *118*(2).
Scott, G. R., & Poulson, S. R. (2012). Stable carbon and nitrogen isotopes of human dental calculus: a potentially new non-destructive proxy for paleodietary analysis. *Journal of Archaeological Science*, *39*(5), 1388–1393.
Segata, N. (2015). Gut Microbiome: Westernization and the Disappearance of Intestinal Diversity. *Current Biology*, *25*(14), R611–R613.
Sender, R., Fuchs, S., & Milo, R. (2016). Revised Estimates for the Number of Human and Bacteria Cells in the Body. *PLoS Biology*, *14*(8), e1002533.
Shillito, L.-M., Blong, J. C., Green, E. J., & van Asperen, E. N. (2020). The what, how and why of archaeological coprolite analysis. *Earth-Science Reviews*, *207*, 103196.
Simon-Soro, A., & Tomás, I. (2013). Microbial geography of the oral cavity. *Journal of Dental Research, 92*(7):616-621.
Skoglund, P., Northoff, B. H., Shunkov, M. V., Derevianko, A. P., Pääbo, S., Krause, J., & Jakobsson, M. (2014). Separating endogenous ancient DNA from modern day contamination in a Siberian Neandertal. *Proceedings of the National Academy of Sciences of the United States of America*, *111*(6), 2229–2234.
Slon, V., Hopfe, C., Weiß, C. L., Mafessoni, F., de la Rasilla, M., Lalueza-Fox, C., Rosas, A., Soressi, M., Knul, M. V., Miller, R., Stewart, J. R., Derevianko, A. P., Jacobs, Z., Li, B., Roberts, R. G., Shunkov, M. V., de Lumley, H., Perrenoud, C., Gušić, I., Kućan, Ž., Rudan, P., Aximu-Petri, A., Essel, E., Nagel, S., Nickel, B., Schmidt, A., Prüfer, K., Kelso, J., Burbano, H. A., Pääbo, S., & Meyer, M. (2017). Neandertal and Denisovan DNA from Pleistocene sediments. *Science*, *356*(6338), 605–608.
Slon, V., Mafessoni, F., Vernot, B., de Filippo, C., Grote, S., Viola, B., Hajdinjak, M., Peyrégne, S., Nagel, S., Brown, S., Douka, K., Higham, T., Kozlikin, M. B., Shunkov, M. V., Derevianko, A. P., Kelso, J., Meyer, M., Prüfer, K., & Pääbo, S. (2018). The genome of the offspring of a Neanderthal mother and a Denisovan father. *Nature*, *561*(7721), 113–116.
Spyrou, M. A., Bos, K. I., Herbig, A., & Krause, J. (2019). Ancient pathogen genomics as an emerging tool for infectious disease research. *Nature Reviews. Genetics*, *20*(6), 323–340.
Steinegger, M., & Salzberg, S. L. (2020). Terminating contamination: large-scale search identifies more than 2,000,000 contaminated entries in GenBank. *Genome Biology*, *21*(1), 115.

Tamura, K., Nakano, K., Hayashibara, T., Nomura, R., Fujita, K., Shintani, S., & Ooshima, T. (2006). Distribution of 10 periodontal bacteria in saliva samples from Japanese children and their mothers. *Archives of Oral Biology*, *51*(5), 371–377.

Tett, A., Huang, K. D., Asnicar, F., Fehlner-Peach, H., Pasolli, E., Karcher, N., Armanini, F., Manghi, P., Bonham, K., Zolfo, M., De Filippis, F., Magnabosco, C., Bonneau, R., Lusingu, J., Amuasi, J., Reinhard, K., Rattei, T., Boulund, F., Engstrand, L., Zink, A., Collado, M. C., Littman, D. R., Eibach, D., Ercolini, D., Rota-Stabelli, O., Huttenhower, C., Maixner, F., & Segata, N. (2019). The *Prevotella copri* Complex Comprises Four Distinct Clades Underrepresented in Westernized Populations. *Cell Host & Microbe*, *26*(5), 666–679.e7.

Thompson, H., Rybalka, A., Moazzez, R., Dewhirst, F. E., & Wade, W. G. (2015). In vitro culture of previously uncultured oral bacterial phylotypes. *Applied and Environmental Microbiology*, *81*(24), 8307–8314.

Tito, R. Y., Knights, D., Metcalf, J., Obregon-Tito, A. J., Cleeland, L., Najar, F., Roe, B., Reinhard, K., Sobolik, K., Belknap, S., Foster, M., Spicer, P., Knight, R., & Lewis, C. M., Jr. (2012). Insights from characterizing extinct human gut microbiomes. *PloS One*, *7*(12), e51146.

Tito, R. Y., Macmil, S., Wiley, G., Najar, F., Cleeland, L., Qu, C., Wang, P., Romagne, F., Leonard, S., Ruiz, A. J., Reinhard, K., Roe, B. A., & Lewis, C. M., Jr. (2008). Phylotyping and functional analysis of two ancient human microbiomes. *PloS One*, *3*(11), e3703.

van Doorn, N. L., Wilson, J., Hollund, H., Soressi, M., & Collins, M. J. (2012). Site-specific deamidation of glutamine: a new marker of bone collagen deterioration. *Rapid Communications in Mass Spectrometry: RCM*, *26*(19), 2319–2327.

Velsko, I. M., Fellows Yates, J. A., Aron, F., Hagan, R. W., Frantz, L. A. F., Loe, L., Martinez, J. B. R., Chaves, E., Gosden, C., Larson, G., & Warinner, C. (2019). Microbial differences between dental plaque and historic dental calculus are related to oral biofilm maturation stage. *Microbiome*, *7*(1), 102.

Velsko, I. M., Frantz, L. A. F., Herbig, A., Larson, G., & Warinner, C. (2018). Selection of Appropriate Metagenome Taxonomic Classifiers for Ancient Microbiome Research. *mSystems*, *3*(4), e00080-18.

Velsko, I. M., Overmyer, K. A., Speller, C., Klaus, L., Collins, M. J., Loe, L., Frantz, L. A. F., Sankaranarayanan, K., Lewis, C. M., Jr, Martinez, J. B. R., Chaves, E., Coon, J. J., Larson, G., & Warinner, C. (2017). The dental calculus metabolome in modern and historic samples. *Metabolomics*, *13*(11), 134.

Velsko, I. M., & Warinner, C. (2017). Bioarchaeology of the Human Microbiome. *Bioarchaeology International*, *1*(1–2), 86–99.

Velsko, I., Skourtanioti, E., & Brandt, G. (2019). Ancient DNA extraction from skeletal material v1. *protocols.io*. ZappyLab, Inc.

Warinner, C., Hendy, J., Speller, C., Cappellini, E., Fischer, R., Trachsel, C., Arneborg, J., Lynnerup, N., Craig, O. E., Swallow, D. M., Fotakis, A., Christensen, R. J., Olsen, J. V., Liebert, A., Montalva, N., Fiddyment, S., Charlton, S., Mackie, M., Canci, A., Bouwman, A., Rühli, F., Gilbert, M. T. P., & Collins, M. J. (2014). Direct evidence of milk consumption from ancient human dental calculus. *Scientific Reports*, *4*, 7104.

Warinner, C., Herbig, A., Mann, A., Fellows Yates, J. A., Weiß, C. L., Burbano, H. A., Orlando, L., & Krause, J. (2017). A Robust Framework for Microbial Archaeology. *Annual Review of Genomics and Human Genetics*, *18*, 321–356.

Warinner, C., Rodrigues, J. F. M., Vyas, R., Trachsel, C., Shved, N., Grossmann, J., Radini, A., Hancock, Y., Tito, R. Y., Fiddyment, S., Speller, C., Hendy, J., Charlton, S., Luder, H. U., Salazar-García, D. C., Eppler, E., Seiler, R., Hansen, L. H., Castruita, J. A. S., Barkow-Oesterreicher, S., Teoh, K. Y., Kelstrup, C. D., Olsen, J. V., Nanni, P., Kawai, T., Willerslev, E., von Mering, C., Lewis Jr, C. M., Collins, M. J., Gilbert, M. T. P., Rühli, F., & Cappellini, E. (2014). Pathogens and host immunity in the ancient human oral cavity. *Nature Genetics*, *46*(4), 336–344.

Warinner, C., Speller, C., Collins, M. J., & Lewis, C. M., Jr. (2015). Ancient human microbiomes. *Journal of Human Evolution*, *79*, 125–136.

Welker, F., Ramos-Madrigal, J., Gutenbrunner, P., Mackie, M., Tiwary, S., Rakownikow Jersie-Christensen, R., Chiva, C., Dickinson, M. R., Kuhlwilm, M., de Manuel, M., Gelabert, P., Martinón-Torres, M., Margvelashvili, A., Arsuaga, J. L., Carbonell, E., Marques-Bonet, T., Penkman, K., Sabidó, E., Cox, J., Olsen, J. V., Lordkipanidze, D., Racimo, F., Lalueza-Fox, C., Bermúdez de Castro, J. M., Willerslev, E., & Cappellini, E. (2020). The dental proteome of *Homo antecessor*. *Nature*, *580*(7802), 235–238.

Weyrich, L. S., Duchene, S., Soubrier, J., Arriola, L., Llamas, B., Breen, J., Morris, A. G., Alt, K. W., Caramelli, D., Dresely, V., Farrell, M., Farrer, A. G., Francken, M., Gully, N., Haak, W., Hardy, K., Harvati, K., Held, P., Holmes, E. C., Kaidonis, J., Lalueza-Fox, C., de la Rasilla, M., Rosas, A., Semal, P., Soltysiak, A., Townsend, G., Usai, D., Wahl, J., Huson, D. H., Dobney, K., & Cooper, A. (2017). Neanderthal behaviour, diet, and disease inferred from ancient DNA in dental calculus. *Nature*, *544*(7650), 357–361.

Weyrich, L. S., Farrer, A. G., Eisenhofer, R., Arriola, L. A., Young, J., Selway, C. A., Handsley-Davis, M., Adler, C. J., Breen, J., & Cooper, A. (2019). Laboratory contamination over time during low-biomass sample analysis. *Molecular Ecology Resources*, *19*(4), 982–996.

Wibowo, M. C., Yang, Z., Borry, M., Hübner, A., Huang, K. D., Tierney, B. T., Zimmerman, S., Barajas-Olmos, F., Contreras-Cubas, C., García-Ortiz, H., Martínez-Hernández, A., Luber, J. M., Kirstahler, P., Blohm, T., Smiley, F. E., Arnold, R., Ballal, S. A., Pamp, S. J., Russ, J., Maixner, F., Rota-Stabelli, O., Segata, N., Reinhard, K., Orozco, L., Warinner, C., Snow, M., LeBlanc, S., & Kostic, A. D. (2021). Reconstruction of ancient microbial genomes from the human gut. *Nature*, *594*(7862), 234–239.

Wilkins, L. J., Monga, M., & Miller, A. W. (2019). Defining Dysbiosis for a Cluster of Chronic Diseases. *Scientific Reports*, *9*(1), 12918.

Yamasaki, Y., Nomura, R., Nakano, K., Naka, S., Matsumoto-Nakano, M., Asai, F., & Ooshima, T. (2012). Distribution of periodontopathic bacterial species in dogs and their owners. *Archives of Oral Biology*, *57*(9), 1183–1188.

Youle, M., Knowlton, N., Rohwer, F., Gordon, J., & Relman, D. A. (2013). Superorganisms and Holobionts. *Microbe* , *8*(4), 152–153.

Ziesemer, K. A., Mann, A. E., Sankaranarayanan, K., Schroeder, H., Ozga, A. T., Brandt, B. W., Zaura, E., Waters-Rist, A., Hoogland, M., Salazar-García, D. C., Aldenderfer, M., Speller, C., Hendy, J., Weston, D. A., MacDonald, S. J., Thomas, G. H., Collins, M. J., Lewis, C. M., Hofman, C., & Warinner, C. (2015). Intrinsic challenges in ancient microbiome reconstruction using 16S rRNA gene amplification. *Scientific Reports*, *5*, 16498.

Ziesemer, K. A., Ramos-Madrigal, J., Mann, A. E., Brandt, B. W., Sankaranarayanan, K., Ozga, A. T., Hoogland, M., Hofman, C. A., Salazar-García, D. C., Frohlich, B., Milner, G. R., Stone,

A. C., Aldenderfer, M., Lewis, C. M., Jr, Hofman, C. L., Warinner, C., & Schroeder, H. (2019). The efficacy of whole human genome capture on ancient dental calculus and dentin. *American Journal of Physical Anthropology*, *168*(3), 496–509.

EIGENSTÄNDIGKEITSERKLÄRUNG

Entsprechend §5 Abs. 4 der Promotionsordnung der Biologisch-Pharmazeutischen Fakultät der Friedrich-Schiller-Universität Jena, hiermit erkläre ich,

(a) dass mir die geltende Promotionsordnung bekannt ist,

(b) dass ich die Dissertation selbst angefertigt habe, keine Textabschnitte eines Dritten oder eigener Prüfungsarbeiten ohne Kennzeichnung übernommen und alle von mir benutzten Hilfsmittel, persönlichen Mitteilungen und Quellen in der Arbeit angegeben habe,

(c) dass ich alle Personen habe, die mir bei der Auswahl und Auswertung sowie bei der Herstellung des Manuskriptes unterstützt haben, in der Autorenliste der Manuskripte und den entsprechenden Danksagungen namentlich erwähnt

(d) dass ich die Hilfe einer kommerziellen Promotionsvermittlung nicht in Anspruch genommen habe und dass Dritte weder unmittelbar noch mittelbar geldwerte Leistungen von mir für Arbeiten erhalten haben, die im Zusammenhang mit dem Inhalt der vorgelegten Dissertation stehen,

(e) dass ich die Dissertation noch nicht als Prüfungsarbeit für eine staatliche oder andere wissenschaftliche Prüfung eingereicht habe,

Kopenhagen, den

_____

Zandra Fagernäs

ACKNOWLEDGEMENTS

This dissertation is brought to you by caffeine and pure stubbornness. And of course, I could never have gotten this far without the support of all the amazing people around me. My biggest thanks go to...

...my supervisor Tina, for guiding me through the maze of ancient microbiome research and the academic world. Thank you for always being encouraging and believing in me, even when I don't believe in myself.

...my mentor Irina, for always being there to answer all my questions and give advice. And for being a fellow ProcrastiKnitter nerd.

...my colleagues in the Microbiome Group and Department of Archaeogenetics, for always being there for both scientific discussions and laughs.

...my friends, who have made my time in Germany amazing. Especially Aida, James and Maia, for always reminding me that there is so much more to life than just work.

...to my family - my mum Carola, dad John, and little brother Martin - for always being there for me and supporting me, no matter what crazy direction I decide to take in my life.

…mummulle, sinun intohimo tieteeseen, luontoon ja historiaan johti minut tälle polulle jo varhaisessa iässä. *(To grandma, your passion for science, nature and history led me onto this path already at an early age).*

...till faffa, utan ditt stöd hade jag högst troligen aldrig ens tagit mig till Tyskland. *(To grandpa, without your support I would most likely never even have made it to Germany).*

...to Tjorven, for being my mental health support and making me happy even when things seem the most tough.

...and last, but definitely not least, to Thiseas, for being my rock through all of this. Without you, I would have descended into complete madness these last years.

APPENDIX

Appendix A: Summary of author contributions

Manuscript Nr. A

Short reference: Fagernäs and Warinner (in press). In Mark Pollard, Ruth-Ann Armitage and Cheryl Makarewicz (eds.), *Handbook of Archaeological Science*, 2nd edition.

Author contributions: ZF and CW designed the structure and topics of the book chapter. ZF wrote a first draft with contributions from CW. ZF and CW refined the manuscript. ZF created illustrations for figures, as well as downloaded data from AncientMetagenomeDir and created the map for Figure 2 in R, including writing the code. CW provided the photo in Figure 1.

Manuscript Nr. B

Short reference: Fagernäs et al. (2020). *J. Archaeol. Sci.*

Author contributions: CS and JH conceived of the unified protocol. ZF, CS, JH, and CW designed the study. MIGC, CS, JH, and CAH provided materials and resources. ZF performed the all the experiments in the laboratory and generated the data for the study. ZF analyzed all the data, including writing scripts for analyses. ZF created all figures. ZF wrote the manuscript with input from CW, IV and the other authors. In total, ZF contributed 100% of all the experimental figures presented in the manuscript (Figure 2-7 and supplementary figures).

Manuscript Nr. C

Short reference: Fagernäs et al. (submitted).

Author contributions: ZF, IMV and CW designed the study. DCSG, AAF, MH, AH, JLM and AO provided materials and resources. ZF, AO and CW performed sampling. ZF performed all the laboratory work, producing the data for the study. ZF performed the analyses, including writing all the scripts required for the analsyes. ZF created all the figures in the manuscript. ZF wrote the manuscript with input from IMV, CW, and the other coauthors. In total, ZF contributed 100% of all the experimental figures presented in the manuscript (Figure 2-7 and supplementary figures).

Manuscript Nr. D

Short reference: Fagernäs et al. (in prep).

Author contributions: ZF, MT, KN, IMV and CW designed the study. MT, SB, HB, GC, HJ, JF, ALT, CMLJ, KN, AO, CP, RS, MS, EW and FV provided materials and resources. ZF, MT and AO performed laboratory work. ZF analyzed the data, including writing all the scripts for the analyses, except for the statistical analysis of the effects on environmental variables on preservation, which was performed by ABR. ZF created all the figures for the manuscript. ZF wrote the manuscript with input from IMV, CW, and the other coauthors. In total, ZF contributed 90% to all experimental figures presented in the manuscript (Figure 1-5 and supplementary figures), by performing all the analyses, based on data partly produced by MT and AO.

Appendix B. Supplementary information for Manuscript B

A unified protocol for simultaneous extraction of DNA and proteins from archaeological dental calculus

Zandra Fagernäs, Maite I. García-Collado, Jessica Hendy, Courtney A. Hofman, Camilla Speller, Irina Velsko, Christina Warinner

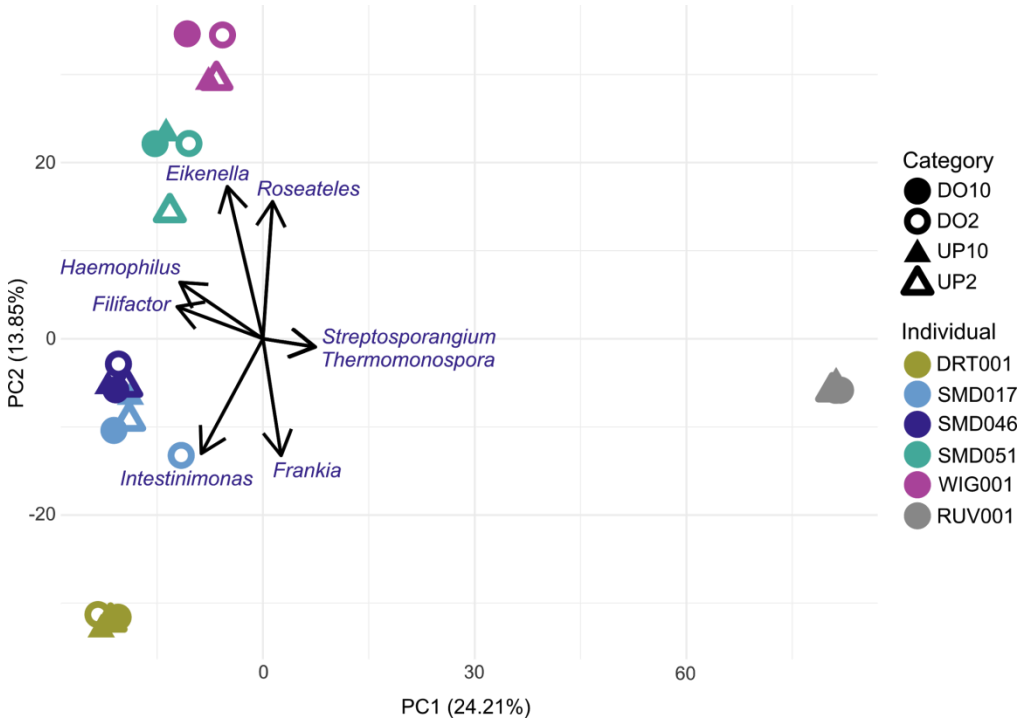


Figure S1. PCA on genus level taxa from metagenomic sequencing data, including individual RUV001. Top genera contributing to separation for each principal component are indicated by arrows and labelled in blue.

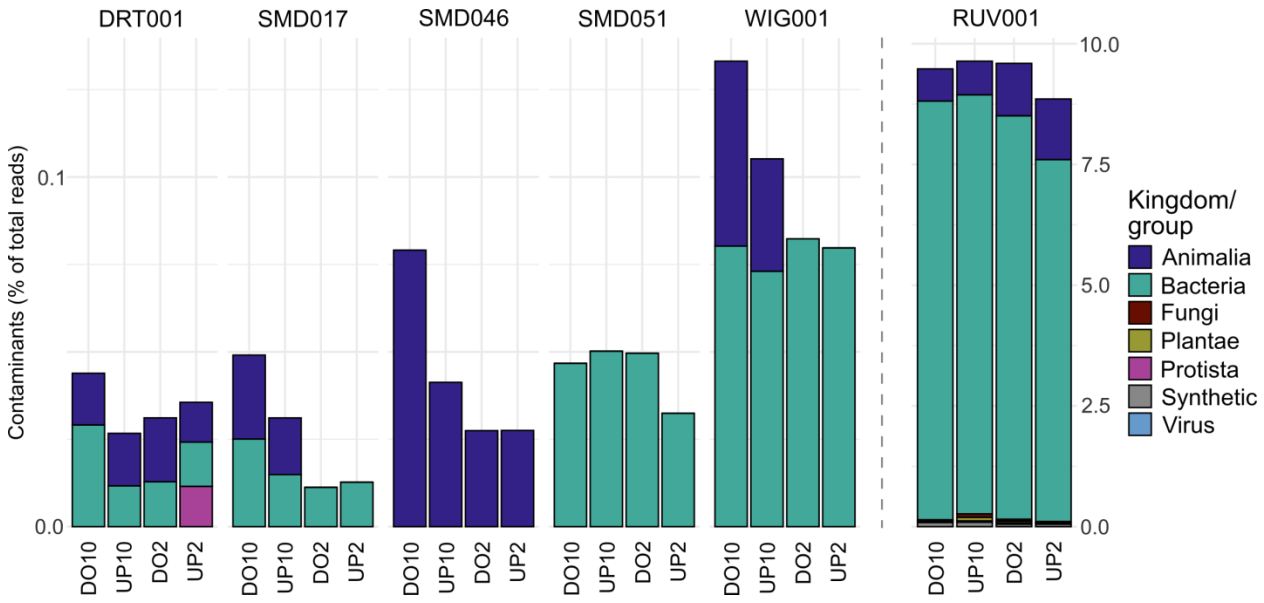


Figure S2. Percentage contaminant reads per protocol, as classified by major taxonomic groups. Note that contamination for individual RUV001 is plotted to a different scale on the right-hand y-axis.

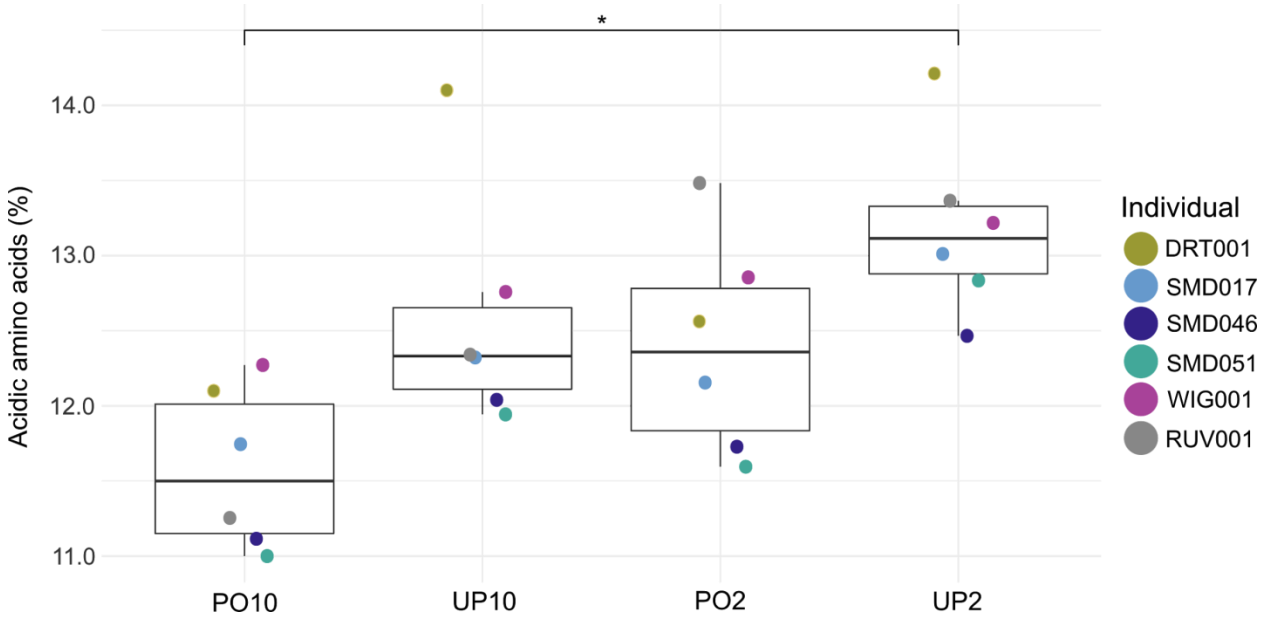


Figure S3. Proportion of acidic amino acids among recovered peptides (* p<0.05; ** p<0.01; *** p<0.001).

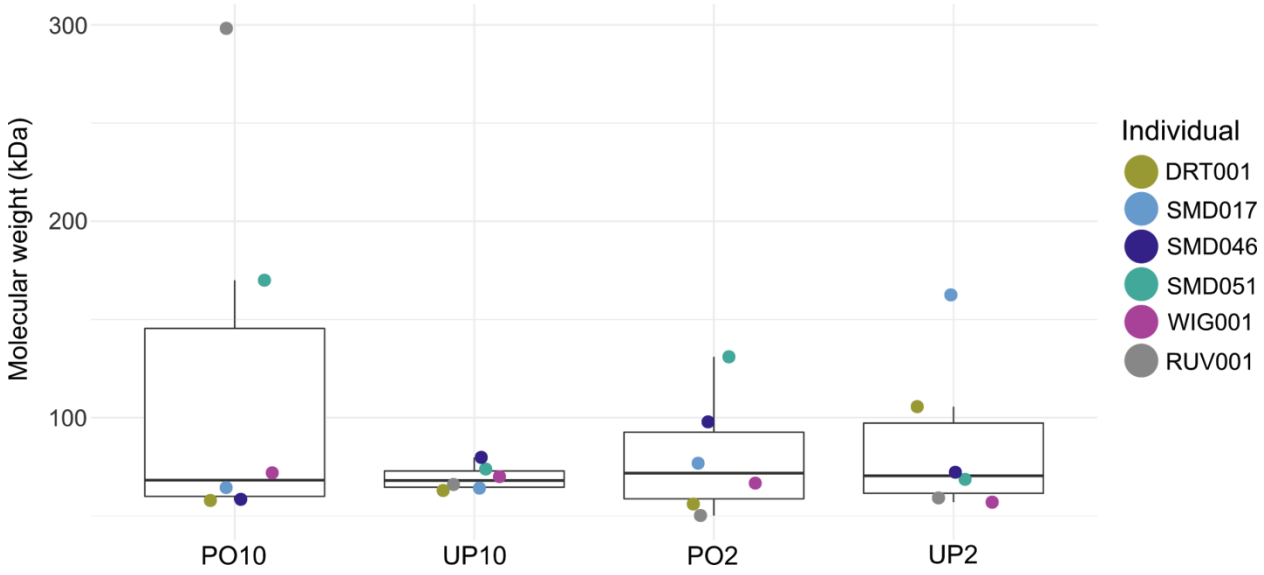


Figure S4. Molecular weight of proteins identified through the different extraction protocols.

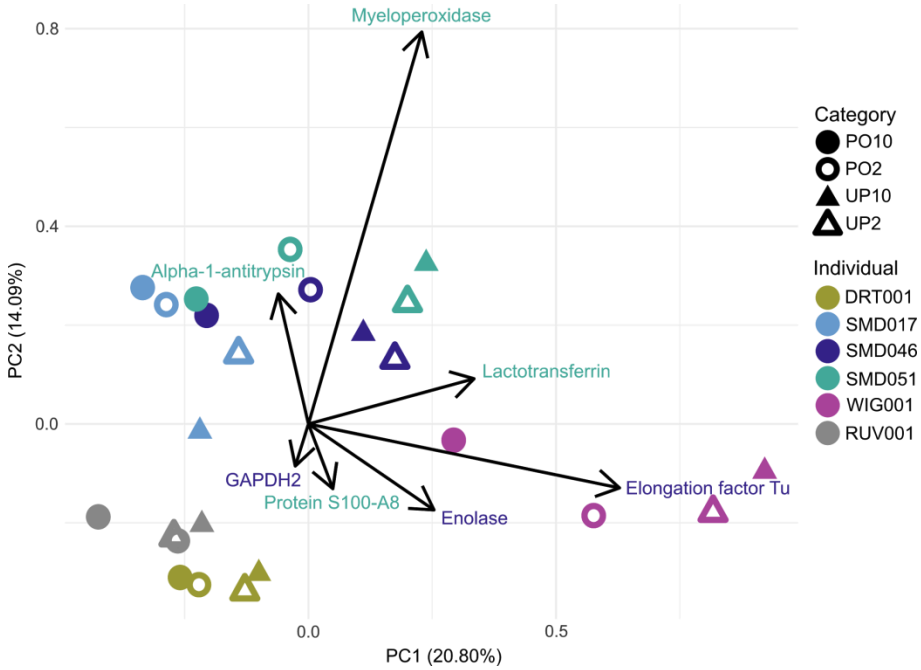


Figure S5. PCA of identified protein clusters, including individual RUV001. The top protein clusters contributing to separation for each principal component are indicated by arrows and labels are colored by organism (blue = prokaryotes, green = eukaryotes).

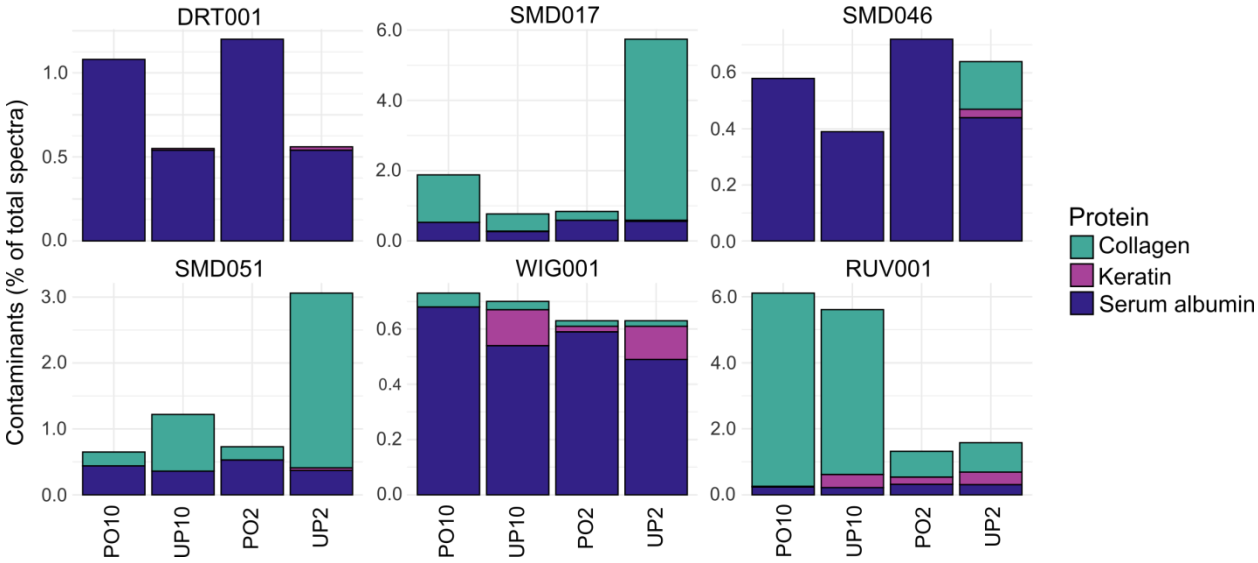
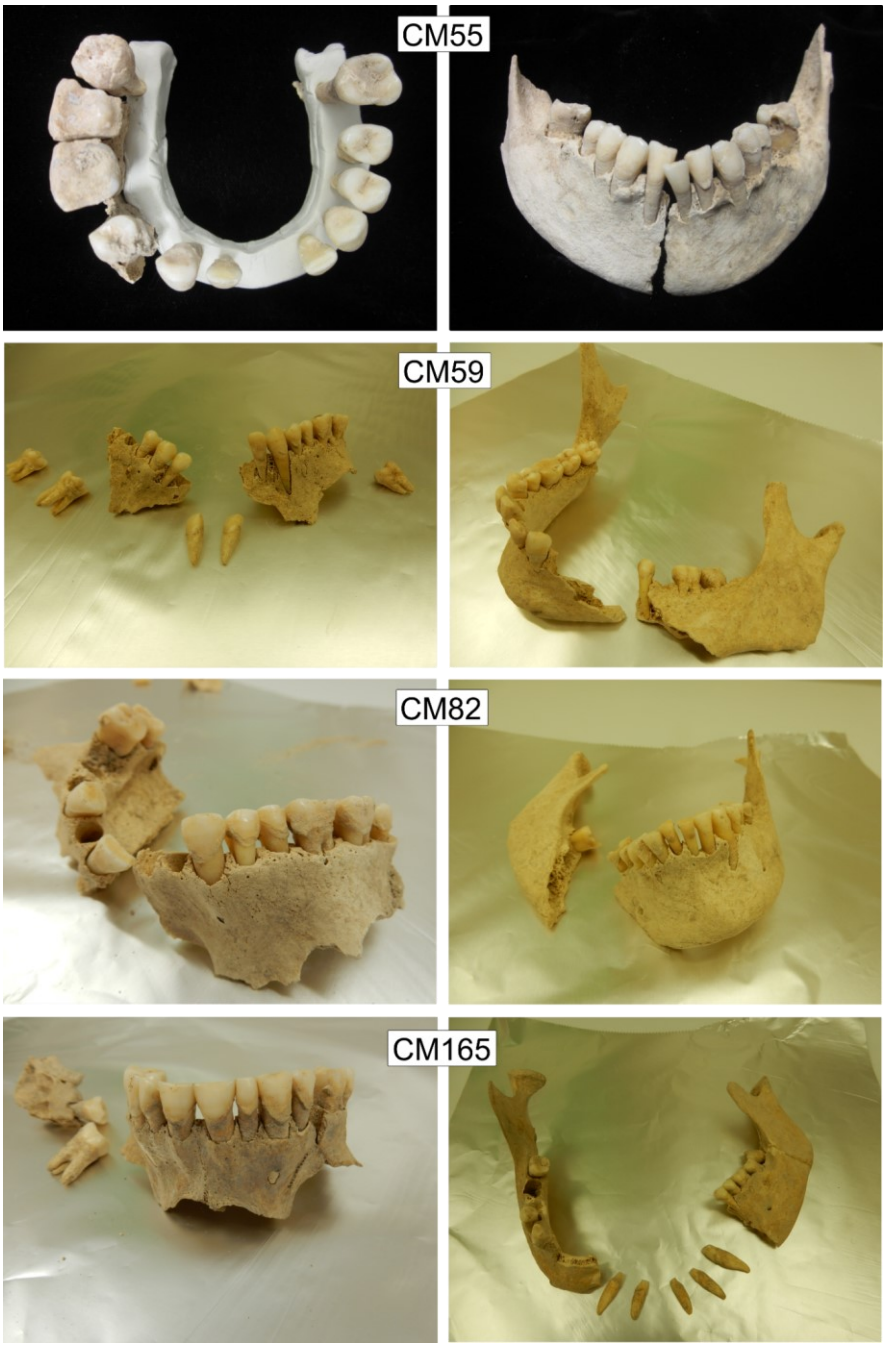


Figure S6. Percentage of contaminant spectra in each sample, colored by protein class.

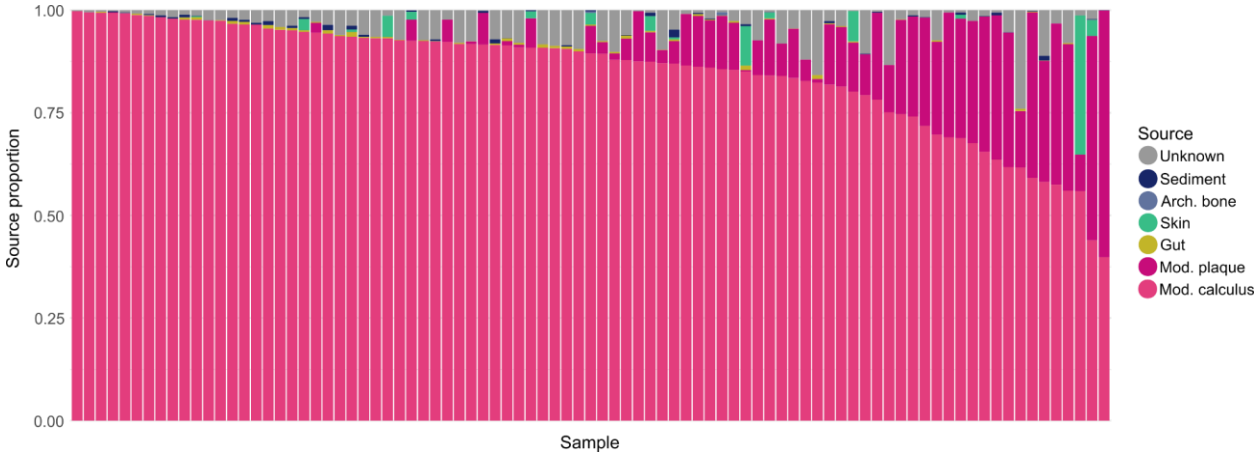
Appendix C: Supplementary information for Manuscript C

“Understanding the microbial biogeography of ancient human dentitions to guide study design and interpretation”

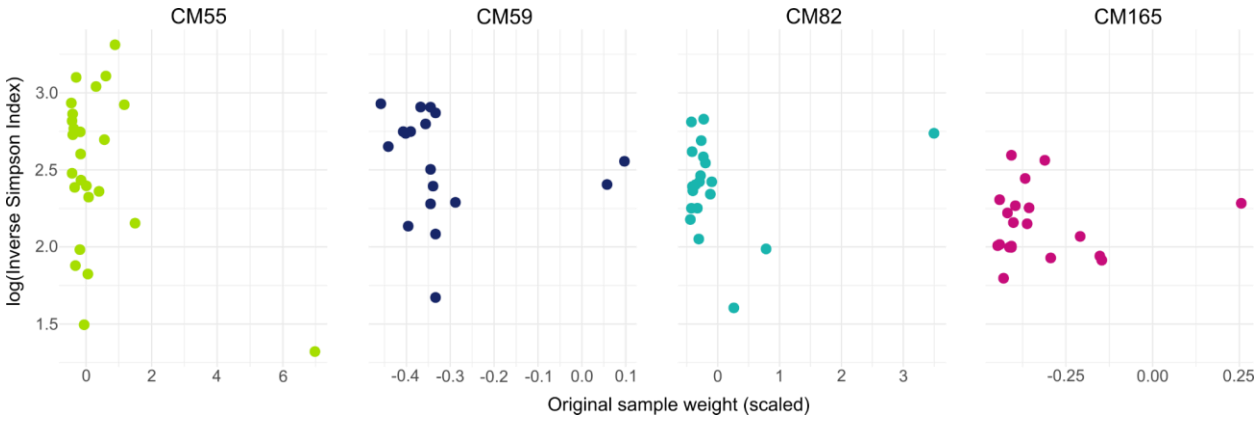
Zandra Fagernäs, Domingo C. Salazar-García, Azucena Avilés, María Haber, Amanda Henry, Joaquín Lomba Maurandi, Andrew Ozga, Irina M Velsko, Christina Warinner



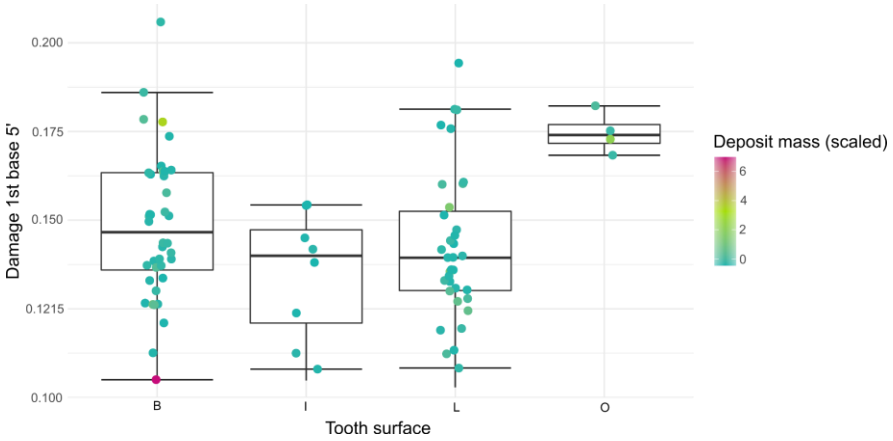
S1. Photos of the entire available dentitions of the four individuals sampled in this study.



S2. SourceTracker results, generated from a genus-level OTU-table.



S3. Inverse Simpson Index by mass of the original calculus deposit.



S4. Damage of first base at the 5' end of fragments mapping to *Tannerella forsythia*.

Appendix D: Supplementary information for Manuscript D

"Exploring the potential of archaeogenetic studies of dental calculus to shed light on past human migrations in the Pacific"

Zandra Fagernäs, Monica Tromp, Stuart Bedford , Hallie Buckley, Geoffrey Clark, John Dudgeon, James Flexner, Anatauarii Leal-Tamarii, Cecil M. Lewis, Jr, Kathrin Nägele, Andrew Ozga, Adam B. Rohrlach, Cosimo Posth, Richard Shing, Matthew Spriggs, Edson Willie, Frédérique Valentin, Irina M. Velsko, Christina Warinner

Table S1. Potential dietary items detected through MALT. 'Individuals' refers to the number of well-preserved samples where reads from the species were detected, and 'Reads' to the mean (± standard deviation) number of reads in the individuals where the species was detected.

Species	Common name	Individuals	Reads
Bos taurus	Cattle	1	442±0
Canis lupus familiaris	Dog	2	550.5±85.5
Dibothriocephalus latus	Broad fish tapeworm	3	408±218.1
Fargesia denudata	Bamboo	3	765.7±406.4
Triticum aestivum	Wheat	4	890.8±423.7

Table S2. Reference genomes used for investigating authenticity of potential dietary finds.

Species	Common name	Assembly	Assembly level
Bos taurus	Cattle	ARS-UCD1.2	Chromosome
Canis lupus familiaris	Dog	ROS_Cfam_1.0	Chromosome
Dibothriocephalus latus	Broad fish tapeworm	D_latum_geneva_0011_upd	Scaffold
Fargesia denudata, chloroplast	Bamboo	NC_034747.1	Complete
Triticum aestivum, mitochondrion	Wheat	NC_036024.1	Complete

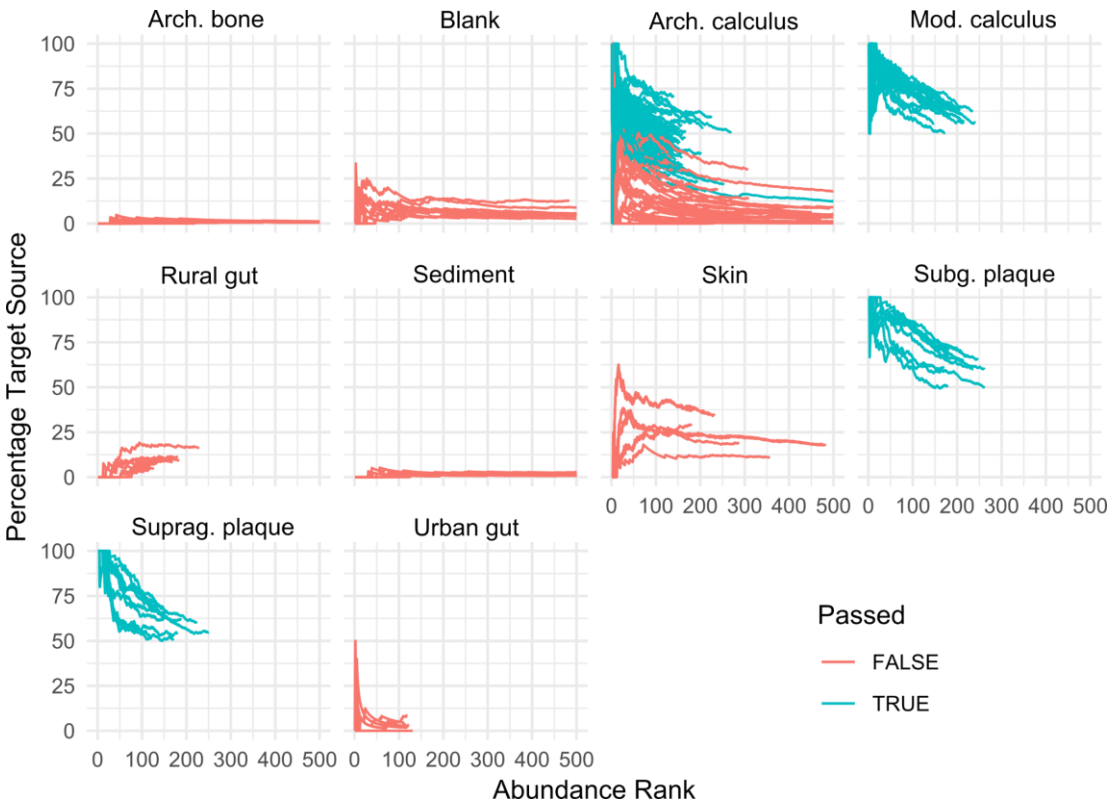


Figure S1. Results of cuperdec analysis, used to assess which samples are well-preserved enough carry on to downstream analyses ('Passed').

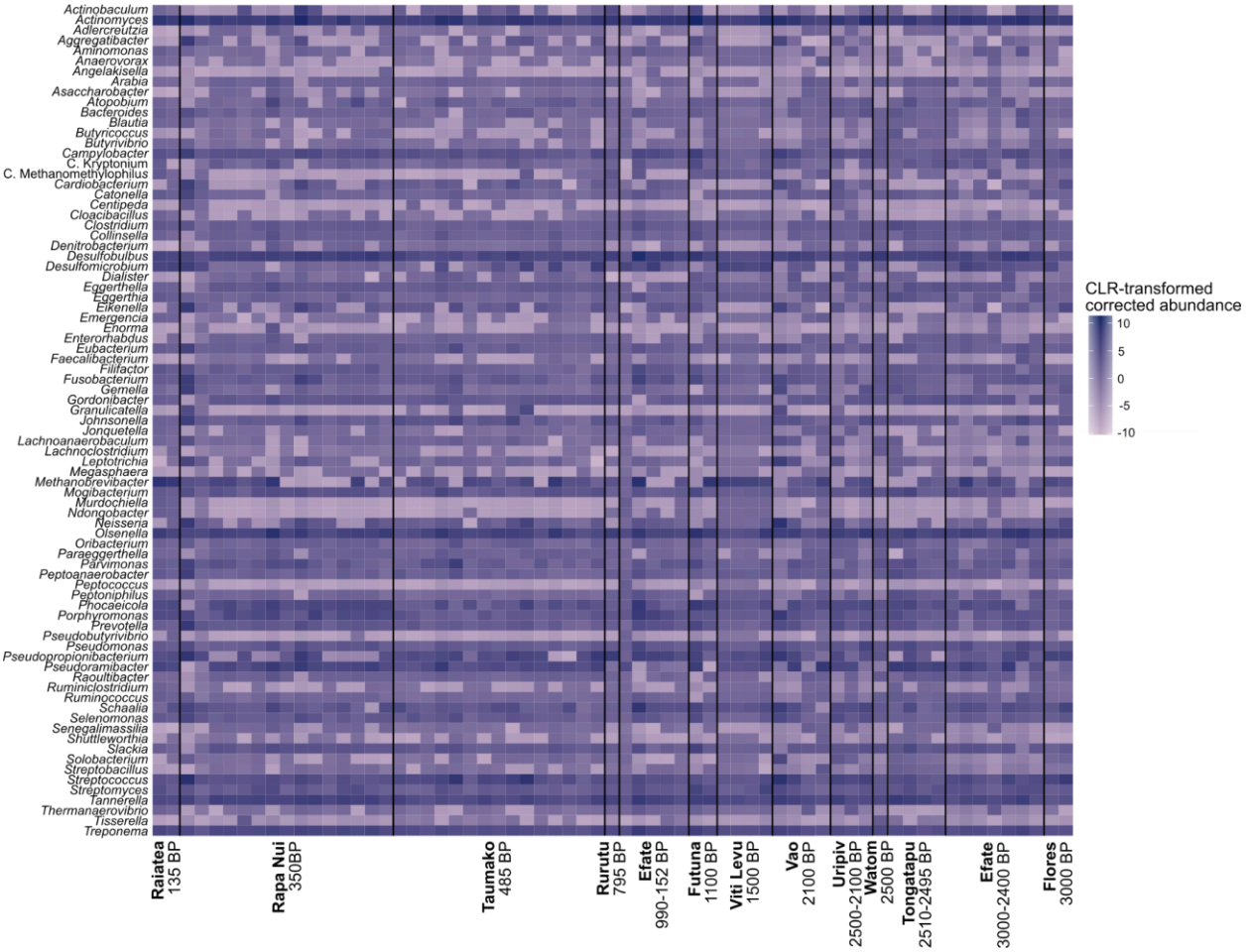


Figure S2. Heatmap of genus-level read counts per genus, after CLR-transformation and removing lab-based bias. Time periods are given as the range of average estimated ages BP of all samples from an island (in cases where a range is not given, all individuals were estimated to the same time period).