



Metagenomic analysis of ancient human and microbial DNA preserved under unusual taphonomic conditions

Kumulative Dissertation

zur Erlangung des akademischen Grades

Doctor rerum naturalium (Dr. rer. nat.)

an der Fakultät für Biologie

vorgelegt von

Dr. agr. Mohamed Sabry Mohamed Sarhan

geb. am 01.03.1993 in Giza, Egypt

Munich 2023

(This page intentionally left blank)

Declaration

Erstgutachter:	PD Dr. Albert R. Zink
Zweitgutachter:	Prof. Dr. Wolfgang Enard
Tag der Abgabe:	26.04.2023
Tag der mündlichen Prüfung:	25.10.2023

Eidesstattliche Erklärung

Ich versichere hiermit an Eid statt, dass die vorgelegte Dissertation von mir selbstständig und ohne unerlaubte Hilfe angefertigt wurde.

Hiermit erkläre ich, dass die Dissertation nicht ganz oder in wesentlichen Teilen einer anderen Prüfungskommission vorgelegt worden ist.

Ich habe mich anderweitig einer Doktorprüfung ohne Erfolg nicht unterzogen.

München, den 26.04.2023

Mohamed Sabry Mohamed Sarhan

The entire work of the presented thesis has been conducted under the supervision of

PD Dr. Albert R. Zink

and Dr. Frank Maixner

at the Eurac Research Institute for Mummy Studies in Bolzano, Italy, and the Faculty of Biology of

Ludwig Maximilian University of Munich.

The research design of this work implemented ethical considerations to the best of the authors'

knowledge.

Table of Contents

Declaration.....	I
List of Publications	V
Summary	VII
Introduction.....	1
Aims of the study	11
Results.....	13
Paper I: Ancient DNA diffuses from human bones to cave stones.....	15
Paper I: Supplementary Information	27
Paper II: Hallstatt miners consumed blue cheese and beer during the Iron Age and retained a non-Westernized gut microbiome until the Baroque period.....	39
Paper II: Supplementary Information.....	63
Paper III: A nontuberculous mycobacterium could solve the mystery of the lady from the Franciscan church in Basel, Switzerland.....	81
Paper III: Supplementary Information	99
Discussion	111
References.....	124
Acknowledgements.....	130
Curriculum vitae	131

(This page intentionally left blank)

List of Publications

Paper I Sarhan MS, Lehmkuhl A, Straub R, Tett A, Wieland G, Francken M, Zink A, Maixner F. Ancient DNA diffuses from human bones to cave stones. *Iscience*. 2021 Dec 17;24(12):103397. <https://doi.org/10.1016/j.isci.2021.103397>

Paper II Maixner F, Sarhan MS, Huang KD, Tett A, Schoenafinger A, Zingale S, Blanco-Míguez A, Manghi P, Cemper-Kiesslich J, Rosendahl W, Kusebauch U, Morrone SR, Hoopmann MR, Rota-Stabelli O, Rattei T, Moritz RL, Oeggl K, Segata N, Zink A, Reschreiter H, Kowarik K. Hallstatt miners consumed blue cheese and beer during the Iron Age and retained a non-Westernized gut microbiome until the Baroque period. *Current Biology*. 2021 Dec 6;31(23):5149-62. <https://doi.org/10.1016/j.cub.2021.09.031>

Paper III Sarhan MS, Wurst C, Tzankov A, Bircher AJ, Wittig H, Briellmann T, Augsburg M, Hotz G, Zink A, Maixner F. A nontuberculous mycobacterium could solve the mystery of the lady from the Franciscan church in Basel, Switzerland. *BMC biology*. 2023 Dec;21(1):1-6. <https://doi.org/10.1186/s12915-022-01509-7>

(This page intentionally left blank)

Summary

In this dissertation, I analyzed different types of ancient organic and inorganic materials preserved under different taphonomic conditions from different time points (~2900-235 BP). During the analysis, I tried to improve the currently used *in situ*, *in vitro*, and *in silico* methods of ancient DNA analysis. The results of this dissertation are summarized and published in three research articles.

In the first article, we analyzed skeletal remains of a late Bronze Age individual found in the Wimsener water caves in Germany. In addition to the skeletal remains, we analyzed the calcite deposits, found surrounding the bones of the individuals, from which we were able to retrieve ancient human DNA fragments, enough to reconstruct the full mitochondrial genome and to assign molecular sex to the individual. We demonstrated the ancient human and microbial DNA fragments diffused from the bones to the calcite stone deposits in the same direction of gravity. This study exemplifies using alternative source for obtaining ancient human and/or microbial DNA without causing destruction to the valuable archeological finding.

In the second article, we analyzed different paleofeces specimens from Hallstatt mines, Austria (dated to the Bronze Age – the Baroque times). We subjected them to microscopic, proteomic, and metagenomic analyses. The collective analysis allowed unveiling the following: i) the molecular sex and the mitochondrial haplogroups of the individuals; ii) consumption of fibrous plant-based diet as well as animal components; iii) non-Westernized gut microbiome composition until the Baroque times; iv) presence of gut parasites; and finally, v) consumption of fermented food (cheese-like) and beverages (beer). During this study, we reconstructed, for the first time, complete ancient fungal genomes of *Saccharomyces cerevisiae* and *Penicillium roqueforti*, and by comparative genomic analyses, we presented different lines of evidence on their being used in beer and cheese fermentation, respectively. The study presented a comprehensive interdisciplinary workflow for the analysis of such precious archeological materials.

In the third article, we analyzed different tissue specimens from the mummy of Anna Catharina Bischoff (ACB), from Basel, Switzerland. Initially, we aimed to find any molecular proof of presence of the syphilis-causing bacterium *Treponema pallidum*, which was not successful. However, by employing *de novo* metagenomic assembly, we were able to reconstruct a complete genome of a pathogen from the brain sample, belonging to the

Mycobacteriaceae family. The genome analysis of the *de novo* detected pathogen supports the assumption of its pathogenicity and was very congruent with the radiological symptoms and its survival in the brain under high concentrations of mercury. The study presented a proof-of-concept on using metagenomic assembly to detect extinct or previously undescribed pathogens.

Overall, throughout this dissertation, I tried to use different analytical methods beyond what is already known and commonly used in the field of ancient DNA.

Introduction

(This page intentionally left blank)

Introduction

In the year 1984, the very first attempt to clone ancient DNA fragments was published, claiming cloning and sequencing small fragments of DNA extracted from a dry tissue the extinct horse species *Equus quagga quagga* (Higuchi et al., 1984). Soon after, Svante Pääbo cloned ancient human DNA from a predynastic Egyptian mummy and used hybridization to prove that. Interestingly, he also claimed that the DNA fragments he cloned contained little or no postmortem modifications (Pääbo, 1985), which he later doubted due to lack of necessary control samples (Gitschier, 2008). Later, polymerase chain reaction (PCR) became widely used to analyze ancient DNA from different archeological sources (Hagelberg et al., 2015), and targeting genetic markers like mitochondrial- and nuclear DNA and to understand human evolution and population history (Cann et al., 1987; Pääbo et al., 1989; van der Kuyl et al., 1995). PCR was even exploited to detect ancient pathogens, like *Plasmodium falciparum* and *Mycobacterium tuberculosis* in human mummies and remains (Hawass et al., 2010; Nerlich et al., 1997; Zink et al., 2003). It was additionally used to analyze ancient DNA of extinct animals, like mammoth (Poulakakis et al., 2006), and archaic humans (Noonan et al., 2006).

Next generation sequencing and metagenomics in ancient DNA

With the huge development in DNA sequencing technologies, particularly the next generation sequencing (NGS), and the completion of the first draft of the human genome, more interest was brought to sequence the ancient DNA in a high throughput manner (Green et al., 2010; Knapp and Hofreiter, 2010). Efforts were then exerted to extract metagenomic ancient DNA fragments and convert them into genomic libraries, prior subjecting them to one of the NGS technologies. The resulting DNA sequencing data are enormous; instead of sequencing few amplified fragments, it became possible to sequence millions and even billions of fragments at one batch. Such advancement in ancient DNA sequencing technology, not only allowed detection and sequencing of marker genes of host but also allowed retrieval of whole genome sequences of hosts, e.g., modern human, Neanderthal, and Denisova (Green et al., 2010; Reich et al., 2010), and their accompanying microbes, like oral- and gut microbiota (Rampelli et al., 2015a; Warinner et al., 2015; Weyrich et al., 2017) and pathogens, like *Mycobacterium tuberculosis* (Jäger et al., 2022).

Ancient DNA damage

Additionally, during the pre-NGS era, it was known that ancient DNA undergo some changes/damages postmortem, however those modifications were not possible to be studied

meticulously until the NGS became used in large scales in paleogenetics (Dabney et al., 2013). Ancient DNA damages can be represented in three different ways (Dabney et al., 2013): i) DNA fragmentations; ii) DNA strand lesions which block enzyme action on DNA molecules; iii) chemical modifications that eventually lead to misincorporation of nucleotides. Later, and in the NGS era, these kinds of modifications were exploited as authentication criteria for ancient DNA and to reduce, computationally, the likelihood of handling modern DNA contamination as ancient (Orlando et al., 2021). For example, different bioinformatics tools have been developed to detect the terminal deamination of Cytosine-to-Thymine to assess the level of aDNA damages (Ginolhac et al., 2011; Neukamm et al., 2021). Other similar tools have been developed to use the previous information to estimate and filter the potential contaminant DNA sequenced reads, given a target reference genome, e.g., human mitochondrial genome (Renaud et al., 2015; Skoglund et al., 2014).

Sources of ancient human and human-associated microbial DNA

One of the inherent challenges of the ancient DNA field, in addition to the high fragmentation rate and deamination of Cytosine to Uracil, is the extraordinarily high amounts of non-target DNA, i.e., background DNA and/or processing contamination. Contaminant DNA is considered as DNA that is introduced during the handling of ancient samples, e.g., during excavation, handling or even in the lab during processing, which sometimes can lead to false results. This was a major topic in the early years of aDNA research and a lot of papers and criteria were published addressing this issue. With NGS this was to a great extent solved, as modern (human) contamination can now be clearly detected.

While background DNA is considered all kind of environmental DNA (mainly microbes) that were introduced during the deposition of the ancient materials in the soil, which has an influence on the amount of endogenous DNA, but not necessarily on the authenticity of the human or pathogen DNA. This issue was brought into focus, particularly when the NGS became more usable in the field, having the vast majority of sequenced reads assigned to environmental microbes and as low as 1% assigned to endogenous DNA. Therefore, efforts were exerted to improve DNA extraction and to remove the contamination prior to DNA extractions. Additionally, by comparing the quality of aDNA extracted from different human bones, it was found that systematically higher endogenous DNA was retrieved when extracted from the petrous part of the temporal bone of human skulls (**Figure 1**), compared to other skeletal bones (Hansen et al., 2017; Pinhasi et al., 2015). Teeth also are considered as

important sources of human DNA and additionally of some pathogens that are expected to be found in blood stream, like *Yersinia pestis* (Drancourt et al., 1998; Margaryan et al., 2018; Prentice et al., 2004).

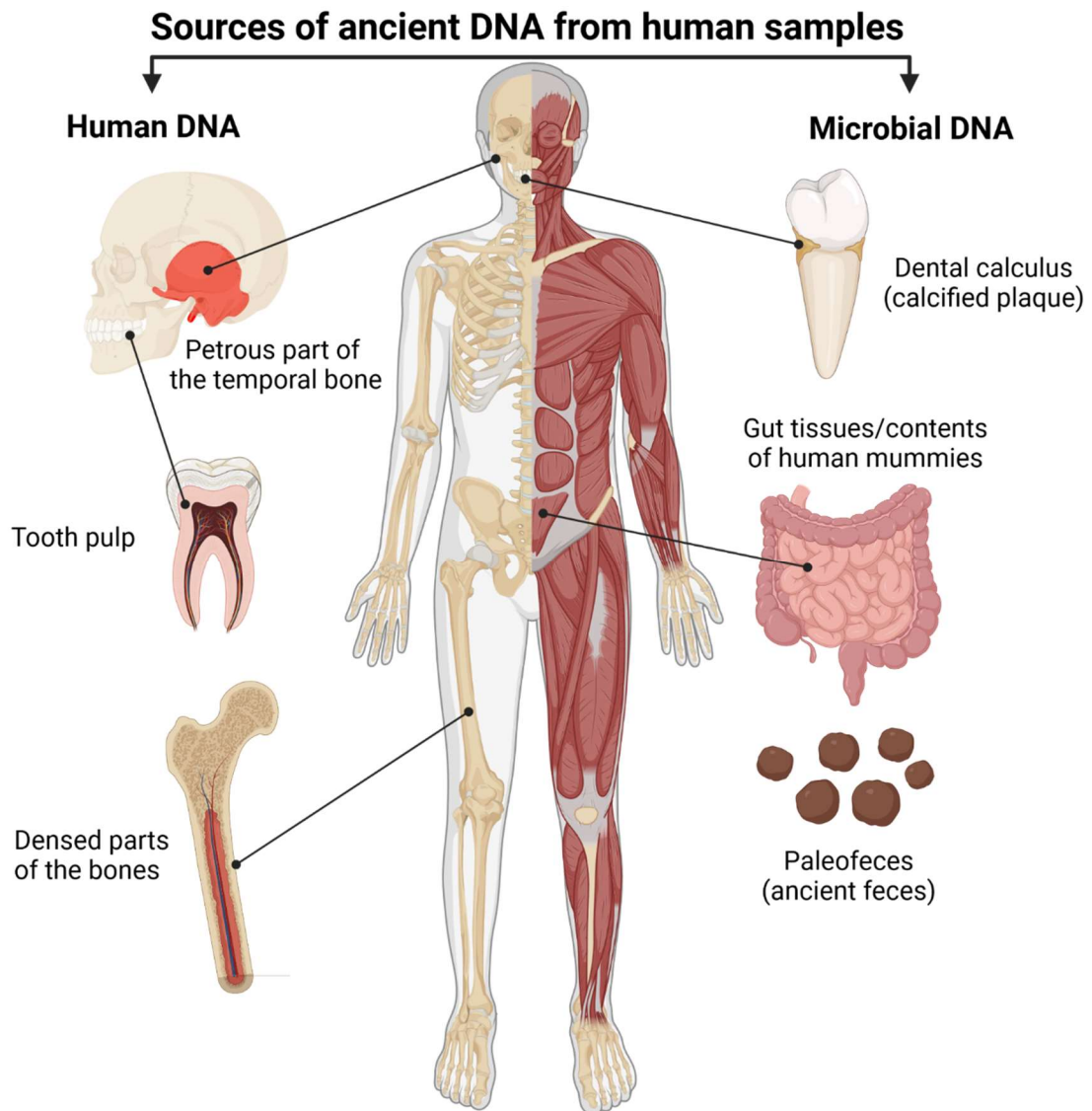


Figure 1: Different sources of ancient DNA from human samples.

While petrous bones became the preferable source of aDNA for those who are interested in endogenous human DNA, microbiologists preferred other types of samples, e.g., paleofeces and dental calculus (Green and Speller, 2017). Despite of their being usually more prone to external contaminations, they can still contain human DNA and maintain additional important ancient biological information. For instance, paleofeces, i.e. ancient feces, and gut content of ancient mummies were found to still contain some DNA of ancient gut microbiota (Tett et al.,

2019; Wibowo et al., 2021), and later were proved even to contain molecular information on the dietary components that were ingested by the ancient populations. Paleofeces also can provide information on the health status, in terms of microbial infections or parasites (Anastasiou et al., 2018; Maixner et al., 2018). Similarly, dental calculus, which is defined as calcified dental plaques that contain complex, mineralized organic biofilms, were found to contain DNA of ancient oral microbiota as well as dietary micro-remains (MacKenzie et al., 2021; Warinner et al., 2015; Weyrich et al., 2017)

Evolution of ancient microbiomes and reconstruction of ancient genomes

Analysis of the ancient dental calculus and paleofeces opened a new avenue to compare the oral and gut microbes/microbiomes of the ancient populations to the oral and gut microbes/microbiomes of the contemporary populations. Recent studies on the human gut microbiomes found that the dietary habits and the lifestyle can reflect on the gut microbiome structures, enriching for some specific taxa and depleting others (Rampelli et al., 2015b; Segata, 2015). For example, analysis of >6500 human gut metagenomes revealed clear clustering of the data into two clusters: Westernized and non-Westernized. The non-Westernized cluster showed presence of 4 different clades/strains of the species *Prevotella copri* complex, while depleted in the Westernized populations (Tett et al., 2019; Tett et al., 2021). Interestingly, analysis of ancient feces metagenomes showed presence of the *Prevotella copri* clades in a pattern similar to the modern non-Westernized populations. Using tip dating including an ancient *Prevotella* genome revealed split times of the clades that predate the human migratory waves out of Africa (Tett et al., 2019; Tett et al., 2021). Such findings might help in the future to understand how such a shift happened and how it was influenced by dietary habits and changes overtime. Despite being slow, but the current expansion of the genomes and marker genes databases of the plant, fungi, and parasites, will boost the progress in this direction.

Evolution of pathogens has been one of the important areas of research of ancient DNA, particularly with the possibility to reconstruct ancient genomes of pathogens. Studies reported reconstruction of ancient human pathogens like *Treponema denticola* (Maixner et al., 2014) and *Mycobacterium tuberculosis* (Jäger et al., 2022), which allowed understanding of infection transmissions and evolution of pathogens.

One of the important milestones of the aDNA field, was in 2013 with the introduction of the capture-based method for enriching the human endogenous aDNA, using biotinylated RNA

baits designed based on the reference genome (Carpenter et al., 2013). This approach was extended later to enrich for target microbes, particularly pathogens, from human mummies and skeletons that showed indicative symptoms of certain diseases (Bos et al., 2015; Duchêne et al., 2020). Employing the enrichment capture method in pathogen detection allowed for recovery of huge genomic data, enough for reconstruction of complete genomes, and hence allowed for genome-level comparison with their modern descendants. It was used also to provide insights onto disease evolution, migration waves, and disease endemics (Duchêne et al., 2020; Spyrou et al., 2019). This allowed for recovery of genomes of ancient prokaryotes, like *Mycobacterium leprae* (Schuenemann et al., 2013), *Mycobacterium tuberculosis* (Sabin et al., 2020), *Treponema pallidum* (Giffin et al., 2020; Majander et al., 2020), *Helicobacter pylori* (Maixner et al., 2016), and *Yersinia pestis* (Keller et al., 2019; Susat et al., 2021). Recently, it was even possible to enrich and reconstruct complete genomes Hepatitis B viruses (Kocher et al., 2021; Mühlemann et al., 2018).

In contrast to prokaryotic and viral genomes, there were not many studies attempted to reconstruct ancient eukaryotic genomes. So far and as of October 2022, there are only 4 reported ancient eukaryotes by reconstructing parts of mitochondrial- or nuclear genomes, namely the plant pathogen oomycetes *Phytophthora infestans* (Martin et al., 2013; Yoshida et al., 2013), the human pathogen *Malassezia restricta* (De-Dios et al., 2020), and the malaria-causing agents *Plasmodium falciparum* (De-Dios et al., 2019; Marciniak et al., 2016) and *Plasmodium vivax* (Gelabert et al., 2016; Van Dorp et al., 2020). The Chagas disease-causing pathogen *Trypanosoma cruzi* has been also reported using DNA probes (Aufderheide et al., 2004).

The reconstruction of ancient genomes, in general, and microbial genomes in particular, is merely a computational hurdle since we do not have plenty of processing options in the laboratory before retrieving the sequencing data. It is commonly done by aligning the reads against the reference genome of interest and then call the variants and assembly the final genome sequence guided by the reference genome (**Figure 2**). This method is known as reference-based or reference-guided genome reconstruction and was exploited successfully to reconstruct the aforementioned ancient microbial genomes.

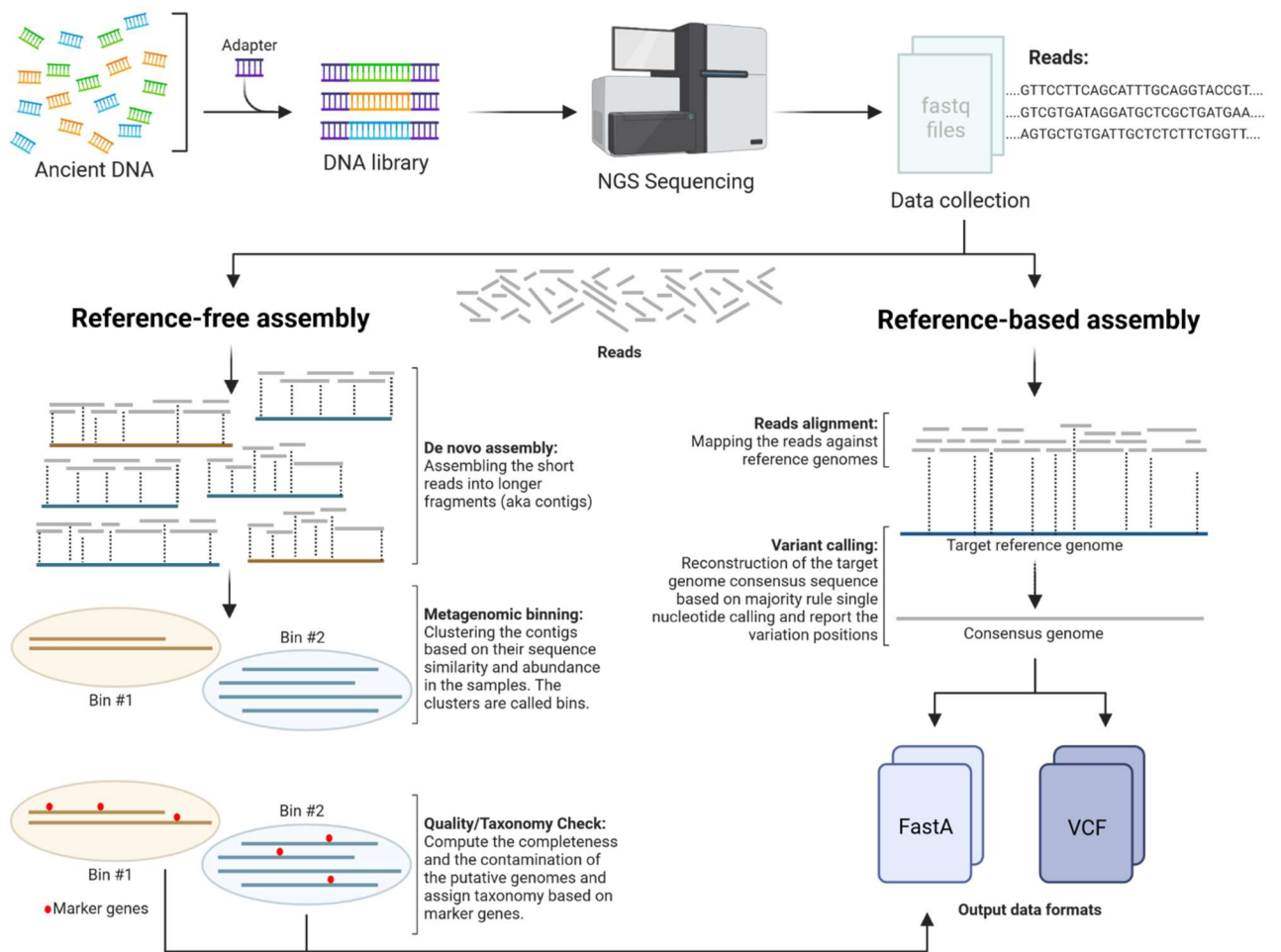


Figure 2: Different approaches to reconstruct microbial genomes from shotgun metagenomic data.

However, this approach implies availability of a modern reference genome in the first place, which require in general availability of an axenic culture (single isolates or pure cultures) of the target microbe. Which means that our potential of obtaining axenic culture of modern microbes is limiting our ability to reconstruct ancient microbial genomes using standard methods. Therefore, we need to understand how such a dilemma was approached in modern microbiology.

In general, modern microbiology field has been facing for decades the challenge of the “*great plate count anomaly*”, which refers to the contradiction between microbial counts as estimated by cultivation from one hand (culture-dependent methods) and microscopy and metagenomics in the other hand (culture-independent methods) (Hugenholtz, 2002; Sarhan et al., 2019). With the progressive development of sequencing technologies, the amount of genomic data produced by the culture-independent methods exceeded with orders of magnitude the amount of data produced by the culture-dependent methods, which created a

huge gap between metagenomes-based diversity and cultivation-based diversity (Achtman and Wagner, 2008). There exist many environmental and clinical microbes that are of biological importance where no cultured type-strain is available, e.g., the phytopathogen *Candidatus Phytoplasma* spp., the human pathogens *Treponema pallidum* and *Mycobacterium leprae*, and the insect symbionts *Wolbachia* spp. (Pallen and Microbiology, 2021). For selected cases, the complete genome sequence of those microbes was possible to be retrieved through tweaking the conventional methods, e.g., by enriching those microbes in eukaryotic hosts to increase their cellular titer or by single-cell isolation (Lasken, 2012; Lewis et al., 2021). While for the more complex environmental communities, it was until recently not possible to retrieve high quality genomes. This changed with the introduction of the new computational methods of *de novo* metagenomic assembly and metagenomic binning, which has been used for years and proved to recover high quality genomes of yet-undescribed taxa. The ground-breaking study used this approach was published in 2015 analyzing water microbes (Brown et al., 2015), followed by many other studies on different environments, including the human gut for example (Almeida et al., 2019; Pasolli et al., 2019).

How metagenomic *de novo* assembly works?

The *de novo* assembly method includes initial assembly of metagenomic reads into longer reads, known as contigs (Saheb Kashaf et al., 2021; Vollmers et al., 2017). The contigs then are to be grouped together into clusters, based on sequence characteristics, like tetranucleotide frequencies, GC content, and/or coverage. The resulting clusters are called bins and the process itself is called metagenomic binning. In theory, each of the resulting bins should represent one genome. Each of the resulting bins is then checked to contain some universal marker genes which are used as a proxy for the whole genome representation. One of the widely used tools for that purpose is CheckM (Parks et al., 2015). CheckM estimates the “completeness” of prokaryotic genome by initially checking for presence of general universal marker genes, which are then used to roughly estimate the taxonomic lineage of the putative genome. Then, it checks for presence of lineage-specific single-copy marker genes (SCG), whose representation percentage is used as a proxy for the whole genome completeness. Then, if the program detects more than one copy of any of the supposed-to-be SCG, it computes the similarity between the multiple copies and gives them as “contamination” estimates and based on how dissimilar they are to each other, it estimates the “strain heterogeneity” of the potential contamination.

In 2017, the genomic standards consortium (GSC) developed standards for reporting prokaryotic genomes (Bowers et al., 2017), published as the “Minimum information about a single amplified genome (MISAG) and a metagenome-assembled genome (MIMAG) of bacteria and archaea” (Table 1).

Table 1: Genome reporting standards for metagenome-assembled genomes (MAGs), as reported (Bowers et al., 2017)

Status	Completeness	Contamination
Finished	100% (Single sequence)	0%
High quality	>90%	<5%
Medium quality	≥50%	<10%
Low quality	<50%	<10%

Aims of the study

In order to improve the research on ancient DNA field, we need to consider the field as a triangle connecting three different aspects. Those aspects are: i) to improve sampling techniques; ii) to improve the sample processing protocols in the lab; and finally, iii) to improve the data analysis tools and methods (**Figure 3**). In this study, I aimed at improving the current methodology in DNA research by touching each angle of the ancient DNA. Using new DNA extraction protocol, I attempted to extent the current repertoire of ancient DNA sources to unusual organic and inorganic materials (Paper I and II in the results section). Additionally, I aimed at using new bioinformatic methods to reconstruct and analyze ancient microbial genomes that have not yet been reported from ancient times (Paper II and III in the results section).

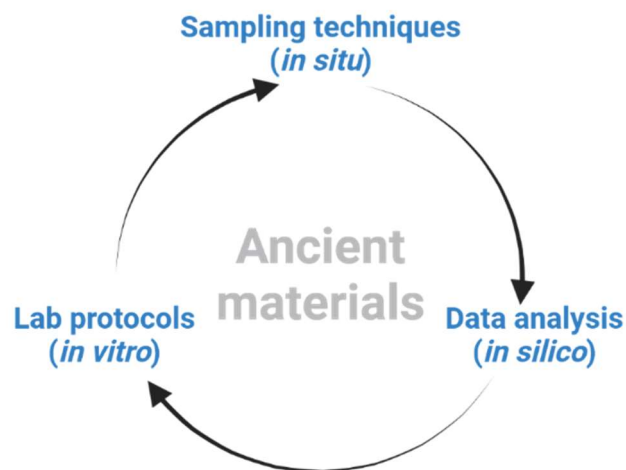


Figure 3: The three aspects of ancient DNA field which can be targeted for further improvements.

(This page intentionally left blank)

Results

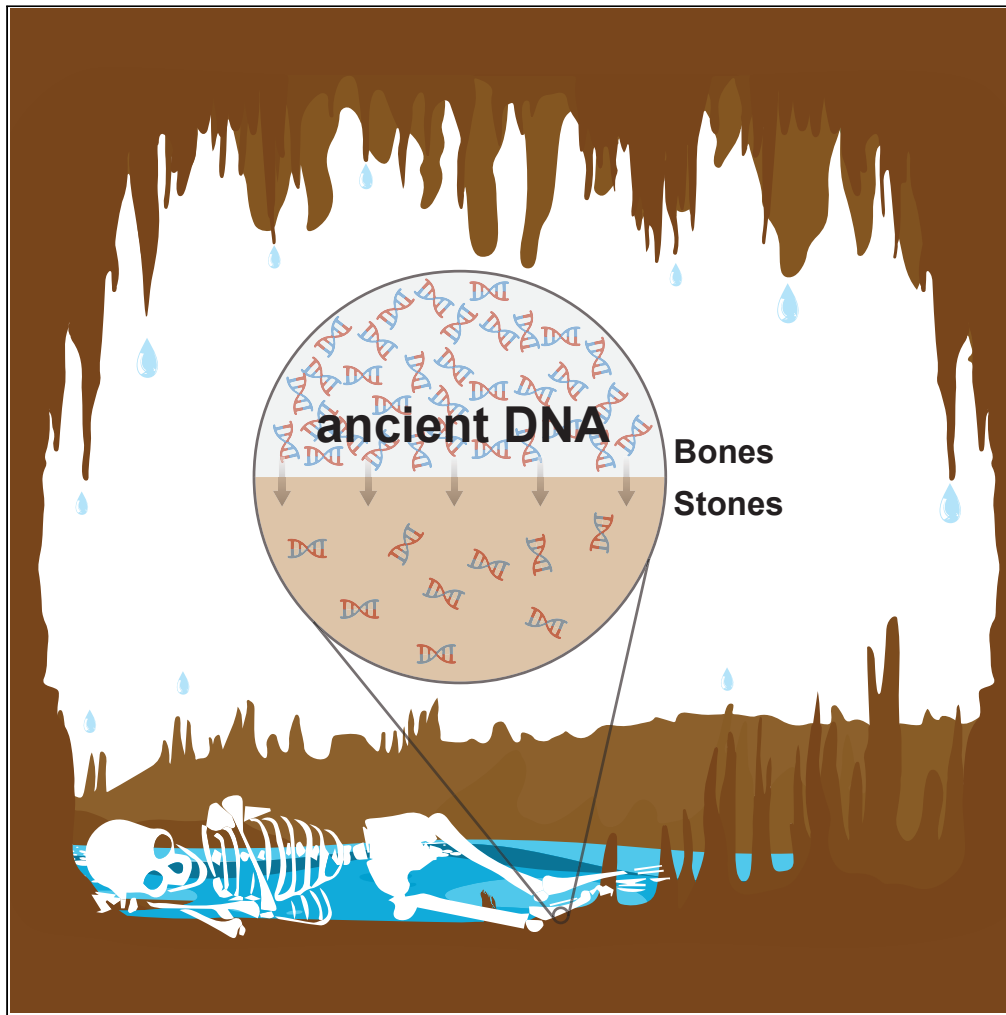
(This page intentionally left blank)

**Paper I: Ancient DNA diffuses from human bones to cave
stones**

(This page intentionally left blank)

Article

Ancient DNA diffuses from human bones to cave stones



Mohamed S. Sarhan, Achim Lehmkühl, Rainer Straub, ..., Michael Francken, Albert Zink, Frank Maixner

frank.maixner@eurac.edu (F.M.)
mohamed.sarhan@eurac.edu (M.S.S.)

Highlights

Bronze Age human skeletal remains were found unburned in Wimsener cave in Germany

The bones were covered with layers of calcite stone deposits

The ancient human and microbial DNA diffused from the bones to the stones

The mitochondrial haplogroup and sex were revealed by analyzing the stones only

Sarhan et al., iScience 24, 103397
December 17, 2021 © 2021 The Authors.
<https://doi.org/10.1016/j.isci.2021.103397>

Article

Ancient DNA diffuses from human bones to cave stones

Mohamed S. Sarhan,^{1,*} Achim Lehmkuhl,² Rainer Straub,³ Adrian Tett,⁴ Günther Wieland,⁵ Michael Francken,⁵ Albert Zink,¹ and Frank Maixner^{1,6,*}

SUMMARY

Recent studies have demonstrated the potential to recover ancient human mitochondrial DNA and nuclear DNA from cave sediments. However, the source of such sedimentary ancient DNA is still under discussion. Here we report the case of a Bronze Age human skeleton, found in a limestone cave, which was covered with layers of calcite stone deposits. By analyzing samples representing bones and stone deposits from this cave, we were able to: i) reconstruct the full human mitochondrial genome from the bones and the stones (same haplotype); ii) determine the sex of the individual; iii) reconstruct six ancient bacterial and archaeal genomes; and finally iv) demonstrate better ancient DNA preservation in the stones than in the bones. Thereby, we demonstrate the direct diffusion of human DNA from bones into the surrounding environment and show the potential to reconstruct ancient microbial genomes from such cave deposits, which represent an additional paleoarcheological archive resource.

INTRODUCTION

Ancient DNA (aDNA) has become a powerful tool to study the ancestral history, not only of hominins, but also other animals, plants, and even microbes (Capo et al., 2021; Epp et al., 2015). In addition to the skeletal and mummified remains also sedimentary materials were recently identified as a promising source of aDNA, that still containing Pleistocene Neanderthal and Denisovan mitochondrial DNA and nuclear DNA (Gelabert et al., 2021; Slon et al., 2017; Vernot et al., 2021). There were different postulations on the source of such sedimentary ancient DNA (sedaDNA), e.g., macrofossils, small bone fragments, excreta, and decayed soft tissues (Haile et al., 2007; Slon et al., 2017; Willerslev et al., 2003). However, we still miss a case where we see a direct link between the human sedaDNA and its source.

Here we report the finding of human skeletal remains (1306-1017 calBCE) dating back to the Urnfield culture of the late Bronze Age within the underwater river cave named either Wimsener Höhle (Straub, 2006; Straub and Lehmkuhl, 2009) or Friedrichshöhle near Hayingen (Swabian Alb, Baden-Württemberg, Germany, Figure 1). The discovery of an unburned skeleton itself is remarkable because the predominant tradition of Urnfield Cultures is cremation of the deceased. The finding of contemporary pottery and other human bone fragments in the entrance lake of the cave suggest the possibility of a cultic site (Straub and Lehmkuhl, 2009). Similar finds from other caves in the Swabian Alb point to a religious phenomenon of the Late Bronze Age, namely burials or ritual acts in caves (Rebay-Salisbury, 2010).

The commingled skeletal remains were found covered with heavy layers of calcite deposits, as a result of continuous water dripping from the cave ceiling, forming bulges toward the direction of the gravity (Figures 2A and 2B). Because most of the skeleton was heavily covered by calcite deposits, only a sample of the more accessible tibia was taken (Figure 1D). To further investigate this archeological finding molecularly, we used a slice of the tibia sample with the surrounding calcite layer (Figure 2B). Different parts of the bone and stone were subsampled and subjected to metagenomic sequencing (Table S1, Figure 2, and STAR Methods). On these remains we could demonstrate the diffusion of ancient human and microbial DNA from the skeleton to the surrounding environment.

¹Eurac Research, Institute for Mummy Studies, 39100 Bolzano\Bozen, Italy

²Staatliches Museum für Naturkunde Stuttgart, Rosenstein 1, 70191 Stuttgart, Germany

³Höhlenforschungsgruppe Ostalb-Kirchheim e.V. (HFGOK), 70771 Leinfelden-Echterdingen, Germany

⁴CUBE - Division of Computational Systems Biology, Centre for Microbiology and Environmental Systems Science, University of Vienna, Althanstraße 14, 1090 Vienna, Austria

⁵State Office for Cultural Heritage Management Baden-Württemberg, 73728 Esslingen, Germany

⁶Lead contact

*Correspondence: frank.maixner@eurac.edu (F.M.), mohamed.sarhan@eurac.edu (M.S.S.)

<https://doi.org/10.1016/j.isci.2021.103397>



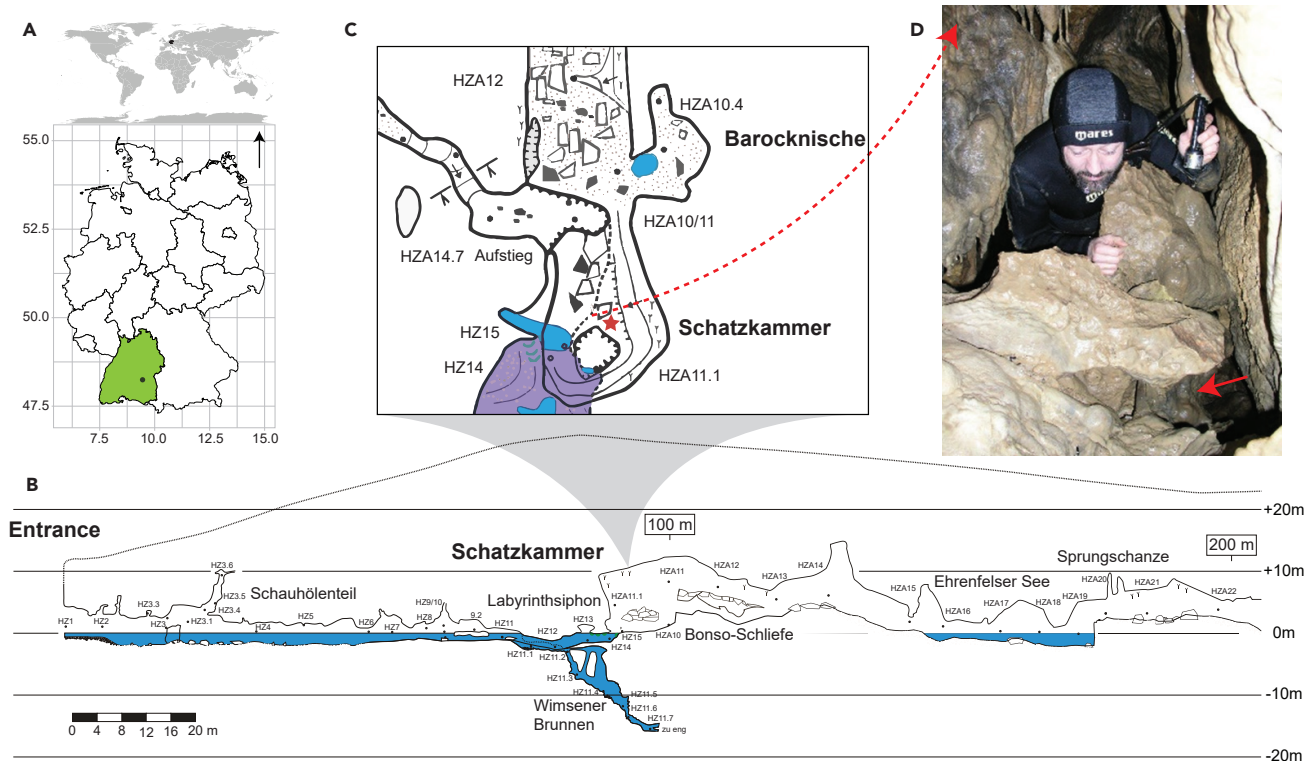


Figure 1. Skeleton finding site

(A) Location of the Wimsener caves in Baden-Württemberg, Germany.

(B) Longitudinal section of the first 200 meters of the Wimsener caves; the scale bar refers to the cave length.

(C) Cross section of the “Schatzkammer”; the red star refers to the location of the tibia.

(D) Photograph showing the difficult accessibility of the finding site; the red arrow indicates the protruding part of the sampled tibia.

RESULTS

Bone and stone contain ancient human DNA of the same individual

Analysis of the human DNA (Table S2), by comparing the DNA against the human reference genome hg19 (Rosenbloom et al., 2015), revealed the presence of human DNA reads, not only in the bone samples, but also in all stone samples. The human DNA of the stones accounted for ~1% of the human DNA found in the bone samples (Figure 2C and Table S2). In general, the human DNA fragment lengths of the stone samples were higher than the bone samples (Figure 2D), with comparable DNA damage levels, i.e., percentage of terminal cytosine deamination to thymine (Figure S1). In addition, the ancient human DNA in the stone and the bone displayed the same molecular sex (male, XY), which would not be possible with classical anthropological investigations. To further confirm that the human DNA in stone belongs to the skeleton, we enriched it for mitochondrial DNA using a hybridization capture assay (STAR Methods). This enabled us to reconstruct the mitochondrial genome from the stone samples, with > 7x coverage (Tables S2 and S3), revealing the identical mitochondrial haplotype of the bone samples (i.e., J1C1). Further and to gain a glimpse into the origin of the individual, we performed principal component analysis (PCA) against selected modern Eurasian individuals and other Bronze Age individuals (Mittnik et al., 2019). The PCA analysis showed our individual falls within the European diversity similar to other Bronze Age individuals from the same region (Figure S2). Overall, we could confidently demonstrate that the human DNA diffused from the bone into the stone and preserved in the calcite for millennia.

Ancient microbial genomes reconstructed from the bone and stone

Based on the human DNA results, we assumed that such calcite deposits could represent a time capsule and still contain ancient molecular information other than human DNA. Therefore, we extended our analysis to the microbial communities of both samples, by performing general microbial profiling and *de novo* assembly of the short metagenomic reads into longer contigs to reconstruct metagenome-assembled

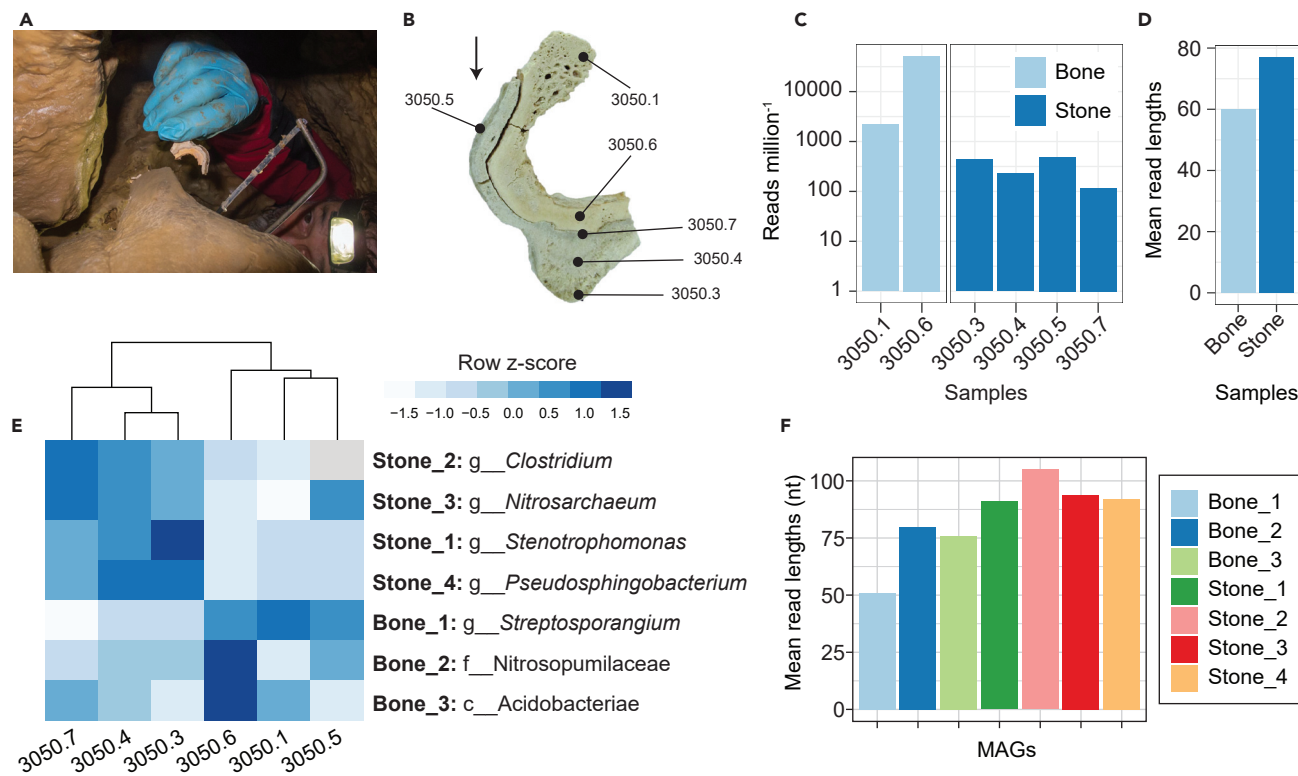


Figure 2. Metagenomic analysis of human and microbial DNA of cave human bone and its surrounding stone deposits

(A) On-site sampling of human tibia covered with a layer of calcite deposits.
 (B) Close-up cross-section of sampled tibia slice, indicating the sampling locations for DNA analysis; the black arrow refers to the direction of gravity.
 (C) Number of metagenomic reads of each sample mapped to the human autosomal DNA (Ref genome hg19).
 (D) Mean read lengths of the mapped reads against the human genome (hg19).
 (E) Heatmap showing the abundance of different metagenome-assembled genomes (MAGs) in different samples (genome copies per million reads). The color intensity indicates row z-scores.
 (F) Mean read lengths of the mapped reads the MAGs. For the (E and F) the names of MAGs refer to the source of the MAGs (stone or bone) e.g., "Stone_1" means that this genome was assembled from the stone sample.

genomes (MAGs) (See STAR Methods and Table S4). In general, the microbial profiles of the bone and stone were similar with a high richness of Archaea and Actinobacteria which is typical for cave environments (Figure S3). The bone and stone samples were then pooled independently for *de novo* assembly, which resulted in three high-quality and four medium-quality prokaryotic MAGs, from both samples (three genomes from the bone and four from the stone, please refer to Table S4 and Figure 2E). All reconstructed MAGs displayed DNA damage, except for one MAG from the bone samples (Figure S4). Similar to what was observed with human DNA, the presence of the MAGs in both the bone and stone indicate microbial DNA diffusion. It was also observed that the read lengths were longer in general in the stone compared to bone (Figures 2F and S4). These genomes represent microbes that may have been involved in initial postmortem decomposition and initiation of secondary stalactite formation. For instance, members of *Clostridium* and *Streptosporangium* are reported among the most abundant postmortem bone degraders (Eriksen et al., 2020; Philips et al., 2017). Moreover, *Streptosporangium* has also been implicated in bone tunneling in waterlogged environments (Eriksen et al., 2020; Kim et al., 2020; Turner-Walker, 2009). Although other microbes, like *Stenotrophomonas* and *Rhodococcus*, have been shown to crystalize calcium carbonate in cave environments (Enyedi et al., 2020), which may suggest a potential role in stalactite formation (Pacton et al., 2013).

DISCUSSION

In this case study, we took bone and calcite stone deposits from a Bronze Age human skeleton and subjected them to DNA analysis. We recovered well-preserved ancient human DNA and microbial DNA

from the stone and bone, enough to determine the sex of the individual, and to reconstruct human mitochondrial- and microbial genomes.

We identified the same human DNA in both bone and stone samples and thereby demonstrated the direct diffusion of human DNA from bones into the surrounding environment. This presents an additional explanation model for the reported presence of ancient human DNA in cave sediments (Slon et al., 2017; Vernot et al., 2021). Earlier, in 2003, Willerslev and colleagues reported the presence of *Euryapteryx curtus* DNA in sand samples collected from the interior of a moa bone in New Zealand (Willerslev et al., 2003). Later in 2007, Haile et al. also demonstrated the potential post-depositional vertical migration of ancient DNA across sedimentary strata (Haile et al., 2007). However, it is yet-to-be determined how the stalactite deposits can better preserve aDNA fragments from the past. In general, DNA can adsorb to different mineralogical elements because of its negative charge (Freeman and Sand, 2020). In addition, cave environments maintain a constant temperature and humidity as well as aphotic zones which allow less exposure to conventional DNA damages (Stahlschmidt et al., 2019; Zepeda Mendoza et al., 2016). It may also be possible that the leached DNA gets less exposed to the postmortem microbial decomposition, because of the fact that the bones are continuously degraded because of their content of organic substrates (e.g., proteins).

In this report we aim to bring attention to such valuable mineralogical deposits that are found attached to archaeological findings in cave environments. We showed that they are not only representing mineralogical deposits, but rather an extension to the archeological findings that retain and preserve its historical information in the form of aDNA. Currently these deposits are mainly used as dating proxies for archeological findings that are beyond the limits of the radiocarbon dating (^{14}C), e.g., using Uranium–Thorium dating (U-Th). Considering our finding, these deposits could help to avoid future destructive sampling of similar archeological remains, offering an additional paleogenetic archive that can be used to reconstruct ancient human genomes and ancient microbial communities. It remains to be explored how far such deposits are capable of preserving, in addition, other ancient biomolecules, e.g., proteins or lipids.

Finally, our message to our community of archeologists and anthropologists is to consider innovative sampling resources (e.g., sediments, mineralogical deposits, or water) during excavations, especially for molecular analyses that often involve destructive procedures.

Limitations of the study

The main limitation of the study is that it is based on a single individual from a single site, and it is also rare to find ancient skeletal samples covered with such calcite stone deposits. In addition, the total amount of human DNA reads obtained from the calcite stone deposits is two order of magnitude less than the bone. This might suggest accompanying the DNA library preparation from similar samples a subsequent application of a human nuclear/mitochondrial DNA capture approach.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
 - Lead contact
 - Material availability
 - Data and code availability
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
- METHOD DETAILS
 - DNA extraction and library preparation
 - DNA sequencing and post-sequencing processing
 - Human DNA analysis
 - Microbial DNA analysis
- QUANTIFICATION AND STATISTICAL ANALYSIS

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at <https://doi.org/10.1016/j.isci.2021.103397>.

ACKNOWLEDGMENTS

Support was provided by the European Regional Development Fund 2014–2020_CALL-FESR 2017 Research and Innovation_Autonomous Province of Bolzano - South Tyrol_Project: FESR1078-MummyLabs. The Sampling campaign was broadcasted in the ZDF documentary “Terra X -Geheimnisse aus der Tiefe” (<https://www.zdf.de/dokumentation/terra-x/geheimnisse-aus-der-tiefe-mit-florian-huber-100.html>). The authors thank the Department of Innovation, Research and University of the Autonomous Province of Bozen/Bolzano, Italy for covering the Open Access publication costs.

AUTHOR CONTRIBUTIONS

Conceptualization, M.S.S., M.F., and F.M.; Methodology, M.S.S. and F.M.; Formal Analysis, M.S.S.; Investigation, M.S.S.; Resources, A.L., R.S., G.W., M.F., A.Z., and F.M.; Writing – Original Draft, M.S.S.; Writing – Review & Editing, M.S.S., G.W., M.F., A.Z. A.T., and F.M.; Visualization, M.S.S.; Supervision, F.M.; Project Administration, M.F., A.Z., and F.M.; Funding Acquisition, M.F., A.Z., and F.M.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: August 19, 2021

Revised: September 30, 2021

Accepted: November 1, 2021

Published: January 1, 2022

REFERENCES

- Alneberg, J., Bjarnason, B.S., de Bruijn, I., Schirmer, M., Quick, J., Ijaz, U.Z., Lahti, L., Loman, N.J., Andersson, A.F., and Quince, C. (2014). Binning metagenomic contigs by coverage and composition. *Nat. Methods* 11, 1144–1146. <https://doi.org/10.1038/nmeth.3103>.
- Andrews, R.M., Kubacka, I., Chinnery, P.F., Lightowlers, R.N., Turnbull, D.M., and Howell, N. (1999). Reanalysis and revision of the Cambridge reference sequence for human mitochondrial DNA. *Nat. Genet.* 23, 147. <https://doi.org/10.1038/13779>.
- Beghini, F., McIver, L.J., Blanco-Míguez, A., Dubois, L., Asnicar, F., Maharjan, S., Mailyan, A., Manghi, P., Scholz, M., and Thomas, A.M. (2021). Integrating taxonomic, functional, and strain-level profiling of diverse microbial communities with bioBakery 3. *Elife* 10, e65088.
- Capo, E., Giguet-Covex, C., Rouillard, A., Nota, K., Heintzman, P.D., Vuillemin, A., Ariztegui, D., Arnaud, F., Belle, S., and Bertilsson, S. (2021). Lake sedimentary DNA research on past terrestrial and aquatic biodiversity: Overview and recommendations. *Quaternary* 4, 6.
- Chaumeil, P.A., Mussig, A.J., Hugenholtz, P., and Parks, D.H. (2019). GTDB-Tk: A toolkit to classify genomes with the genome taxonomy database. *Bioinformatics*. <https://doi.org/10.1093/bioinformatics/btz848>.
- Chen, S., Zhou, Y., Chen, Y., and Gu, J. (2018). fastp: An ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics* 34, i884–i890. <https://doi.org/10.1093/bioinformatics/bty560>.
- Enyedi, N.T., Makk, J., Kótai, L., Berényi, B., Klébert, S., Sebestyén, Z., Molnár, Z., Borsodi, A.K., Leél-Óssy, S., and Demény, A. (2020). Cave bacteria-induced amorphous calcium carbonate formation. *Sci. Rep.* 10, 1–12.
- Epp, L.S., Gussarova, G., Boessenkool, S., Olsen, J., Haile, J., Schröder-Nielsen, A., Ludikova, A., Hassel, K., Stenøien, H.K., and Funder, S. (2015). Lake sediment multi-taxon DNA from North Greenland records early post-glacial appearance of vascular plants and accurately tracks environmental changes. *Quat. Sci. Rev.* 117, 152–163.
- Eriksen, A.M.H., Nielsen, T.K., Matthiesen, H., Carøe, C., Hansen, L.H., Gregory, D.J., Turner-Walker, G., Collins, M.J., and Gilbert, M.T.P. (2020). Bone biodeterioration—the effect of marine and terrestrial depositional environments on early diagenesis and bone bacterial community. *PLoS One* 15, e0240512.
- Freeman, C., and Sand, K. (2020). Survival of environmental DNA in natural environments: Surface charge and topography of minerals as driver for DNA storage. *bioRxiv*. 2020.2001.2028.922997. <https://doi.org/10.1101/2020.01.28.922997>.
- Gelabert, P., Sawyer, S., Bergström, A., Margaryan, A., Collin, T.C., Meshveliani, T., Belfer-Cohen, A., Lordkipanidze, D., Jakeli, N., Matskevich, Z., et al. (2021). Genome-scale sequencing and analysis of human, wolf, and bison DNA from 25,000-year-old sediment. *Curr. Biol.* <https://doi.org/10.1016/j.cub.2021.06.023>.
- Haile, J., Holdaway, R., Oliver, K., Bunce, M., Gilbert, M.T.P., Nielsen, R., Munch, K., Ho, S.Y., Shapiro, B., and Willerslev, E. (2007). Ancient DNA chronology within sediment deposits: Are paleobiological reconstructions possible and is DNA leaching a factor? *Mol. Biol. Evol.* 24, 982–989.
- Jónsson, H., Ginolhac, A., Schubert, M., Johnson, P., and Orlando, L. (2013). mapDamage2.0: Fast approximate Bayesian estimates of ancient DNA damage parameters. *Bioinformatics* 29, 1682–1684.
- Kang, D.D., Li, F., Kirton, E., Thomas, A., Egan, R., An, H., and Wang, Z. (2019). MetaBAT 2: An adaptive binning algorithm for robust and efficient genome reconstruction from metagenome assemblies. *PeerJ* 7, e7359. <https://doi.org/10.7717/peerj.7359>.
- Kim, H., Cho, Y., Lee, J., Kim, H.S., Jung, J.Y., and Kim, E.S. (2020). Metagenomic analysis of postmortem-bone using next-generation sequencing and forensic microbiological application §. *Microbiol. Soc. Korea* 56, 10–18.
- Langmead, B., and Salzberg, S.L. (2012). Fast gapped-read alignment with Bowtie 2. *Nat. Methods* 9, 357. <https://doi.org/10.1038/nmeth.1923>. <https://www.nature.com/articles/nmeth.1923#supplementary-information>.
- Li, D., Liu, C.-M., Luo, R., Sadakane, K., and Lam, T.-W. (2015). MEGAHIT: An ultra-fast single-node solution for large and complex metagenomics assembly via succinct de Bruijn graph. *Bioinformatics* 31, 1674–1676. <https://doi.org/10.1093/bioinformatics/btv033>.
- Li, H., and Durbin, R. (2010). Fast and accurate long-read alignment with Burrows-Wheeler transform. *Bioinformatics* 26, 589–595.
- Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., and Durbin, R. (2009). The sequence alignment/map format and SAMtools. *Bioinformatics* 25, 2078–2079.
- Meyer, M., and Kircher, M. (2010). Illumina sequencing library preparation for highly multiplexed target capture and sequencing. *Cold Spring Harb Protoc.* 2010. pdb prot5448. <https://doi.org/10.1101/pdb.prot5448>.

- Mittnik, A., Massy, K., Knipper, C., Wittenborn, F., Friedrich, R., Pfrengle, S., Burri, M., Carlich-Witjes, N., Deeg, H., Furtwangler, A., et al. (2019). Kinship-based social inequality in Bronze Age Europe. *Science* 366, 731–734. <https://doi.org/10.1126/science.aax6219>.
- Neukamm, J., Peltzer, A., and Nieselt, K. (2021). DamageProfiler: Fast damage pattern calculation for ancient DNA. *Bioinformatics*. <https://doi.org/10.1093/bioinformatics/btab190>.
- Nurk, S., Meleshko, D., Korobeynikov, A., and Pevzner, P.A. (2017). metaSPAdes: A new versatile metagenomic assembler. *Genome Res.* 27, 824–834. <https://doi.org/10.1101/gr.213959.116>.
- Okonechnikov, K., Conesa, A., and Garcia-Alcalde, F. (2016). Qualimap 2: Advanced multi-sample quality control for high-throughput sequencing data. *Bioinformatics* 32, 292–294. <https://doi.org/10.1093/bioinformatics/btv566>.
- Pacton, M., Breitenbach, S.F., Lechleitner, F.A., Vaks, A., Rolhion-Bard, C., Gutareva, O., Osintcev, A., and Vasconcelos, C. (2013). The role of microorganisms in the formation of a stalactite in Botovskaya Cave, Siberia—paleoenvironmental implications. *Biogeosciences* 10, 6115–6130.
- Parks, D.H., Imelfort, M., Skennerton, C.T., Hugenholtz, P., and Tyson, G.W. (2015). CheckM: Assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. *Genome Res.* 25, 1043–1055. <https://doi.org/10.1101/gr.186072.114>.
- Philips, A., Stolarek, I., Kuczkowska, B., Juras, A., Handschuh, L., Piontek, J., Kozłowski, P., and Figlerowicz, M. (2017). Comprehensive analysis of microorganisms accompanying human archaeological remains. *GigaScience* 6, gix044.
- Rebay-Salisbury, K. (2010). Cremations: Fragmented Bodies in the Bronze and Iron Ages. In *Body Parts and Bodies Whole: Changing Relations and Meanings*, K. Rebay-Salisbury, M.L.S. Sørensen, and J. Hughes, eds. (Oxford: Oxbow), pp. 64–71.
- Renaud, G., Slon, V., Duggan, A.T., and Kelso, J. (2015). Schmutzi: Estimation of contamination and endogenous mitochondrial consensus calling for ancient DNA. *Genome Biol.* 16, 224. <https://doi.org/10.1186/s13059-015-0776-0>.
- Rosenbloom, K.R., Armstrong, J., Barber, G.P., Casper, J., Clawson, H., Diekhans, M., Dreszer, T.R., Fujita, P.A., Guruvadoo, L., Haeussler, M., et al. (2015). The UCSC genome browser database: 2015 update. *Nucleic Acids Res.* 43, D670–D681. <https://doi.org/10.1093/nar/gku1177>.
- Shen, W., Le, S., Li, Y., and Hu, F. (2016). SeqKit: A cross-platform and ultrafast toolkit for FASTA/Q file manipulation. *PLoS One* 11, e0163962. <https://doi.org/10.1371/journal.pone.0163962>.
- Skoglund, P., Storå, J., Götherström, A., and Jakobsson, M. (2013). Accurate sex identification of ancient human remains using DNA shotgun sequencing. *J. Archaeol. Sci.* 40, 4477–4482.
- 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., et al. (2017). Neandertal and denisovan DNA from pleistocene sediments. *Science* 356, 605–608. <https://doi.org/10.1126/science.aam9695>.
- Stahlschmidt, M.C., Collin, T., Fernandes, D., Bar-Oz, G., Belfer-Cohen, A., Gao, Z., Jakeli, N., Matskevich, Z., Meshveliani, T., and Pritchard, J. (2019). Ancient mammalian and plant DNA from late quaternary stalagmite layers at Solkota Cave, Georgia. *Sci. Rep.* 9, 1–10.
- Straub, R. (2006). Archäologische Funde aus der Wimsener Höhle (Kat. Nr.-7722/01), Schwäbische Alb—Frühgeschichtliche Nutzung oder Quellenkult vor 3.200 Jahren. *Mitt. Verb. dt. Höhlen- u. Karstforscher* 52, 45–51.
- Straub, R., and Lehmkuhl, A. (2009). Spätbronzezeitlicher Skelettfund aus der Wimsener Höhle, Schwäbische Alb. *Baden-Württemberg. Mitteilungen des Verbands der deutschen Höhlen- und Karstforscher* 553, 86–90.
- Tang, J.-N., Zeng, Z.-G., Wang, H.-N., Yang, T., Zhang, P.-J., Li, Y.-L., Zhang, A.-Y., Fan, W.-Q., Zhang, Y., and Yang, X. (2008). An effective method for isolation of DNA from pig faeces and comparison of five different methods. *J. Microbiol. Methods* 75, 432–436.
- Turner-Walker, G. (2009). *Degradation Pathways and Conservation Strategies for Ancient Bone from Wet, Anoxic Sites* (Nederlandse Archeologische Rapporten (NAR)), pp. 659–675.
- Uritskiy, G.V., DiRuggiero, J., and Taylor, J. (2018). MetaWRAP—a flexible pipeline for genome-resolved metagenomic data analysis. *Microbiome* 6, 158. <https://doi.org/10.1186/s40168-018-0541-1>.
- Vernot, B., Zavala, E.I., Gómez-Olivencia, A., Jacobs, Z., Slon, V., Mafessoni, F., Romagné, F., Pearson, A., Petr, M., and Sala, N. (2021). Unearthing neanderthal population history using nuclear and mitochondrial DNA from cave sediments. *Science* 372, eabf1667.
- Weissensteiner, H., Pacher, D., Kloss-Brandstatter, A., Forer, L., Specht, G., Bandelt, H.J., Kronenberg, F., Salas, A., and Schonherr, S. (2016). HaploGrep 2: Mitochondrial haplogroup classification in the era of high-throughput sequencing. *Nucleic Acids Res.* 44, W58–W63. <https://doi.org/10.1093/nar/gkw233>.
- Willerslev, E., Hansen, A.J., Binladen, J., Brand, T.B., Gilbert, M.T., Shapiro, B., Bunce, M., Wiuf, C., Gilichinsky, D.A., and Cooper, A. (2003). Diverse plant and animal genetic records from Holocene and Pleistocene sediments. *Science* 300, 791–795. <https://doi.org/10.1126/science.1084114>.
- Wu, Y.W., and Singer, S.W. (2021). Recovering individual genomes from metagenomes using MaxBin 2.0. *Curr. Protoc.* 1, e128. <https://doi.org/10.1002/cpz1.128>.
- Zepeda Mendoza, M.L., Lundberg, J., Ivarsson, M., Campos, P., Nylander, J.A., Sallstedt, T., and Dalen, L. (2016). Metagenomic analysis from the interior of a speleothem in Tjuv-Ante's cave, northern Sweden. *PLoS One* 11, e0151577.
- Zhou, C., Xu, Q., He, S., Ye, W., Cao, R., Wang, P., Ling, Y., Yan, X., Wang, Q., and Zhang, G. (2020). GTDB: An integrated resource for glycosyltransferase sequences and annotations. *Database (Oxford)* 2020. <https://doi.org/10.1093/database/baaa047>.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Biological samples		
Bone sample from the tibia 1	This study	3050.1
Bone sample from the tibia 2	This study	3050.6
Calcite deposit sample 1	This study	3050.3
Calcite deposit sample 2	This study	3050.4
Calcite deposit sample 3	This study	3050.5
Calcite deposit sample 4	This study	3050.7
Chemicals, peptides, and recombinant proteins		
EDTA disodium salt dihydrate	Carl Roth	Cat #8043.2
Proteinase K	Promega	Cat #MC5005
Critical commercial assays		
myBaits Mito – Human	Daicel Arbor Biosciences	Cat #303008.v4
Deposited data		
Metagenomic shotgun datasets	This study	ENA: PRJEB47715
1240K SNPs + Human origin datasets		https://reich.hms.harvard.edu/datasets
Software and algorithms		
Fastp	(Chen et al., 2018)	https://github.com/OpenGene/fastp
SeqKit	(Shen et al., 2016)	https://bioinf.shenwei.me/seqkit/
SequenceTools	N/A	https://github.com/stschiff/sequenceTools
EIGENSOFT	N/A	https://github.com/DReichLab/EIG
mapDamage2	(Jónsson et al., 2013)	https://ginolhac.github.io/mapDamage/
MetaPhlan3.0	(Beghini et al., 2021)	https://github.com/biobakery/MetaPhlan/wiki/MetaPhlan-3.0
BWA	(Li and Durbin, 2010)	http://bio-bwa.sourceforge.net/
QualiMap	(Okonechnikov et al., 2016)	http://qualimap.conesalab.org/
Schmutzi	(Renaud et al., 2015)	https://github.com/grenaud/schmutzi
DeDup tool	N/A	https://github.com/apeltzer/DeDup
DamageProfiler	(Neukamm et al., 2021)	https://damageprofiler.readthedocs.io/en/latest/index.html
Molecular sex determination	(Skoglund et al., 2013)	https://github.com/pontussk/ry_compute
SAMtools	(Li et al., 2009)	http://samtools.github.io/
HaploGrep2.0	(Weissensteiner et al., 2016)	https://haplogrep.i-med.ac.at/
bowtie2	(Langmead and Salzberg, 2012)	http://bowtie-bio.sourceforge.net/bowtie2/
SPAdes	(Nurk et al., 2017)	https://cab.spbu.ru/software/spades/
MEGAHIT	(Li et al., 2015)	https://github.com/voutcn/megahit
MetaBAT2	(Kang et al., 2019)	https://bitbucket.org/berkeleylab/metabat/src/master/
MaxBin2	(Wu and Singer, 2021)	https://sourceforge.net/projects/maxbin2/
CONCOCT	(Alneberg et al., 2014)	https://concoct.readthedocs.io/en/latest/
metaWRAP	(Uritskiy et al., 2018)	https://github.com/bxlab/metaWRAP
CheckM	(Parks et al., 2015)	https://ecogenomics.github.io/CheckM/
GTDB-tk	(Chaumeil et al., 2019; Zhou et al., 2020)	https://github.com/GenomeMicrobiology/GenomeTaxonAnnotator
R-Studio	N/A	https://www.rstudio.com/

RESOURCE AVAILABILITY

Lead contact

- Further information on materials, datasets, and protocols should be directed to and will be fulfilled by the Lead Contact, Frank Maixner (frank.maixner@eurac.edu).

Material availability

- This study did not generate new unique reagents.

Data and code availability

- The metagenomic shotgun sequencing data have been deposited at ENA: PRJEB47715 and are publicly available as of the date of publication.
- All codes used in this study and other previously published genomic data is available at the sources referenced in the [key resources table](#).
- Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

In this study, we metagenomically analyzed samples of human skeletal remains (1306-1017 calBCE) dating back to the Urnfield culture of the late Bronze Age within the underwater river cave named either Wimsener Höhle or Friedrichshöhle near Hayingen (Swabian Alb, Baden-Württemberg, Germany). The bone samples were found covered with layers of calcite deposits, which was also metagenomically analyzed.

METHOD DETAILS

DNA extraction and library preparation

Amounts of 20-100 mg of bone/stone were sampled at the locations indicated in [Figure 2B](#) (Please refer to [Table S1](#) for further details). EDTA/Proteinase K mixtures were added to the samples and incubated for 24 h at 40°C, followed by purification and recovery modified protocol from Tang et al. ([Tang et al., 2008](#)). DNA extracts were quantified using QUANTUS (Promega, USA), then 20 µl of each sample were converted into double-indexed pair-end DNA libraries following a special protocol for highly degraded ancient DNA ([Meyer and Kircher, 2010](#)). All previous steps were carried out in the ancient DNA laboratory of the institute of mummy studies at EURAC research in Bolzano, Italy. Further, selected indexed samples were pooled and enriched for human mitochondrial genome (myBaits®, <https://arborbiosci.com>), following the manufacturer's instructions (Selected samples are indicated in [Table S2](#)).

DNA sequencing and post-sequencing processing

The indexed libraries were pooled and subjected to next generation DNA sequencing using HiSeqX (2 × 150 pair end). After demultiplexing, we used the tool fastp ([Chen et al., 2018](#)) to trim the adapters and low quality sequences, and to merge pair-end reads with at least 10 nucleotides overlap. The overall workflow is described in the [Figure S5](#).

Human DNA analysis

First, for each sample we used the tool SeqKit ([Shen et al., 2016](#)) to deduplicate the merged sequences, and to remove sequences shorter than 25 nucleotides. Next, we mapped these reads against the human reference genome (build hg19) ([Rosenbloom et al., 2015](#)) and the human reference mtDNA genome (rCRS) ([Andrews et al., 1999](#)) using Burrows-Wheeler Aligner (BWA) ([Li and Durbin, 2010](#)). SAMtools was used to filter for minimum mapping quality of 30. We used QualiMap to generate basic mapping statistics ([Okonechnikov et al., 2016](#)). To check the authenticity of the mapped reads being of ancient origin and not modern contamination we used the tool mapDamage 2.0 ([Jónsson et al., 2013](#)) to quantify the percentages of C to T and G to A substitution. In the case of mitochondrial DNA, we additionally used option “-rescale” to rescale the quality scores of the damaged mis-incorporated sites, and the tool Schmutzi to estimate the level of contamination based on deamination patterns ([Renaud et al., 2015](#)). For haplogroup assignment, we first converted the rescaled bam files into variant calling format (VCF) and then employed HaploGrep 2.0 ([Weissensteiner et al., 2016](#)).

To compare our sample to other individuals, we used SAMtools mpileup and pileupCaller (<https://github.com/stschiff/sequenceTools>) to call the 1240K targeted SNPs (David Reich lab, <https://reich.hms.harvard.edu/datasets>), with pseudodiploid method. Then, we used the EIGENSOFT tools “mergeit” to merge the data and “smartpca” to perform the principal component analysis (PCA). Finally, our ancient sample as well as selected Bronze Age individuals from Mittnik et al. (Mittnik et al., 2019) were projected to the modern Eurasian and Middle Eastern individuals.

To determine the sex of the individual, we used the mapped human DNA reads to compute the karyotype frequency of X and Y chromosomes, using a Maximum likelihood method (Skoglund et al., 2013).

Microbial DNA analysis

For general microbiome profiling, we applied MetaPhlan 3.0 on the merged reads, with options of “-min_mapq_val 30” and “-read_min_len 25” (Beghini et al., 2021).

Additionally, we used the quality-filtered unmerged reads to perform *de novo* sequence assembly, following a co-assembly approach by combining samples into two groups, bone (n = 2) and stone (n = 4), and using the MetaWRAP pipeline (Uritskiy et al., 2018). In detail, we used the metagenomic assembler MEGAHIT (Li et al., 2015) and SPAdes with “-meta” option (Nurk et al., 2017), to assemble the pair-end reads into longer contigs. All contigs shorter than 1000 nt were excluded from downstream analyses. We used three different metagenomic binning tools, MetaBAT2 (Kang et al., 2019), MaxBin2 (Wu and Singer, 2021), and CONCOCT (Alneberg et al., 2014), to cluster the contigs into bins based on abundance and kmer frequency. Then, we checked the completeness and contamination of the resulting bins using CheckM (Parks et al., 2015), and kept only the high- (completeness > 90% and contamination < 5%) and medium (completeness 50-90% and contamination < 10%) quality bins for further analysis. Further, we used the “reassemble_bins” module to remap the raw reads against each of the bins, then reassembled the mapped reads using SPAdes assembler. This step was proved to increase the completeness and reduce the contamination of the bins. In order to assess the abundance of each bin in each of the samples, we implemented the MetaWRAP module “quant_bins”, which quantifies the abundance of each bin and expresses it as a unit of “genome copies per million reads”.

To taxonomically classify each bin, we used GTDB-Tk v1.5.0 and the module “classify_wf” (Chaumeil et al., 2019; Zhou et al., 2020), which identifies 120 marker genes in bacteria and 122 marker genes in archaea, performs a multiple sequence alignment of the markers, and finally assigns taxonomy based on the phylogenetic placement with known reference genomes. Average nucleotide identity (ANI)/relative evolutionary divergence between reference genomes and metagenomic bins (henceforth referred to as metagenomes-assembled genomes, MAGs) was also performed using GTDB-Tk.

Finally, to differentiate the MAGs of ancient origin from those modern ones, we used Bowtie2 (Langmead and Salzberg, 2012) to map the quality filtered raw reads against these MAGs, then we sorted and indexed the resulted bam files using SAMtools (Li et al., 2009), and implemented DamageProfiler (Neukamm et al., 2021) to check for ancient DNA damage patterns, i.e. C-to-T and G-to-A substitutions, resulting from cytosine deamination.

QUANTIFICATION AND STATISTICAL ANALYSIS

The abundance of MAGs in each of the samples were shown as row z-scores calculated and visualized, in R-Studio (<https://www.rstudio.com/>) using the pheatmap package.

Paper I: Supplementary Information

(This page intentionally left blank)

iScience, Volume 24

Supplemental information

Ancient DNA diffuses from human bones to cave stones

**Mohamed S. Sarhan, Achim Lehmkuhl, Rainer Straub, Adrian Tett, Günther
Wieland, Michael Francken, Albert Zink, and Frank Maixner**

Supplementary Information

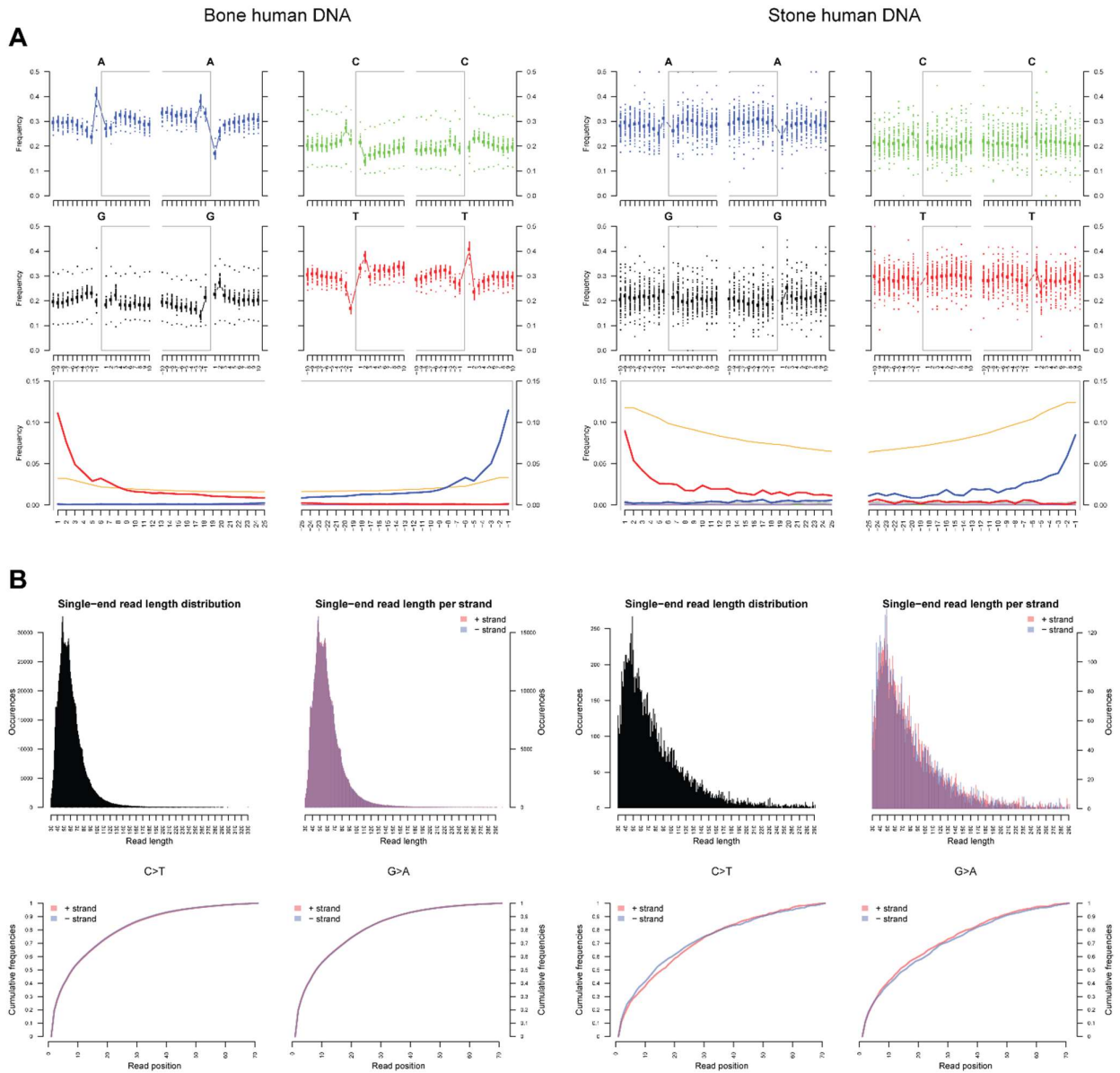


Figure S1: DNA damage of human DNA extracted from bone and stone samples, as inferred by mapDamage (Jonsson et al., 2013) , Related to Figure 2C and 2D. (A) DNA damage plots of bone- (left) and stone (right) human DNA. (B) Fragment lengths distribution of bone- (left) and stone (right) human DNA.

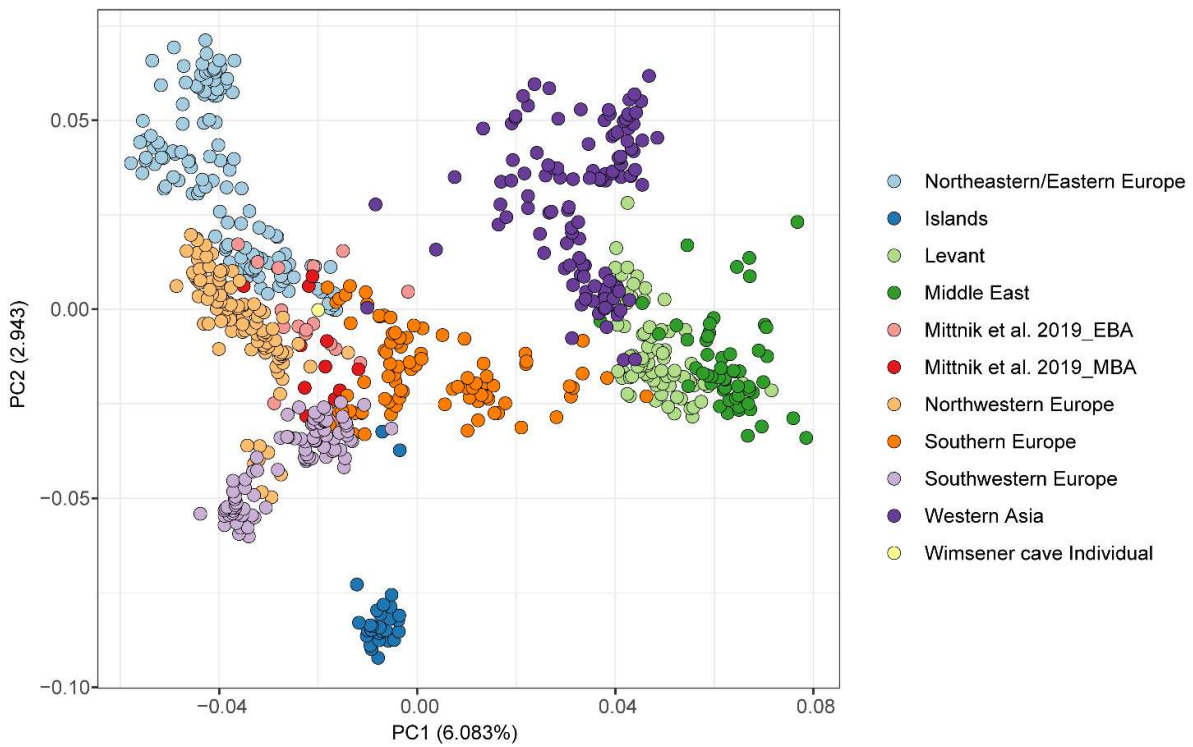


Figure S2: Principal component analysis (PCA) plot of modern west Eurasian individuals; the Wimsener cave ancient individual, as well as selected Middle Bronze Age (MBA) individuals and Early Bronze Age (EBA) individuals (Mittnik et al., 2019) , Related to Figure 2C.

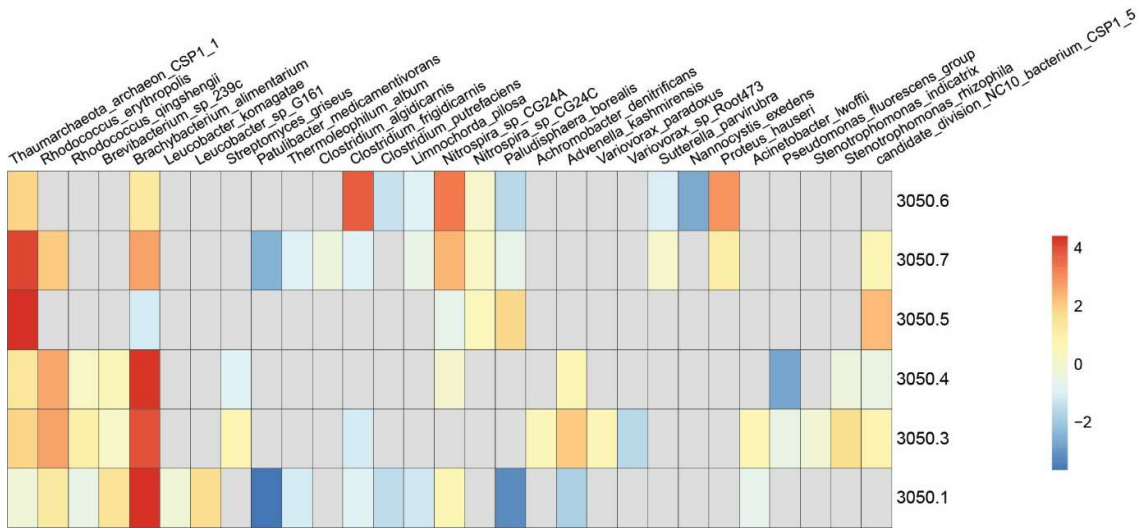


Figure S3: Microbial profiling of the bone- (3050.1 and 3050.6) and stone (3050.3, 3050.4, 3050.5, and 3050.7) samples as inferred by MetaPhlan 3.0, Related to Figure 2E. The heatmap represent the row-z scores of the relative abundances of the microbial species.

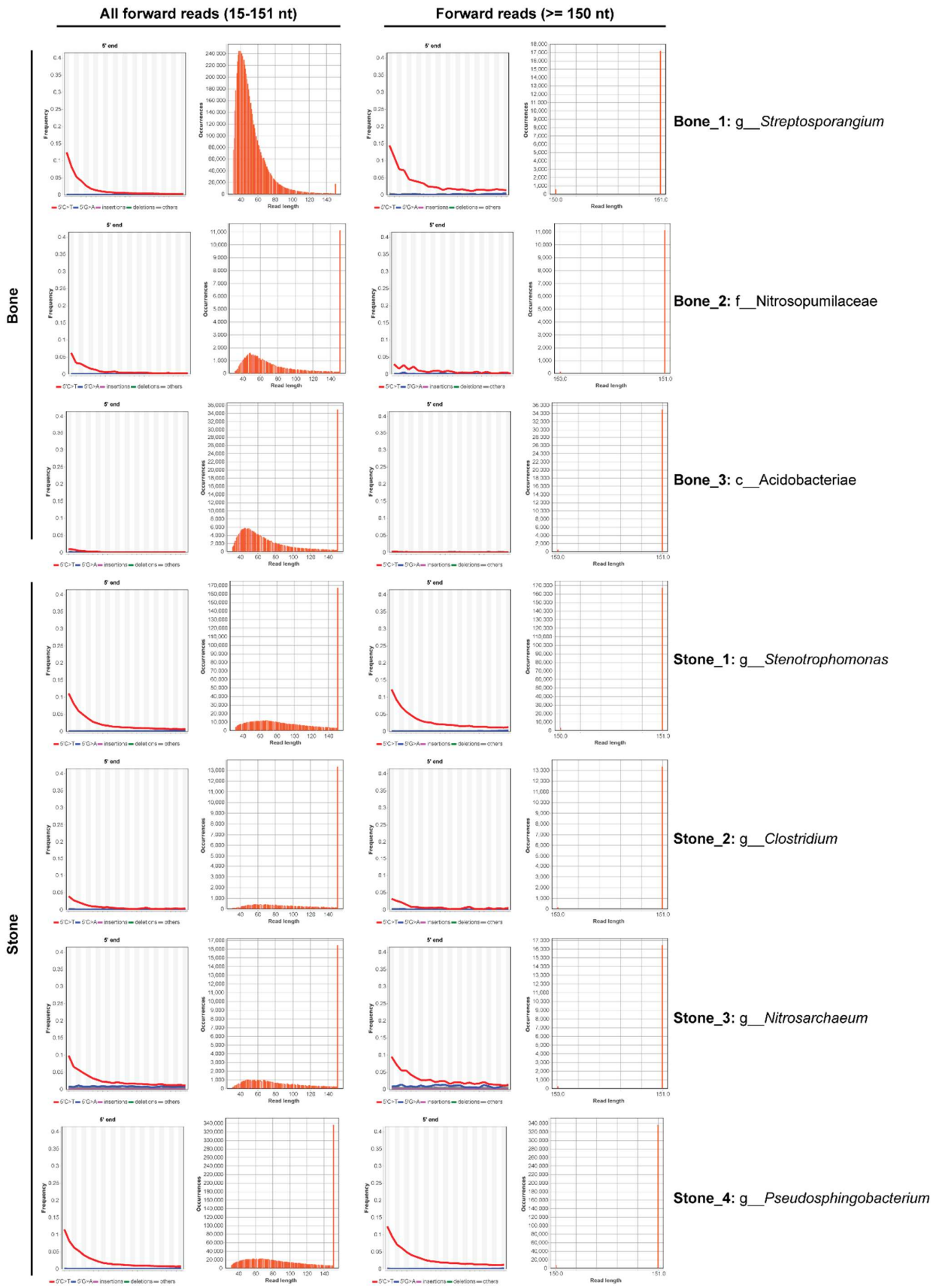


Figure S4: DNA damage of microbial DNA reads mapped to the reconstructed metagenome-assembled genomes (MAGs) from bone- and stone samples, as inferred by DamageProfiler

(Neukamm et al., 2021) , Related to Figure 2E and 2F. The left panel shows the DNA damage and read length distribution of all forward reads, while the right panel shows the DNA damage of reads of ≥ 150 nt length, which we assume failed to merge with their corresponding reverse read due to insufficient overlapped nucleotides. Notice in both the long and short reads comparable damage patterns with both displaying a typical ancient DNA read-length distribution. Taxonomic labels to the right refer to the classification assigned by GTDB-Tk (Chaumeil et al., 2019).

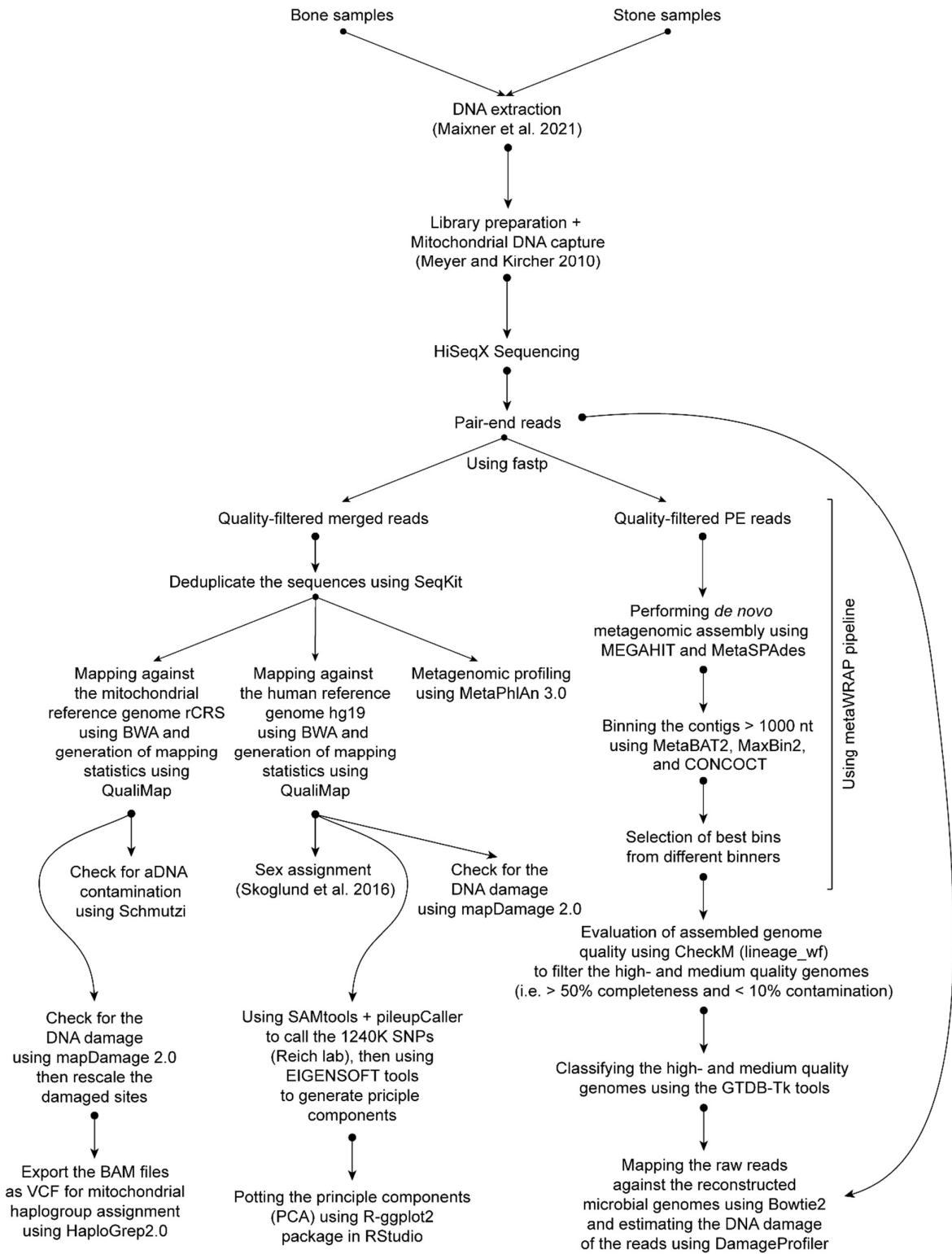


Figure S5: The methodological workflow used in this study; DNA extraction, library preparation, and the shotgun metagenomic data analysis, Related to STAR Methods.

Table S1: Total number of shotgun reads generated from the bone and stone samples, Related to Figure 2B and STAR Methods

ID	position	DNA extraction		raw_reads	quality filtered reads			quality filtered reads after removing PCR duplicates			
		sample weight (mg)	DNA_conc.(ng/μl)		merged	unmerged	merging_rate	merged	forward	reverse	Duplication rate
3050.1	bone	85	4.53	16930821	16517550	297168	97.56%	13596797	247201	287848	17.68%
3050.3	stone	42	0.235	12758786	11933560	706049	93.53%	9326305	567985	674023	21.85%
3050.4	stone	24	0.361	12117084	11560379	470498	95.41%	9401534	386832	450627	18.67%
3050.5	stone	93	1.021	8877178	7957401	736903	89.64%	6571267	595709	712381	17.42%
3050.6	bone	20	1.65	23772913	23379227	213249	98.34%	18886233	186108	199428	19.22%
3050.7	stone	100	0.714	24338373	23570444	637652	96.84%	19001671	545281	585695	19.38%

Table S2: Human autosomal and mitochondrial DNA mapping statistics, Related to Figure 2C and Figure 2D

ID	position	# of reads	Human autosomal DNA				Mitochondrial DNA	
			# of mapped reads	# of reads / million	mean	std_dev	coverage (x)	Enrichment
3050.1	bone	13596797	37223	2738	67	32	53.14	Yes
3050.3	stone	9326305	4588	492	96	52	2.44	Yes
3050.4	stone	9401534	2274	242	64	46	2.68	Yes
3050.5	stone	6571267	3397	517	63	32	2.1	Yes
3050.6	bone	18886233	1042717	55210	60	20	5.77	No
3050.7	stone	19001671	3008	158	71	38	0.1023	No
Bone	bone	32483030	1079944	33246	60	20	58.9	-
Stone	stone	44300777	13264	299	77	46	7.32	-

Table S3: Mitochondrial DNA alternative alleles of the Wimsener cave individual, compared to the reference mitochondrial genome (rCRS), Related to Figure 2C

CHROM	POS	REF	Bone		Stone	
			ALT	Depth	ALT	Depth
chrMT	73	A	G	DP=43	G	DP=1
chrMT	185	G	A	DP=57	A	DP=2
chrMT	263	A	G	DP=76	G	DP=2
chrMT	295	C	T	DP=64	ND	ND
chrMT	309*	C	CCT	INDEL	ND	ND
chrMT	310*	T	TTT	INDEL	C	DP=2
chrMT	462	C	T	DP=64	T	DP=11
chrMT	482	T	C	DP=66	C	DP=10
chrMT	489	T	C	DP=67	C	DP=10
chrMT	750	A	G	DP=52	G	DP=2
chrMT	1438	A	G	DP=40	ND	ND
chrMT	2706	A	G	DP=36	ND	ND
chrMT	3010	G	A	DP=54	A	DP=7
chrMT	3394	T	C	DP=50	C	DP=1
chrMT	4216	T	C	DP=54	C	DP=1
chrMT	4769	A	G	DP=52	G	DP=2
chrMT	7028	C	T	DP=58	T	DP=3
chrMT	8860	A	G	DP=68	G	DP=2
chrMT	10398	A	G	DP=61	G	DP=4
chrMT	11251	A	G	DP=58	G	DP=3
chrMT	11719	G	A	DP=52	A	DP=6
chrMT	12612	A	G	DP=60	ND	ND
chrMT	13708	G	A	DP=54	A	DP=2
chrMT	14766	C	T	DP=69	T	DP=4
chrMT	14798	T	C	DP=59	C	DP=6
chrMT	15326	A	G	DP=75	G	DP=3
chrMT	15452	C	A	DP=67	A	DP=5
chrMT	16069	C	T	DP=65	T	DP=8
chrMT	16126	T	C	DP=67	C	DP=3
chrMT	16264	C	T	DP=65	T	DP=8

*There is a homopolymeric C stretch at this position of the mitochondrial genome which is difficult to analyse and align with short aDNA reads.

Table S4: Taxonomic classification and quality estimations of the assembled microbial genome as inferred by GTDB-tk and CheckM, respectively, Related to Figure 2E.

MAG_ID	Methods	Description/Result
Bone_1	GTDB taxonomy	d_Bacteria;p_Actinobacteriota;c_Actinomycetia;o_Streptosporangiales;f_Streptosporangiaceae;g_Streptosporangium;s_
	fastANI_taxonomy	d_Bacteria;p_Actinobacteriota;c_Actinomycetia;o_Streptosporangiales;f_Streptosporangiaceae;g_Streptosporangium;s_Streptosporangium violaceochromogenes
	pplacer_taxonomy	d_Bacteria;p_Actinobacteriota;c_Actinomycetia;o_Streptosporangiales;f_Streptosporangiaceae;g_Streptosporangium;s_
	taxonomy method	taxonomic classification defined by topology and ANI
	CheckM_Completeness (%)	92.59
	CheckM_Contamination (%)	1.74
Bone_2	GTDB taxonomy	d_Archaea;p_Thermoproteota;c_Nitrososphaeria;o_Nitrososphaerales;f_Nitrosopumilaceae;g_UBA8516;s_
	fastANI_taxonomy	d_Archaea;p_Thermoproteota;c_Nitrososphaeria;o_Nitrososphaerales;f_Nitrosopumilaceae;g_UBA8516;s_UBA8516 sp8516u
	pplacer_taxonomy	d_Archaea;p_Thermoproteota;c_Nitrososphaeria;o_Nitrososphaerales;f_Nitrosopumilaceae;g_UBA8516;s_UBA8516 sp8516u
	taxonomy method	taxonomic classification defined by topology and ANI
	CheckM_Completeness (%)	55.05
	CheckM_Contamination (%)	0.97
Bone_3	GTDB taxonomy	d_Bacteria;p_Acidobacteriota;c_Acidobacteriae;o_UBA7541;f_UBA7541;g_UBA7541;s_UBA7541
	fastANI_taxonomy	N/A
	pplacer_taxonomy	d_Bacteria;p_Acidobacteriota;c_Acidobacteriae;o_UBA7541;f_UBA7541;g_UBA7541;s_UBA7541
	taxonomy method	taxonomic classification fully defined by topology
	CheckM_Completeness (%)	59.95
	CheckM_Contamination (%)	6.57
Stone_1	GTDB taxonomy	d_Bacteria;p_Proteobacteria;c_Gammaproteobacteria;o_Xanthomonadales;f_Xanthomonadaceae;g_Stenotrophomonas;s_Stenotrophomonas lactitubi
	fastANI_taxonomy	d_Bacteria;p_Proteobacteria;c_Gammaproteobacteria;o_Xanthomonadales;f_Xanthomonadaceae;g_Stenotrophomonas;s_Stenotrophomonas lactitubi
	pplacer_taxonomy	d_Bacteria;p_Proteobacteria;c_Gammaproteobacteria;o_Xanthomonadales;f_Xanthomonadaceae;g_Stenotrophomonas;s_
	taxonomy method	taxonomic classification defined by topology and ANI
	CheckM_Completeness (%)	96.75
	CheckM_Contamination (%)	1.46
Stone_2	GTDB taxonomy	d_Bacteria;p_Firmicutes_A;c_Clostridia;o_Clostridiales;f_Clostridiaceae;g_Clostridium;s_
	fastANI_taxonomy	d_Bacteria;p_Firmicutes_A;c_Clostridia;o_Clostridiales;f_Clostridiaceae;g_Clostridium;s_Clostridium botulinum_B
	pplacer_taxonomy	d_Bacteria;p_Firmicutes_A;c_Clostridia;o_Clostridiales;f_Clostridiaceae;g_Clostridium;s_
	taxonomy method	taxonomic classification defined by topology and ANI
	CheckM_Completeness (%)	51.9
	CheckM_Contamination (%)	5.25
Stone_3	GTDB taxonomy	d_Archaea;p_Thermoproteota;c_Nitrososphaeria;o_Nitrososphaerales;f_Nitrosopumilaceae;g_Nitrosarchaeum;s_
	fastANI_taxonomy	d_Archaea;p_Thermoproteota;c_Nitrososphaeria;o_Nitrososphaerales;f_Nitrosopumilaceae;g_Nitrosarchaeum;s_Nitrosarchaeum sp004297665
	pplacer_taxonomy	d_Archaea;p_Thermoproteota;c_Nitrososphaeria;o_Nitrososphaerales;f_Nitrosopumilaceae;g_Nitrosarchaeum;s_
	taxonomy method	taxonomic classification defined by topology and ANI
	CheckM_Completeness (%)	68.46
	CheckM_Contamination (%)	2.02
Stone_4	GTDB taxonomy	d_Bacteria;p_Bacteroidota;c_Bacteroidia;o_Sphingobacteriales;f_Sphingobacteriaceae;g_Pseudosphingobacterium;s_
	fastANI_taxonomy	d_Bacteria;p_Bacteroidota;c_Bacteroidia;o_Sphingobacteriales;f_Sphingobacteriaceae;g_Pseudosphingobacterium;s_Pseudosphingobacterium domesticum
	pplacer_taxonomy	d_Bacteria;p_Bacteroidota;c_Bacteroidia;o_Sphingobacteriales;f_Sphingobacteriaceae;g_Pseudosphingobacterium;s_
	taxonomy method	taxonomic classification defined by topology and ANI
	CheckM_Completeness (%)	100
	CheckM_Contamination (%)	0.48

Supplementary references

Chaumeil, P.A., Mussig, A.J., Hugenholtz, P., and Parks, D.H. (2019). GTDB-Tk: a toolkit to classify genomes with the Genome Taxonomy Database. *Bioinformatics*. 10.1093/bioinformatics/btz848.

Jonsson, H., Ginolhac, A., Schubert, M., Johnson, P.L., and Orlando, L. (2013). mapDamage2.0: fast approximate Bayesian estimates of ancient DNA damage parameters. *Bioinformatics* 29, 1682-1684. 10.1093/bioinformatics/btt193.

Mittnik, A., Massy, K., Knipper, C., Wittenborn, F., Friedrich, R., Pfrengle, S., Burri, M., Carlich-Witjes, N., Deeg, H., Furtwangler, A., et al. (2019). Kinship-based social inequality in Bronze Age Europe. *Science* 366, 731-734. 10.1126/science.aax6219.

Neukamm, J., Peltzer, A., and Nieselt, K. (2021). DamageProfiler: Fast damage pattern calculation for ancient DNA. *Bioinformatics*. 10.1093/bioinformatics/btab190.

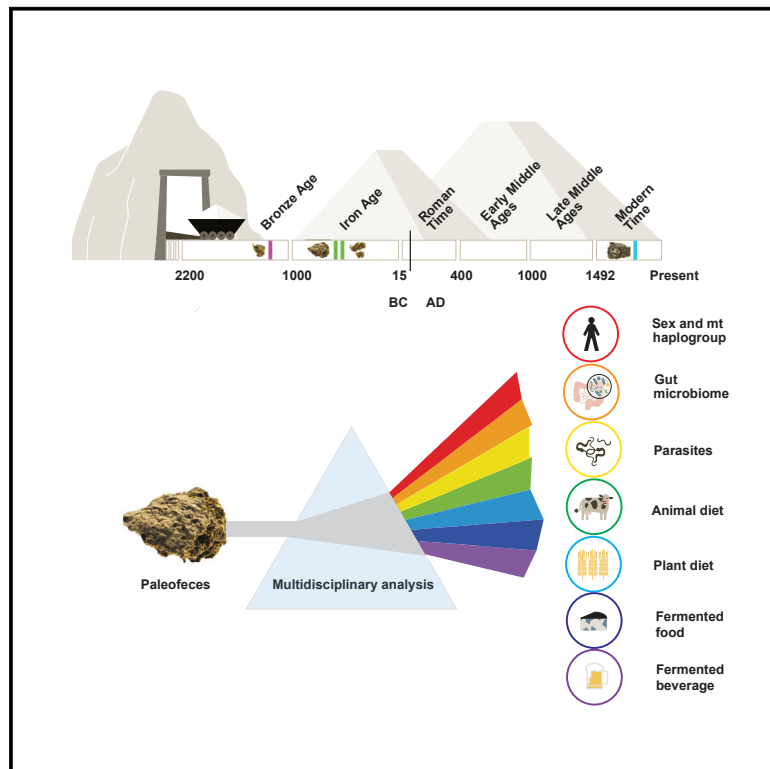
**Paper II: Hallstatt miners consumed blue cheese and beer during
the Iron Age and retained a non-Westernized gut microbiome
until the Baroque period**

(This page intentionally left blank)

Current Biology

Hallstatt miners consumed blue cheese and beer during the Iron Age and retained a non-Westernized gut microbiome until the Baroque period

Graphical abstract



Authors

Frank Maixner, Mohamed S. Sarhan, Kun D. Huang, ..., Albert Zink, Hans Reschreiter, Kerstin Kowarik

Correspondence

frank.maixner@eurac.edu (F.M.), kerstin.kowarik@nhm-wien.ac.at (K.K.)

In brief

Maixner et al. describe the gut microbiome and diet of European salt miners using paleofeces dating from the Bronze Age to the Baroque period. This analysis provides evidence for recent changes in the gut microbiome due to industrialization and for the consumption of fermented food and beverages in Iron Age Europe.

Highlights

- Gut microbiome and diet of European salt miners determined using paleofeces
- Until the Baroque, the microbiome resembled that of modern non-Westernized people
- Food-fermenting fungi in Iron Age feces indicates blue cheese and beer consumption



Article

Hallstatt miners consumed blue cheese and beer during the Iron Age and retained a non-Westernized gut microbiome until the Baroque period

Frank Maixner,^{1,12,13,16,17,*} Mohamed S. Sarhan,^{1,12,16} Kun D. Huang,^{2,3,16} Adrian Tett,^{2,4} Alexander Schoenafinger,^{1,5,12} Stefania Zingale,^{1,12} Aitor Blanco-Míguez,² Paolo Manghi,² Jan Cemper-Kiesslich,⁶ Wilfried Rosendahl,^{7,8} Ulrike Kusebauch,⁹ Seamus R. Morrone,⁹ Michael R. Hoopmann,⁹ Omar Rota-Stabelli,¹⁰ Thomas Rattej,⁴ Robert L. Moritz,⁹ Klaus Oeggl,⁵ Nicola Segata,^{2,14,16} Albert Zink,^{1,16} Hans Reschreiter,^{11,16} and Kerstin Kowarik^{11,15,16,*}

¹Institute for Mummy Studies, EURAC Research, Viale Druso 1, 39100 Bolzano, Italy

²Department CIBIO, University of Trento, Via Sommarive 9, 38123 Povo (Trento), Italy

³Department of Sustainable Agro-Ecosystems and Bioresources, Fondazione Edmund Mach, Via Edmund Mach 1, 38010 San Michele all'Adige (TN), Italy

⁴CUBE (Division of Computational Systems Biology), Centre for Microbiology and Environmental Systems Science, University of Vienna, Althanstraße 14, 1090 Vienna, Austria

⁵Institute of Botany, University of Innsbruck, Sternwartestraße 15, 6020 Innsbruck, Austria

⁶Interfaculty Department of Legal Medicine & Department of Classics, University of Salzburg, Ignaz-Harrer-Straße 79, 5020 Salzburg, Austria

⁷Reiss-Engelhorn-Museen, Zeughaus C5, 68159 Mannheim, Germany

⁸Curt-Egelhorn-Zentrum Archäometrie, D6,3, 61859 Mannheim, Germany

⁹Institute for Systems Biology, 401 Terry Avenue North, Seattle, WA 98109, USA

¹⁰Center Agriculture Food Environment (C3A), University of Trento, 38010 San Michele all'Adige (TN), Italy

¹¹Prehistoric Department, Museum of Natural History Vienna, Burgring 7, 1010 Vienna, Austria

¹²Twitter: @EuracMummy

¹³Twitter: @FrankMaixner

¹⁴Twitter: @cibiocm

¹⁵Twitter: @KowarikKerstin

¹⁶These authors contributed equally

¹⁷Lead contact

*Correspondence: frank.maixner@eurac.edu (F.M.), kerstin.kowarik@nhm-wien.ac.at (K.K.)

<https://doi.org/10.1016/j.cub.2021.09.031>

SUMMARY

We subjected human paleofeces dating from the Bronze Age to the Baroque period (18th century AD) to in-depth microscopic, metagenomic, and proteomic analyses. The paleofeces were preserved in the underground salt mines of the UNESCO World Heritage site of Hallstatt in Austria. This allowed us to reconstruct the diet of the former population and gain insights into their ancient gut microbiome composition. Our dietary survey identified bran and glumes of different cereals as some of the most prevalent plant fragments. This highly fibrous, carbohydrate-rich diet was supplemented with proteins from broad beans and occasionally with fruits, nuts, or animal food products. Due to these traditional dietary habits, all ancient miners up to the Baroque period have gut microbiome structures akin to modern non-Westernized individuals whose diets are also mainly composed of unprocessed foods and fresh fruits and vegetables. This may indicate a shift in the gut community composition of modern Westernized populations due to quite recent dietary and lifestyle changes. When we extended our microbial survey to fungi present in the paleofeces, in one of the Iron Age samples, we observed a high abundance of *Penicillium roqueforti* and *Saccharomyces cerevisiae* DNA. Genome-wide analysis indicates that both fungi were involved in food fermentation and provides the first molecular evidence for blue cheese and beer consumption in Iron Age Europe.

INTRODUCTION

Paleofeces are naturally preserved ancient feces found in dry caves, desert areas, waterlogged environments, and frozen habitats. Specific environmental processes such as desiccation or freezing prevent their deterioration in mummies, ancient latrines, bogs, and soils.¹ Previous studies have shown that paleofecal material still contains plant macro- and microfossils, parasite

eggs, and even ancient biomolecules (DNA, proteins, metabolites).² Ancient paleofeces have therefore recently been used as a source of information to study prehistoric nutrition patterns^{3–5} and health^{6,7} and to analyze single representatives^{8,9} or the overall composition of the intestinal microbiome of our ancestors.^{10–12}

One of the few archaeological sites where well-preserved paleofeces can be found is the protohistoric salt mines of the



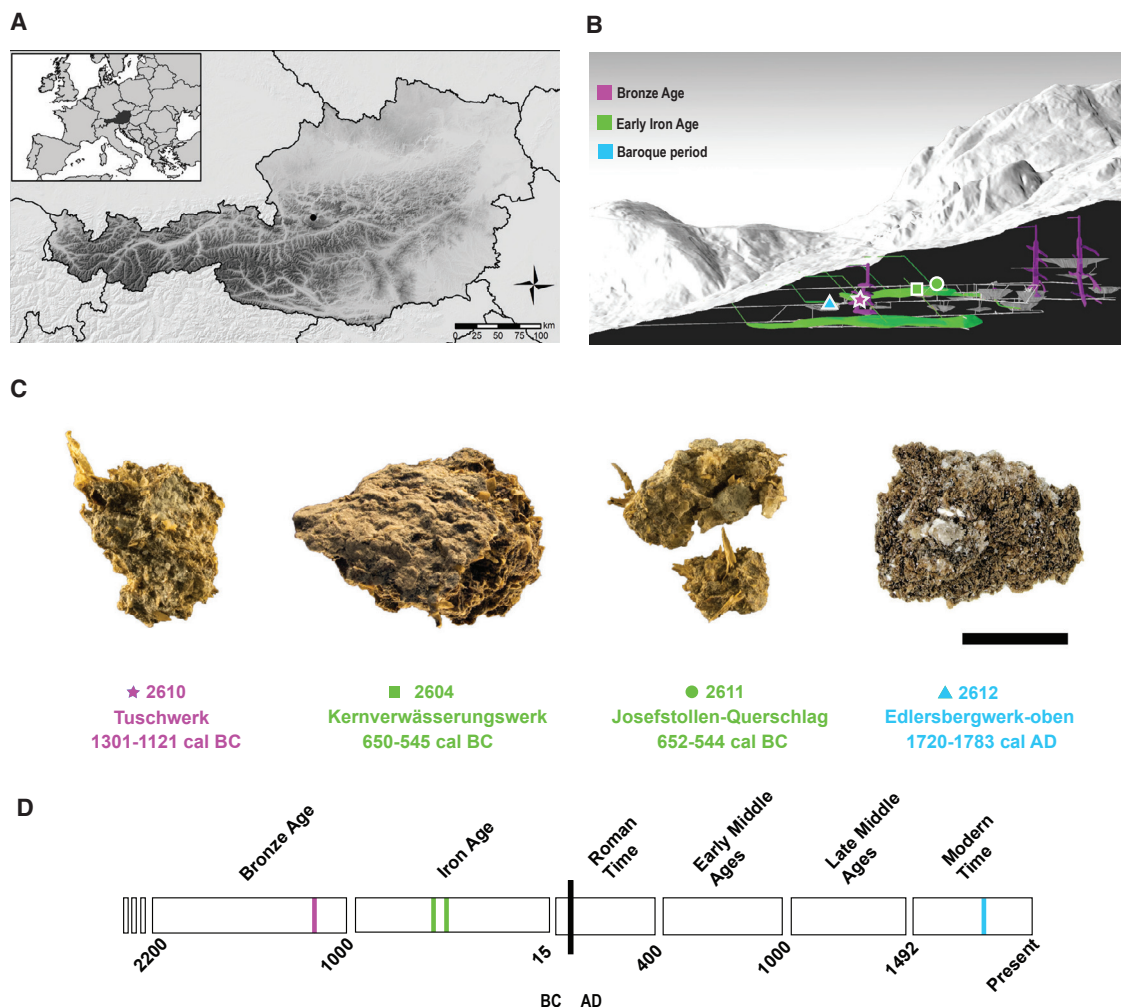


Figure 1. The Hallstatt salt mine and radiocarbon-dated paleofeces samples used in this study

(A) The salt mines are located in Upper Austria.

(B) Finding sites of the four paleofeces samples in the Bronze Age, Iron Age, and Baroque mining area. The symbol color corresponds to the radiocarbon date of the paleofeces.

(C) Macroscopic appearance of the four paleofeces samples. The scale bar corresponds to 1 cm of length. The sample description includes the sample ID, the mine workings name, and the radiocarbon date. The provided radiocarbon date range corresponds to the Cal 2-sigma values with the highest probability.

(D) Temporal assignment of the radiocarbon-dated paleofeces to the major European time periods from the Bronze Age onward.

See also [Figure S1](#) for details of the salt crystals surrounding sample 2612. [Data S1A](#) and [S1B](#) provide additional information about the samples and the radiocarbon dating.

Austrian UNESCO World Heritage area Hallstatt-Dachstein/Salzkammergut. Protohistoric salt mines offer ideal preservation conditions for organic materials. The high salt concentrations and the constant annual temperature of around 8°C inside the isolated mine workings preserve organic artifacts very well. The Hallstatt salt mines located in the Eastern Alps ([Figure 1](#)) offer one of the world's oldest and most continuous record of underground salt mining.^{13–15} Large-scale underground mining in the Hallstatt salt mountains dates back at least to the 14th century BC (late Bronze Age). Several protohistoric (Bronze, Iron Age) and historic (14th century AD to present) mining phases are well documented. The site also gave name to the early period of the Iron Age in Europe, the so-called Hallstatt Period (800 to 400 BC). Dense layers of production waste reaching several

meters of thickness were excavated from the protohistoric Bronze Age and Iron Age mine workings of Hallstatt, uncovering thousands of wooden tools and construction elements, implements made from fur, rawhide, hundreds of woolen textile fragments, grass, bast ropes, and human excrements.¹⁶ These objects provide insights into the daily life of a Bronze Age and Iron Age mining community ranging from mining technology, organization of production, and resource management to human health, dietary habits, social organization of production processes, and social status within a mining system. These aspects have been studied extensively in Hallstatt based on a combination of data sources encompassing the protohistoric mine working, Bronze Age meat-curing facilities, and large Iron Age cemetery.^{14,16,17}

Here we focus on the question of the structure and evolution of dietary habits as well as the human gut microbiome in one of Europe's most important early production communities. We used microscopic, metagenomic, and proteomic analysis to characterize nutrition patterns of the protohistoric (Bronze Age, early Iron Age) and historic (Baroque period, 18th century AD) miners and metagenomic analysis to determine the structure and evolution of the gut microbiome. Our findings will enhance the understanding of early European dietary habits (especially the production and consumption of processed foodstuffs) and provide further evidence of the recent change in gut microbiome structure as result of industrialization and Westernization processes.

RESULTS

Paleofeces from Bronze Age to Baroque Period contain ancient endogenous DNA

In this study, we initially subjected four paleofeces samples, collected from Bronze Age and Iron Age Hallstatt mine workings, to radiocarbon dating, then to in-depth microscopic and molecular analysis (Figures 1A–1C; STAR Methods; Data S1A). The four paleofeces samples can be macroscopically differentiated into three samples containing a high amount of fibrous plant material (2610, 2604, 2611) and one more homogeneous sample (2612) that does not contain any visible larger plant fragments (Figures 1 and S1A). Radiocarbon analyses date the roughly structured samples to the Late Bronze Age (2610) and Iron Age (2604, 2611), which is in perfect accordance with the proposed period of usage of the mine workings where the paleofeces have been found.¹⁵ In contrast, the fine-textured paleofeces 2612 sampled in an Iron Age minedates to the Baroque period (18th century AD) (Figure 1C; Data S1B). For this part of the salt mine, however, it is historically documented that the mine workings had started to be reused from the beginning of 18th century onward.¹⁸ Independently of the paleofeces' age, their storage time since excavation (some samples were recovered in the year 1983), or the mode of excavation (wet sieving versus direct sampling) (Figure S1), we could retrieve biomolecules (DNA and protein) from all samples for the subsequent molecular analysis (Data S1 and S2; STAR Methods). Proteomics analysis provided the first evidence for the presence of endogenous biomolecules in the paleofeces material. The most abundant peptides were assigned to human intestinal tract proteins that are involved in food digestive processes (Data S2B, S2D, S2F, and S2H). The DNA of the paleofeces material was further subjected to a deep shotgun sequencing approach resulting in 57,130,584 to 221,314,691 quality-filtered reads (Data S1C). A first taxonomic overview using DIAMOND against the NCBI NR database revealed that the majority of reads in the samples are assigned to *Bacteria* (93.9% to 78.9% of all assigned reads), with *Firmicutes* and *Bacteroidetes* being the most abundant phyla of this kingdom (Figures S2A–S2D). Less than 7.5% of the reads were eukaryotic, with up to 6.7% fungal reads in sample 2604. The *Metazoa* and *Viridiplantae* reads, important for the molecular reconstruction of the diet, comprised 0.5% to 0.01% of all assigned reads. Further analysis of the human DNA in the paleofeces revealed an endogenous DNA content between 0.26% and 0.06%, sufficient for molecular sex and mitochondrial haplogroup assignment

(Data S1D). The highly fragmented human reads display a very low deamination pattern at the 5' ends (Figures S2E–S2H). In the most recent sample (2612), the reads appear even less fragmented and display almost no DNA damage. Considering the age of these samples, the DNA damage is exceptionally low. This high preservation is most likely due to the rapid desiccation of the samples in the salt mine, which may result in reduced hydrolytic damage of the biomolecules. Our analyses show that the four paleofeces come from male individuals that carry distinct mitogenomes with low contamination estimates (1% to 2%), indicating that each sample represents unique unmixed ancient feces.

Ancient paleofeces display a gut microbiome structure similar to modern non-Westernized individuals

We compared the microbiome structure of the paleofeces to a large number of contemporary metagenomes ($n = 823$) (Data S1E). Principal coordinate analysis (PCoA) performed on a species-level taxonomic composition shows that the paleofeces from the Bronze Age to the Baroque period cluster with stool samples from contemporary non-Westernized individuals (Figure 2A) with diets mainly consisting of unprocessed foods and fresh fruits and vegetables.¹⁹ This clustering is similarly observed for encoded metabolic pathways (Figure 2B). All the paleofecal samples were distinct from the oral and, more importantly, from the soil samples, suggesting little evidence of soil contamination, which is sometimes observed in ancient metagenomics studies.²⁰ The source prediction analysis further supports the sample preservation (Data S1F).

To further assess the paleofeces samples, we analyzed the prevalence of the top 15 most abundant species in the paleofeces compared to 8,968 gut microbiomes of healthy Westernized and non-Westernized adults (Data S1G). As a result, 13 out of the 15 most abundant species were identified to be associated with human gut environment, of which 11 species were found to be more prevalent in modern non-Westernized compared to Westernized populations. Five of these species, *Bifidobacterium angulatum*, *Lactobacillus ruminis*, *Catenibacterium mitsuokai*, *Prevotella copri*, and *Clostridium ventriculi*, were over twice as prevalent in non-Westernized populations (Figure 2C; Data S1H). One of the two species not associated with the human gut is the halophilic archaeon *Halococcus morrhuae*, which survives on a high concentration of salt.²² It was observed in low abundance in the paleofeces sample 2612, the only sample that was not subjected to wet sieving. Therefore, we assume that the archaeon was introduced from the environment via the salt crystals (Figure S1B). In the paleofeces samples 2604 and 2611, we also identified *Clostridium perfringens*, a known intestinal foodborne pathogen,²³ that also occurs free living in the soil and other environments.²⁴ Since an infection with *C. perfringens* causes acute diarrhea and the paleofeces does not indicate any characteristics pointing to that disease, we assume that the presence of this bacterium is due to an environmental contaminant rather than a remnant of food spoilage in the miners' gut. When all paleofecal microbiome members were considered in population prevalence analysis, 100 out of 158 species were found in $\geq 5\%$ stool samples from modern healthy adult individuals and 65% of these species are overrepresented in non-Westernized populations in comparison with

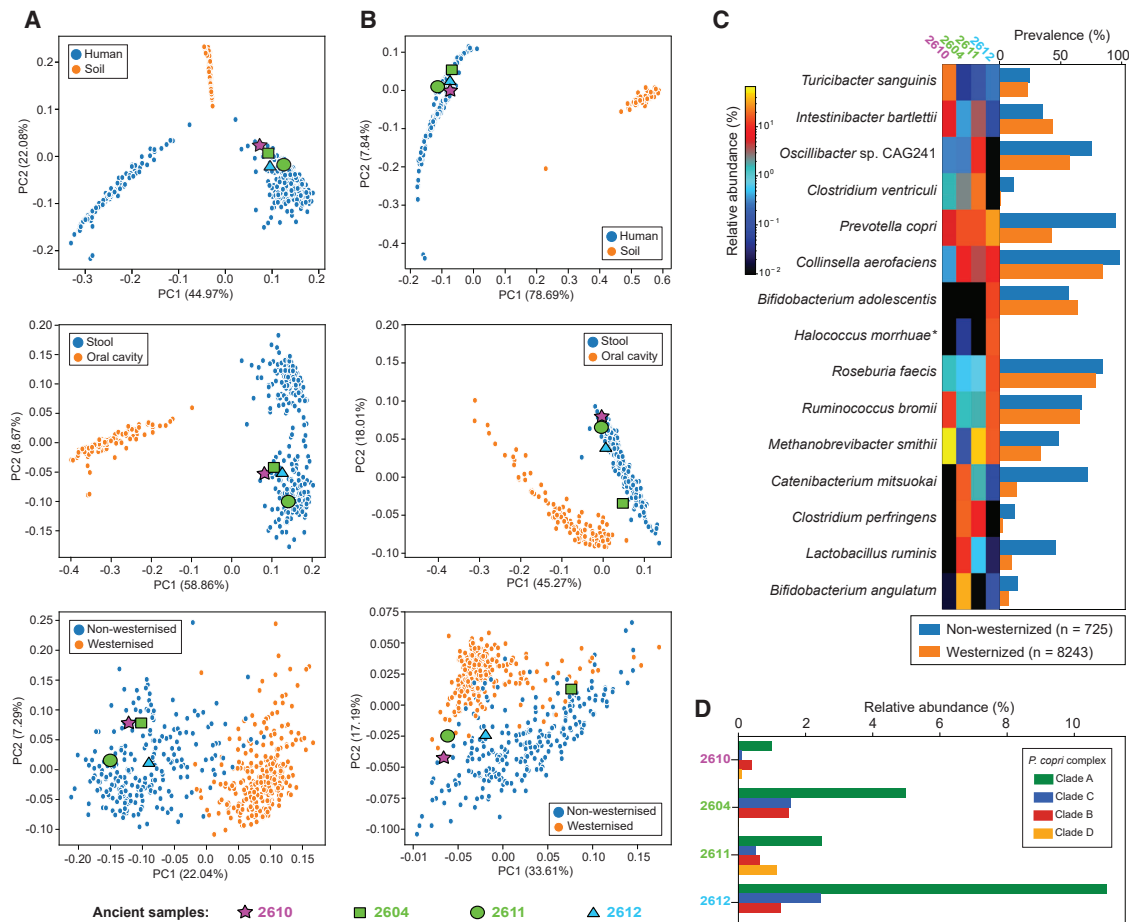


Figure 2. Overview of microbial composition and metabolic pathways of paleofeces samples in comparison to a large collection of contemporary metagenomes

(A) Principal coordinate analysis (PCoA) based on microbial abundance profiled using MetaPhlan 3.0²¹ between four paleofeces samples and 823 contemporary samples characterized by sampling environment, body site, and non-Westernized lifestyle.

(B) Principal coordinate analysis (PCoA) based on metabolic pathway abundance profiled using HUMAnN 3.0²¹ between four paleofeces samples and the same contemporary samples used in (A).

(C) Prevalence of the top 15 most enriched species of four paleofeces samples in non-Westernized and Westernized datasets comprising 8,968 stool samples from healthy adult individuals. Asterisk indicates species that is likely from external contamination.

(D) Relative abundance of *P. copri* four clades estimated using MetaPhlan 3.0²¹ in each paleofecal sample.

See also Figure S1 for additional microbial profiles in the DNA “wash-out” experiment. Data S1 contains additional information about the comparative datasets and the results obtained by the prevalence and abundance analysis.

the Westernized equivalent (Data S1I). A similar result was obtained when we decreased the prevalence threshold to 1% (Data S1J).

The *Prevotella copri* complex, which is highly prevalent in non-Westernized populations and prevalent in previously investigated ancient samples,⁹ was identified in all paleofecal samples, representing, on average, 7.3% (~1.6%–14.7%) of the relative abundance (Figure 2C; Data S1K). Consistent with previous findings,⁹ we found multiple clades of the complex to be present in each of the paleofeces samples (Figure 2D) with the exception of Clade D, which was barely detectable in sample 2604 and 2612. All other clades were detected with relative abundances > 0.01% in all samples (Figure 2D; Data S1L). Of note, in contrast to other samples, the sample 2604 displayed higher abundance of bacterial species such as *Lactobacillus brevis*,

Bifidobacterium merycicum, *Bifidobacterium angulatum*, and *Lactobacillus plantarum* (Data S1K) that are known to be of probiotic activities or involved in processing of dairy products.²⁵

Microscopic and molecular reconstruction of the Hallstatt miners’ diet

Next, we aimed to reconstruct the dietary components in the paleofeces using both a microscopic and a molecular survey. The above-mentioned structural differences between the paleofeces became even more evident in the microscopic analyses. The Baroque period sample 2612 was much finer textured than all other samples from protohistory (Figure 1C). This was also reflected in the macro-remain composition of the paleofeces, showing that samples 2610, 2604, and 2611 contained a lot of seeds contrary to 2612, which consisted of frequent tissues of

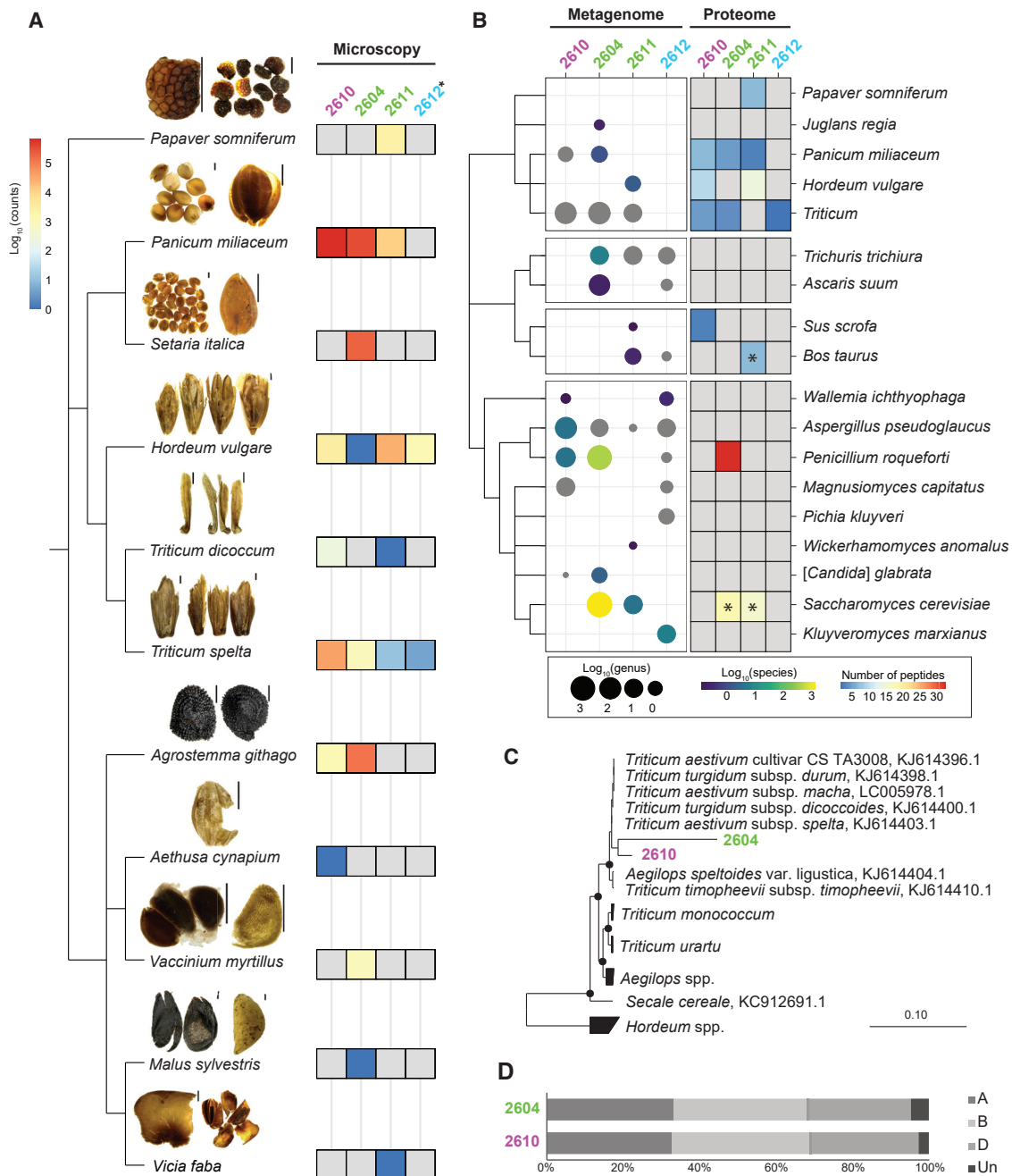


Figure 3. Microscopic and molecular dietary analysis of the Hallstatt paleofeces

(A) Plant macro-remains microscopically detected in the four paleofeces samples. The scale bar indicates 1 mm of length. The heatmap shows the log-scale macro-remain counts normalized to 3.7 g sample. The sample with asterisk was assessed in a semiquantitative manner. For further details, please refer to [Data S1](#).

(B) Most abundant taxa (plants, nematodes, animals, fungi) detected in the four paleofeces metagenomes and proteomes. The circle size and circle color correspond to log₁₀ “normalized” number of reads per million at genus and species levels, respectively. The asterisks in the proteome heatmap mean the peptides were assigned only to genus level.

(C) Phylogenetic assignment of two partial *Triticum* chloroplast genomes in the 2604 and 2610 metagenomes. The comparative dataset included complete chloroplast genomes of selected members of the *Triteae* tribe (NCBI accession numbers are provided in the figure). The tree was calculated using the maximum-likelihood algorithm (PhyML) based on 136,160 informative positions. Black circles symbolize parsimony and neighbor joining bootstrap support (>90%) based on 100 and 1,000 iterations, respectively. The scale bar indicates 10% estimated sequence divergence.

(legend continued on next page)

fruit husks and seed coats (Figure 3A; Data S1M). Generally, all samples displayed a predominance of cereal remains.

Microscopic analysis revealed that the Bronze Age sample (2610) consisted more or less exclusively of cereal remains, which originated from barley (*Hordeum vulgare*), spelt (*Triticum spelta*), some emmer (*Triticum dicoccum*), proso millet (*Panicum miliaceum*), and a few weeds, e.g., corn cockle (*Agrostemma githago*) and poison parsley (*Aethusa cynapium*). The Iron Age samples (2604, 2611) were characterized by a predominance of cereal remains from barley (*Hordeum vulgare*), spelt (*Triticum spelta*), millets (*P. miliaceum*, *Setaria italica*), and a little emmer (*T. dicoccum*). Furthermore, in sample 2611, testa remains of broad beans (*Vicia faba*) and seeds of opium poppies (*Papaver somniferum*) were observed. Crab apples (*Malus sylvestris*) and bilberries/cranberries (*Vaccinium myrtillus/vitis-idea*) in sample 2604 document the consumption of gathered wild fruits. Striking in the Iron Age sample 2604 was the contamination with weeds, in particular corn cockle (*Agrostemma githago*). In the sample 2612 from the Baroque time, the microscopic pattern was notably different from the other samples. The plant material was finely ground, and entire fruits were missing apart from a digested mericarp of anise (*Pimpinella anisum*). The plant tissue belonged to bran (fragments of cereal testa, pericarp, hairs, hilum, endosperm) of wheat (*Triticum* sp.). The precise species was unidentifiable, but according to the rare occurrence of tube cells in the pericarp fragments, a member of the tetra- or hexaploid wheat group is suggested. Bran of barley (*H. vulgare*) was also observed in minor quantities. Furthermore, the consumption of legumes is documented by testa remains of garden bean (*Phaseolus vulgare*) in this sample.

In addition to the microscopic analysis, we subjected paleofeces biomolecules (DNA and proteins) to molecular dietary analyses. Both metagenomic and proteomic analyses included a homology search against different databases, followed by strict filtering steps of the obtained hits and a subsequent in-depth analysis of selected identified taxa (STAR Methods; Figure S3; Data S1N, S1O, S2B, S2D, S2F, and S2H). For the plant diet, we could confirm the presence of the most abundant domesticated plant macro-remains, including broomcorn millet (*P. miliaceum*), barley (*H. vulgare*), and wheat (*Triticum* spp.) (Figure 3B). In addition, we found DNA-based evidence of the presence of walnut (*Juglans regia*) in the sample 2604 and protein-based support for the occurrence of opium poppy seeds (*P. somniferum*) in the sample 2611. In addition to the foxtail millet (*S. italica*), which appeared with high grain number, all the low abundant wild plants unveiled by the microscopic investigation were not identified in our molecular survey, which has undergone strict filtering to minimize the false positives (STAR Methods). Further phylogenetic analysis assigned the *Triticum* spp. chloroplast genomes of the samples 2604 and 2610 closest to the chloroplasts of tetraploid (emmer, durum) and hexaploid (spelt wheat, bread wheat) wheat varieties, respectively (Figures 3C,

S3A, and S3B; Data S1P). Additional comparison with the bread wheat genome revealed an equal subgenome (A, B, and D) representation in the 2604 and 2610 metagenomic reads, which suggests, in combination with the microscopic identification of numerous characteristic grains, glumes, and spikelets, the presence of hexaploid spelt wheat (*T. spelta*) in these paleofeces (Figure 3D). Beside the plant diet, we obtained molecular evidence for the consumption of cattle (*Bos taurus*) and swine meats (*Sus scrofa*) throughout all investigated time periods (Figures 3B and S3D). Interestingly, the most abundant cattle proteins in sample 2611 (hemoglobin and coagulation proteins) indicate the plant diet was supplemented by blood-rich animal tissues (e.g., muscle, liver) (Data S2F). The molecular analyses revealed in addition, that individuals from both the Iron Age (2611) and the Baroque (2612) suffered from intestinal infections of whipworms (*Trichuris trichuria*) and roundworm (*Ascaris* spp.) (Figure 3B and S3D–S3G). Finally, all samples showed a continuous low background with fungal DNA mainly coming from different Ascomycota.

Molecular evidence for blue cheese and beer consumption during the Iron Age

In contrast to all other samples, the Iron Age sample 2604 displayed an exceptionally high abundance of *Penicillium roqueforti* and *Saccharomyces cerevisiae* proteins (Data S2D) and DNA (Data S1N), making up to 7%–22% of total eukaryotic reads. This was characteristic of this sample as compared with the other samples that did not show such prevalence—even the sample 2611, which was taken from a similar context and dated back to the same time point. To authenticate the data and gain further insights into their potential ecological significance, we mapped the high-quality reads of sample 2604 against the reference genomes of these two fungi (Figure S4A; Data S1O). With 11–13× coverage, we were able to reconstruct >92% of both genomes, displaying even coverage and SNP distribution. To confirm whether these two fungi are of ancient origin and not modern contaminants, we initially checked the ancient DNA damage pattern of the mapped reads. Both fungi displayed typical ancient DNA damage patterns, with levels comparable to the human endogenous DNA (Figures S4B and S4C). Hence, and considering their extraordinarily high abundance and exclusive incidence in this sample, we assumed their endogenous originality to the coprolite microbial community. Additionally, both fungi are commonly used nowadays in food processing: *P. roqueforti* is used for cheese fermentation, and *S. cerevisiae* is used for fermenting bread and alcoholic beverages including beer, mead, and wine. Therefore, we assume that they could have been involved in food processing at that time. To test this assumption, we used the reconstructed genomes for further comparative phylogeny and population genetic analyses to infer whether they had been truly involved in food processing or were just transient environmental microbes.

(D) Wheat subgenome (A, B, and D) representation in the 2604 and 2610 metagenomes (Data S1), aligned to the modern hexaploid bread wheat reference genome (accession number GCA_900519105). Both the wheat chloroplast and nuclear reads were highly fragmented and display aDNA-specific damage patterns (Figures S3H–S3K).

See also Figure S3 for details about the comparative analysis, phylogenetic assignment, and damage pattern of selected plant, animal, and parasite DNA. Data S1 provides further details of the macro-remains, comparative datasets, dietary DNA, and mapping statistics. Data S2 provides additional information about the comparative datasets and proteomics results.

First, we compared our putative *P. roqueforti* strain to 33 other sequenced modern *P. roqueforti* strains coming from different functional niches.²⁶ The comparative dataset included 18 cheese-fermenting and 15 non-cheese-fermenting strains (Data S1Q), in addition to *Penicillium psychrosexualis* and *Penicillium carneum* as an outgroup. After mapping the raw reads of all strains to the reference *P. roqueforti* genome FM164 and data filtering, we resolved 120,337 SNPs, which were used for inferring maximum likelihood (ML) phylogenetic relationships among the tested strains (STAR Methods). Consistent with the original publication of Dumas and colleagues,²⁶ the resulting phylogeny revealed four distinct clades: a Roquefort cheese clade, a non-Roquefort cheese clade containing blue cheeses others than Roquefort, a silage/food spoilage clade, and a wood/food spoilage clade (Figure 4A). Initially, the phylogenetic analysis separated the non-Roquefort cheese clade from the other, then the Roquefort cheese clade was diverged from the other food spoilage clades. The ancient *P. roqueforti* strain showed highest similarity to the non-Roquefort cheese strains, being clustered together with their corresponding clade as an earlier divergent. The reason behind such early divergence might be attributed to the recent acquisition of some genomic regions—most importantly, *CheesyTer* and *Wallaby*—by the non-Roquefort cheese strains. This gene acquisition most likely happened via repeated multiplication of selected spores of the best cheeses on bread used as a growth medium in the late 19th century and early 20th century before the advent of microbiological *in vitro* culturing techniques.²⁷ Thereby, modern non-Roquefort strains were exposed to extensive selection coupled to horizontal gene transfer events from other cheese-producing *Penicillium* spp. or even other genera.^{26,27–29} Importantly, our Iron Age strain did not contain any of those recently acquired fragments, which comes in congruence with the hypothesis that such domestication events occurred during the last two centuries.

We further used the tool ADMIXTURE to infer the degree of admixture among the strains. By assuming the presence of 3 ancestries ($K = 3$), we could clearly distinguish the non-Roquefort, Roquefort, and food spoilage strains (Figure 4B). Our putative strain displayed ~70% cheese-producing ancestry (60% of the non-Roquefort and 10% of Roquefort cheese) and ~30% food-spoiler ancestry. Both the phylogenetic placement and ADMIXTURE profile indicate that the ancient *P. roqueforti* has already been under positive selection toward the non-Roquefort cheese cluster, a selection process that most likely occurred during the process of cheese production. Some archeological findings excavated from the mines might have been used for that purpose (Figure 4C), as they showed some traces of fatty food products.

Next, we compared the ancient *S. cerevisiae* genome to 157 recent strains coming from different ecological niches, i.e., food, alcoholic beverages (e.g., beer, wine, sake, and spirits), biofuels, and laboratories, as well as wild strains (Data S1R). ML phylogenetic analysis, based on 375,629 SNP positions, distinguished 2 main clades. The first main clade splits into two subclades, with one containing most of the beer strains (beer 1 clade) that show a successive sub-clustering based on the origin of the strains (Figure 4D). The other subclade (henceforth referred to as “mixed” clade) included a mixture of bread, wine, beer, and spirit strains. The second main clade is composed of two subclades: a wine clade and another beer clade (beer 2). All other

wild, laboratory, and sake strains fall to the base of the whole phylogeny. The ancient *S. cerevisiae* strain clustered basal to the second main clade, which includes the wine and beer 2 strains. Further population structure analysis displayed high admixture in our putative strain, resembling primarily the wine ancestral population (47%), followed by 29.2% beer ancestries (Figure S4D) and only 19% wild strain ancestry. Therefore, and considering the ML phylogenetic assignment, we assume that the possibility of our strain to be of wild origin is unlikely. The results rather indicate higher similarity to wine and beer strains. Principal component analysis (PCA) provided further indication for the domestication of our strain in alcoholic beverage fermentation. Along the PC1 that explains 25.42% of the variation, our strain clustered closer to the strains of beer 2 than to the strains of the wine clade (Figure 4E). This was further supported with proteomic analysis (Data S2D) that unveiled that most of the peptides assigned to the genus *Saccharomyces* derived from proteins involved in alcohol fermentation pathways (e.g., glycolysis). To further narrow down the possible routes of domestication, we decided to differentiate the strains based on functional marker genes (Data S1S). According to recent literature,^{31,32,33} the genes *RTM1*, *BIO1/BIO6*, and the chromosomal regions *A/B/C* can be used to differentiate yeast strains based on their functional niches. The gene *RTM1* is a strong domestication marker responsible for conferring resistance against the toxicity of molasses and other rich-sugar substrates and is assumed to be positively selected in beer yeast strains.^{32,34} The genes *BIO1* and *BIO6*, which are involved in *de novo* biosynthesis of biotin, are highly selected in sake fermenting yeasts, due to lack of biotin in the fermentation substrates, such as rice.³⁵ The regions *A*, *B*, and *C* are horizontally transferred genomic regions from other yeast genera, e.g., *Kluyveromyces*, *Pichia*, and *Zygosaccharomyces*.³⁶ These regions contain 39 genes distributed over 3 different chromosomes and are assumed to play a role in wine fermentation.

Therefore, we searched the presence of these marker genes in our comparative dataset, including our ancient strain. In accordance with the literature, almost all beer strains—either of clade 1 or clade 2—were positive for *RTM1*, while the wine clade was mainly positive for the genomic regions *A/B/C*. The mixed clade contained both *RTM1* and the regions *A/B/C*. The sake clade exclusively contained the *BIO1/BIO6* genes and partially the *RTM1* gene. The wild strains, isolated from cacao in Africa, clustered in the basal clade and did not contain any of these marker genes (Figure 4D; Data S1T), contrary to our strain that contained the *RTM1* and lacked the *BIO1/BIO6* genes and the regions *A/B/C*.

Based on the previous findings—i.e., (1) the ancient DNA damage profile, (2) the high prevalence of *S. cerevisiae* reads, (3) the presence of fermentable cereal substrates such as wheat and barley, (4) the phylogenetic assignment of the ancient yeast strain, (5) the yeast admixture profile, and (6) the distribution of marker genes—we assume that this yeast is of ancient origin and has been involved in beer fermentation, although the mode of fermentation is unknown (i.e., bottom, top, or spontaneous fermentation).

DISCUSSION

Our interdisciplinary analyses of the samples have given detailed insight into the microbiome evolution and dietary habits and food

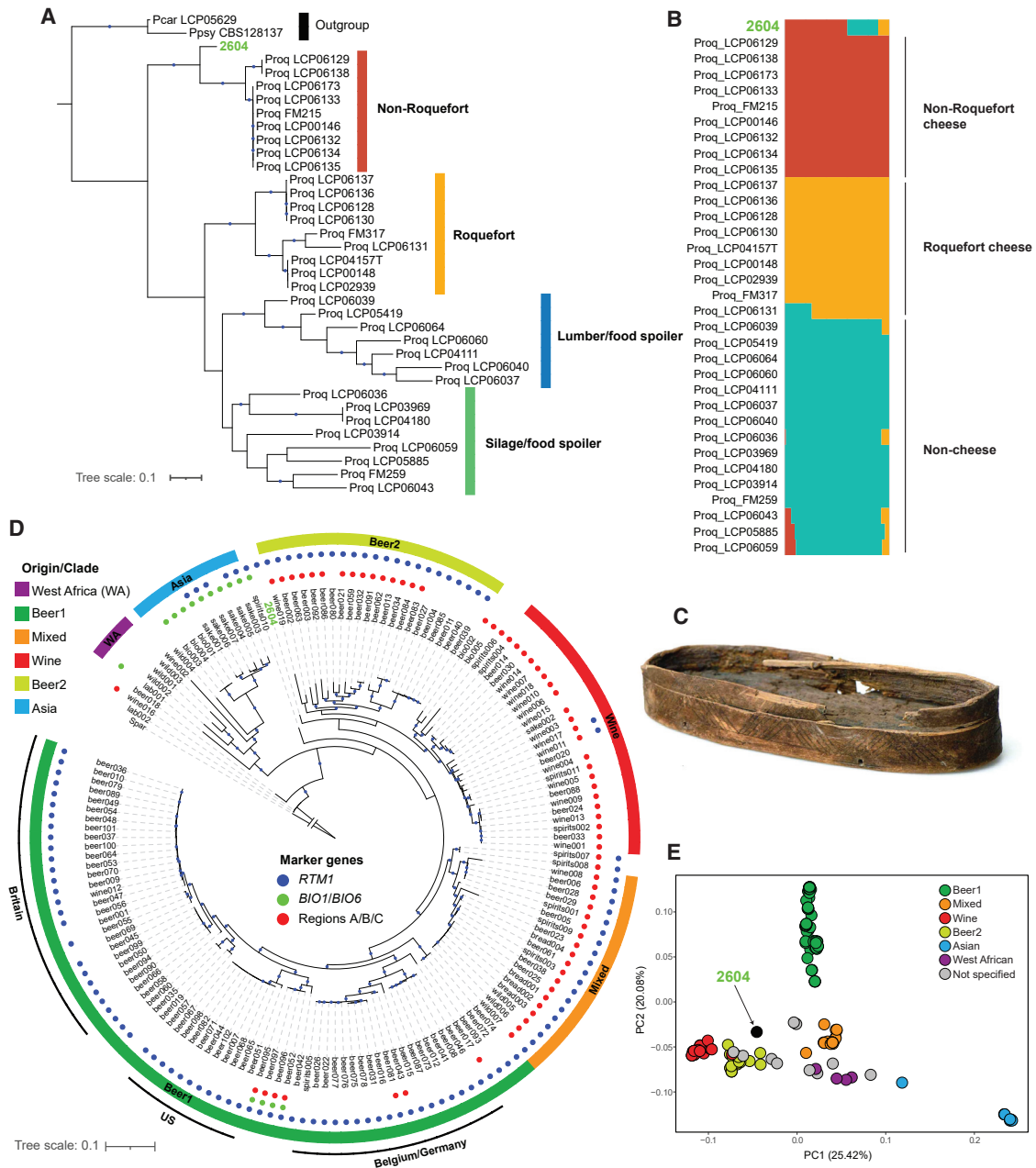


Figure 4. Genome-wide SNP analysis of ancient fungal “strains” versus modern industrial and wild/environmental strains

(A) Maximum likelihood (ML) phylogenetic analysis of the *Penicillium roqueforti* genome assembled from the sample 2604 in addition to other previously published *P. roqueforti* genomes.²⁶ A total number of 120,359 SNP positions were used for the analysis. *P. roqueforti* FM164 was used as a reference, while *P. carneum* and *P. psychrosexualis* were used as outgroups. The scale bar depicts 0.1 substitutions per residue. Colored strips indicate the *P. roqueforti* population as previously inferred.²⁶ For further information on the comparison dataset, please refer to [Data S1Q](#).

(B) Population structure analysis of *P. roqueforti* 2604 with the same previous dataset, considering 3 ancestries (K = 3 with lowest cross-validation error), based on 120,337 SNPs. The order of labels corresponds to the clustering in panel (A).

(C) Wooden containers that have been found among other archeological findings in the mines and assumed to be used as cheese strainers

(D) ML phylogenetic analysis of *Saccharomyces cerevisiae* genome assembled from the sample 2604 compared with other published *S. cerevisiae* genomes.³⁰ The dataset for the analysis included 375,629 SNPs. The *Saccharomyces paradoxus* CBS432 was used as an outgroup. The scale bar depicts 0.1 substitutions per residue. The colored strips indicate the clade/origin as reported previously.³⁰ The colored dots at the tree edges refer to the presence/absence of functional marker genes.³¹ Blue dots in (A) and (C) indicate bootstrap support >80% based on 1,000 bootstrap replicates.

(E) Principal component analysis based on 136,712 SNP, of the *S. cerevisiae* strains.

For additional information on the coverage and SNP density, DNA damage, and ADMIXTURE, please refer to [Figure S4](#). [Data S1](#) provides further details about the comparative datasets, mapping results, and functional marker analysis.

processing techniques of the Hallstatt miners over the past three millennia. Molecular and microscopic investigations revealed that the miner's diet was mainly composed of cereals, such as domesticated wheats (emmer and spelt), barley, common millets, and foxtail millets. This carbohydrate-rich diet was supplemented with proteins from broad beans and occasionally with fruits, nuts, or animal products. The food remains in the protohistoric sample with abundant entire fruits and seeds were less processed than those of the Baroque sample, which consisted of finely ground wheat. This suggests that the protohistoric miners consumed the cereals and legumes in a sort of gruel or porridge,³⁷ whereas miners in the 18th century AD ate their cereals in a more processed form, e.g., as a bread or biscuit.

In general, such carbohydrate-rich fibrous dietary components as observed in the Bronze Age and Iron Age samples are typical for traditional communities and are considered to be the main drivers of the non-Westernized microbiome structure.^{38,39} Consistent with this observation, our analysis showed that the Hallstatt paleofecal samples contain microbial features similar to gut microbiomes of modern non-Westernized populations (Figures 2A and 2B). Species identified in the samples, such as *Lactobacillus ruminis*, *Catenibacterium mitsuokai*, and *Prevotella copri*, were also found to be highly prevalent in present-day individuals with a more traditional lifestyle (Figure 2C). Furthermore, our paleofecal samples were rich in the *P. copri* complex (Figure 2D), including the four clades that are nearly ubiquitous and co-present in non-Westernized populations.⁹ Of particular interest, *P. copri* members have been shown to be associated with the digestion of complex carbohydrates,^{9,40–42} which are the major component of an unprocessed fibrous plant diet. Finding paleofeces highly resembling that of non-Westernized individuals, in terms of microbiome structure, supports previous observations.^{10,12} It also adds weight to the hypothesis that the modern industrialized human gut microbiome has diverged from an ancestral state, probably due to modern lifestyle, diet, or medical advances. Interestingly, this non-Westernized microbiome structure has been observed in all four paleofeces dating from the Bronze Age to the Baroque period, which would indicate quite a recent change in the gut community. However, to spot the critical time points when this shift in the human gut microbiome began requires more ancient samples spanning a wider time range; of particular interest would be samples from the past two or three centuries, when major dietary and medical changes occurred. Overall, our results support the theory that the shift from traditional to an industrial Westernized lifestyle might be the driving force for changing the human gut microbiome from its ancestral state.^{9,43–45}

In one of the Iron Age samples (2604), the molecular analyses indicated consumption of fermented food and beverages. The fungal analysis revealed a high prevalence of *Penicillium roqueforti* and *Saccharomyces cerevisiae*, which are nowadays involved in fermenting blue cheeses and alcoholic beverages, respectively, with clear signs of domestication. Following rapidly in the wake of ruminant animal domestication (mainly cattle, sheep, and goat), cheese production represents one of the oldest and widespread food preservation techniques developed by humans.⁴⁶ The oldest evidence for milk use dates back to the Neolithic in the Fertile Crescent yet provides only indirect evidence for fermentation.⁴⁷ The oldest reported chemical evidence for processing of milk into fermented products (i.e., kefir dairy) is

dated back to the Early Bronze Age in Western China.⁴⁸ Other indications, including actual preserved pieces of cheese, whey strainers, and recipes for cheese production, were found in Northern Europe, the Middle and Near East, and the Mediterranean basin.^{49–52} Here, we report evidence for the domestication of the fungus *Penicillium roqueforti* in the course of food processing in the 1st millennium BC that would likely produce a cheese resembling a blue cheese (non-Roquefort cheese clade, in Figure 4A). To our knowledge, this represents the earliest known evidence for directed cheese ripening and affinage in Europe, adding a crucial aspect to an emerging picture of highly sophisticated culinary traditions in European protohistory.³⁷ Importantly, the production of blue cheese today involves a surface application of dry salt; therefore, it is characterized by a high salt content of up to 7.5% (w/w).⁵³ The cheese curd could have been collected, desiccated, and inoculated with the fungi in wooden cheese containers like the ones excavated in the Hallstatt mines (Figure 4C). The presence of *P. roqueforti* indicates a major step in ruminant milk processing from fresh to ripened cheese, which could have offered, in addition to new flavors, several advantages to the Hallstatt miners including longer storage (i.e., months) and less lactose content in the fermented dairy product.⁵⁴ The reduced lactose content may have helped the ancient minors to better digest milk products, living in a time when lactose persistence frequencies only started to rise in Europe.⁵⁵ The presence of salt as well as the constant temperature (8°C) and humidity inside the Hallstatt mine workings represent ideal conditions for blue cheese production, following the current cheese production standards.⁵⁶ It is noteworthy that the early discovery of the Roquefort cheese was linked to Roquefort-sur-Soulzon caves in France, which maintain a temperature of 10°C and ~90% humidity over the year. With such conditions protecting the cheese from desiccation, these caves have been used exclusively for centuries for ripening and aging of the “Roquefort” cheese.^{57,58}

Indications for the production of fermented alcoholic beverages in protohistory are abundant, albeit frequently ambiguous,⁵⁹ and can be found in the Near East, Middle East, Far East, and Europe.^{60–67} Evidence for the production of grape wine in Europe and viticulture in the Near East dates back to the 6th and 7th millennium BC.⁶⁸ Such evidence was mainly based on chemical residue analysis or archaeobotanical analysis or indicated in ancient inscriptions. Recently scientists claimed that they were able to revive an ancient yeast strain from Egyptian potteries and used it to ferment beer.⁶⁹ Here we were able to reconstruct >90% of the *S. cerevisiae* genome from an Iron Age-dated paleofecal sample. We used different molecular analysis at the genome level to infer the possible routes of domestication for this yeast. Our results suggested it was used in beer fermentation. Together with the results of the dietary analysis that showed presence of different fermentable cereals, e.g., wheat, barley, and millets, we can envisage how the fermentation was carried out.

It might be assumed that the fermentation was carried out in a spontaneous manner—i.e., by adding water to wort and allowing the fermentation process to take place by the wild air-borne yeasts or the constitutional microbiota of the used cereals.⁷⁰ We do not see, however, indications for other yeasts species, such as *Brettanomyces bruxellensis*, that co-occur in spontaneously fermented beers.⁷¹ In addition, we see clear indications of

domestication and continuous supply of new admixture components to this yeast, which might suggest that fermentation vessels were repeatedly used for this purpose or the inoculation of the fermentation batches has been done by back-slopping (i.e., inoculation of new fermentation batches with portions of previous batches).⁷² Albeit varied evidence for beer production in protohistoric Europe exists,^{59,65–67} these beers could not be preserved for longer time periods and would have had to be consumed rapidly after production,⁶⁶ which also presupposes that the beer would have had to be produced either in Hallstatt itself or in the very near surroundings.

Considering the constant temperature of 8°C inside the Hallstatt mines, it might be expected that this yeast was used for production of lager-like beer, when fermentation is carried out at low temperatures (also known as bottom-fermentation) and results in a beer that can be stored for longer time periods.⁷³ Historically, however, the bottom-fermentation was most likely developed after the year 1553, when the Duke of Bavaria Albrecht V forbade brewing during summer months.⁷⁴ Additionally, Gonçalves and colleagues demonstrated that *Saccharomyces pastorianus* strains, which are hybrids of *S. cerevisiae* and another *Saccharomyces* species and are used for production of lager beers, belong to the main beer clade (Figure 4D).³¹ Therefore, we postulate that the beer produced at that time is similar to what would nowadays be known as pale beer, produced mainly by top-fermenting *S. cerevisiae* strains.

Paleofeces material displays an archaeological information source that provides insights into the diet and gut microbiome composition of ancestors. Here, we had access to four paleofeces samples from the Hallstatt salt mines dating from the Bronze Age to the Baroque period. The constant low annual temperature and high salt concentrations inside the mine preserved both plant macro-remains and biomolecules (DNA and protein) in the paleofeces. We demonstrate the indispensable complementarity of using microscopic and molecular approaches in resolving the paleofecal dietary residual components and to reconstruct the ancient gut microbiome. Furthermore, we extended our paleofeces microbiome analysis to focusing on key microbes that are involved in food processing, which opens new avenues in understanding fermentation history. In the future, additional samples from different time points will provide a more fine-scaled diachronic picture, which may help us to understand the role of dietary changes in shaping our gut microbiome and how much this was further influenced by modern lifestyles or medical advances recently introduced through industrialization and Westernization.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
 - Lead contact
 - Materials availability
 - Data and code availability
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
- METHOD DETAILS

- Paleofeces samples, radiocarbon dating
- Microscopic analysis of the paleofeces
- Molecular analysis of the paleofeces
- Proteomic analysis
- QUANTIFICATION AND STATISTICAL ANALYSIS

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at <https://doi.org/10.1016/j.cub.2021.09.031>.

A video abstract is available at <https://doi.org/10.1016/j.cub.2021.09.031#mmc5>.

ACKNOWLEDGMENTS

We acknowledge the following funding sources: Programma Ricerca Budget prestazioni Eurac 2017 of the Province of Bolzano, Italy, and the South Tyrolean grant legge 14 (F.M., M.S.S., S.Z., and A.Z.). Additional support was provided by the European Regional Development Fund 2014-2020_CALL-FESR 2017 Research and Innovation_Autonomous Province of Bolzano South Tyrol_Project: FESR1078-MummyLabs. The authors thank the Department of Innovation, Research and University of the Autonomous Province of Bozen/Bolzano for covering the Open Access publication costs. We thank Dr. John Wilson (ProtiFi, USA) for helpful discussions regarding proteomics sample preparation. This work was in addition supported by the European Research Council grant ERC-STG Project MetaPG (N.S.); the US National Institutes of Health, National Institute for General Medical Sciences under grant no. GM087221 and the Office of the Director 1S10OD026936; and the US National Science Foundation award 1920268 (R.L.M.). We would like to thank Eva-Maria Geigl and the two anonymous reviewers for their insightful comments that helped to improve the manuscript.

AUTHOR CONTRIBUTIONS

F.M., K.O., N.S., A.Z., H.R., and K.K. conceived the investigation. F.M., A.S., R.L.M., K.O., N.S., A.Z., H.R., and K.K. designed experiments. K.O., A.S., F.M., A.Z., H.R., and K.K. were involved in the sampling campaign. F.M., A.S., U.K., M.R.H., and K.O. conducted the experiments. F.M., M.S.S., K.D.H., A.T., A.S., S.Z., A.B.-M., P.M., J.C.-K., W.R., U.K., S.R.M., M.R.H., O.R.-S., T.R., and K.O. performed analyses. F.M. and M.S.S. wrote the manuscript with contributions from K.D.H., A.T., A.B.-M., J.C.-K., M.R.H., O.R.-S., T.R., R.L.M., K.O., N.S., A.Z., H.R., and K.K.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: May 27, 2021

Revised: August 16, 2021

Accepted: September 14, 2021

Published: October 13, 2021

REFERENCES

1. Fry, G. (1985). Analysis of fecal material. In *The Analysis of Prehistoric Diets*, R.J. Gilbert, and J. Mielke, eds. (Academic Press), pp. 127–154.
2. Shillito, L.-M., Blong, J.C., Green, E.J., and van Asperen, E.N. (2020). The what, how and why of archaeological coprolite analysis. *Earth Sci. Rev.* 207, 103196.
3. Gilbert, M.T.P., Jenkins, D.L., Götherstrom, A., Naveran, N., Sanchez, J.J., Hofreiter, M., Thomsen, P.F., Binladen, J., Higham, T.F.G., Yohe, R.M., 2nd., et al. (2008). DNA from pre-Clovis human coprolites in Oregon, North America. *Science* 320, 786–789.
4. Maixner, F., Turaev, D., Cazenave-Gassiot, A., Janko, M., Krause-Kyora, B., Hoopmann, M.R., Kusebauch, U., Sartain, M., Guerriero, G.,

- O'Sullivan, N., et al. (2018). The iceman's last meal consisted of fat, wild meat, and cereals. *Curr. Biol.* 28, 2348–2355.e9.
5. Poinar, H.N., Kuch, M., Sobolik, K.D., Barnes, I., Stankiewicz, A.B., Kuder, T., Spaulding, W.G., Bryant, V.M., Cooper, A., and Pääbo, S. (2001). A molecular analysis of dietary diversity for three archaic Native Americans. *Proc. Natl. Acad. Sci. USA* 98, 4317–4322.
 6. Mitchell, P.D. (2017). Human parasites in the Roman World: health consequences of conquering an empire. *Parasitology* 144, 48–58.
 7. Reinhard, K.J., Ferreira, L.F., Bouchet, F., Sianto, L., Dutra, J.M.F., Iniguez, A., Leles, D., Le Bailly, M., Fugassa, M., Pucu, E., and Araújo, A. (2013). Food, parasites, and epidemiological transitions: a broad perspective. *Int. J. Paleopathol.* 3, 150–157.
 8. Maixner, F., Krause-Kyora, B., Turaev, D., Herbig, A., Hoopmann, M.R., Hallows, J.L., Kusebauch, U., Vigil, E.E., Malferteiner, P., Megraud, F., et al. (2016). The 5300-year-old *Helicobacter pylori* genome of the Iceman. *Science* 351, 162–165.
 9. Tett, A., Huang, K.D., Asnicar, F., Fehlner-Peach, H., Pasolli, E., Karcher, N., Armanini, F., Manghi, P., Bonham, K., Zolfo, M., et al. (2019). The *Prevotella copri* complex comprises four distinct clades underrepresented in Westernized populations. *Cell Host Microbe* 26, 666–679.e7.
 10. Tito, R.Y., Knights, D., Metcalf, J., Obregon-Tito, A.J., Cleeland, L., Najjar, F., Roe, B., Reinhard, K., Sobolik, K., Belknap, S., et al. (2012). Insights from characterizing extinct human gut microbiomes. *PLoS ONE* 7, e51146.
 11. Borry, M., Cordova, B., Perri, A., Wibowo, M., Prasad Honap, T., Ko, J., Yu, J., Britton, K., Girdland-Flink, L., Power, R.C., et al. (2020). CoproID predicts the source of coprolites and paleofeces using microbiome composition and host DNA content. *PeerJ* 8, e9001.
 12. 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., et al. (2021). Reconstruction of ancient microbial genomes from the human gut. *Nature* 594, 234–239.
 13. Harding, A. (2013). *Salt in Prehistoric Europe* (Sidestone Press).
 14. Reschreiter, H., and Kowarik, K. (2019). Bronze Age mining in Hallstatt. A new picture of everyday life in the salt mines and beyond. *Archaeologia Austriaca* 103, 99–136.
 15. Grabner, M., Wächter, E., Nicolussi, K., Bolka, M., Sormaz, T., Steier, P., Wild, E.M., Barth, F.E., Kern, A., Rudorfer, J., et al. (2021). Prehistoric salt mining in Hallstatt, Austria. New chronologies out of small wooden fragments. *Dendrochronologia* 66, 125814.
 16. Kowarik, K. (2019). Mining and landscape: synthesis. In *Hallstätter Beziehungsgeschichten. Wirtschaftsstrukturen und Umfeldbeziehungen der bronze- und ältereisenzeitlichen Salzbergbaue von Hallstatt/OÖ.*, K. Kowarik, ed. (Oberösterreichisches Landesmuseum), pp. 245–266.
 17. Festi, D., Brandner, D., Grabner, M., Knierzinger, W., Reschreiter, H., and Kowarik, K. (2021). 3500 years of environmental sustainability in the large-scale alpine mining district of Hallstatt, Austria. *Journal of Archaeological Science: Reports* 35, 102670.
 18. *Wochenberichte* (1723). *Hofschreiberamt Hallstatt*. <https://www.landesarchiv-ooe.at/>.
 19. Brewster, R., Tamburini, F.B., Asimwe, E., Oduaran, O., Hazelhurst, S., and Bhatt, A.S. (2019). Surveying gut microbiome research in Africans: toward improved diversity and representation. *Trends Microbiol.* 27, 824–835.
 20. Key, F.M., Posth, C., Krause, J., Herbig, A., and Bos, K.I. (2017). Mining metagenomic data sets for ancient DNA: recommended protocols for authentication. *Trends Genet.* 33, 508–520.
 21. Beghini, F., Mclver, L.J., Blanco-Míguez, A., Dubois, L., Asnicar, F., Maharjan, S., Mailyan, A., Manghi, P., Scholz, M., Thomas, A.M., et al. (2021). Integrating taxonomic, functional, and strain-level profiling of diverse microbial communities with bioBakery 3. *eLife* 10, e65088.
 22. Grant, W.D. (2015). *Halococcus*. In *Bergey's Manual of Systematics of Archaea and Bacteria*, M.E. Trujillo, S. Dedysh, P. DeVos, B. Hedlund, P. Kämpfer, F.A. Rainey, and W.B. Whitman, eds. (John Wiley and Sons).
 23. García, S., Vidal, J.E., Heredia, N., and Juneja, V.K. (2019). *Clostridium perfringens*. In *Food Microbiology: Fundamentals and Frontiers*, 5th Edition, M.P. Doyle, F. Diez-Gonzalez, and C. Hill, eds. (American Society of Microbiology).
 24. Voidarou, C., Bezirtzoglou, E., Alexopoulos, A., Plessas, S., Stefanis, C., Papadopoulos, I., Vavias, S., Stavropoulou, E., Fotou, K., Tzora, A., and Skoufos, I. (2011). Occurrence of *Clostridium perfringens* from different cultivated soils. *Anaerobe* 17, 320–324.
 25. Pasolli, E., De Filippis, F., Mauriello, I.E., Cumbo, F., Walsh, A.M., Leech, J., Cotter, P.D., Segata, N., and Ercolini, D. (2020). Large-scale genome-wide analysis links lactic acid bacteria from food with the gut microbiome. *Nat. Commun.* 11, 2610.
 26. Dumas, E., Feurtey, A., Rodríguez de la Vega, R.C., Le Prieur, S., Snirc, A., Coton, M., Thierry, A., Coton, E., Le Piver, M., Roueyre, D., et al. (2020). Independent domestication events in the blue-cheese fungus *Penicillium roqueforti*. *Mol. Ecol.* 29, 2639–2660.
 27. Ropars, J., Rodríguez de la Vega, R.C., López-Villavicencio, M., Gouzy, J., Sallet, E., Dumas, É., Lacoste, S., Debuchy, R., Dupont, J., Branca, A., and Giraud, T. (2015). Adaptive horizontal gene transfers between multiple cheese-associated fungi. *Curr. Biol.* 25, 2562–2569.
 28. Cheeseman, K., Ropars, J., Renault, P., Dupont, J., Gouzy, J., Branca, A., Abraham, A.-L., Ceppi, M., Conseiller, E., Debuchy, R., et al. (2014). Multiple recent horizontal transfers of a large genomic region in cheese making fungi. *Nat. Commun.* 5, 2876.
 29. Ropars, J., Didot, E., Rodríguez de la Vega, R.C., Bennetot, B., Coton, M., Poirier, E., Coton, E., Snirc, A., Le Prieur, S., and Giraud, T. (2020). Domestication of the emblematic white cheese-making fungus *Penicillium camemberti* and its diversification into two varieties. *Curr. Biol.* 30, 4441–4453.e4.
 30. Gallone, B., Steensels, J., Prahil, T., Soriaga, L., Saels, V., Herrera-Malaver, B., Merlevede, A., Roncoroni, M., Voordeckers, K., Miraglia, L., et al. (2016). Domestication and divergence of *Saccharomyces cerevisiae* beer yeasts. *Cell* 166, 1397–1410.e16.
 31. Gonçalves, M., Pontes, A., Almeida, P., Barbosa, R., Serra, M., Libkind, D., Hutzler, M., Gonçalves, P., and Sampaio, J.P. (2016). Distinct domestication trajectories in top-fermenting beer yeasts and wine yeasts. *Curr. Biol.* 26, 2750–2761.
 32. Pontes, A., Hutzler, M., Brito, P.H., and Sampaio, J.P. (2020). Revisiting the taxonomic synonyms and populations of *Saccharomyces cerevisiae*—phylogeny, phenotypes, ecology and domestication. *Microorganisms* 8, 903.
 33. Marsit, S., Leducq, J.-B., Durand, É., Marchant, A., Filteau, M., and Landry, C.R. (2017). Evolutionary biology through the lens of budding yeast comparative genomics. *Nat. Rev. Genet.* 18, 581–598.
 34. Ness, F., and Aigle, M. (1995). RTM1: a member of a new family of telomeric repeated genes in yeast. *Genetics* 140, 945–956.
 35. Wu, H., Ito, K., and Shimoi, H. (2005). Identification and characterization of a novel biotin biosynthesis gene in *Saccharomyces cerevisiae*. *Appl. Environ. Microbiol.* 71, 6845–6855.
 36. Novo, M., Bigey, F., Beyne, E., Galeote, V., Gavory, F., Mallet, S., Cambon, B., Legras, J.-L., Wincker, P., Casaregola, S., and Dequin, S. (2009). Eukaryote-to-eukaryote gene transfer events revealed by the genome sequence of the wine yeast *Saccharomyces cerevisiae* EC1118. *Proc. Natl. Acad. Sci. USA* 106, 16333–16338.
 37. Heiss, A.G., Jakobitsch, T., Wiesinger, S., and Trebsche, P. (2021). Dig out, dig in! Plant-based diet at the Late Bronze Age copper production site of Priggltitz-Gasteil (Lower Austria) and the relevance of processed foodstuffs for the supply of Alpine Bronze Age miners. *PLoS ONE* 16, e0248287.
 38. Makki, K., Deehan, E.C., Walter, J., and Bäckhed, F. (2018). The impact of dietary fiber on gut microbiota in host health and disease. *Cell Host Microbe* 23, 705–715.

39. Crittenden, A.N., and Schnorr, S.L. (2017). Current views on hunter-gatherer nutrition and the evolution of the human diet. *Am. J. Phys. Anthropol.* *162* (Suppl 63), 84–109.
40. De Filippis, F., Pasolli, E., Tett, A., Tarallo, S., Naccarati, A., De Angelis, M., Neviani, E., Cocolin, L., Gobbetti, M., Segata, N., and Ercolini, D. (2019). Distinct genetic and functional traits of human intestinal *Prevotella copri* strains are associated with different habitual diets. *Cell Host Microbe* *25*, 444–453.e3.
41. Gálvez, E.J.C., Iljazovic, A., Amend, L., Lesker, T.R., Renault, T., Thiemann, S., Hao, L., Roy, U., Gronow, A., Charpentier, E., and Strowig, T. (2020). Distinct polysaccharide utilization determines interspecies competition between intestinal *Prevotella* spp. *Cell Host Microbe* *28*, 838–852.e6.
42. Fehner-Peach, H., Magnabosco, C., Raghavan, V., Scher, J.U., Tett, A., Cox, L.M., Gottsegen, C., Watters, A., Wiltshire-Gordon, J.D., Segata, N., et al. (2019). Distinct polysaccharide utilization profiles of human intestinal *Prevotella copri* isolates. *Cell Host Microbe* *26*, 680–690.e5.
43. Blaser, M.J. (2017). The theory of disappearing microbiota and the epidemics of chronic diseases. *Nat. Rev. Immunol.* *17*, 461–463.
44. Pasolli, E., Asnicar, F., Manara, S., Zolfo, M., Karcher, N., Armanini, F., Beghini, F., Manghi, P., Tett, A., Ghensi, P., et al. (2019). Extensive unexplored human microbiome diversity revealed by over 150,000 genomes from metagenomes spanning age, geography, and lifestyle. *Cell* *176*, 649–662.e20.
45. Sonnenburg, E.D., and Sonnenburg, J.L. (2019). The ancestral and industrialized gut microbiota and implications for human health. *Nat. Rev. Microbiol.* *17*, 383–390.
46. Fox, P.F., Guinee, T.P., Cogan, T.M., and McSweeney, P.L. (2017). Cheese: historical aspects. *Fundamentals of Cheese Science* (Springer), pp. 1–10.
47. Evershed, R.P., Payne, S., Sherratt, A.G., Copley, M.S., Coolidge, J., Urem-Kotsu, D., Kotsakis, K., Ozdoğan, M., Ozdoğan, A.E., Nieuwenhuys, O., et al. (2008). Earliest date for milk use in the Near East and southeastern Europe linked to cattle herding. *Nature* *455*, 528–531.
48. Yang, Y., Shevchenko, A., Knaust, A., Abuduresule, I., Li, W., Hu, X., Wang, C., and Shevchenko, A. (2014). Proteomics evidence for kefir dairy in Early Bronze Age China. *J. Archaeol. Sci.* *45*, 178–186.
49. Bottéro, J. (1985). The cuisine of ancient Mesopotamia. *The Biblical Archaeologist* *48*, 36–47.
50. Greco, E., El-Aguizy, O., Ali, M.F., Foti, S., Cunsolo, V., Saletti, R., and Ciliberto, E. (2018). Proteomic analyses on an ancient Egyptian cheese and biomolecular evidence of brucellosis. *Anal. Chem.* *90*, 9673–9676.
51. McClure, S.B., Magill, C., Podrug, E., Moore, A.M.T., Harper, T.K., Culleton, B.J., Kennett, D.J., and Freeman, K.H. (2018). Fatty acid specific $\delta^{13}\text{C}$ values reveal earliest Mediterranean cheese production 7,200 years ago. *PLoS ONE* *13*, e0202807.
52. Salque, M., Bogucki, P.I., Pyzel, J., Sobkowiak-Tabaka, I., Grygiel, R., Szmyt, M., and Evershed, R.P. (2013). Earliest evidence for cheese making in the sixth millennium BC in northern Europe. *Nature* *493*, 522–525.
53. Cakmakci, S., Gundogdu, E., Hayaloglu, A.A., Dagdemir, E., Gurses, M., Cetin, B., and Tahmas-Kahyaoglu, D. (2012). Chemical and microbiological status and volatile profiles of mouldy C ivil cheese, a T urkish mould-ripened variety. *Int. J. Food Sci. Technol.* *47*, 2405–2412.
54. Monti, L., Pelizzola, V., Povoio, M., Fontana, S., and Contarini, G. (2019). Study on the sugar content of blue-veined “Gorgonzola” PDO cheese. *Int. Dairy J.* *95*, 1–5.
55. Burger, J., Link, V., Blöcher, J., Schulz, A., Sell, C., Pochon, Z., Diekmann, Y., Žegarac, A., Hofmanová, Z., Winkelbach, L., et al. (2020). Low prevalence of lactase persistence in Bronze Age Europe indicates ongoing strong selection over the last 3,000 years. *Curr. Biol.* *30*, 4307–4315.e13.
56. Cantor, M.D., van den Tempel, T., Hansen, T.K., and Ardö, Y. (2017). Blue cheese. In *Cheese*, P.L.H. McSweeney, P.F. Fox, P.D. Cotter, and D.W. Everett, eds. (Elsevier), pp. 929–954.
57. Capozzi, V., and Spano, G. (2011). Food microbial biodiversity and “microbes of protected origin”. *Front. Microbiol.* *2*, 237.
58. Desmasures, N. (2014). Mold-ripened varieties. In *Encyclopedia of Food Microbiology*, C. Batt, ed. (Elsevier), pp. 409–416.
59. Heiss, A.G., Azorín, M.B., Antolín, F., Kubiak-Martens, L., Marinova, E., Arendt, E.K., Biliaderis, C.G., Kretschmer, H., Lazaridou, A., Stika, H.-P., et al. (2020). Mashers to mashes, crust to crust. Presenting a novel microstructural marker for malting in the archaeological record. *PLoS ONE* *15*, e0231696.
60. Zarnkow, M., Otto, A., and Einwag, B. (2011). Interdisciplinary investigations into the brewing technology of the ancient Near East and the potential of the cold mashing process. In *Liquid Bread: Beer and Brewing in Cross-cultural Perspective*, W. Schiefenhövel, and H. Macbeth, eds. (Berghahn Books), pp. 47–54.
61. Liu, L., Wang, J., Rosenberg, D., Zhao, H., Lengyel, G., and Nadel, D. (2018). Fermented beverage and food storage in 13,000 y-old stone mortars at Raqefet Cave, Israel: Investigating Natufian ritual feasting. *Journal of Archaeological Science: Reports* *21*, 783–793.
62. Wang, J., Liu, L., Ball, T., Yu, L., Li, Y., and Xing, F. (2016). Revealing a 5,000-y-old beer recipe in China. *Proc. Natl. Acad. Sci. USA* *113*, 6444–6448.
63. Farag, M.A., Elmassry, M.M., Baba, M., and Friedman, R. (2019). Revealing the constituents of Egypt’s oldest beer using infrared and mass spectrometry. *Sci. Rep.* *9*, 16199.
64. Damerow, P. (2012). Sumerian beer: the origins of brewing technology in ancient Mesopotamia. *Cuneiform Digital Library Journal* *2*, 1–20.
65. Stika, H.P. (2011). Beer in prehistoric Europe. In *Liquid Bread: Beer and Brewing in Cross-Cultural Perspective*, W. Schiefenhövel, and H. Macbeth, eds. (Berghahn Books), pp. 55–62.
66. Rosenstock, E., and Scheibner, A. (2017). Fermentierter Brei und vergorenes Malz: Bier in der Vorgeschichte Südwestasiens und Europas. *Mitteilungen der Anthropologischen Gesellschaft Wien* *147*, 31–62.
67. Guerra-Doce, E. (2015). The origins of inebriation: archaeological evidence of the consumption of fermented beverages and drugs in prehistoric Eurasia. *J. Archaeol. Method Theory* *22*, 751–782.
68. McGovern, P., Jalabadze, M., Batiuk, S., Callahan, M.P., Smith, K.E., Hall, G.R., Kvavadze, E., Maghradze, D., Rusishvili, N., Bouby, L., et al. (2017). Early neolithic wine of Georgia in the South Caucasus. *Proc. Natl. Acad. Sci. USA* *114*, E10309–E10318.
69. Aouizerat, T., Gutman, I., Paz, Y., Maeir, A.M., Gadot, Y., Gelman, D., Szitenberg, A., Drori, E., Pinkus, A., Schoemann, M., et al. (2019). Isolation and characterization of live yeast cells from ancient vessels as a tool in bio-archaeology. *MBio* *10*, e00388, e19.
70. Basso, R.F., Alcarde, A.R., and Portugal, C.B. (2016). Could non-Saccharomyces yeasts contribute on innovative brewing fermentations? *Food Res. Int.* *86*, 112–120.
71. Spitaels, F., Wieme, A.D., Janssens, M., Aerts, M., Daniel, H.-M., Van Landschoot, A., De Vuyst, L., and Vandamme, P. (2014). The microbial diversity of traditional spontaneously fermented lambic beer. *PLoS ONE* *9*, e95384.
72. Spitaels, F., Wieme, A.D., Snauwaert, I., De Vuyst, L., and Vandamme, P. (2017). Microbial ecology of traditional beer fermentations. In *Brewing Microbiology: Current Research, Omics and Microbial Ecology*, N.A. Bokulich, and C.W. Bamforth, eds. (Caister Academic Press), pp. 179–196.
73. Baker, E.P., Peris, D., Moriarty, R.V., Li, X.C., Fay, J.C., and Hittinger, C.T. (2019). Mitochondrial DNA and temperature tolerance in lager yeasts. *Sci. Adv.* *5*, v1869.
74. Dornbusch, H.D. (1998). *Prost!: The Story of German Beer* (Brewers Publications).

75. NCBI Resource Coordinators (2017). Database resources of the National Center for Biotechnology Information. *Nucleic Acids Res.* **45** (D1), D12–D17.
76. Banchi, E., Ametrano, C.G., Greco, S., Stanković, D., Muggia, L., and Pallavicini, A. (2020). PLANITS: a curated sequence reference dataset for plant ITS DNA metabarcoding. *Database (Oxford)* **2020**, baz155.
77. Nilsson, R.H., Larsson, K.-H., Taylor, A.F.S., Bengtsson-Palme, J., Jeppesen, T.S., Schigel, D., Kennedy, P., Picard, K., Glöckner, F.O., Tedersoo, L., et al. (2019). The UNITE database for molecular identification of fungi: handling dark taxa and parallel taxonomic classifications. *Nucleic Acids Res.* **47** (D1), D259–D264.
78. Ratnasingham, S., and Hebert, P.D. (2007). bold: The Barcode of Life Data System (<http://www.barcodinglife.org>). *Mol. Ecol. Notes* **7**, 355–364.
79. Buchfink, B., Reuter, K., and Drost, H.-G. (2021). Sensitive protein alignments at tree-of-life scale using DIAMOND. *Nat. Methods* **18**, 366–368.
80. Huson, D.H., Beier, S., Flade, I., Górski, A., El-Hadidi, M., Mitra, S., Ruscheweyh, H.J., and Tappu, R. (2016). MEGAN Community Edition - interactive exploration and analysis of large-scale microbiome sequencing data. *PLoS Comput. Biol.* **12**, e1004957.
81. Ondov, B.D., Bergman, N.H., and Phillippy, A.M. (2011). Interactive metagenomic visualization in a Web browser. *BMC Bioinformatics* **12**, 385.
82. Li, H., and Durbin, R. (2010). Fast and accurate long-read alignment with Burrows-Wheeler transform. *Bioinformatics* **26**, 589–595.
83. Neukamm, J., Peltzer, A., and Nieselt, K. (2021). DamageProfiler: fast damage pattern calculation for ancient DNA. *Bioinformatics* **btab190**. <https://doi.org/10.1093/bioinformatics/btab190>.
84. Renaud, G., Slon, V., Duggan, A.T., and Kelso, J. (2015). Schmutzi: estimation of contamination and endogenous mitochondrial consensus calling for ancient DNA. *Genome Biol.* **16**, 224.
85. Skoglund, P., Storå, J., Götherström, A., and Jakobsson, M. (2013). Accurate sex identification of ancient human remains using DNA shotgun sequencing. *J. Archaeol. Sci.* **40**, 4477–4482.
86. Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., and Durbin, R.; 1000 Genome Project Data Processing Subgroup (2009). The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078–2079.
87. Weissensteiner, H., Pacher, D., Kloss-Brandstätter, A., Forer, L., Specht, G., Bandelt, H.J., Kronenberg, F., Salas, A., and Schönherr, S. (2016). HaploGrep 2: mitochondrial haplogroup classification in the era of high-throughput sequencing. *Nucleic Acids Res.* **44** (W1), W58–63.
88. Langmead, B., and Salzberg, S.L. (2012). Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **9**, 357–359.
89. Korneliussen, T.S., Albrechtsen, A., and Nielsen, R. (2014). ANGSD: analysis of next generation sequencing data. *BMC Bioinformatics* **15**, 356.
90. Katoh, K., Misawa, K., Kuma, K., and Miyata, T. (2002). MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Res.* **30**, 3059–3066.
91. Ludwig, W., Strunk, O., Westram, R., Richter, L., Meier, H., Yadukumar, Buchner, A., Lai, T., Steppi, S., Jobb, G., et al. (2004). ARB: a software environment for sequence data. *Nucleic Acids Res.* **32**, 1363–1371.
92. Guindon, S., and Gascuel, O. (2003). A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Syst. Biol.* **52**, 696–704.
93. Altschul, S.F., Gish, W., Miller, W., Myers, E.W., and Lipman, D.J. (1990). Basic local alignment search tool. *J. Mol. Biol.* **215**, 403–410.
94. Cook, D.E., and Andersen, E.C. (2017). VCF-kit: assorted utilities for the variant call format. *Bioinformatics* **33**, 1581–1582.
95. Stamatakis, A. (2014). RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* **30**, 1312–1313.
96. Letunic, I., and Bork, P. (2021). Interactive Tree Of Life (iTOL) v5: an online tool for phylogenetic tree display and annotation. *Nucleic Acids Res.* **49** (W1), W293–W296.
97. Alexander, D.H., and Lange, K. (2011). Enhancements to the ADMIXTURE algorithm for individual ancestry estimation. *BMC Bioinformatics* **12**, 246.
98. Kessner, D., Chambers, M., Burke, R., Agus, D., and Mallick, P. (2008). ProteoWizard: open source software for rapid proteomics tools development. *Bioinformatics* **24**, 2534–2536.
99. Deutsch, E.W., Mendoza, L., Shteynberg, D., Slagel, J., Sun, Z., and Moritz, R.L. (2015). Trans-Proteomic Pipeline, a standardized data processing pipeline for large-scale reproducible proteomics informatics. *Proteomics Clin. Appl.* **9**, 745–754.
100. Keller, A., Nesvizhskii, A.I., Kolker, E., and Aebersold, R. (2002). Empirical statistical model to estimate the accuracy of peptide identifications made by MS/MS and database search. *Anal. Chem.* **74**, 5383–5392.
101. Nesvizhskii, A.I., Keller, A., Kolker, E., and Aebersold, R. (2003). A statistical model for identifying proteins by tandem mass spectrometry. *Anal. Chem.* **75**, 4646–4658.
102. Hoopmann, M.R., Winget, J.M., Mendoza, L., and Moritz, R.L. (2018). StPeter: seamless label-free quantification with the trans-proteomic pipeline. *J. Proteome Res.* **17**, 1314–1320.
103. Eng, J.K., Jahan, T.A., and Hoopmann, M.R. (2013). Comet: an open-source MS/MS sequence database search tool. *Proteomics* **13**, 22–24.
104. Gurdeep Singh, R., Tanca, A., Palomba, A., Van der Jeugt, F., Verschaffelt, P., Uzzau, S., Martens, L., Dawyndt, P., and Mesuere, B. (2019). Unipept 4.0: functional analysis of metaproteome data. *J. Proteome Res.* **18**, 606–615.
105. Stuiver, M., and Polach, H.A. (1977). Discussion reporting of ¹⁴C data. *Radiocarbon* **19**, 355–363.
106. Reimer, P.J., Austin, W.E.N., Bard, E., Bayliss, A., Blackwell, P.G., Bronk Ramsey, C., Butzin, M., Cheng, H., Edwards, R.L., Friedrich, M., et al. (2020). The IntCal20 Northern Hemisphere radiocarbon age calibration curve (0–55 cal kBP). *Radiocarbon* **62**, 725–757.
107. Pearsall, D.M. (2015). *Paleoethnobotany: A Handbook of Procedures* (Left Coast Press).
108. Cappers, R.T., Bekker, R.M., and Jans, J.E. (2006). *Digitale zadenatlas van Nederland* (Barkhuis Publishing).
109. Jacomet, S. (2006). *Bestimmung von Getreidefunden aus archäologischen Ausgrabungen* (Universität Basel).
110. Neef, R., Cappers, R.T., and Bekker, R.M. (2012). *Digital Atlas of Economic Plants in Archaeology* (Barkhuis).
111. Hohmann, B., Deutschmann, F., and Gassner, G. (1989). *Mikroskopische Untersuchungen pflanzlicher Lebensmittel: mit einem Kapitel über die mikroskopische Untersuchung der wichtigsten als Futtermittel verwendeten tierischen Produkte* (Gustav Fischer Verlag).
112. Tang, J.N., Zeng, Z.G., Wang, H.N., Yang, T., Zhang, P.J., Li, Y.L., Zhang, A.Y., Fan, W.Q., Zhang, Y., Yang, X., et al. (2008). An effective method for isolation of DNA from pig faeces and comparison of five different methods. *J. Microbiol. Methods* **75**, 432–436.
113. Kircher, M., Sawyer, S., and Meyer, M. (2012). Double indexing overcomes inaccuracies in multiplex sequencing on the Illumina platform. *Nucleic Acids Res.* **40**, e3.
114. Meyer, M., Kircher, M., Gansauge, M.T., Li, H., Racimo, F., Mallick, S., Schraiber, J.G., Jay, F., Prüfer, K., de Filippo, C., et al. (2012). A high-coverage genome sequence from an archaic Denisovan individual. *Science* **338**, 222–226.
115. Rosenbloom, K.R., Armstrong, J., Barber, G.P., Casper, J., Clawson, H., Diekhans, M., Dreszer, T.R., Fujita, P.A., Guruvadoo, L., Haeussler, M., et al. (2015). The UCSC Genome Browser database: 2015 update. *Nucleic Acids Res.* **43**, D670–D681.
116. Andrews, R.M., Kubacka, I., Chinnery, P.F., Lightowlers, R.N., Turnbull, D.M., and Howell, N. (1999). Reanalysis and revision of the Cambridge reference sequence for human mitochondrial DNA. *Nat. Genet.* **23**, 147.

117. 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. *Nat. Methods* 8, 761–763.
118. Marchione, D.M., Ilieva, I., Devins, K., Sharpe, D., Pappin, D.J., Garcia, B.A., Wilson, J.P., and Wojcik, J.B. (2020). HYPERsol: high-quality data from archival FFPE tissue for clinical proteomics. *J. Proteome Res.* 19, 973–983.
119. Martens, L., Chambers, M., Sturm, M., Kessner, D., Levander, F., Shofstahl, J., Tang, W.H., Römpp, A., Neumann, S., Pizarro, A.D., et al. (2011). mzML—a community standard for mass spectrometry data. *Mol. Cell. Proteomics* 10, 000133.
120. Moosa, J.M., Guan, S., Moran, M.F., and Ma, B. (2020). Repeat-preserving decoy database for false discovery rate estimation in peptide identification. *J. Proteome Res.* 19, 1029–1036.
121. Shteynberg, D., Deutsch, E.W., Lam, H., Eng, J.K., Sun, Z., Tasman, N., Mendoza, L., Moritz, R.L., Aebersold, R., and Nesvizhskii, A.I. (2011). iProphet: multi-level integrative analysis of shotgun proteomic data improves peptide and protein identification rates and error estimates. *Mol. Cell. Proteomics* 10, 007690.
122. Perez-Riverol, Y., Csordas, A., Bai, J., Bernal-Llinares, M., Hewapathirana, S., Kundu, D.J., Inuganti, A., Griss, J., Mayer, G., Eisenacher, M., et al. (2019). The PRIDE database and related tools and resources in 2019: improving support for quantification data. *Nucleic Acids Res.* 47 (D1), D442–D450.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Biological samples		
Paleofeces sample from the Bronze Age	This study	2610
Paleofeces sample from the Iron Age	This study	2604
Paleofeces sample from the Iron Age	This study	2611
Paleofeces sample from the Baroque time	This study	2612
Chemicals, peptides, and recombinant proteins		
Sodium phosphate	Sigma-Aldrich	Cat #342483
Trypsin gold	Promega	Cat # V528A
Critical commercial assays		
S-Trap	ProtiFi	Cat #K02-mini-10
Deposited data		
Paleofeces metagenomic shotgun datasets	This study	ENA: PRJEB44507
NCBI-nr database	75	https://www.ncbi.nlm.nih.gov/protein/
NCBI-nt database	75	https://www.ncbi.nlm.nih.gov/nucleotide/
PlantITS database	76	https://github.com/apallavicini/PLANIITS
Chloroplast genome database	N/A	https://www.ncbi.nlm.nih.gov/genome/organelle
Fungal ITS database	77	https://unite.ut.ee/
BOLD system databases	78	https://v3.boldsystems.org/
Software and algorithms		
MetaPhlAn	21	https://github.com/biobakery/MetaPhlAn/wiki/MetaPhlAn-3.0
DIAMOND	79	https://github.com/bbuchfink/diamond
MEGAN6	80	https://www.wsi.uni-tuebingen.de/lehrstuehle/algorithms-in-bioinformatics/software/megan6
Krona tool	81	https://github.com/marbl/Krona/wiki
BWA	82	http://bio-bwa.sourceforge.net/
DeDup tool	N/A	https://github.com/apeltzer/DeDup
DamageProfiler	83	https://damageprofiler.readthedocs.io/en/latest/index.html
Schmutzi	84	https://github.com/grenaud/schmutzi
Molecular sex determination	85	https://github.com/pontussk/ry_compute
SAMtools	86	http://samtools.github.io/
HaploGrep	87	https://haplogrep.i-med.ac.at/
HUMAnN	21	https://github.com/biobakery/humann
python package scikit-bio	N/A	http://scikit-bio.org/
bowtie2	88	http://bowtie-bio.sourceforge.net/bowtie2/
FastQ Screen	N/A	https://github.com/StevenWingett/FastQ-Screen
ANGSD tool	89	https://github.com/ANGSD/angsd
MAFFT	90	https://mafft.cbrc.jp/alignment/software/
ARB software package	91	http://www.arb-home.de/
PhyML	92	https://github.com/stephaneguindon/phyml
BLASTn	93	N/A
Picard tools	N/A	https://broadinstitute.github.io/picard/
GATK4	N/A	https://gatk.broadinstitute.org/hc

(Continued on next page)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
vcf-kit	94	https://vcf-kit.readthedocs.io/en/latest/
RAxML	95	https://github.com/stamatak/standard-RAxML
Interactive Tree of Life (iTOL)	96	https://itol.embl.de/
ADMIXTURE	97	http://dalexander.github.io/admixture/download.html
R-Studio	N/A	https://www.rstudio.com/
ProteoWizard	98	https://proteowizard.sourceforge.io/
Trans-Proteomic Pipeline	99	http://tools.proteomecenter.org/software.php
PeptideProphet	100	http://peptideprophet.sourceforge.net/
ProteinProphet	101	http://proteinprophet.sourceforge.net/
StPeter	102	http://tools.proteomecenter.org/wiki/index.php?title=Software:StPeter
Comet	103	http://comet-ms.sourceforge.net/
Unipept 4.0	104	https://unipept.ugent.be/

RESOURCE AVAILABILITY

Lead contact

Further information on materials, datasets, and protocols should be directed to and will be fulfilled by the lead contact, Frank Maixner (frank.maixner@eurac.edu).

Materials availability

This study did not generate new unique reagents.

Data and code availability

Sequencing data are available at the European Nucleotide Archive (ENA) under ENA: PRJEB44507. Mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE partner repository, PRIDE: PXD027613. All code used in this study and other previously published genomic data is available at the sources referenced in the [key resources table](#).

EXPERIMENTAL MODEL AND SUBJECT DETAILS

In this study, we subjected four paleofeces samples to in-depth microscopic and molecular analysis. The paleofeces material stems from Bronze Age and Iron Age mine workings in the Hallstatt salt mines in Upper Austria (Figures 1A–1C). See [Data S1A](#) for additional details to the samples used in this study.

METHOD DETAILS

Paleofeces samples, radiocarbon dating

Three paleofeces (2610, 2604, 2611) were recovered in 1983, 1989 (2604) and 2003 (2610) through wet-sieving of larger blocks of protohistoric production debris excavated in the mine workings. One additional paleofeces (2612) was sampled in 2019 *in situ* at the site Edlersbergwerk-oben with sterile sampling tools. This sample was not subjected to wet-sieving. All four samples (using approx. 500 mg each) were subjected to radiocarbon dating at the Curt-Engelhorn-Centre for Archaeometry in Mannheim, Germany (Figure 1C; [Data S1B](#)). The ^{14}C content was measured with an AMS system of the MICADAS type. The isotopic ratios $^{14}\text{C}/^{12}\text{C}$ and $^{13}\text{C}/^{12}\text{C}$ of the samples, the calibration standards (oxalic acid II), the blanks and the control standards were measured simultaneously in the AMS system. The determined ^{14}C ages are based on $^{13}\text{C} = -25\text{‰}^{105}$ and were calibrated to calendar ages with the dataset INTCAL20¹⁰⁶ and the software SwissCal (L.Wacker, ETH-Zurich). The remaining material has been used for microscopic and molecular analyses.

Microscopic analysis of the paleofeces

Before chemical treatment of the paleofeces the surface of each sample was stripped off to avoid contamination. Then rehydration in a 0.5% solution of trisodium phosphate for 72 h.¹⁰⁷ After olfactory and visual testing, the liquid was screened through 500, 250, 125, and 63 μm steel meshes, and the outwash was kept for further microfossil studies. Plant macro-remains were picked out of the fractionated residues in the steel meshes and then identified and quantified under a stereo microscope with magnifications up to $\times 63$. The

identification of the plant remains was conducted with identification keys^{108–111} and the reference collection of the Department of Botany, Innsbruck University, was consulted for comparative purposes.

Molecular analysis of the paleofeces

DNA extraction, library preparation and sequencing

The molecular analysis of the paleofeces samples was conducted at the ancient DNA laboratory of the EURAC Institute for Mummy Studies in Bolzano, Italy. Sample documentation, sample preparation and DNA extraction were performed in a dedicated pre-PCR area following the strict procedures required for studies of ancient DNA: use of protective clothing, UV-light exposure of the equipment and bleach sterilization of surfaces, use of PCR workstations and filtered pipette tips. The DNA was extracted from the paleofeces samples (200 mg) using a chloroform-based DNA extraction method according to the protocol of Tang and colleagues with minor modifications.¹¹² To test whether the wet-sieving step during the classical archaeological excavation results in an “wash-out” effect of DNA from the paleofeces, we re-hydrated additional 200 mg material of two samples (2611, 2612) in 1 ml of 0.5N Na₃PO₄ (trisodium phosphate)-buffer for 18 h at RT.¹ After re-hydration, the samples were centrifuged (5000 g, 5 min) and both the supernatant (2611RW, 2612RW) and the pelleted paleofeces (2611R, 2612R) were subjected to DNA extraction as described above. From all DNA extracts double-indexed libraries were generated for the sequencing runs with a modified protocol for Illumina multiplex sequencing.^{113,114} Libraries were first shallow sequenced on an Illumina MiSeq platform and then deep sequenced on Illumina HiSeq2500 and HiSeqX10 platforms using 101–base pair (HiSeq2500) and 151–base pair (MiSeq, HiSeqX) paired-end sequencing kits. For details to the metagenomic shotgun datasets please refer to [Data S1C](#). Data are available from the European Nucleotide Archive under accession number PRJEB44507.

Pre-processing of the sequence data, general taxonomic overview, human DNA analysis

The paired Illumina reads were first quality-checked and processed (adaptor removal and read merging) as previously described in Maixner et al.⁴ Initially we tested for the above-described possible DNA wash-out effect by assigning taxonomically the microbial reads using MetaPhlAn 3.0²¹ in the shallow sequenced MiSeq datasets of the untreated samples (2611, 2612), the washed and pelleted paleofeces (2611R, 2612R), and the supernatant (2611RW, 2612RW). For all subsequent analyses, we used the combined sequencing data available for each sample ([Data S1C](#)). First, we assessed a general taxonomic profile of the sequencing reads using DIAMOND (v2.0.7) blastx search⁷⁹ against the NCBI nr database (Release 237, April 2020). The DIAMOND tables were converted to rma6 (blast2rma tool) format (–minPercentIdentity 97), imported into MEGAN6 software,⁸⁰ and subsequently visualized using the Krona tool.⁸¹ Next, we assessed the endogenous human DNA content in the paleofeces samples by aligning the sequence reads against the human genome (build hg19)¹¹⁵ and the human mtDNA reference genome (rCRS)¹¹⁶ using BWA⁸² with a seed length of 1,000. The minimum mapping and base quality were both 30. To deduplicate the mapped reads, we used the DeDup tool (<https://github.com/apeltzer/DeDup>). For details to the mapping results please refer to [Data S1D](#). The resulting bam files were used to check for characteristic aDNA nucleotide misincorporation frequency patterns and for the fragment length distribution using the DamageProfiler tool.⁸³ Mitochondrial human contamination rates were assessed using Schmutzi.⁸⁴ The sex of the mapped human reads was assigned using a Maximum likelihood method, based on the karyotype frequency of X and Y chromosomal reads.⁸⁵ Variants in the mitochondrial genome were called using SAMtools mpileup and bcftools⁸⁶ with stringent filtering options (quality > 30). The haplogroup was identified by submitting the variant calling file to the HaploGrep website.⁸⁷

Comparing paleofeces microbiome structure to contemporary metagenomic datasets

To compare the paleofeces microbiome structure with contemporary individuals', we downloaded 9,368 publicly available shotgun metagenomes representing modern-day human populations (n = 9,207), and, as a control, environmental soil sample (n = 161). Each metagenome was characterized by source (human or soil), and in the case of human samples, by body-site, and whether from a Westernized or non-Westernized population. The term “non-Westernized” describes a population that follows a traditional non-urbanized lifestyle encompassing factors such as diet, hygiene, and with no or limited access to medical healthcare and pharmaceuticals (e.g., antibiotics) as previously described.^{9,19,44} Afterward, we performed profiling of the microbial composition of each metagenomic sample including the four paleofeces samples from this study using MetaPhlAn 3.0²¹ using default settings. To survey paleofeces microbiome members in different populations, microbial abundance profiles of four paleofeces were first merged using `merge_metaphlan_tables.py`, a utility script in MetaPhlAn 3.0. The merged abundance table was then visualized in the form of a hierarchically-clustered heatmap using `hclust2` (<https://github.com/SegataLab/hclust2>) with parameters `–ftop 15–f_dist_f braycurtis–s_dist_f braycurtis`, with a result of displaying only 15 most abundant microbial species in the four paleofeces samples. We further analyzed the prevalence of these top 15 ancient-sample enriched species in the context of global populations using a subset (n = 8,968) of shotgun metagenomes that are stool samples from healthy adult individuals characterized by non-Westernization ([Data S1H](#)). A species was counted as presence in a sample if the relative abundance was above 0.01%. We repeated the prevalence analysis for all species detected in the four paleofeces samples to take low-abundance members into account as well ([Data S1I](#) and [S1J](#)). Next, we randomly selected 662 metagenomes (132 oral cavity and 530 stool samples) from all publicly available human metagenomes, along with 161 soil metagenomes, for profiling of the metabolic pathways using HUMAnN 3.0²¹ with default settings. The Python script `humann_renorm_table.py` as part of the pipeline was used to normalize the metabolic pathways to relative abundance, followed by converting individual profiles into a merged abundance table using the utility script `humann_join_tables.py`. For the same selected metagenomic samples, the individual profiles of microbial composition generated by MetaPhlAn 3.0 as described above were similarly merged into a single abundance table using a python script `merge_metaphlan_tables.py`. Subsequently, the merged abundance tables were used for calculating Bray-Curtis's distance with logarithmic transformation based on which a Principal

Coordinates Analysis was performed using python package scikit-bio (version 0.5.6; <http://scikit-bio.org/>). Lastly, to predict the source of microbial communities in the paleofeces samples SourceTracker2¹¹⁷ was used as described in a recent study,¹² comparing the paleofeces samples' abundance profiles with those of different sources (a subset of samples randomly selected from the aforementioned Principal Coordinates Analysis: 16 non-Westernized oral samples, 20 non-Westernized stool samples, 26 soil samples, 10 Westernized oral samples and 30 Westernized stool samples). For details to the comparative datasets please refer to [Data S1E](#) and S1G.

Molecular characterization of ancient diet

The merged reads were searched for homology, using BLAST search, against different specific DNA barcoding databases ([Data S1N](#)). Specifically, the databases of plant Internal transcribed spacers,⁷⁶ RuBisCo-large subunit and Maturase K genes (*rbcL/matK*, <https://v3.boldsystems.org/>), and chloroplast genomes (<https://www.ncbi.nlm.nih.gov/genome/organelle>) were used to identify the plant dietary components. The BOLD database of the cytochrome c oxidase subunit I gene (*COI*, <https://v3.boldsystems.org/>) was used to specifically target the animal diet and intestinal parasites. The UNITE database of the fungal ITS gene database (<https://unite.ut.ee/>) was used to identify potential food processing fungi. Finally, to target all the eukaryotic dietary and diet-related components, the reads were aligned against the currently available full mitochondrial and chloroplast genomes of the NCBI database⁷⁵ using BWA⁸² with default parameters. Subsequently, to obtain a taxonomic overview of the aligned reads we performed a sequence similarity search using BLASTn⁹³ with default parameters against the complete NCBI-nt database.⁷⁵

Both the resulting BLAST and the previously created DIAMOND tables were converted to rma6 (blast2rma tool) format and imported into MEGAN6 software. The read counts at genus- and species-level were exported as csv files and imported into R-Studio software for further analysis.

To distinguish the true hits from the false hits, we applied two filters as follows: i) For the BLAST and DIAMOND tables, we only considered hits of $\geq 90\%$ identity and $\geq 90\%$ coverage of each query read; then ii) We considered the sample positive for a certain component only if it shows incidence in majority of databases of that component, e.g., the sample 2604 was considered positive for *Bos taurus*, because it contained true hits for mitochondrial-, *COI*-, and nr-databases; and finally iii) for the plant dietary components, we used the plantITS database as a primary filter, since it is highly curated and recently updated. So, we initially filtered out all the plantITS negative taxa then considered the same majority rule.

Finally, we used transformed sums of counts of the genera and species that passes the aforementioned filters into normalized log base 2 (\log_{10} reads million⁻¹) and plotted them as dot-plots proportional to their abundances. The dendrogram to the left of the dot-plot represent the phylogeny based on NCBI taxonomy (<https://www.ncbi.nlm.nih.gov/Taxonomy/Browser/wwwtax.cgi>).

Selected identified taxa were further subjected to in-depth analysis by mapping the quality-filtered sequence reads against organellar (mitochondrial and chloroplast) and nuclear genomes using bowtie2 (v1.2.1.1) and the parameter “end-to-end”⁸⁸ (for details to the references used please refer to the [Data S1O](#)). In two cases where the previously described molecular dietary analysis resulted in an assignment down to the genus level (*Triticum* spp., *Ascaris* spp.) we assessed an appropriate reference sequence by mapping the sequence reads against selected chloroplast and mitochondrial genomes using BWA⁸² (with default parameters) implemented in the program FastQ Screen (<https://github.com/StevenWingett/FastQ-Screen>), and selected as reference the species that belongs to one of these two genera and has the most specific hits. After the bowtie2 alignment against the reference sequences the mapped reads were deduplicated and checked for damage patterns as described above for the human DNA. For details to the mapping results please refer to [Data S1O](#) and S1P. Organellar sequences that showed higher than 70% sequence coverage were further subjected to phylogenetic analysis as previously detailed.⁴ In brief, a consensus FASTA sequence was generated using the ANGSD tool⁸⁹ and together with other comparative datasets subjected to a multiple alignment using the MAFFT multiple sequence alignment program.⁹⁰ Phylogenetic analysis were performed by applying distance-matrix, maximum-parsimony, and maximum-likelihood methods implemented in the ARB software package:⁹¹ neighbor-joining (using the Jukes-Cantor algorithm for nucleic acid correction with 1000 bootstrap iterations), DNA parsimony (PHYLIP version 3.66 with 100 bootstrap iterations), and DNA maximum-likelihood [PhyML⁹² with the HKY substitution model]. The number of informative nucleotide positions used for the phylogenetic analysis and the bootstrap support is indicated in the respective figure captions.

Genome-level analysis of ancient fungi - Variant calling

The sample 2604 showed high prevalence of *Penicillium roqueforti* and *Saccharomyces cerevisiae* DNA allowing further genome-level comparative analysis with modern datasets ([Data S1Q](#) and S1R). As comparative datasets, we included the recently published genomic data of 34 *P. roqueforti*²⁶ as well as 157 *S. cerevisiae*.³⁰

We mapped the quality-filtered reads against the reference genomes of *Penicillium roqueforti* FM164 and *Saccharomyces cerevisiae* S288c, independently, using bowtie2 (v1.2.1.1) and the parameter “end-to-end.”⁸⁸ Using samtools,⁸⁶ the mapping qualities of <30 were filtered out and the bam files were sorted and indexed. The read duplicates were marked and removed using the DeDup tool (<https://github.com/apeltzer/DeDup/>) and new read-groups were assigned using the “AddOrReplaceReadGroups” of Picard tools (<https://broadinstitute.github.io/picard/>). Then, we used the Genome Analysis Toolkit 4 (GATK4) (<https://gatk.broadinstitute.org/hc>) to call genome variants, following the best practices workflow as follows: i) Variants within each of the single bam files were called using the tool “HaplotypeCaller”; ii) The resulting VCF files were combined into a single file using the tool “CombineGVCFs”; iii) The combined VCF file was then genotyped using the tool “GenotypeGVCFs”; iv) InDels were filtered out and only single nucleotide polymorphisms (SNPs) were kept using the tool “SelectVariants”; and finally v) the VCF file was filtered to include SNPs with minor allele frequency (MAF) of ≥ 0.05 -0.1 and $\leq 10\%$ missing data. A total of 120359 and 375629 SNPs were resolved for *P. roqueforti* and

S. cerevisiae, respectively. The total number of positions with alternative alleles to the reference genome accounted for 0.97% in case of *P. roqueforti* and 10.05% in case of *S. cerevisiae*. More than 99% of the positions with alternative alleles have a major allele frequency of > 90%, which minimize the likelihood of having chimeric alignments. In addition, we applied majority rule selection for the positions with multiple alleles, thereby analyzing only the major allele type.

Phylogenetic- and population structure analysis

To analyze the phylogenetic relationship of the strains, we used the vcf-kit to convert the SNP datasets (vcf files) to fastA alignment format.⁹⁴ Then we used the Fasta2phylip.pl script to create PHYLIP alignment files (<https://mullinslab.microbiol.washington.edu/>), which were then subjected to maximum likelihood (ML) phylogenetic analysis using RAxML (v8.2.12), following generalized time reversible substitution model (GTR) and GAMMA model of rate heterogeneity.⁹⁵ Best-scoring ML trees were searched after 1000 bootstrap replicates and visualized and annotated using the interactive tree of life, iTOL tool (<https://itol.embl.de>).

To infer the degree of admixture and the number of populations in the analyzed datasets (120,337 SNPs for *P. roqueforti* and 136,712 *S. cerevisiae*), we carried out unsupervised population structure analysis using ADMIXTURE (v1.3.0), testing K in the range of 3 to 22.⁹⁷ The best K values were determined based on the lowest cross-validation error (cv) and maximum likelihood.

Functional marker genes

To infer the potential functions of the modern and ancient *Saccharomyces cerevisiae* yeast strains based on their genome sequence, we searched for the presence of functional marker genes (*RTM1*, *BIO1/BIO6*, and Regions A/B/C) that were recently described by Cheeseman et al.²⁸ The *RTM1* gene is responsible for conferring resistance to the toxicity of molasses, and therefore it is assumed to be positively selected in beer yeast strains.³⁴ While the genes *BIO1* and *BIO6*, which are involved in *de novo* synthesis of biotin, are highly selected in sake fermenting yeasts.³⁵ The regions A, B, and C are horizontally transferred genomic regions from other yeast genera, e.g., *Kluyveromyces*, *Pichia*, and *Zygosaccharomyces*.³⁶ These regions contain 39 genes distributed over 3 different chromosomes (Data S1S) and are assumed to play a role in wine fermentation. To search for these marker genes, we used bowtie2 to map the quality filtered reads, of our sample as well as the recently published *S. cerevisiae* comparative genomic data,³⁰ against the individual genes. Then, we considered a gene as present if it is covered $\geq 90\%$ with 3x depth (Data S1T). For phylogenetic tree annotation (Figure 4D), the regions A/B/C were considered present if any of the genes was positive, while *BIO1/BIO6* were considered positive if both genes were positive.

Proteomic analysis

Sample preparation

Paleofeces samples (ID 2610, 2604, 2611, 2612) and a blank control sample to highlight any contamination were resuspended in 800 μ L 5% (w/v) sodium dodecyl sulfate (SDS), 50 mM triethylammonium bicarbonate (TEAB, pH 8.5) and disrupted at 4°C using three 2.8 mm ceramic beads (QIAGEN, USA) and a Precellys 24 homogenizer (Bertin Corp, USA) at 6500 rpm for 30 s. Samples were transferred to Eppendorf tubes and centrifuged at 4,000 rpm for 7 min, the supernatant removed and centrifuged twice at 13,000 g for 10 min. The supernatant was subjected to the S-Trap mini spin column digestion protocol (ProtiFi, USA).¹¹⁸ Briefly, samples were acidified with 12% phosphoric acid in water (pH ≤ 1). Binding/wash buffer (100 mM TEAB (final) in 90% methanol) was added to each sample at 6.4 x the volume of each sample. Samples were vortexed and applied to the S-Trap column in ≤ 500 μ L aliquots until the entire sample was loaded (J. Wilson, personal communication). After each sample load the S-Trap column was centrifuged at 4,000 g for 30 s. Samples were washed to remove any residual SDS by adding 3x 400 μ L binding/wash buffer and centrifugation at 4,000 g for 30 s after each wash, and 1 min after the last wash. Proteins were digested on the S-Trap at 37°C for 13 h by adding 125 μ L digestion buffer containing trypsin gold (Promega, USA) in 125 mM NH_4HCO_3 . The S-Trap was rehydrated with 100 μ L 125 mM NH_4HCO_3 and peptides eluted with 80 μ L 125 mM NH_4HCO_3 , 80 μ L 0.2% (v/v) formic acid in water, and 80 μ L 50% (v/v) acetonitrile in water. Peptide eluates were pooled and dried under centrifugal evaporation (Savant, Thermo-Fisher Scientific, USA).

Liquid chromatography - mass spectrometry analysis (LC-MS/MS)

Samples were analyzed by high-resolution nano LC-MS/MS on an Orbitrap-Eclipse Tribid mass spectrometer (Thermo-Fisher Scientific, USA) equipped with an Easy-nLC 1000 (Thermo-Fisher Scientific, USA). Peptides were resolubilized in 50 μ L 0.1% (v/v) formic acid in water. 5 μ L of each sample were loaded onto a 2 cm Acclaim PepMap 100 trap column (75 μ m ID, C18 3 μ m (dp), Thermo-Fisher Scientific, USA) and separated using a 50 cm C18 2 μ m (dp) Easy Spray column (ES903, Thermo-Fisher Scientific, USA) using a 120-min gradient. Mobile phase A consisted of 0.1% (v/v) formic acid in water, and mobile phase B consisted of 0.1% (v/v) formic acid in acetonitrile. The gradient used 5% to 22% mobile phase B over 105 min, followed by separation from 22% to 32% mobile phase B over the remaining 15 min. The column was washed and equilibrated at 95% and 5% mobile phase B, respectively, over 40 min following each run. The flow rate was set at 300 nL/min and the column was heated at 45°C. Mass spectra were acquired on the Orbitrap-Eclipse mass spectrometer using data dependent acquisition (DDA) with dynamic exclusion. The precursor scan range was from 375-1500 *m/z* at 120,000 resolution and 100% normalized target AGC (4e5) with a 50 ms maximum injection time. The duty cycle was set to acquire as many MS/MS spectra as possible over a three second period following each precursor scan. A 1.2 *m/z* selection window was used to acquire MS/MS spectra at 30,000 resolution, a normalized AGC target of 100% (5e4), 54 ms maximum injection time, and fragmented using HCD with a normalized collision energy of 30. Dynamic exclusion was set to 60 s, with peptide match set to preferred and isotope exclusion turned on. Charge exclusion was set to 1 and greater than 7.

Protein data analysis

The raw mass spectra files were converted to mzML¹¹⁹ format using msConvert from ProteoWizard⁹⁸ and analyzed using the Trans-Proteomic Pipeline (TPP v6.0.0-rc16 Noctilucent).⁹⁹ The analysis pipeline consisted of database searching with Comet (version 2018.01 rev. 4)¹⁰³ against a 17-species UniProt FASTA database (Data S2A) and shuffled decoy sequences.¹²⁰ The 17 species were selected based on the microscopic and metagenomics analyses described above. An equal number of decoy sequences were appended to the FASTA database for downstream statistical analysis and validation of any observed peptide sequences, as described below. For details to the comparative datasets used in this study and to the protein analysis results please refer to Data S2. Comet parameters included variable modifications of +15.994915 Da (Met, Cys, His, Tyr, Trp, and Phe), and +0.984016 (Asn and Gln), as these types of protein degradations are commonly found in aged samples. A precursor tolerance of 20 ppm was set, a fragment bin tolerance of 0.2 and fragment bin offset of 0. Semi-enzymatic cleavage with up to 3 missed cleavages was allowed, and an isotope error tolerance of 3. Peptide-spectrum matches (PSMs) were validated using PeptideProphet¹⁰⁰ and iProphet.¹²¹ Protein inference was performed using ProteinProphet.¹⁰¹ Relative quantitation of protein groups was performed using StPeter.¹⁰² Observed protein groups for each sample were derived from the set of protein groups validated with ProteinProphet at an estimated 1% false discovery rate (FDR). Those protein groups were then curated by removing any group that intersected with the set of protein groups identified in the blank control analysis. A list of observed peptides was created from the peptide sequences that were used to infer the final set of protein groups. The peptide list was then filtered to contain only those peptides with a probability below a 1% FDR threshold as determined from the iProphet analysis. The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE¹²² partner repository with the dataset identifier PXD027613.

The identified peptides, evaluated by statistical analysis to 1% false discovery rate, were further taxonomically screened for homologous peptides and subsequently filtered for peptide sequences that can be unambiguously assigned to certain plants, animals, and fungi. Therefore, the peptide sequences of each sample were imported into Unipept 4.0 Desktop version,¹⁰⁴ and subjected to LCA assignment, and filtered out the following: i) peptides assigned to bacteria; ii) peptides assigned to multiple species; iii) single peptide incidences that uniquely infer a species presence (Data S2). Finally, functional annotation of the samples was carried out on different taxonomic levels to infer the molecular function of the obtained peptides.

QUANTIFICATION AND STATISTICAL ANALYSIS

Phylogenetic analyses were carried out based on the maximum-likelihood method using the tool Randomized Axelerated Maximum Likelihood (RAxML) with 1000 bootstrap replicates.⁹⁵ For the dietary analysis, majority rule-based selection was implemented using in-house R-script in R-Studio (<https://www.rstudio.com/>).

(This page intentionally left blank

Paper II: Supplementary Information

(This page intentionally left blank)

Supplemental Information

**Hallstatt miners consumed blue cheese and beer
during the Iron Age and retained a non-Westernized
gut microbiome until the Baroque period**

Frank Maixner, Mohamed S. Sarhan, Kun D. Huang, Adrian Tett, Alexander Schoenafinger, Stefania Zingale, Aitor Blanco-Míguez, Paolo Manghi, Jan Cemper-Kiesslich, Wilfried Rosendahl, Ulrike Kusebauch, Seamus R. Murrone, Michael R. Hoopmann, Omar Rota-Stabelli, Thomas Rattai, Robert L. Moritz, Klaus Oeggl, Nicola Segata, Albert Zink, Hans Reschreiter, and Kerstin Kowarik

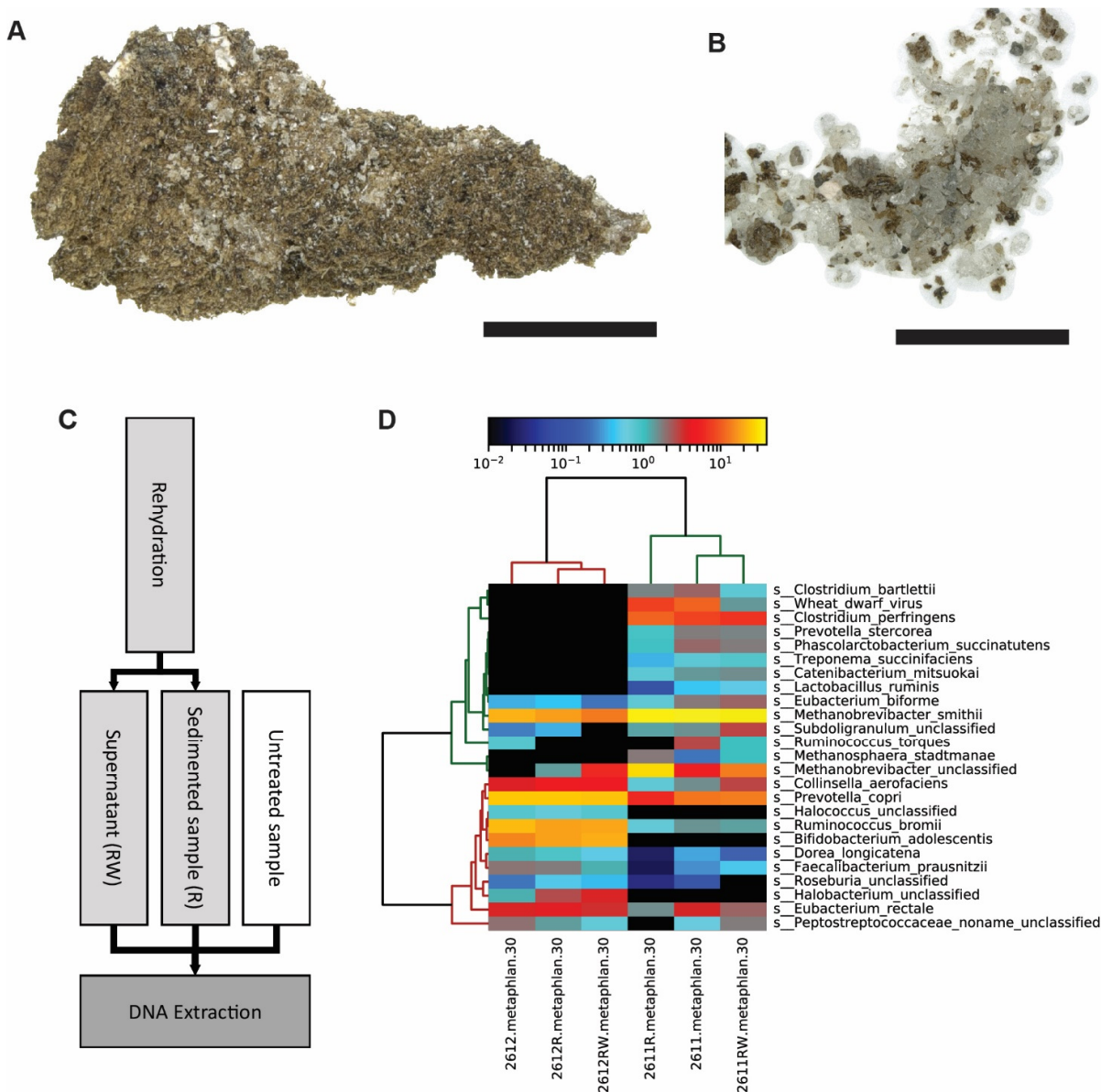


Figure S1: Wet-sieving step during the classical archaeological excavation results in an “wash-out” effect of DNA from the paleofeces. Related to Figure 1 and 2. Before subjecting the material to further in-depth molecular analysis, we tested whether the currently used wet-sieving step during the archaeological excavation and recovery of the paleofeces from the salt mine may have a negative “wash-out” effect on the DNA that is initially present in the material. We tested this effect on the paleofeces sample 2611, that has been recovered in the year 1983 via wet-sieving, and sample 2612 (A) that has not been subjected to wet-sieving is still covered by salt crystals (B). (C) Workflow to test the DNA “wash out effect. Both the “fresh” sample 2612 and the wet-sieved sample 2611 were subjected to a re-hydration treatment and DNA has been extracted from the untreated sample, the sedimented sample (R) and the supernatant (RW). (D) Taxonomic analysis of the microbial DNA of

the untreated, rehydrated (R) paleofeces samples and the supernatant (RW). MetaPhlAn 3.0 heatmap of the top 25 species found in the different MiSeq datasets. Our “wash-out” test analyzing the microbial DNA of shallow sequenced data of two re-hydrated paleofeces (2611, 2612) revealed that “free” endogenous DNA becomes lost into the aqueous phase when the material gets into contact with water (Figure S2). Importantly, however, even after 18h of re-hydration, the paleofeces material still retains sufficient DNA that resembles the bacterial taxonomic profile of the original untreated sample. Therefore, we also consider the wet-sieved samples (2604, 2610, 2611) which were considerably shorter in contact with water as an important information source for our molecular analysis. For the future, however, we strongly recommend to directly sample this precious material in the mine workings without subjecting it to additional wash-sieving treatment. For details to the samples please refer to STAR Methods and Data S1A.

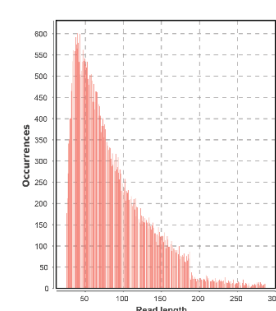
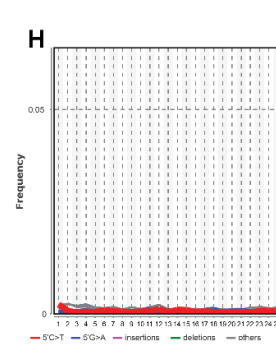
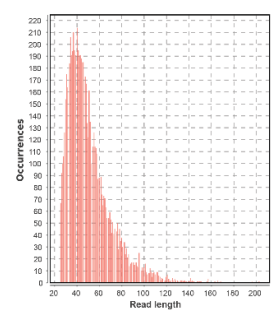
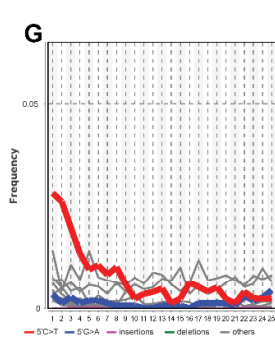
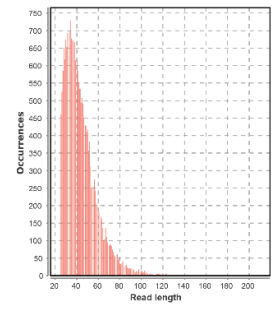
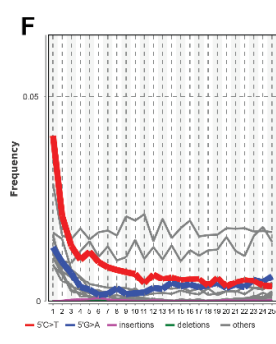
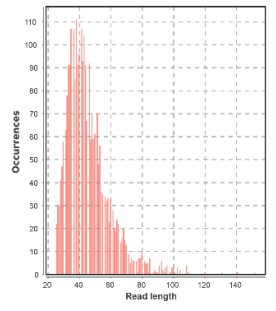
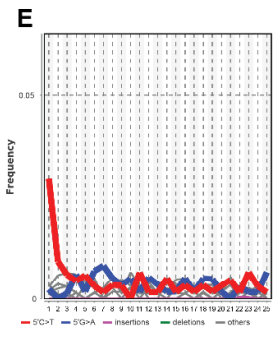
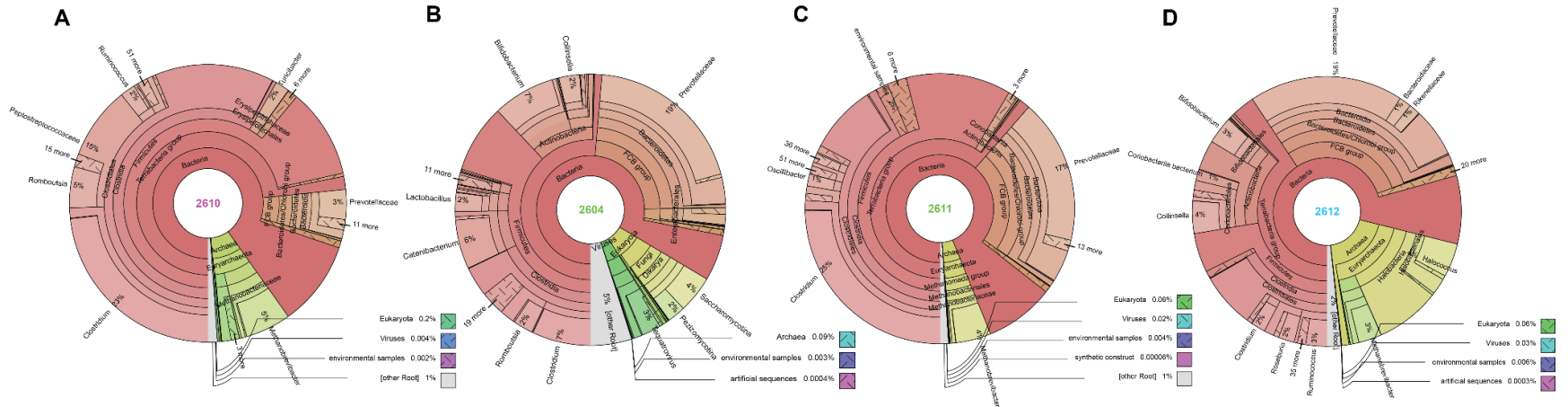


Figure S2: General taxonomic overview of the metagenomic reads and sequence characteristics of the mitochondrial human reads in the Hallstatt paleofeces samples. Related to Figure 2 and 3. (A to D) Taxonomic overview of the sequence reads in the merged shotgun datasets of the paleofeces samples 2610 (A), 2604 (B), 2611 (C), and 2612 (D). The metagenomic reads were taxonomically assigned using the Diamond tool^{S1} against the NCBI nr database. **(E to H)** DNA damage patterns (up) and read length distribution (down) of the mitochondrial human reads in the paleofeces samples 2610 (E), 2604 (F), 2611 (G), and 2612 (H). The frequency of C to T base misincorporations at the 5' end of the reads and the length distribution of the aligned reads has been assessed using the DamageProfiler tool. For methodological details please refer to STAR Methods. For the details to the Illumina shotgun datasets and the human mapping results please refer to Data S1C and S1D.

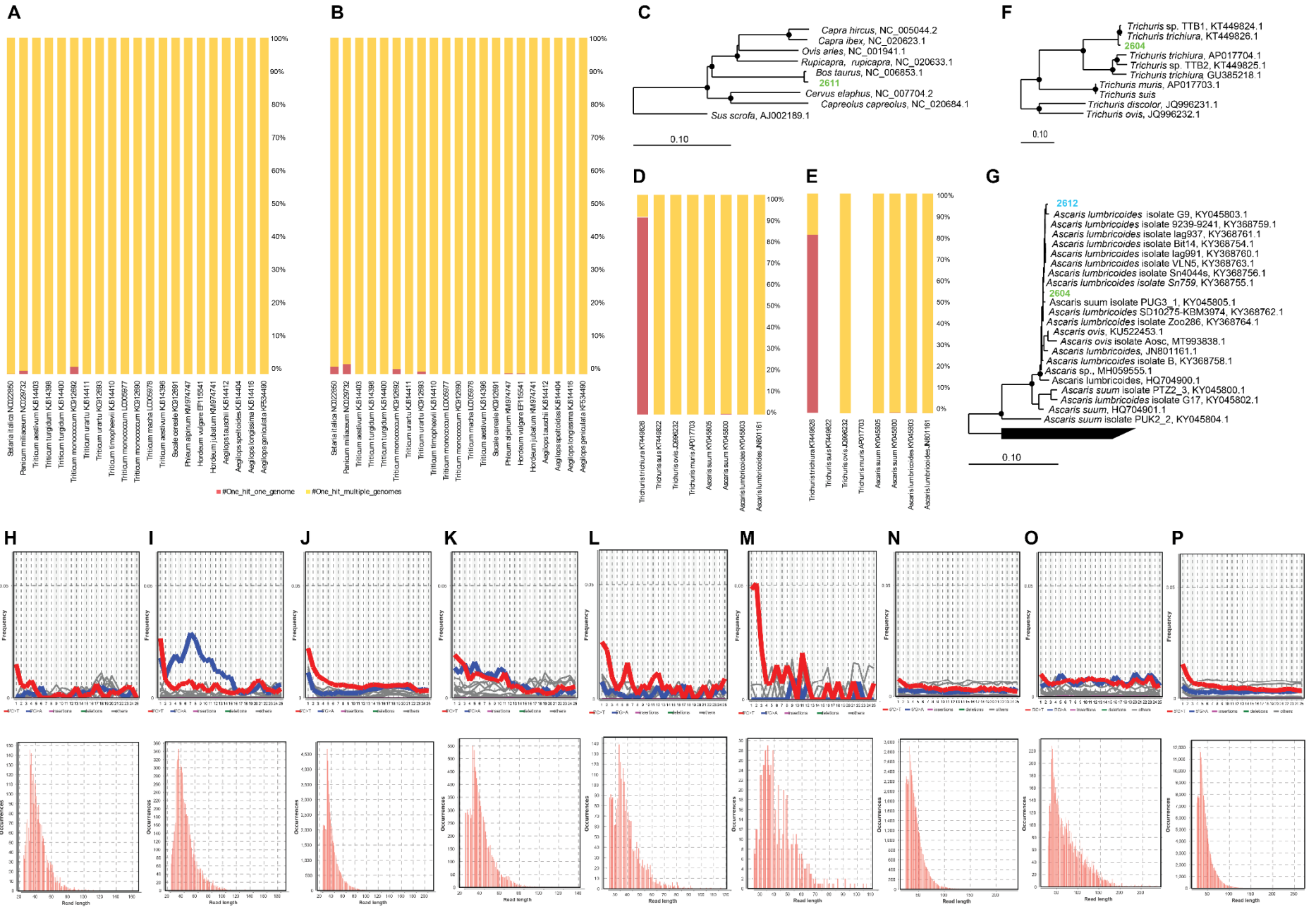


Figure S3: Comparative sequence-based analysis of selected plant, animal, and parasite DNA in the Hallstatt paleofeces samples. Related to Figure 3. (A and B) FastQ-Screen analyses of the ancient plant dietary components in the samples 2610 and 2604. FastQ Screen analysis of the quality-filtered sequence reads of the paleofeces samples 2610 (A) and 2604 (B) against the chloroplast genomes of selected members of the family *Poaceae*. FastQ Screen^{S2} was used with default parameters using BWA^{S3} as the alignment tool. **(C)** Phylogenetic analysis of the ancient animal dietary component in the samples 2611. Phylogenetic assignment of the partial cattle (*Bos taurus*) mitochondrial genome recovered from the paleofeces sample 2611. The displayed tree was calculated using the maximum-likelihood algorithm (PhyML) based on 16,412 informative positions. Black circles symbolize parsimony and neighbor joining bootstrap support (>90%) based on 100 and 1000 iterations, respectively. The scale bar depicts 0.1 substitutions per residue. **(D and E)** FastQ Screen analyses of the ancient parasites of the samples 2604 and 2612. FastQ Screen analysis of the quality-filtered sequence reads of the paleofeces samples 2604 (D) and 2612 (E) against the mitochondrial genomes of selected members of the genus *Trichuris* spp. and *Ascaris* spp with the same settings as described above. Please refer to the figure legend in subfigure A and B. **(F and G)** Phylogenetic analysis of the ancient parasites of the samples 2604 and 2612. Phylogenetic assignment of the partial mitochondrial genomes of *Trichuris* spp. (F) and *Ascaris* spp. (G) recovered from the paleofeces samples 2604 and 2612. The displayed trees were calculated using the maximum-likelihood algorithm (PhyML) based on 13,189 (*Trichuris* spp.) and 14,265 (*Ascaris* spp.) informative positions. Black circles symbolize parsimony and neighbor joining bootstrap support (>90%) based on 100 and 1000 iterations, respectively. The scale bars depict 0.1 substitutions per residue. **(H to P)** DNA damage patterns (up) and read length distribution (down) of the wheat (*Triticum* spp.) autosomal and chloroplast (H,J) reads and chloroplast reads only (I,K) in the paleofeces sample 2610 (H,I) and 2604 (J,K), of the cattle (*Bos taurus*) autosomal and mitochondrial (L) reads and mitochondrial reads only (M) in the paleofeces sample 2611, and of the roundworm (*Ascaris* spp.) autosomal and mitochondrial reads in the paleofeces sample 2604 (N) and 2612 (O) and of the whipworm (*Trichuris trichuria*) autosomal and mitochondrial reads in the paleofeces sample 2604 (P). The frequency of C to T base misincorporations at the 5' end of the reads and the length distribution of the aligned reads has been assessed using the DamageProfiler tool. For the details to the reference sequence used and mapping results please refer to the Data S10.

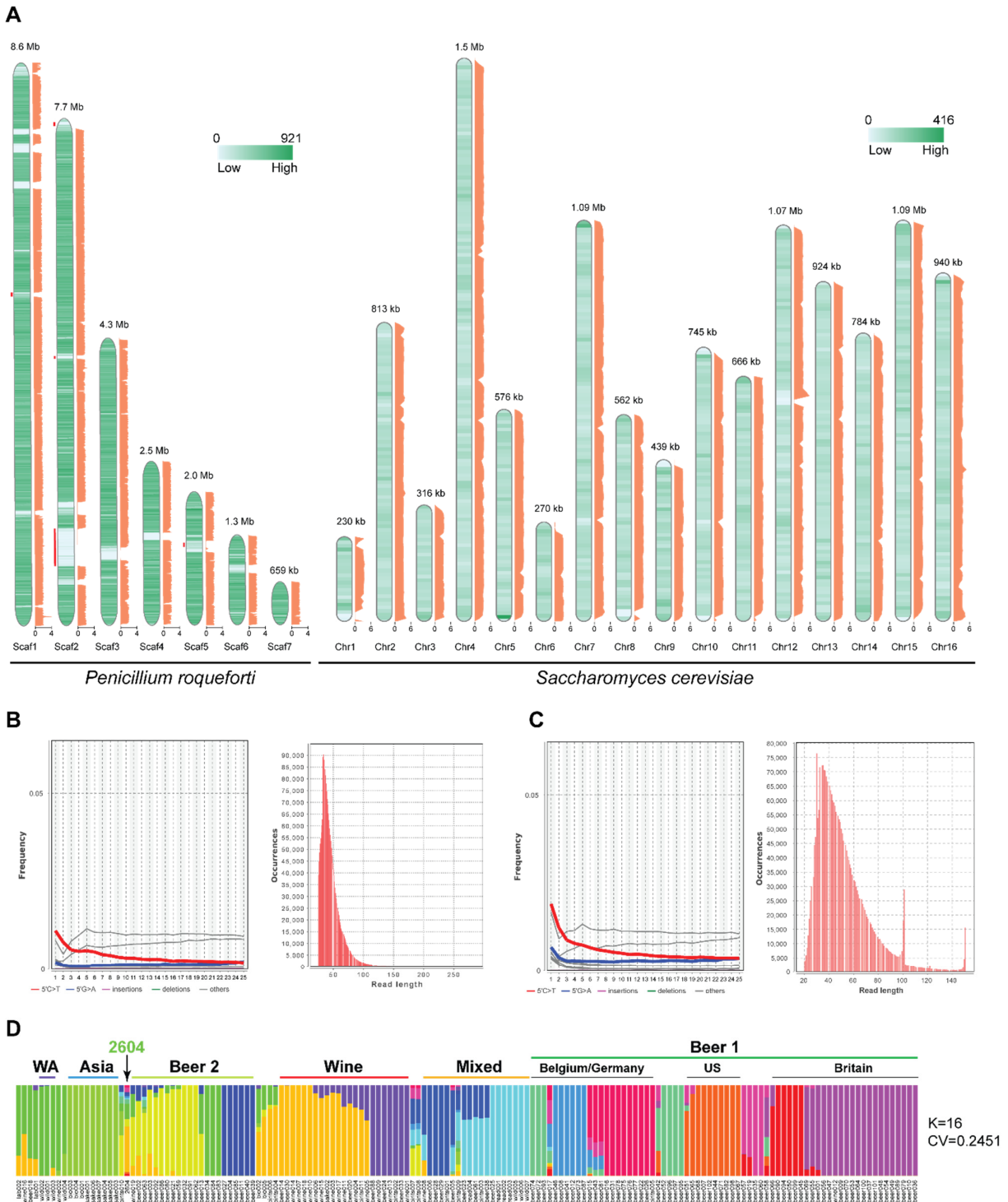


Figure S4: Genomic analyses of the food processing strains resolved from the sample 2604. Related to Figure 4. (A) Coverage plots (log₂ scale, orange color) and SNP-density heatmap (number of SNPs per window size of 10000 nt) of the fungi *Penicillium roqueforti* and *Saccharomyces cerevisiae*. The chromosome/scaffold sizes are indicated on the top of each map. For the *Penicillium roqueforti*, only scaffolds of > 0.5 Mb were shown, which represent ~ 94% of the *Penicillium roqueforti* genome. The red lines on the left of the *Penicillium roqueforti* scaffolds refer to the

putative horizontally-transferred regions from other *Penicillium* spp., that occurred during the last centuries^{S4}. The total number of positions with alternative alleles accounted for 0.97% in case of *P. roqueforti* and 10.05% in case of *S. cerevisiae*. **(B and C)** DNA damage patterns and read length distribution of the autosomal reads of *Penicillium roqueforti* and *Saccharomyces cerevisiae*, respectively. The frequency of C to T base misincorporations at the 5' end of the reads and the length distribution of the aligned reads has been assessed using the DamageProfiler tool. For the details to the reference sequences used please refer to the Data S10. **(D)** Population admixture profile of the *Saccharomyces cerevisiae* as inferred by ADMIXTURE tool. The analysis was carried out assuming K=2 to 22, and only K=16 of lowest cross validation error (cv) values is shown. A total of 136712 SNPs were included in the final analyzed dataset.

Supplemental References

- S1. Buchfink, B., Xie, C., and Huson, D.H. (2015). Fast and sensitive protein alignment using DIAMOND. *Nat Methods* 12, 59-60.
- S2. Andrew, S. (2011). FastQ Screen. Available online at: http://www.bioinformatics.babraham.ac.uk/projects/fastq_screen/.
- S3. Li, H., and Durbin, R. (2010). Fast and accurate long-read alignment with Burrows-Wheeler transform. *Bioinformatics* 26, 589-595.
- S4. Ropars, J., de la Vega, R.C.R., López-Villavicencio, M., Gouzy, J., Sallet, E., Dumas, É., Lacoste, S., Debuchy, R., Dupont, J., and Branca, A. (2015). Adaptive horizontal gene transfers between multiple cheese-associated fungi. *Current Biology* 25, 2562-2569.
- S5. Brito, I.L., Yilmaz, S., Huang, K., Xu, L., Jupiter, S.D., Jenkins, A.P., Naisilisili, W., Tamminen, M., Smillie, C.S., Wortman, J.R., et al. (2016). Mobile genes in the human microbiome are structured from global to individual scales. *Nature* 535, 435-439.
- S6. Castro-Nallar, E., Bendall, M.L., Pérez-Losada, M., Sabuncyan, S., Severance, E.G., Dickerson, F.B., Schroeder, J.R., Yolken, R.H., and Crandall, K.A. (2015). Composition, taxonomy and functional diversity of the oropharynx microbiome in individuals with schizophrenia and controls. *PeerJ* 3, e1140.
- S7. Kaur, K., Khatri, I., Akhtar, A., Subramanian, S., and Ramya, T.N.C. (2020). Metagenomics analysis reveals features unique to Indian distal gut microbiota. *PLOS ONE* 15, e0231197.
- S8. Rosa, B.A., Supali, T., Gankpala, L., Djuardi, Y., Sartono, E., Zhou, Y., Fischer, K., Martin, J., Tyagi, R., Bolay, F.K., et al. (2018). Differential human gut microbiome assemblages during soil-transmitted helminth infections in Indonesia and Liberia. *Microbiome* 6.
- S9. Rubel, M.A., Abbas, A., Taylor, L.J., Connell, A., Tanes, C., Bittinger, K., Ndze, V.N., Fonsah, J.Y., Ngwang, E., Essiane, A., et al. (2020). Lifestyle and the presence of helminths is associated with gut microbiome composition in Cameroonians. *Genome biology* 21, 122.
- S10. Schirmer, M., Smeekens, S.P., Vlamakis, H., Jaeger, M., Oosting, M., Franzosa, E.A., Horst, R.T., Jansen, T., Jacobs, L., Bonder, M.J., et al. (2016). Linking the Human Gut Microbiome to Inflammatory Cytokine Production Capacity. *Cell* 167, 1897.
- S11. Yu, J., Feng, Q., Wong, S.H., Zhang, D., Liang, Q.Y., Qin, Y., Tang, L., Zhao, H., Stenvang, J., Li, Y., et al. (2017). Metagenomic analysis of faecal microbiome as a tool towards targeted non-invasive biomarkers for colorectal cancer. *Gut* 66, 70-78.
- S12. Bulgarelli, D., Garrido-Oter, R., Münch, P.C., Weiman, A., Dröge, J., Pan, Y., McHardy, A.C., and Schulze-Lefert, P. (2015). Structure and function of the bacterial root microbiota in wild and domesticated barley. *Cell Host Microbe* 17, 392-403.
- S13. Orellana, L.H., Chee-Sanford, J.C., Sanford, R.A., Löffler, F.E., and Konstantinidis, K.T. (2018). Year-Round Shotgun Metagenomes Reveal Stable Microbial Communities in Agricultural Soils and Novel Ammonia Oxidizers Responding to Fertilization. *Appl. Environ. Microbiol.* 84.
- S14. Wilhelm, R.C., Cardenas, E., Maas, K.R., Leung, H., McNeil, L., Berch, S., Chapman, W., Hope, G., Kranabetter, J.M., Dubé, S., et al. (2017). Biogeography and organic matter removal shape long-term effects of timber harvesting on forest soil microbial communities. *ISME J.* 11, 2552-2568.
- S15. Asnicar, F., Manara, S., Zolfo, M., Truong, D.T., Scholz, M., Armanini, F., Ferretti, P., Gorfer, V., Pedrotti, A., Tett, A., et al. (2017). Studying Vertical Microbiome Transmission from Mothers to Infants by Strain-Level Metagenomic Profiling. *mSystems* 2.

- S16. Asnicar, F., Berry, S.E., Valdes, A.M., Nguyen, L.H., Piccinno, G., Drew, D.A., Leeming, E., Gibson, R., Le Roy, C., Khatib, H.A., et al. (2021). Microbiome connections with host metabolism and habitual diet from 1,098 deeply phenotyped individuals. *Nat. Med.* *27*, 321-332.
- S17. Bäckhed, F., Roswall, J., Peng, Y., Feng, Q., Jia, H., Kovatcheva-Datchary, P., Li, Y., Xia, Y., Xie, H., Zhong, H., et al. (2015). Dynamics and Stabilization of the Human Gut Microbiome during the First Year of Life. *Cell Host Microbe* *17*, 852.
- S18. Bengtsson-Palme, J., Angelin, M., Huss, M., Kjellqvist, S., Kristiansson, E., Palmgren, H., Larsson, D.G.J., and Johansson, A. (2015). The Human Gut Microbiome as a Transporter of Antibiotic Resistance Genes between Continents. *Antimicrob. Agents Chemother.* *59*, 6551-6560.
- S19. Chu, D.M., Ma, J., Prince, A.L., Antony, K.M., Seferovic, M.D., and Aagaard, K.M. (2017). Maturation of the infant microbiome community structure and function across multiple body sites and in relation to mode of delivery. *Nat. Med.* *23*, 314-326.
- S20. Costea, P.I., Coelho, L.P., Sunagawa, S., Munch, R., Huerta-Cepas, J., Forslund, K., Hildebrand, F., Kushugulova, A., Zeller, G., and Bork, P. (2017). Subspecies in the global human gut microbiome. *Mol. Syst. Biol.* *13*, 960.
- S21. De Filippis, F., Pasolli, E., Tett, A., Tarallo, S., Naccarati, A., De Angelis, M., Neviani, E., Cocolin, L., Gobetti, M., Segata, N., et al. (2019). Distinct Genetic and Functional Traits of Human Intestinal *Prevotella copri* Strains Are Associated with Different Habitual Diets. *Cell Host Microbe* *25*, 444-453.e443.
- S22. Dhakan, D.B., Maji, A., Sharma, A.K., Saxena, R., Pulikkan, J., Grace, T., Gomez, A., Scaria, J., Amato, K.R., and Sharma, V.K. (2019). The unique composition of Indian gut microbiome, gene catalogue, and associated fecal metabolome deciphered using multi-omics approaches. *GigaScience* *8*.
- S23. Feng, Q., Liang, S., Jia, H., Stadlmayr, A., Tang, L., Lan, Z., Zhang, D., Xia, H., Xu, X., Jie, Z., et al. (2015). Gut microbiome development along the colorectal adenoma-carcinoma sequence. *Nat. Commun.* *6*, 6528.
- S24. Ferretti, P., Pasolli, E., Tett, A., Asnicar, F., Gorfer, V., Fedi, S., Armanini, F., Truong, D.T., Manara, S., Zolfo, M., et al. (2018). Mother-to-Infant Microbial Transmission from Different Body Sites Shapes the Developing Infant Gut Microbiome. *Cell Host Microbe* *24*, 133-145.e135.
- S25. Gupta, A., Dhakan, D.B., Maji, A., Saxena, R., P K, V.P., Mahajan, S., Pulikkan, J., Kurian, J., Gomez, A.M., Scaria, J., et al. (2019). Association of *Flavonifractor plautii*, a Flavonoid-Degrading Bacterium, with the Gut Microbiome of Colorectal Cancer Patients in India. *mSystems* *4*.
- S26. Consortium, T.H.M.P., and The Human Microbiome Project, C. (2012). Structure, function and diversity of the healthy human microbiome. *Nature* *486*, 207-214.
- S27. Schirmer, M., Franzosa, E.A., Lloyd-Price, J., McIver, L.J., Schwager, R., Poon, T.W., Ananthakrishnan, A.N., Andrews, E., Barron, G., Lake, K., et al. (2018). Dynamics of metatranscription in the inflammatory bowel disease gut microbiome. *Nat Microbiol* *3*, 337-346.
- S28. Hall, A.B., Yassour, M., Sauk, J., Garner, A., Jiang, X., Arthur, T., Lagoudas, G.K., Vatanen, T., Fornelos, N., Wilson, R., et al. (2017). A novel *Ruminococcus gnavus* clade enriched in inflammatory bowel disease patients. *Genome Med.* *9*, 103.

- S29. Hannigan, G.D., Duhaime, M.B., Ruffin, M.T.t., Koumpouras, C.C., and Schloss, P.D. (2018). Diagnostic Potential and Interactive Dynamics of the Colorectal Cancer Virome. *MBio* 9.
- S30. Hansen, L.B.S., Roager, H.M., Søndertoft, N.B., Gøbel, R.J., Kristensen, M., Vallès-Colomer, M., Vieira-Silva, S., Ibrügger, S., Lind, M.V., Mærkedahl, R.B., et al. (2018). A low-gluten diet induces changes in the intestinal microbiome of healthy Danish adults. *Nat. Commun.* 9, 4630.
- S31. Heintz-Buschart, A., May, P., Laczny, C.C., Lebrun, L.A., Bellora, C., Krishna, A., Wampach, L., Schneider, J.G., Hogan, A., de Beaufort, C., et al. (2016). Integrated multi-omics of the human gut microbiome in a case study of familial type 1 diabetes. *Nat Microbiol* 2, 16180.
- S32. Ijaz, U.Z., Quince, C., Hanske, L., Loman, N., Calus, S.T., Bertz, M., Edwards, C.A., Gaya, D.R., Hansen, R., McGrogan, P., et al. (2017). The distinct features of microbial 'dysbiosis' of Crohn's disease do not occur to the same extent in their unaffected, genetically-linked kindred. *PLoS One* 12, e0172605.
- S33. Jie, Z., Xia, H., Zhong, S.-L., Feng, Q., Li, S., Liang, S., Zhong, H., Liu, Z., Gao, Y., Zhao, H., et al. (2017). The gut microbiome in atherosclerotic cardiovascular disease. *Nat. Commun.* 8, 845.
- S34. Keohane, D.M., Ghosh, T.S., Jeffery, I.B., Molloy, M.G., O'Toole, P.W., and Shanahan, F. (2020). Microbiome and health implications for ethnic minorities after enforced lifestyle changes. *Nat. Med.* 26, 1089-1095.
- S35. David, L.A., Weil, A., Ryan, E.T., Calderwood, S.B., Harris, J.B., Chowdhury, F., Begum, Y., Qadri, F., LaRocque, R.C., and Turnbaugh, P.J. (2015). Gut microbial succession follows acute secretory diarrhea in humans. *MBio* 6, e00381-00315.
- S36. Le Chatelier, E., Nielsen, T., Qin, J., Prifti, E., Hildebrand, F., Falony, G., Almeida, M., Arumugam, M., Batto, J.-M., Kennedy, S., et al. (2013). Richness of human gut microbiome correlates with metabolic markers. *Nature* 500, 541-546.
- S37. Li, J., Jia, H., Cai, X., Zhong, H., Feng, Q., Sunagawa, S., Arumugam, M., Kultima, J.R., Prifti, E., Nielsen, T., et al. (2014). An integrated catalog of reference genes in the human gut microbiome. *Nat. Biotechnol.* 32, 834-841.
- S38. Li, J., Zhao, F., Wang, Y., Chen, J., Tao, J., Tian, G., Wu, S., Liu, W., Cui, Q., Geng, B., et al. (2017). Gut microbiota dysbiosis contributes to the development of hypertension. *Microbiome* 5, 14.
- S39. Li, S.S., Zhu, A., Benes, V., Costea, P.I., Hercog, R., Hildebrand, F., Huerta-Cepas, J., Nieuwdorp, M., Salojärvi, J., Voigt, A.Y., et al. (2016). Durable coexistence of donor and recipient strains after fecal microbiota transplantation. *Science* 352, 586-589.
- S40. Zhernakova, A., Kurilshikov, A., Bonder, M.J., Tigchelaar, E.F., Schirmer, M., Vatanen, T., Mujagic, Z., Vila, A.V., Falony, G., Vieira-Silva, S., et al. (2016). Population-based metagenomics analysis reveals markers for gut microbiome composition and diversity. *Science* 352, 565-569.
- S41. Liu, W., Zhang, J., Wu, C., Cai, S., Huang, W., Chen, J., Xi, X., Liang, Z., Hou, Q., Zhou, B., et al. (2016). Unique Features of Ethnic Mongolian Gut Microbiome revealed by metagenomic analysis. *Sci. Rep.* 6, 34826.
- S42. Lokmer, A., Cian, A., Froment, A., Gantois, N., Viscogliosi, E., Chabé, M., and Ségurel, L. (2019). Use of shotgun metagenomics for the identification of protozoa in the gut microbiota

- of healthy individuals from worldwide populations with various industrialization levels. *PLoS One* 14, e0211139.
- S43. Louis, S., Tappu, R.-M., Damms-Machado, A., Huson, D.H., and Bischoff, S.C. (2016). Characterization of the Gut Microbial Community of Obese Patients Following a Weight-Loss Intervention Using Whole Metagenome Shotgun Sequencing. *PLoS One* 11, e0149564.
- S44. Mehta, R.S., Abu-Ali, G.S., Drew, D.A., Lloyd-Price, J., Subramanian, A., Lochhead, P., Joshi, A.D., Ivey, K.L., Khalili, H., Brown, G.T., et al. (2018). Stability of the human faecal microbiome in a cohort of adult men. *Nat Microbiol* 3, 347-355.
- S45. Nielsen, H.B., Almeida, M., Juncker, A.S., Rasmussen, S., Li, J., Sunagawa, S., Plichta, D.R., Gautier, L., Pedersen, A.G., Le Chatelier, E., et al. (2014). Identification and assembly of genomes and genetic elements in complex metagenomic samples without using reference genomes. *Nat. Biotechnol.* 32, 822-828.
- S46. 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., et al. (2015). Subsistence strategies in traditional societies distinguish gut microbiomes. *Nat. Commun.* 6, 6505.
- S47. Pasolli, E., Asnicar, F., Manara, S., Zolfo, M., Karcher, N., Armanini, F., Beghini, F., Manghi, P., Tett, A., Ghensi, P., et al. (2019). Extensive Unexplored Human Microbiome Diversity Revealed by Over 150,000 Genomes from Metagenomes Spanning Age, Geography, and Lifestyle. *Cell* 176, 649-662.e620.
- S48. Pehrsson, E.C., Tsukayama, P., Patel, S., Mejía-Bautista, M., Sosa-Soto, G., Navarrete, K.M., Calderon, M., Cabrera, L., Hoyos-Arango, W., Bertoli, M.T., et al. (2016). Interconnected microbiomes and resistomes in low-income human habitats. *Nature* 533, 212-216.
- S49. Qin, J., Li, Y., Cai, Z., Li, S., Zhu, J., Zhang, F., Liang, S., Zhang, W., Guan, Y., Shen, D., et al. (2012). A metagenome-wide association study of gut microbiota in type 2 diabetes. *Nature* 490, 55-60.
- S50. Qin, N., Yang, F., Li, A., Prifti, E., Chen, Y., Shao, L., Guo, J., Le Chatelier, E., Yao, J., Wu, L., et al. (2014). Alterations of the human gut microbiome in liver cirrhosis. *Nature* 513, 59-64.
- S51. Rampelli, S., Schnorr, S.L., Consolandi, C., Turrioni, S., Severgnini, M., Peano, C., Brigidi, P., Crittenden, A.N., Henry, A.G., and Candela, M. (2015). Metagenome Sequencing of the Hadza Hunter-Gatherer Gut Microbiota. *Curr. Biol.* 25, 1682-1693.
- S52. Raymond, F., Ouameur, A.A., Déraspe, M., Iqbal, N., Gingras, H., Dridi, B., Leprohon, P., Plante, P.-L., Giroux, R., Bérubé, È., et al. (2016). The initial state of the human gut microbiome determines its reshaping by antibiotics. *ISME J.* 10, 707-720.
- S53. 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., et al. (2015). Gut Microbiome Diversity among Cheyenne and Arapaho Individuals from Western Oklahoma. *Curr. Biol.* 25, 3161-3169.
- S54. Shao, Y., Forster, S.C., Tsaliki, E., Vervier, K., Strang, A., Simpson, N., Kumar, N., Stares, M.D., Rodger, A., Brocklehurst, P., et al. (2019). Stunted microbiota and opportunistic pathogen colonization in caesarean-section birth. *Nature* 574, 117-121.
- S55. Smits, S.A., Leach, J., Sonnenburg, E.D., Gonzalez, C.G., Lichtman, J.S., Reid, G., Knight, R., Manjurano, A., Changalucha, J., Elias, J.E., et al. (2017). Seasonal cycling in the gut microbiome of the Hadza hunter-gatherers of Tanzania. *Science* 357, 802-806.
- S56. Tett, A., Huang, K.D., Asnicar, F., Fehlner-Peach, H., Pasolli, E., Karcher, N., Armanini, F., Manghi, P., Bonham, K., Zolfo, M., et al. (2019). The *Prevotella copri* Complex Comprises

- Four Distinct Clades Underrepresented in Westernized Populations. *Cell Host Microbe* 26, 666-679 e667.
- S57. Thomas, A.M., Manghi, P., Asnicar, F., Pasolli, E., Armanini, F., Zolfo, M., Beghini, F., Manara, S., Karcher, N., Pozzi, C., et al. (2019). Metagenomic analysis of colorectal cancer datasets identifies cross-cohort microbial diagnostic signatures and a link with choline degradation. *Nat. Med.* 25, 667-678.
- S58. Vincent, C., Miller, M.A., Edens, T.J., Mehrotra, S., Dewar, K., and Manges, A.R. (2016). Bloom and bust: intestinal microbiota dynamics in response to hospital exposures and *Clostridium difficile* colonization or infection. *Microbiome* 4, 12.
- S59. Vogtman, E., Hua, X., Zeller, G., Sunagawa, S., Voigt, A.Y., Hercog, R., Goedert, J.J., Shi, J., Bork, P., and Sinha, R. (2016). Colorectal Cancer and the Human Gut Microbiome: Reproducibility with Whole-Genome Shotgun Sequencing. *PLoS One* 11, e0155362.
- S60. Wampach, L., Heintz-Buschart, A., Fritz, J.V., Ramiro-Garcia, J., Habier, J., Herold, M., Narayanasamy, S., Kaysen, A., Hogan, A.H., Bindl, L., et al. (2018). Birth mode is associated with earliest strain-conferred gut microbiome functions and immunostimulatory potential. *Nat. Commun.* 9, 5091.
- S61. Wen, C., Zheng, Z., Shao, T., Liu, L., Xie, Z., Le Chatelier, E., He, Z., Zhong, W., Fan, Y., Zhang, L., et al. (2017). Quantitative metagenomics reveals unique gut microbiome biomarkers in ankylosing spondylitis. *Genome biology* 18, 142.
- S62. Wirbel, J., Pyl, P.T., Kartal, E., Zych, K., Kashani, A., Milanese, A., Fleck, J.S., Voigt, A.Y., Palleja, A., Ponnudurai, R., et al. (2019). Meta-analysis of fecal metagenomes reveals global microbial signatures that are specific for colorectal cancer. *Nat. Med.* 25, 679-689.
- S63. Xie, H., Guo, R., Zhong, H., Feng, Q., Lan, Z., Qin, B., Ward, K.J., Jackson, M.A., Xia, Y., Chen, X., et al. (2016). Shotgun Metagenomics of 250 Adult Twins Reveals Genetic and Environmental Impacts on the Gut Microbiome. *Cell Syst* 3, 572-584.e573.
- S64. Yachida, S., Mizutani, S., Shiroma, H., Shiba, S., Nakajima, T., Sakamoto, T., Watanabe, H., Masuda, K., Nishimoto, Y., Kubo, M., et al. (2019). Metagenomic and metabolomic analyses reveal distinct stage-specific phenotypes of the gut microbiota in colorectal cancer. *Nat. Med.* 25, 968-976.
- S65. Yassour, M., Lim, M.Y., Yun, H.S., Tickle, T.L., Sung, J., Song, Y.-M., Lee, K., Franzosa, E.A., Morgan, X.C., Gevers, D., et al. (2016). Sub-clinical detection of gut microbial biomarkers of obesity and type 2 diabetes. *Genome Med.* 8, 17.
- S66. Yassour, M., Jason, E., Hogstrom, L.J., Arthur, T.D., Tripathi, S., Siljander, H., Selvenius, J., Oikarinen, S., Hyöty, H., Virtanen, S.M., et al. (2018). Strain-Level Analysis of Mother-to-Child Bacterial Transmission during the First Few Months of Life. *Cell Host Microbe* 24, 146-154.e144.
- S67. Ye, Z., Zhang, N., Wu, C., Zhang, X., Wang, Q., Huang, X., Du, L., Cao, Q., Tang, J., Zhou, C., et al. (2018). A metagenomic study of the gut microbiome in Behcet's disease. *Microbiome* 6, 135.
- S68. Zeevi, D., Korem, T., Zmora, N., Israeli, D., Rothschild, D., Weinberger, A., Ben-Yacov, O., Lador, D., Avnit-Sagi, T., Lotan-Pompan, M., et al. (2015). Personalized Nutrition by Prediction of Glycemic Responses. *Cell* 163, 1079-1094.
- S69. Zeller, G., Tap, J., Voigt, A.Y., Sunagawa, S., Kultima, J.R., Costea, P.I., Amiot, A., Böhm, J., Brunetti, F., Habermann, N., et al. (2014). Potential of fecal microbiota for early-stage detection of colorectal cancer. *Mol. Syst. Biol.* 10, 766.

- S70. Zhu, F., Ju, Y., Wang, W., Wang, Q., Guo, R., Ma, Q., Sun, Q., Fan, Y., Xie, Y., Yang, Z., et al. (2020). Metagenome-wide association of gut microbiome features for schizophrenia. *Nat. Commun.* *11*, 1612.
- S71. Dumas, E., Feurtey, A., Rodríguez de la Vega, R.C., Le Prieur, S., Snirc, A., Coton, M., Thierry, A., Coton, E., Le Piver, M., and Roueyre, D. (2020). Independent domestication events in the blue-cheese fungus *Penicillium roqueforti*. *Molecular ecology* *29*, 2639-2660.
- S72. Gallone, B., Steensels, J., Prahl, T., Soriaga, L., Saels, V., Herrera-Malaver, B., Merlevede, A., Roncoroni, M., Voordeckers, K., and Miraglia, L. (2016). Domestication and divergence of *Saccharomyces cerevisiae* beer yeasts. *Cell* *166*, 1397-1410. e1316.

(This page intentionally left blank)

**Paper III: A nontuberculous mycobacterium could solve the
mystery of the lady from the Franciscan church in Basel,
Switzerland**


(This page intentionally left blank)

RESEARCH ARTICLE

Open Access



A nontuberculous mycobacterium could solve the mystery of the lady from the Franciscan church in Basel, Switzerland

Mohamed S. Sarhan^{1*} , Christina Wurst¹, Alexandar Tzankov², Andreas J. Bircher^{3,4}, Holger Wittig⁵, Thomas Briellmann⁶, Marc Augsburger⁷, Gerhard Hotz^{8,9}, Albert Zink¹ and Frank Maixner^{1*}

Abstract

Background In 1975, the mummified body of a female has been found in the Franciscan church in Basel, Switzerland. Molecular and genealogic analyses unveiled her identity as Anna Catharina Bischoff (ACB), a member of the upper class of post-reformed Basel, who died at the age of 68 years, in 1787. The reason behind her death is still a mystery, especially that toxicological analyses revealed high levels of mercury, a common treatment against infections at that time, in different body organs. The computed tomography (CT) and histological analysis showed bone lesions in the femurs, the rib cage, and the skull, which refers to a potential syphilis case.

Results Although we could not detect any molecular signs of the syphilis-causing pathogen *Treponema pallidum* subsp. *pallidum*, we realized high prevalence of a nontuberculous mycobacterium (NTM) species in brain tissue sample. The genome analysis of this NTM displayed richness of virulence genes and toxins, and similarity to other infectious NTM, known to infect immunocompromised patients. In addition, it displayed potential resistance to mercury compounds, which might indicate a selective advantage against the applied treatment. This suggests that ACB might have suffered from an atypical mycobacteriosis during her life, which could explain the mummy's bone lesion and high mercury concentrations.

Conclusions The study of this mummy exemplifies the importance of employing differential diagnostic approaches in paleopathological analysis, by combining classical anthropological, radiological, histological, and toxicological observations with molecular analysis. It represents a proof-of-concept for the discovery of not-yet-described ancient pathogens in well-preserved specimens, using de novo metagenomic assembly.

Keywords Nontuberculous mycobacteria (NTM), Ancient DNA (aDNA), Bacteriophage, Syphilis, Brain infections, Mycobacteriosis, Franciscan church mummy, Anna Catharina Bischoff (ACB), De novo assembly

*Correspondence:

Mohamed S. Sarhan
mohamed.sarhan@eurac.edu
Frank Maixner
frank.maixner@eurac.edu

¹ Eurac Research – Institute for Mummy Studies, 39100 Bolzano, Italy

² Institute of Medical Genetics and Pathology, University Hospital Basel, University of Basel, 4031 Basel, Switzerland

³ Department of Allergology, University Hospital Basel, 4031 Basel, Switzerland

⁴ Faculty of Biomedical Sciences, Università della Svizzera italiana, Lugano, Switzerland

⁵ Department of Biomedical Engineering, Institute of Forensic Medicine, University of Basel, 4056 Basel, Switzerland

⁶ Citizen Science Basel; formerly Institute of Forensic Medicine, Forensic Chemistry and Toxicology, University of Basel, 4056 Basel, Switzerland

⁷ University Center of Legal Medicine, Lausanne, Geneva, Switzerland

⁸ Natural History Museum Basel, 4051 Basel, Switzerland

⁹ Integrative Prehistory and Archaeological Science, University of Basel, 4056 Basel, Switzerland



© The Author(s) 2023. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

Background

In 1975, a mummified corpse of a female individual was found in the Franciscan church (also known as the Barfüsser church) in Basel, Switzerland, during an excavation by the Archäologischen Bodenforschung Basel-Stadt in the church (Fig. 1A, B). The mummy’s coffin was encountered in a brick grave at a prominent position in the church (Fig. 1C, D), in front of the choir, along with another coffin that contained another human skeleton [1]. Genealogic

studies and molecular analyses unveiled the mummy’s identity identifying her as Anna Catharina Bischoff (ACB, 29.03.1719-30.08.1787), a member of the upper class of post-reformed Basel, who died at the age of 68 years [2, 3]. By checking historical records of the church, it turned out that the mummy had been discovered earlier during the nineteenth century; then, due to ethical concerns, it has been reburied where it was again found in 1975 [2]. During this reburial, the coffin had been covered with soil (Fig. 1C).

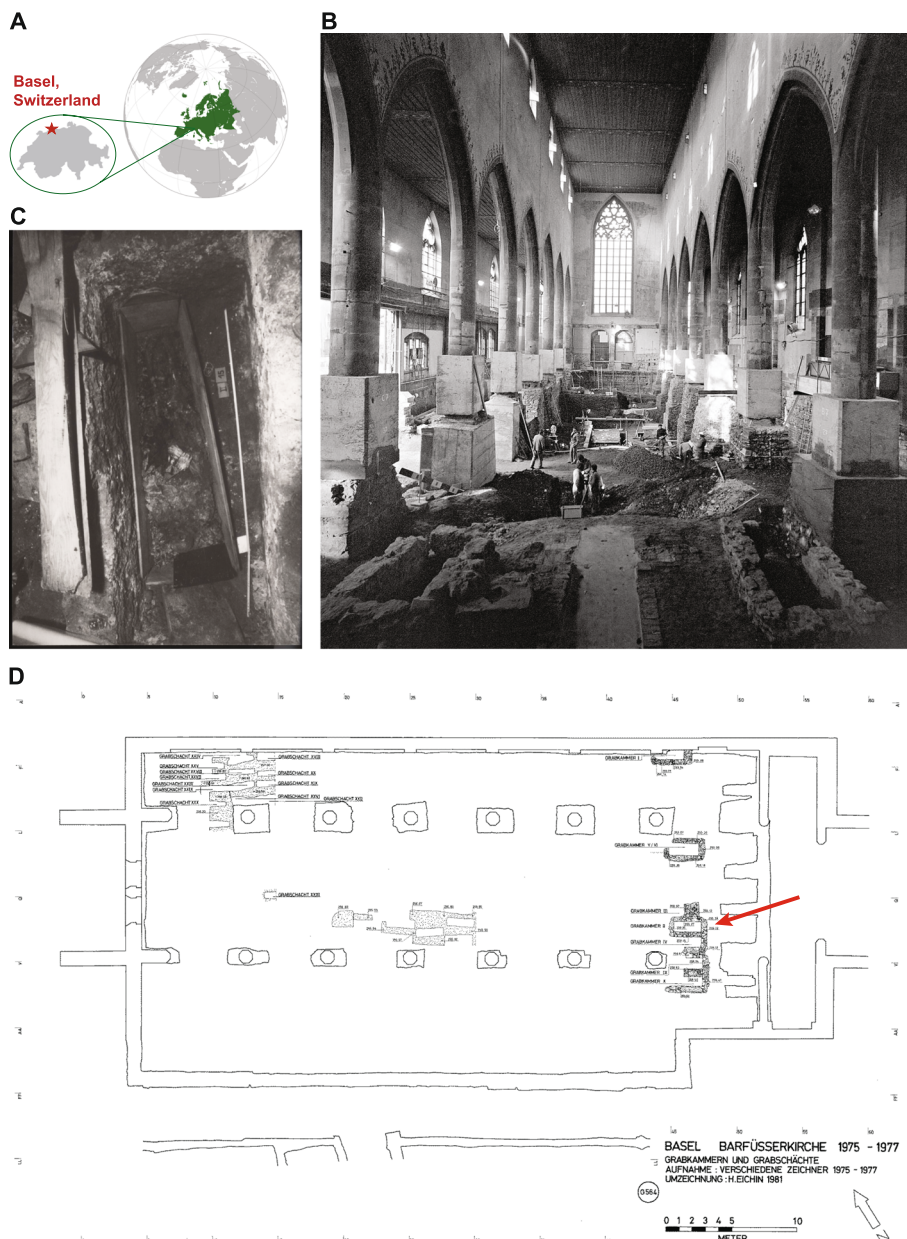


Fig. 1 Description of the mummy’s finding site. **A** Map of Europe with zoom-in on Basel in Switzerland. **B** The Franciscan church during the renovation in 1975. **C** Photograph showing the first glance on the mummy tomb; recognizable are the overlapped hands. **D** The location of the burial chamber (indicated by red arrow) inside the Franciscan church (© Archäologische Bodenforschung, 1975/6 – Plan “Grabkammer und Grabschächte (G564)”, modified by H. Eichin 1981)

The diseases and probable causes of death of ACB are still unknown, especially since the toxicological analyses of the mummy revealed high levels of mercury in different body organs [4], which is assumed to be the reason behind the mummification of the corpse (the antibacterial effect of mercury can slow down putrefaction), in addition to the dry conditions in the masonry grave shaft and the high salt concentration of the surrounding soils. Since mercury inhalation was a common treatment against infections, particularly syphilis [5], it was believed that she might have suffered from syphilis during her lifetime. This assumption was further supported by the computed tomography (CT) and histological results, which showed suggestive bone lesions in the skull, but degenerative lesions in femurs and rib cage. However, these lesions were histologically considered equivocal respecting late sequelae of syphilis infection, which led to additional differential diagnoses of a possible tuberculosis (TB) or Paget's disease [6].

In this study, we aimed to investigate molecularly whether the mummy could have suffered from syphilis, by carrying out a comprehensive shotgun metagenomic analysis on different body organs, in order to detect possible DNA traces of the causing pathogen, *Treponema pallidum* subsp. *pallidum*, or other pathogens that might have led to the bone lesions or might be linked to the mercury treatment.

Results

Human mitochondrial DNA analysis of the samples confirms their origin

Initially, we analyzed the mitochondrial DNA of all tissues in order to confirm that they contain authentic DNA. All the samples taken from the mummy showed the same mitochondrial haplotype (i.e., U5a1+!16192) as previously reported [3], except for the skull bone sample, which showed additional background contamination with other human DNA (Additional file 1: Table S1) [3]. Interestingly, analysis of the human DNA of the maggots (Sample ID 3169), which were collected from underneath the mummy, revealed the same mitochondrial haplotype as of the mummy, indicating initial feasting on the decayed flesh of the mummy [7] or potential diffusion of mummy's DNA into the surroundings [8]. The other control samples, coming from the skeletons of other individuals from the same tomb, showed two different mitochondrial haplotypes (Additional file 1: Table S1).

Metagenomic analysis did not reveal any *Treponema* genomes but unusual high prevalence of *Mycobacteriaceae* in brain tissue

Driven by the radiological and histological observations and the toxicological analysis (Fig. 2A, B), we carried

out a shotgun metagenomic screening of different samples from different body parts of the mummy, representing tissues where the infection can be expected to occur (please refer to Additional file 1: Table S1 for details). Although most of the tissues displayed overwhelming prevalence of postmortem microbial communities, e.g., Clostridia, we were still able to spot some tissue-specific taxa, particularly in the gut (*Turicibacter sanguinis* and *Ruminococcus gnavus*) and the tooth (e.g., *Prevotella denticola* and *Actinomyces dentalis*) samples (Additional file 1: Table S2). Additionally, we did not find any metagenomic reads assigned to the syphilis-causing pathogen *Treponema pallidum* subsp. *pallidum* (Additional file 2: Figure S1), nor even the containing-family *Spirochaetaceae*, except for the tooth sample, which displayed presence of other *Treponema* species, e.g., *T. socranskii*, *T. denticola*, and *T. maltophilum*, which are all known to be linked to periodontitis and being members of the oral microflora [9, 10].

Unexpectedly, the brain tissue displayed exceptional high abundance of the family *Mycobacteriaceae*, representing more than 80% of the total microbial metagenomic reads (Fig. 2C, Additional file 2: Figure S2) [11–13]. In this respect, it is important to mention that the toxicological analysis displayed the highest mercury concentrations in the brain samples, i.e., up to $28 \text{ ng} \times \text{mg}^{-1}$ tissue material (Fig. 2B), since the brain is the target organ in the uptake or administration of elemental mercury. This opens a question on whether these two observations are correlated.

De novo assembly revealed a nontuberculous mycobacterium (NTM) in the brain

Based on the high abundance of *Mycobacteriaceae* in the brain, we performed a de novo metagenomic assembly on the brain metagenomic reads (for further information on assembly, please refer to the "Methods" section). We were able to resolve a high-quality metagenome-assembled genome (MAG, 99.5 % completeness and < 0.5 % contamination, as estimated by CheckM), belonging to a nontuberculous mycobacterium (NTM) species (Fig. 3A, B, Additional file 1: Table S3). Interestingly, more than 57% of the brain metagenomic reads were mapped against the assembled genome (Additional file 1: Table S4). The genome constituted of 66 contigs of total size of $\sim 4.8 \text{ Mb}$ and mean coverage of $185.9 \pm 45.4 \times$ (Additional file 1: Table S5).

Considering that the genus *Mycobacterium* has undergone major phylogenomic-based taxonomic reassignments and rearrangements [16, 17], we used species-representative genomes of the whole *Mycobacteriaceae* family in order to gain an in-depth taxonomic characterization of our discovered genome within it,

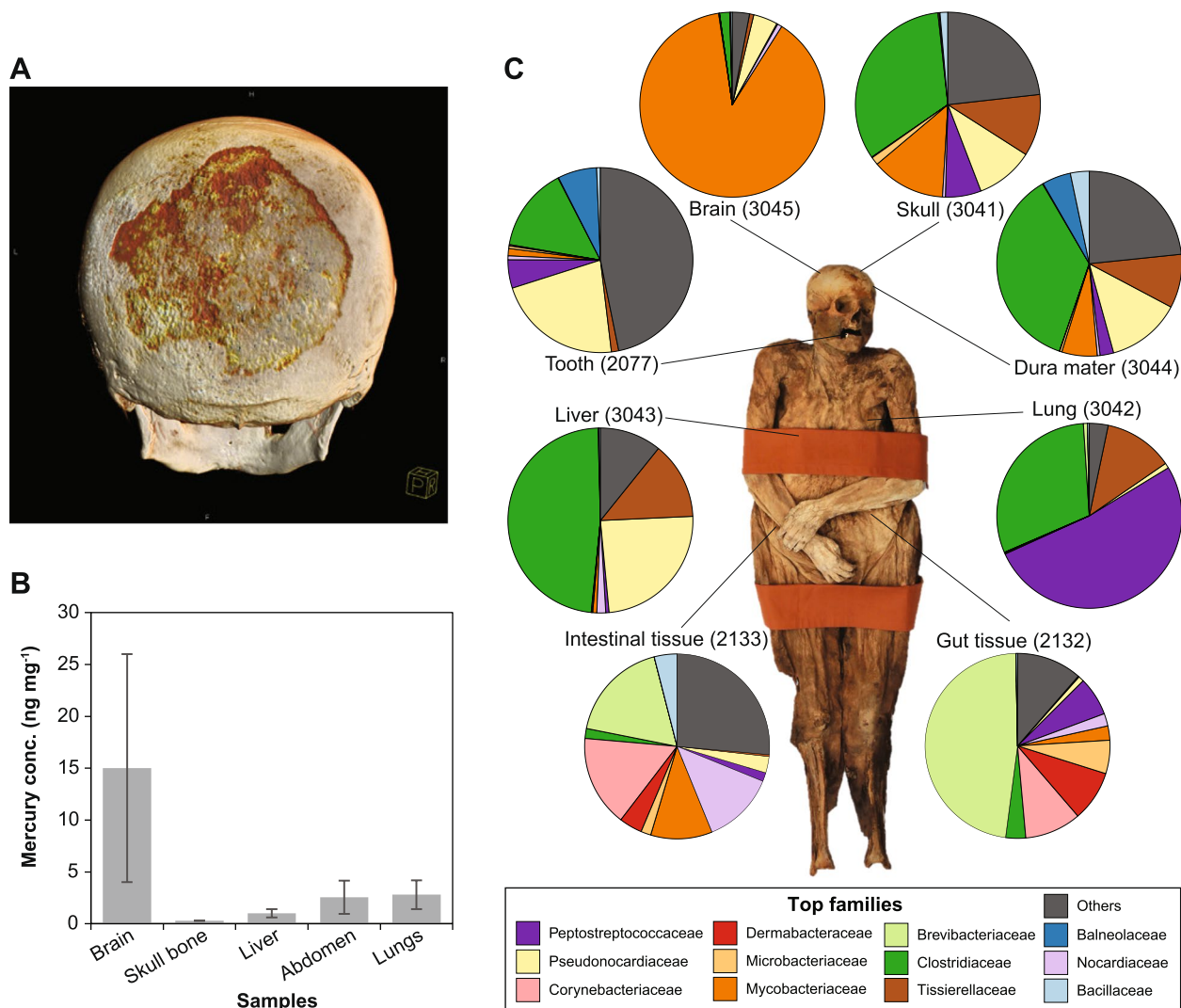


Fig. 2 Overview on the radiological, toxicological, and microbiological characteristics of the ACB mummy. **A** Computed tomography (CT)-based three-dimensional reconstruction of the skull. Notice the darker colors which represent lower x-ray densities than healthy bone. Copyright: Holger Wittig, Institute of Forensic Medicine, University of Basel. **B** Concentration of elemental mercury in different body samples, where the error bars refer to the standard errors. **C** Relative abundances of the top 12 microbial families on different body tissues as inferred by number of shotgun-metagenomic reads compared against the nr-database (please refer to the “Methods” section for details). Numbers in parentheses refer to sample IDs (please refer to Additional file 1: Table S1 for further details)

using PhyloPhlAn marker genes (please refer to the “Methods” section). The resulting phylogeny was highly congruent with the recently proposed taxonomy [16], having main clades representing the classical human pathogenic *Mycobacterium* spp., *Mycobacteroides* spp., *Mycolicibacter* spp., and finally *Mycolicibacterium* spp., where our genome falls within (Fig. 3B). Further, we compared our genome with all characterized species within the genus *Mycolicibacterium* to find the closest relative within the genus. For this purpose, we carried out a pairwise genomic comparison between all species including ours, using the Mash distance tool [18]. The

mash distances of the brain NTM genome is relatively distant from all the characterized species within the genus (minimum distance = 20.03%) and clusters close to *M. thermoresistibile*, *M. hassiacum*, *M. agri*, and *M. moriokaense* (Fig. 3C). Although the four species are known to be, like other members of the genus, environmental bacteria, they have previously been reported to cause infections in humans (Fig. 3C, and for detailed examples, please refer to Additional file 1: Table S6) [19–94]. Moreover, most of *Mycolicibacterium* spp. have been previously isolated from hosts (Additional file 1: Table S6).

Genome-wide analysis indicates potential virulence of the brain NTM

Before subjecting the brain genome to further functional analysis, we checked for the authenticity of this bacterium (henceforth referred to as “brain NTM”), i.e., being ancient on the one hand, and on the other being exclusively present in the brain, not being a contaminant from other tissues or even from the environment surrounding the burial site. Therefore, we tested the terminal deamination levels on the metagenomic brain reads mapped to the genome of the brain NTM (the “Methods” section). We noticed very low levels of terminal C-to-T and G-to-A substitutions, even less than the human DNA damage in this tissue (Fig. 3D, Additional file 2: Figure S3), although the fragment length distribution of the brain NTM was lower than of the human autosomal DNA (Fig. 3D, Additional file 2: Figure S3). When we further compared the human autosomal DNA from different organs, we realized variable levels of ancient DNA (aDNA) damages, correlating with the variable concentrations of mercury in different organs.

To further assess the possibility of the brain NTM being an external contamination, we investigated in addition to the mummy tissues more samples representing the following (Additional file 2: Figure S4) [2]: (i) bone and tooth samples from other skeletons found in the same burial site (upper coffin); (ii) textile sample and maggots that were found on the mummy; and finally, (iii) soil samples that were found covering the upper parts of the mummy. After mapping all metagenomic reads of each of the aforementioned samples against the brain NTM genome and considering threshold of minimum 3x coverage (please refer to the “Methods” section for details), we did not find any sufficient breadth (i.e., > 50% of the genome covered at least 3 times) for any of the mummy’s samples, except for the skull bone and dura mater

samples, which appeared to carry the bacterium, having average breadth values of > 70% (Fig. 4A). Thereby, we excluded the possibility of the external contamination and continued with the functional analyses.

We checked the overall virulence potential of the bacterium, by comparing all coding sequences (CDS) against the virulence factor database (VFDB) [95]. The genome contained three different clusters of type VII secretion systems (T7SSs), which are responsible for effector proteins in pathogenic and non-pathogenic mycobacteria, and help to survive in the host by evading the immune system [96]. Considering the genome of *M. tuberculosis* as a reference, our genome contained the core genes of the ESX-1 system as well as the full genes of the ESX-3 and ESX-4, with the same exact synteny arrangement (Additional file 2: Figure S5).

Since the NTM genome was exclusively present in the brain, we checked for the ability of the bacterium to invade the brain and cross the blood brain barrier (BBB). Be and colleagues identified the genes Rv0311, Rv0931, Rv0986, and Rv0805 (CpdA) in *M. tuberculosis* to be significantly involved in brain invasion and survival [97]. Indeed, we detected homologs of the four genes scattered throughout the genome (Fig. 4A). Moreover, it has been also previously reported that NTM can invade brain within circulating macrophages [98].

We additionally checked the brain NTM genome for the presence of toxin-antitoxin (TA) systems that are assumed to be contributing to the virulence of mycobacteria [99]. In comparison with other members of the genus *Mycobacterium*, and representatives of other known pathogenic mycobacteria, we realized that the brain NTM genome together with *M. tuberculosis* H37Rv were on the top in richness of TA systems (Additional file 2: Figure S6). Generally, TA systems are typically composed of a protein (toxin) and another antagonistic protein (antitoxin). Under stress conditions, the antitoxin,

(See figure on next page.)

Fig. 3 Genome-level taxonomic assignment of the brain bacterium. **A** Taxonomic assignment of the brain NTM contigs as assigned by searching against the NCBI-nt database, using BLASTn. The number of the assigned contigs is shown next the taxon names based on the lowest common ancestor (LCA) as determined by MEGAN (please refer to the “Methods” section for further details). **B** Unrooted phylogenetic tree of the family *Mycobacteriaceae*, including a single representative genome of each species, based on PhyloPhlAn marker genes. The background colors of the clades refer to: red, *Mycobacterium* spp.; yellow, *Mycolicibacter* spp.; green, *Mycobacteroides* spp.; and blue, *Mycolicibacterium* spp. **C** Heatmap-based on MASH distances of all characterized species’ genomes within the genus *Mycolicibacterium* including the genome of the brain bacterium (highlighted in bold red font). The heatmap annotations to the left of the heatmap refers to whether the microbe was isolated from a host or was reported as a human pathogen. For further details on the isolation sources, please refer to the Additional file 1: Table S6. **D** Damage plots of human DNA of different tissues as well as the brain NTM. The damage plots were generated considering the mapped reads of the indicated tissues different body tissues (i.e., tooth, intestinal tissues, skull, dura mater, and brain) against human genome (hg19) while the brain NTM was generated considering the brain metagenomic reads mapped against the brain NTM assembled genome. Ancient DNA damage represented by the terminal substitution of Cytosine to Thymine at the 5’ ends of the DNA fragments. The labels in parenthesis refer to sample ID and the mercury concentrations \pm standard error. For further information on the read lengths distribution please refer to Additional file 2: Figure S3 [14, 15]. Note: The human DNA are showing variable levels of ancient DNA damages, despite of being of the same individual. The lowest levels of the human DNA damages are in the brain and dura mater samples, which goes with the abundance of the brain NTM and the mercury concentrations as well

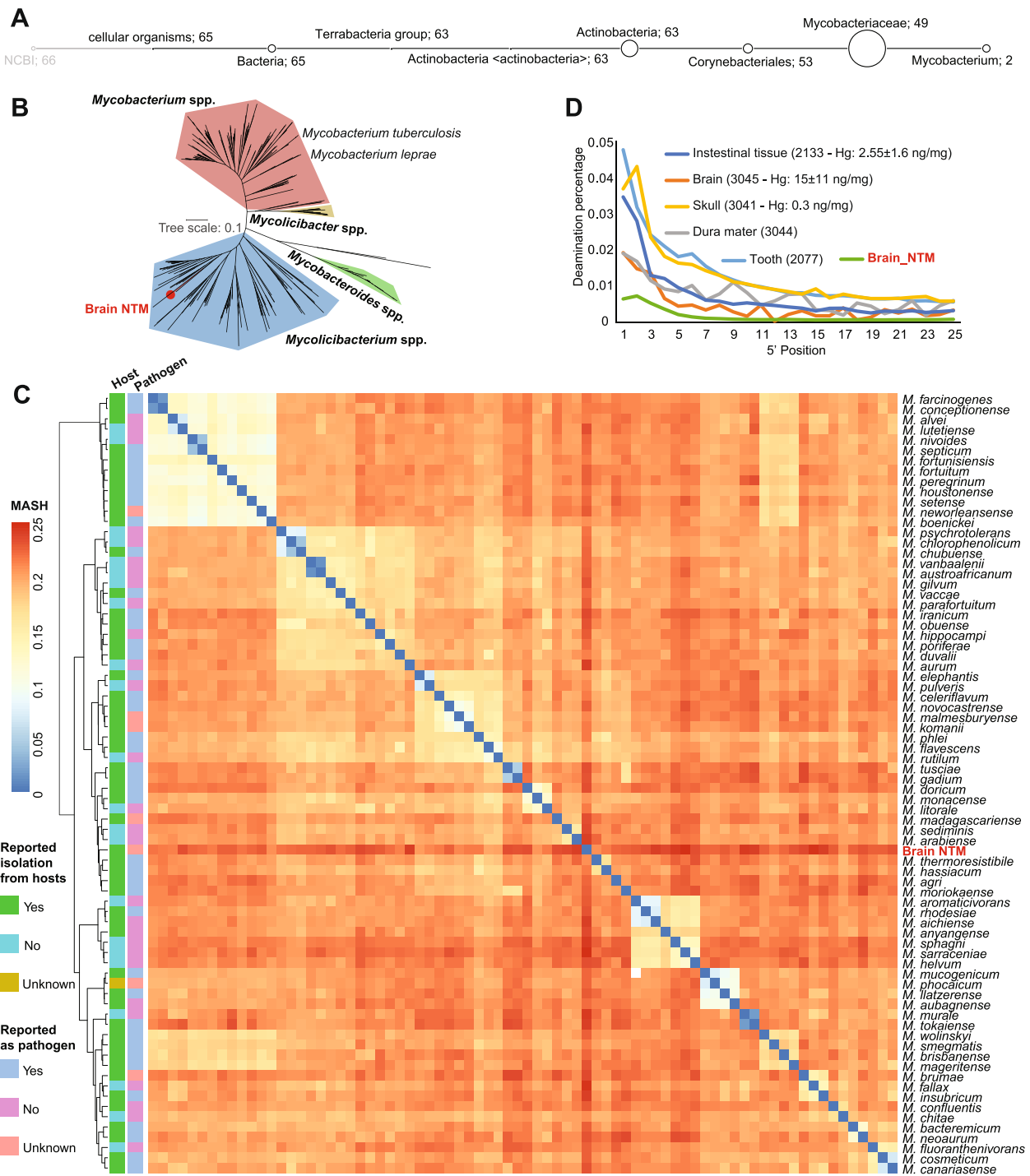


Fig. 3 (See legend on previous page.)

which is a labile molecule gets degraded, allowing the toxin molecule to halt essential cellular processes, leading to growth inhibition of bacteria [99]. Such process is reversible, which means that as soon as the conditions become favorable again, the antitoxin binds to the toxin and maintains essential cellular processes are restored.

It is postulated that such TA systems play a great role in bacterial persistence under stress conditions, particularly in *M. tuberculosis* infections, which sometimes mediates dormant states tolerant to antibiotic treatment and the host's immune response, i.e., latent tuberculosis. Analysis of transcriptomic profiles of antibiotic-induced

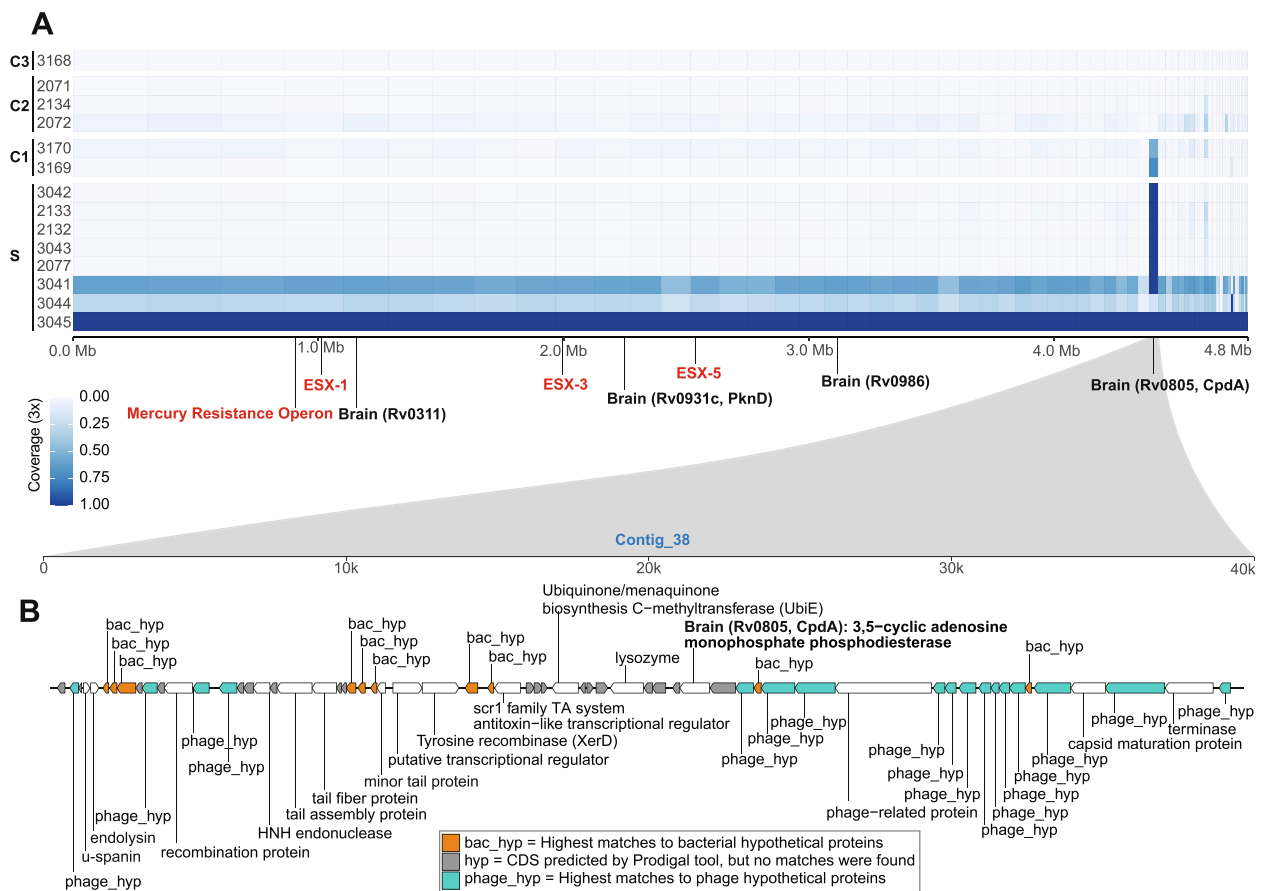


Fig. 4 Genetic map of the brain bacterium genome. **A** The heatmap shows distribution of the DNA of the brain NTM in the mummy's samples as well as other control samples; The numbers to the left refer to the sample ID as referred to in Additional file 1: Table S1, while the letters refer to the samples group as follows: S, the mummy's tissue samples; C1, control samples that were in contact with the mummy; C2, samples that were taken from other individuals from the same burial site; C3, soil sample that was collected from inside the coffin of the mummy (please refer to Additional file 1: Table S1 and Additional file 2: Figure S4 for further details). The presence is shown by breadth values with minimum coverage of 3x per site. The contigs are arranged in a descending order from left to right (for details on the contigs, please refer to Additional file 1: Table S5). The loci referred to in bold red fonts are gene clusters/operons and are further detailed in Additional file 2: Figure S5, while those in black bold are genes that are known to be involved in crossing blood brain barrier (BBB) and brain invasion. **B** The genetic map of the contig 38, which is assumed to be a phage genome, as inferred by the tool PHASTER (for further details on the phage annotation, please refer to Additional file 1: Table S7)

persistence in *M. tuberculosis* revealed upregulation of 10 TA systems [100], which supports the assumptions of them being involved in pathogen persistence. In general, *M. tuberculosis* and ACB's brain NTM contain remarkably higher numbers of TA systems, compared with other mycobacteria, so we can deduce that it may have caused a persistent infection, by entertaining a privilege over other microbes in terms of survivability and competitiveness in extreme conditions.

Finally, and in relation to the high mercury concentrations in the brain, we checked for the mercury resistance genes in the genome of the brain NTM in comparison with the mercury resistance operon of the *M. marinum* plasmid. We found in addition to the regulator protein gene and the core genes, i.e., mercuric reductase and

alkylmercury lyase, other heavy metal resistance genes and antibiotic resistance genes (Additional file 2: Figure S5). The heavy metal resistance genes were mainly responsible for assimilation of copper, zinc, and arsenic, while the antibiotic resistance genes were directed against vancomycin, tetracycline, and beta-lactam antibiotics (Additional file 2: Figure S5).

Evidence for a pre-infection transduction event that possibly increased the virulence

Although all mummy's samples (except the brain, dura mater, and skull bones) seemed to be void of the brain NTM, they all displayed remarkable presence of one contig, i.e., contig 38, which indicates spread throughout the body, likely through the blood stream. To check the

possibility that this contig might have been misplaced during the binning step, we went back to the metagenomic contigs and clustered them based on the coverage and the GC content, to whether this contig is an outlier to the rest of the genome. Considering both assembly methods, the contig was placed perfectly within the defined cluster of the brain NTM, showing similar coverage and GC content (Additional file 2: Figure S7). As further confirmation on the authenticity of the brain NTM, the contig 38 was completely absent from the samples of the other individuals and the soil sample. To explain the presence of this contig throughout the body, we hypothesized that it could be a circulating phage, since bacteriophages were shown to be present in the stool of patients with pulmonary diseases and to migrate through blood stream [101, 102]. Therefore, we screened for presence of viral sequences within our genome using the tool PHASTER (see the “Methods” section) and realized that the contig_38, which showed high prevalence in other mummy tissues, highly resembles phage genomes (Fig. 4B). Further, the CheckV tool displayed a high-quality viral genome with 91.35 completeness (Additional file 1: Table S8), having additionally other host/non-viral genes within the genome. By deeply annotating the contig, we found the basic structural proteins of phages, e.g., capsids, tail fibers, and portal proteins, in addition to other functional proteins, e.g., terminase, recombination proteins, and endolysin (Fig. 4B).

Interestingly, within this contig, there was the homologous gene of Rv0805 (CpdA), one of the genes that is involved in brain invasion and persistence. The presence of prophages has been recently reported in the genomes of clinical mycobacteria and assumed to aid in their virulence evolution. Moreover, when compared with environmental mycobacteriophages from PhagesDB, bacterial virulence genes were found to be enriched in the clinical mycobacterial prophages [103, 104]. This might indicate that the brain NTM has undergone a natural pre-infection transduction process that increased its virulence and enabled invasion and microbial survival in the brain.

Discussion

The mummy of the Franciscan Church from Basel, Switzerland (ACB, 1719–1787) represents a unique example of multidisciplinary studied mummy [2]. Since the mummy is dated back to the eighteenth century, extensive historical, genealogical, and molecular investigations were necessary that resulted in reconstructing a family tree over 22 generations including nowadays living relatives [3]. The analyses showed that she belonged to a wealthy upper-class family, which might indicate that she might have had access to an advanced and special healthcare. The initial radiological and histological analyses of

her mummified body indicated combination of different diseases, such as atherosclerosis and gallstones, and the toxicological analysis showed high concentrations of mercury in different body parts. Thus, studying the paleopathological status and the treatments she received provide insights into an important time in the European history, as she lived during the scientific revolution and before the onset of the industrial revolution.

One of the main intriguing findings was the high mercury concentration that has been revealed even at the mummy’s discovery time in 1975 [105] and was assumed to be one of the main reason for mummification. The presence of mercury also triggered some questions regarding her health status and in which context she had been exposed to these high concentrations of mercury. Historically, the beginning of mercury usage in the medical fields in Europe dates probably back to 1495, at the beginning of the first documented syphilis outbreak [106]. During that time, mercury had been already in use in the Arabic medicine to treat skin infections and leprosy [5, 107]. Therefore, it was adopted in Europe to treat syphilis and other infectious diseases by different methods of treatments, such as topical application, inhalation of mercury vapors, or even by ingestion of mercury salts [5, 108–110]. From the toxicological point of view, humans can get exposed to different forms of mercury, i.e., inorganic (e.g., mercury chloride), organic (e.g., methylmercury), or elemental (e.g., mercury vapors) [111]. The elemental form is absorbed by inhalation, passes from the lungs into the blood, crosses the blood-brain barrier and accumulates in the brain [112, 113]. While the inorganic forms, due to their lower liposolubility, were hypothesized to need to be accompanied with selenium (Se) that neutralizes the mercury’s toxicity to be able to access and cumulate in the brain [114]. On the other hand, mercury in its ionic forms is mainly accumulated in the kidney and liver. Considering that we found higher mercury concentrations in the brain than in the other body samples, we postulate that ACB was exposed to mercury vapors as a treatment for an extended time period. However, we cannot exclude that ACB was exposed to a treatment with an ointment containing colloidal elemental mercury or another inorganic mercury treatment, considering the correlation between the concentrations of the selenium and mercury, particularly in the brain of ACB (Additional file 2: Figure S8).

The classical explanation for the presence of mercury in the mummy would be the syphilis, which is mainly a sexually transmitted bacterial disease caused by *Treponema pallidum* subsp. *pallidum* [115]. Its symptoms are highly variable and can range from small chancres and nodular granulomatous skin lesion to late-onset prominent bone lesions [116]. Syphilis is also known as one of the “great

imitators,” as it causes symptoms similar to other diseases making its diagnosis challenging [117]. For instance, in 1885, a bacillus bacterium was isolated from a syphilitic chancres [118, 119], which was assumed to be the causing pathogen, but later and by further research turned out to be the non-tuberculous *M. smegmatis* that causes very similar ulcers to those of *T. pallidum* subsp. *pallidum* [77]. In our case, we could not find any metagenomic reads that can be assigned to *T. pallidum*, and the only partially supporting clues would be the bone lesions unveiled by the radiological analysis with yet equivocal histopathology and the existence of mercury itself, where both are indirect evidence on syphilis. However, it might be useful in the future to consider employing a target enrichment capture approach particularly for those hardly to recover pathogens [120].

The other possible explanation for the mercury treatment can be deduced from the extraordinarily high abundance of *Mycobacteriaceae* in the brain of ACB as well as scattered incidence in the dura mater and the skull bone samples. By employing de novo metagenomic assembly, we resolved a near-full NTM genome. In general, NTM are considered as environmental bacteria, which inhabit different niches like soils and water. Many of the NTM were reported to cause human infections but remained neglected for decades until their recent remarkable emergence in clinical infections [121, 122]. NTM affect primarily human lungs, causing pulmonary disease, but can also infect other body parts, particularly in immunocompromised patients [123]. In the ancient metagenomics and paleopathology fields, NTM are usually mentioned within environmental contexts or as contaminants, due to the difficulty to confidently assign them as true pathogens like *M. tuberculosis* [124].

However, in this study, we opted to extensively analyze the genome of the brain NTM to explain their unusual existence and high abundance in the brain. The primary approach to check the DNA of a genome of whether it is ancient, is to analyze the ancient DNA damage pattern, i.e., terminal 5'C-to-T deamination. When we checked the damage pattern of the brain NTM, we found low damage levels (Fig. 3D). Interestingly, when we checked the damage of the human DNA of different organs of the mummy, we realized variability in the damage levels, negatively correlating with the mercury concentrations. For instance, the highest mercury concentrations were found in the brain samples, where we found the lowest human DNA damage (please refer to Fig. 3D for further examples). Despite of being a rare finding, this observation could explain the low damage levels on the brain NTM and highlights the inter-body variability of the DNA damages, which might open a discussion on using the ancient DNA damage as a primary tool for assessing

modern contamination, particularly when dealing with unusually mummified materials.

In our case, the NTM contamination possibility can be toned down as soon as we consider the following: (i) the extraordinarily high abundance in the brain; (ii) the exclusive incidence in the brain and the close-by tissues and the absence from the other control samples as well as other organs of the mummy; (iii) the distribution of the putative phage within the mummy's tissues and the absence from the control samples; (iv) the virulence potential of the genome particularly including genes that can enable crossing the BBB; (v) the richness of TA-systems that presumably enabled persistence under stress and unfavorable conditions; and, finally, (vi) the mercury-, heavy metal-, and antibiotic resistance genes of the bacterium. Therefore, we can gain confidence on the origin of the bacterium and its potential to be a brain pathogen that survived despite of the mercury treatment.

Overall, and based on the radiological, histological, toxicological, and molecular analyses of the ACB's mummy, the current situation might suggest two different scenarios that explain her health status in relation to the high mercury concentrations. First, ACB suffered from syphilis and was exposed to a mercury treatment successfully, and *T. pallidum* was completely eradicated, while the NTM brain infection occurred later. Second, she suffered from an NTM, and probably showed symptoms similar to syphilis, and was therefore subjected to mercury treatment, but the pathogen survived the mercury due to its content of mercury and heavy metal resistance genes. The first scenario can be supported by the radiological findings, i.e., the bone lesions of the skull and the ribs; however, it is not supported at the molecular level. While the second scenario is highly supported by the molecular analysis, and radiological signs could also support this hypothesis.

Conclusions

In conclusion, this study spots the light on one NTM as one of the, not only nowadays, cause of neglected diseases and infections, which might have been misdiagnosed as syphilis or tuberculosis during the eighteenth century and may still be overlooked or misinterpreted nowadays in paleopathological studies due to the guided interest in more common and better-known diseases. The study of this mummy exemplifies the importance of employing differential diagnostic approaches in paleopathological analysis by combining classical anthropological and radiological observations with molecular investigations. It also demonstrates the value of employing comparative metagenomic approaches in analyzing multiple samples from one individual. Finally, the study shows the value of employing de novo metagenomic assembly in recovering

extinct and not-yet described ancient genomes, in well-preserved specimens, and gaining better insights on their phylogenetic and functional characteristics that are often beyond the potential of other tools.

Methods

Mummy sampling

The analyzed tissues and bones have been sampled at the Natural History Museum of Basel (Additional file 1: Table S1). The sampled tissues were checked visually and histologically to confirm their tissue of origin. Additional control samples were taken from the soil surrounding the mummy and the textile found underneath the body. Moreover, bone and teeth samples were taken from the other skeletons found in another coffin in the same burial site (please refer to Additional file 2: Figure S4).

DNA extraction, library preparation, and DNA sequencing

Amounts of 9–200 mg of the ancient materials or the control samples were used for DNA extraction, following the protocol described by Maixner and colleagues [125]. DNA extracts were quantified using QUANTUS (Promega, USA); then, 20 μ l of each extract were converted into double-indexed DNA libraries following a special protocol for highly degraded ancient DNA [126]. All previous steps were carried out in the ancient DNA laboratory of the Eurac Research Institute for mummy studies in Bolzano, Italy. The double-indexed libraries were subjected to next generation DNA sequencing using HiSeqX (2 \times 150 paired-end), resulting in sequencing depths of 9–51 \times 10⁶ paired-end reads per sample (Additional file 1: Table S1).

Processing of ancient DNA sequences

We used the tool fastp [127] to trim the adapters and low-quality reads and to merge paired-end reads with at least 10 nucleotides overlap. Then, the quality filtered merged reads were de-duplicated and filtered for minimum sequence length of 25 nucleotides, using SeqKit [128].

Human DNA analysis

To check the authenticity of the analyzed materials, we included, as a reference, in our analyses the shotgun metagenomic dataset generated from a tooth sample that was used to reveal the identity and the mitochondrial haplogroup of the mummy (<https://www.ebi.ac.uk/ena/browser/view/PRJEB44723>) [3]. For the analysis, we mapped the quality-filtered deduplicated merged reads against the human reference genome (build hg19) [129] and the human reference mtDNA genome (rCRS) [130] using Burrows-Wheeler Aligner (BWA) [131]. Then, we used SAMtools to filter for minimum mapping quality

of 30, QualiMap to generate basic mapping statistics [132], and mapDamage 2.0 [14] to quantify the percentages of C-to-T and G-to-A substitution of the mapped ancient DNA reads. In the case of mitochondrial DNA, we additionally used option “--rescale” to rescale the quality scores of the damaged mis-incorporated sites and the tool Schmutzi to estimate the level of contamination based on deamination patterns [133]. For haplogroup assignment, we first converted the rescaled bam files into variant calling format (VCF) and then employed HaploGrep 2.0 [134].

Moreover, and to confirm the sex of the mummy samples, we used the mapped human DNA reads to compute the karyotype frequency of X and Y chromosomes, using a Maximum likelihood method [135].

General microbial profiling

To have an overview on the microbial composition of the samples, we used the search tool BLASTX of DIAMOND v2.0.13 [136] to compare our metagenomic reads, with the default parameters, against the nr-protein database [137]. Then, we used MEGAN v6.21.16 [11] to process the DIAMOND outputs and to assign the metagenomic reads to their lowest common ancestor (LCA), with the parameters “--minPercentIdentity 97” and “--minSupport 10”. Since the BLASTX compares nucleotides against amino acid sequences, we confined our LCA filters to the family-level to minimize the false positives. The number and percentages of assigned reads are summarized in Additional file 1: Table S4.

De novo assembly of metagenomic reads

We used the quality-filtered unmerged reads of the brain sample (Eurac ID 3045, 26411722 pair-end reads) to perform de novo sequence assembly, using the metagenomic assemblers MEGAHIT [138] and SPAdes with “--meta” option [139]. All contigs shorter than 1000 nt were excluded from the downstream analyses. The metagenomic bidders MetaBAT2 [140], MaxBin2 [141], and CONCOCT [142] were used to resolve potential genomes, based on similarity of abundance and tetra-nucleotide frequency. Then, the DAS Tool was used to resolve consensus bins from each assembler, independently (Additional file 2: Figure S7).

Further, we checked the completeness and contamination of the resulting bins using CheckM [143] and kept only the high-quality genomes (completeness > 90% and contamination < 5%). CheckM checks for the presence of single-copy genes (SCG) specific for each lineage. If the program finds more than one copy of any of those SCG, it considers this as a potential contamination. To calculate the strain heterogeneity, the program calculates the relatedness between the multiple copies of those SCG, based

on amino acid similarity. Then, the function “*classify_wf*” of GTDB-Tk v1.5.0 (April 23, 2021) has been used, with default parameters, to classify the resulting bins taxonomically [144, 145].

As a result, two of the bins displayed identical taxonomic classification (*g_Mycobacterium* sp.) and MASH distance of < 0.01, each is resulting from different assembler, i.e., MEGAHIT and SPAdes. Therefore, we used SeqMan tool of DNASTAR [146] to reassemble the contigs of both bins, by aligning all contigs vs all contigs, in order to improve the assembly quality, e.g., N50 value and number of contigs. Finally, we ended up with near-full genome with 99.55% completeness and 0.45% contamination (Additional file 1: Table S3).

To check the taxonomy of the resulting contigs, we searched all the contigs against the NCBI-nt database [137]. Then, the module “*blast2rma*” of MEGAN v6.21.16 [11] with the parameters “*--minPercentIdentity 80*” has been applied to assign the contigs to their lowest common ancestor (LCA).

Phylogenetic analysis

To classify the resulting genome phylogenetically, we used the PhyloPhlAn 3.0 tool [147] to retrieve representative genomes of each species within the family Mycobacteriaceae. Then, we built the phylogeny based on the 400 universal PhyloPhlAn marker genes using the option “*--diversity low --accurate*”. The configuration file included the following tools with the default parameters set by PhyloPhlAn: DIAMOND v2.0.13 [136], MAFFT v7.427 [148], trimAl v1.4.1 [149], FastTree v2.1.11 [150], and RAxML v8.2.12 [151].

Based on the PhyloPhlAn phylogenetic assignment, we retrieved representative genomes of all characterized species belonging to the genus *Mycolicibacterium* and added our brain bacterium genome, then calculated MASH distances all vs. all, using the “*phylophlan_metagenomic*” module of PhyloPhlAn 3.0 [18].

Checking the brain bacterium in the control samples

To check the presence of the brain bacterium in other tissues than the brain or in the control samples, we mapped the quality-filtered reads of each sample against the sequence of the brain bacterium, using Bowtie2 [152], with the option “*--end-to-end*.” Then, we sorted and indexed the resulted bam files using SAMtools [153], including a minimum mapping quality of 30. We used the script CMSeq (<https://github.com/SegataLab/cmseq>) to calculate the depth and breadth of the mapping. Positions were considered as true covered positions if they were covered at least 3 times. Afterwards, we calculated the breadth per contig and plotted the breadth as a heatmap (Fig. 4).

Finally, we used the tool DamageProfiler [15] to check for ancient DNA damage patterns, i.e., C-to-T and G-to-A substitutions, resulting from cytosine deamination.

Bacterial genome annotation

In addition to the genomes of the *Mycolicibacterium* spp., we included the genome of *Mycobacterium tuberculosis* H37Rv from the Genbank databases (NCBI Reference Sequence: NC_018143.2) and *Mycobacteroides abscessus* (NCBI Reference Sequence: NZ_CP034180.1), in the genome annotation analyses. We used the Prokka pipeline to annotate the genomes [154], implemented Prodigal for gene prediction [155], and RNAmmer to find ribosomal RNA genes [156].

To search for the virulence genes, we compared all coding sequences against the virulence factor database (VFDB) [95], using mmseqs2 with its default parameters. To annotate the type VII secretion system gene clusters, we manually extracted the regions from *M. tuberculosis* H37Rv and our brain bacterium genome, based on Prokka annotation and compared them in a pairwise manner, using the BLASTp (all vs. all). Visualization of gene synteny was done using Clinker [157]. Following the same previous approach, we compared the mercury resistance operon in our brain bacterium genome to the well-characterized mercury resistance operon of *M. abscessus*.

To find the genes that are potentially involved in brain invasion [97], we manually retrieved the genes (*Rv0311*, *Rv0805*, *Rv0931c*, and *Rv0986*) from the Mycobrowser database (<https://mycobrowser.epfl.ch>). The tool OrthoFinder [158] was used to find the homologous sequences in our brain bacterium genome as well as the other *Mycolicibacterium* spp. genomes, including the *M. tuberculosis* and *M. abscessus*.

To check the presence of toxin/antitoxin (TA) systems in the analyzed genomes, we compared all coding sequences inferred by Prokka against the TA database (TADB 2.0), and confined the analysis to the experimentally validated type II TA loci [159]. We used mmseq2 for comparison using default parameters.

Viral genome annotation

To search for viral sequences within the genome of the brain bacterium, we used the tool PHASTER (PHAge Search Tool Enhanced Release) [160], which predicts and annotates phage genes with comparison to curated phage and bacterial gene databases (<https://phaster.ca>). Additionally, we used the tool CheckV [161] to check the quality and completeness of the potential phage genomes. Then, we used different approaches to perform functional annotation of the contig which has been assigned as phages: (i) we used Prokka standard annotation as

described previously; (ii) we used the PHASTER annotation tool; (iii) we used the tool MicrobeAnnotator [162] implementing a DIAMOND search against different databases (i.e., KEGG Orthology, KO; Enzyme Commission, E.C.; Gene Ontology, GO; Pfam; and InterPro); and finally (iv) we used the MMseqs2 protein search tool [163] against the IMG/VR v3 database, which includes genomes of cultivated and uncultivated viruses [164].

Mercury determination

The concentrations of mercury were measured in samples by inductively coupled plasma system coupled to mass spectrometry (ICP-MS; 7700 Series; Agilent, Palo Alto) at the University center of legal medicine (Geneva, Switzerland). Prior to analysis, samples were diluted with aqua regia to dissolve even poorly soluble mercury salts. The solution contained 10 ng/mL Rhodium (Rh) and 10 ng/mL Indium (In) as internal standards. In addition, each analytical batch of study samples was processed with laboratory controls, including method blanks and standard reference materials to continuously monitor method performance.

Graphical plotting

The following R packages were used to plot the data: “ggplot2”, “pheatmap”, “webR”, and “moonBook” in R-studio (<https://www.rstudio.com>).

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12915-022-01509-7>.

Additional file 1: Table S1. Sequencing basic statistics and mitochondrial haplogroup assignments. **Table S2.** Taxonomic classification on the bacterial species-level as estimated by mapping the metagenomic reads against the NCBI-nr database using DIAMOND/BLASTx mode and lowest common ancestor (LCA) assignment using MEGAN. **Table S3.** Basic assembly statistics and CheckM results of different assembly approaches. **Table S4.** Comparison of total number of bacterial-assigned reads, as estimated using DIAMOND search against NCBI-nr database and Brain_NTM-assigned reads by mapping. **Table S5.** The lengths, coverage, and GC content of the contigs of the brain NTM after combining SPAdes and MEGAHIT assemblies. **Table S6.** Metadata on the characterized members of the genus *Mycobacterium*. **Table S7.** Combined annotation of the contig_38. **Table S8.** CheckV results of the contig_38.

Additional file 2: Figure S1. NCBI taxonomy-based cladogram of the metagenomic reads mapped to the reference genome of *Treponema pallidum* subsp. *pallidum* (NZ_CP010561.1). **Figure S2.** Top bacterial families in the brain tissue as inferred by different taxonomic classifiers. **Figure S3.** Read lengths distribution of human DNA of different tissues as well as the brain NTM. **Figure S4.** Description of the analyzed samples. **Figure S5.** Virulence genes of the brain NTM. **Figure S6.** Toxin/Antitoxin (TA) Systems in the analyzed mycobacterial genomes. **Figure S7.** Metagenomic binning of Anna Catharina Bischoff's (ACB) brain sample. **Figure S8.** Correlation analysis between concentrations of mercury (Hg) and Selenium (Se) in different body parts.

Acknowledgements

We thank the Citizen Science Team of Basel (CSB) for the genealogical studies and the research in the City Town archives, which led to the identification of the Mummy. We also thank the CSB-team of genealogists of CSB, in particular Diana Gysin, Marie-Louise Gamma, Odette Haas, Marina Zulauf-Semmler, and Jürgen Rauber and the CSB-team of medical history research in the town ship archives of Basel, in particular Ursula Hirter-Trüb, Erika Borner, Regine Dendler, Dascha Herber, Ludwig Huber, Hans-Ulrich Fiechter, Brigitte Kuhn, Christof Meissburger, Gitta Reinhardt-Fehrenbach, Willy Ruess, Martin Schneider, Albert Spycher, and Helena Vogler. We thank Lena Granehäll for her assistance with processing the samples. We are grateful to the organizing team of Uppsala University “Snakemake bring-your-own-code (BYOC)” workshop for assisting in improving the metagenomic analysis workflow, and to the members of SPAAM community who helped during the peer-reviewing of the article.

Authors' contributions

MSS and FM designed the study and performed the analysis. MSS processed the samples, analyzed and visualized the data, and drafted manuscript. FM coordinated the molecular investigations. CW revised the human DNA analysis. AZ, TB, AJB, and GH were involved in the sampling campaign. AT and AJB performed the histopathological analyses. HW was responsible for the computed tomography imaging. MA and TB were responsible for the toxicological analyses. FM and AZ were responsible for the funding acquisition. GH was responsible for the curation of the mummy and was directing all anthropological, genealogical, and medical history research. MSS, CW, AZ, TB, AJB, GH, AZ, and FM edited the manuscript. All authors read and approved the final manuscript.

Funding

Support was provided by the European Regional Development Fund 2014-2020_CALL-FESR 2017 Research and Innovation_Autonomous Province of Bolzano - South Tyrol_Project: FESR1078-MummyLabs. The authors thank the Department of Innovation, Research and University of the Autonomous Province of Bozen/Bolzano for covering the Open Access publication costs.

Availability of data and materials

Raw sequencing data are publicly available on the European Nucleotide Archive (ENA, Project ID: PRJEB44723) [165].

Declarations

Ethics approval and consent to participate

All human samples used in this study are more than 70 years old and anonymous. Therefore, ethical approval for the genetic analysis is not required under current Swiss law (<https://www.admin.ch/opc/de/classified-compilation/20061313/index.html>), but ethical guidelines were considered during sampling and storage, following the recommendations of the International Council of Museums (ICOM). The mummy is housed at the Natural History Museum of Basel (NHM), Switzerland, and the sample collection for this study has been performed in the framework of a scientific collaboration. The museum fully supported the study, and the curator is included as coauthor in the study. The consent is available upon request.

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests.

Received: 7 March 2022 Accepted: 15 December 2022

Published online: 07 February 2023

References

- Hotz G, Augsburg M, Briellmann T, Bircher A, Castella V, Fiechter R, et al. Der rätselhafte Mumienfund aus der Barfüsserkirche in Basel. Ein aussergewöhnliches Beispiel interdisziplinärer Familienforschung.

- Jahrbuch der Schweizerischen Gesellschaft für Familienforschung. 2018;2018:1–30.
2. Hotz G, Opitz-Belakhal C, Anna Catharina Bischoff. Die Mumie aus der Barfüsserkirche. Basel: Christoph Merian Verlag; 2021.
 3. Wurst C, Maixner F, Castella V, Cipollini G, Hotz G, Zink A. The lady from Basel's Barfüsserkirche - molecular confirmation of the mummy's identity through mitochondrial DNA of living relatives spanning 22 generations. *Forensic Sci Int Genet*. 2022;56:102604.
 4. Briellmann T, Hotz G, Augsburg M, Lenglet S. Toxikologische Untersuchungen. In: Hotz G, Opitz-Belakhal C, editors. Anna Catharina Bischoff die Mumie aus der Barfüsserkirche (Rekonstruktion einer Basler Frauenbiografie des 18 Jahrhunderts). Basel: Christoph Merian Verlag; 2021.
 5. Tampa M, Sarbu I, Matei C, Benea V, Georgescu SR. Brief history of syphilis. *J Med Life*. 2014;7(1):4–10.
 6. Tzankov A, Bircher AJ. Mikroskopische Untersuchung. In: Hotz G, Opitz-Belakhal C, editors. Anna Catharina Bischoff die Mumie aus der Barfüsserkirche (Rekonstruktion einer Basler Frauenbiografie des 18 Jahrhunderts). Basel: Christoph Merian Verlag; 2021.
 7. Njau DG, Muge E, Kinyanjui P, Omwandho C, Mukwana S. STR analysis of human DNA from maggots fed on decomposing bodies: assessment of the time period for successful analysis. *Egypt J Forensic Sci*. 2016;6(3):261–9.
 8. Sarhan MS, Lehmkühl A, Straub R, Tett A, Wieland G, Francken M, et al. Ancient DNA diffuses from human bones to cave stones. *iScience*. 2021;24(12):103397.
 9. You M, Mo S, Leung WK, Watt RM. Comparative analysis of oral treponemes associated with periodontal health and disease. *BMC Infect Dis*. 2013;13(1):1–13.
 10. Verma D, Garg PK, Dubey AK. Insights into the human oral microbiome. *Arch Microbiol*. 2018;200(4):525–40.
 11. Huson DH, Beier S, Flade I, Gorska A, El-Hadidi M, Mitra S, et al. MEGAN Community edition - interactive exploration and analysis of large-scale microbiome sequencing data. *PLoS Comput Biol*. 2016;12(6):e1004957.
 12. Lu J, Rincon N, Wood DE, Breitwieser FP, Pockrandt C, Langmead B, et al. Metagenome analysis using the kraken software suite. *Nat Protoc*. 2022;17:2815–39.
 13. Herbig A, Maixner F, Bos K, Zink A, Krause J, Huson DH. MALT: fast alignment and analysis of metagenomic DNA sequence data applied to the Tyrolean iceman. *bioRxiv*. 2016:050559. <https://www.biorxiv.org/content/10.1101/050559v1>.
 14. Jónsson H, Ginolhac A, Schubert M, Johnson P, Orlando L. mapDamage2.0: fast approximate Bayesian estimates of ancient DNA damage parameters. *Bioinformatics*. 2013;29(13):1682–4.
 15. Neukamm J, Peltzer A, Nieselt K. DamageProfiler: fast damage pattern calculation for ancient DNA. *Bioinformatics*. 2021. <https://academic.oup.com/bioinformatics/article/37/20/3652/6247758>.
 16. Gupta RS, Lo B, Son J. Phylogenomics and comparative genomic studies robustly support division of the genus *Mycobacterium* into an emended genus *Mycobacterium* and four novel genera. *Front Microbiol*. 2018;9:67. <https://doi.org/10.3389/fmicb.2018.00067>.
 17. Meehan CJ, Barco RA, Loh Y-HE, Cogneau S, Rigouts L. Reconstituting the genus *Mycobacterium*. *Int J Syst Evol Microbiol*. 2021;71(9):004922.
 18. Ondov BD, Treangen TJ, Melsted P, Mallonee AB, Bergman NH, Koren S, et al. Mash: fast genome and metagenome distance estimation using MinHash. *Genome Biol*. 2016;17(1):1–14.
 19. Khan AA, Kim S-J, Paine DD, Cerniglia CE. Classification of a polycyclic aromatic hydrocarbon-metabolizing bacterium, *Mycobacterium* sp. strain PYR-1, as *Mycobacterium vanbaalenii* sp. nov. *Int J Syst Evol Microbiol*. 2002;52(6):1997–2002.
 20. Stanford J, Gunthorpe W. A study of some fast-growing scotochromogenic mycobacteria including species descriptions of *Mycobacterium gilvum* (new species) and *Mycobacterium duvalii* (new species). *Br J Exp Pathol*. 1971;52(6):627.
 21. Hennessee CT, Seo J-S, Alvarez AM, Li QX. Polycyclic aromatic hydrocarbon-degrading species isolated from Hawaiian soils: *Mycobacterium crocinum* sp. nov., *Mycobacterium pallens* sp. nov., *Mycobacterium rutilum* sp. nov., *Mycobacterium rufum* sp. nov. and *Mycobacterium aromaticivorans* sp. nov. *Int J Syst Evol Microbiol*. 2009;59(2):378–87.
 22. Cooksey RC, de Waard JH, Yakus MA, Rivera I, Chopite M, Toney SR, et al. *Mycobacterium cosmeticum* sp. nov., a novel rapidly growing species isolated from a cosmetic infection and from a nail salon. *Int J Syst Evol Microbiol*. 2004;54(6):2385–91.
 23. Chamoiseau G. M. *Farcinogenes agent causal du farcin du boeuf en Afrique*. In: *Annales de Microbiologie*; 1973.
 24. Toro A, Adekambi T, Cheynet F, Fournier P-E, Drancourt M. *Mycobacterium setense* infection in humans. *Emerg Infect Dis*. 2008;14(8):1330.
 25. Gomila M, Ramirez A, Gasco J, Lalucat J. *Mycobacterium Ilatzerense* sp. nov., a facultatively autotrophic, hydrogen-oxidizing bacterium isolated from haemodialysis water. *Int J Syst Evol Microbiol*. 2008;58(12):2769–73.
 26. Greninger AL, Langelier C, Cunningham G, Keh C, Melgar M, Chiu CY, et al. Two rapidly growing mycobacterial species isolated from a brain abscess: first whole-genome sequences of *Mycobacterium immunogenium* and *Mycobacterium Ilatzerense*. *J Clin Microbiol*. 2015;53(7):2374–7.
 27. Gcebe N, Michel A, Gey van Pittius NC, Rutten V: comparative genomics and proteomic analysis of four non-tuberculous *Mycobacterium* species and *Mycobacterium tuberculosis* complex: occurrence of shared immunogenic proteins. *Front Microbiol*. 2016;7:795.
 28. Schinsky MF, Morey RE, Steigerwalt AG, Douglas MP, Wilson RW, Floyd MM, et al. Taxonomic variation in the *Mycobacterium fortuitum* third biovariant complex: description of *Mycobacterium boenickei* sp. nov., *Mycobacterium houstonense* sp. nov., *Mycobacterium neworleansense* sp. nov. and *Mycobacterium brisbanense* sp. nov. and recognition of *Mycobacterium porcinum* from human clinical isolates. *Int J Syst Evol Microbiol*. 2004;54(5):1653–67.
 29. Okamori S, Asakura T, Nishimura T, Tamizu E, Ishii M, Yoshida M, et al. Natural history of *Mycobacterium fortuitum* pulmonary infection presenting with migratory infiltrates: a case report with microbiological analysis. *BMC Infect Dis*. 2018;18(1):1–6.
 30. Shojaei H, Goodfellow M, Magee J, Freeman R, Gould F, Brignall C. *Mycobacterium novocastrense* sp. nov., a rapidly growing photochromogenic mycobacterium. *Int J Syst Evol Microbiol*. 1997;47(4):1205–7.
 31. Apajalahti JH, Kärpänoja P, Salkinoja-Salonen MS. *Rhodococcus chlorophenolicus* sp. nov., a chlorophenol-mineralizing actinomycete. *Int J Syst Evol Microbiol*. 1986;36(2):246–51.
 32. Poh M-E, Liam C-K, Ng K-P, Tan R. *Mycobacterium brisbanense* species nova isolated from a patient with chronic cavitary lung infection. *Chest*. 2014;145(4):858–60.
 33. Franco MMJ, Paes AC, Ribeiro MG, de Figueiredo Pantoja JC, Santos ACB, Miyata M, et al. Occurrence of mycobacteria in bovine milk samples from both individual and collective bulk tanks at farms and informal markets in the southeast region of Sao Paulo, Brazil. *BMC Vet Res*. 2013;9(1):1–8.
 34. Shahrahi AH, Çavuşoğlu C, Borroni E, Heidarieh P, Koksalan OK, Cabibbe AM, et al. *Mycobacterium celeriflavum* sp. nov., a rapidly growing scotochromogenic bacterium isolated from clinical specimens. *Int J Syst Evol Microbiol*. 2015;65(Pt_2):510–5.
 35. Heidarieh P, Shojaei H, Hashemi A, Feizabadi MM, Daei-Naser A, Ataei B. First report of isolation of *Mycobacterium elephantis* from bronchial lavage of a patient in Asia. *JRSM Short Rep*. 2011;2(4):1–3.
 36. Sethi S, Gupta V, Bhattacharyya S, Sharma M. Post-laparoscopic wound infection caused by scotochromogenic nontuberculous *Mycobacterium*. *Jpn J Infect Dis*. 2011;64(5):426–7.
 37. Allen DM, Chng HH. Disseminated *Mycobacterium flavescens* in a probable case of chronic granulomatous disease. *J Inf Secur*. 1993;26(1):83–6.
 38. Brown-Elliott BA, Wallace RJ Jr, Petti CA, Mann LB, McGlasson M, Chihara S, et al. *Mycobacterium neoaurum* and *Mycobacterium bacteremicum* sp. nov. as causes of mycobacteremia. *J Clin Microbiol*. 2010;48(12):4377–85.
 39. Tsukamura M, Mizuno S, Gane N, Mills A, King L: *Mycobacterium rhodesiae* sp. nov. A new species of rapid-growing scotochromogenic mycobacteria. *Japanese J Microbiol*. 1971;15(5):407–16.
 40. Tortoli E, Kroppenstedt RM, Bartoloni A, Caroli G, Jan I, Pawlowski J, et al. *Mycobacterium tusciae* sp. nov. *Int J Syst Evol Microbiol*. 1999;49(4):1839–44.
 41. Jiménez MS, Campos-Herrero MI, García D, Luquin M, Herrera L, García MJ. *Mycobacterium canariensis* sp. nov. *Int J Syst Evol Microbiol*. 2004;54(5):1729–34.

42. Shojaei H, Daley C, Gitti Z, Hashemi A, Heidarieh P, Moore ER, et al. *Mycobacterium iranica* sp. nov., a rapidly growing scotochromogenic species isolated from clinical specimens on three different continents. *Int J Syst Evol Microbiol*. 2013;63(Pt_4):1383–9.
43. Chen Y-C, Jou R, Huang W-L, Huang S-T, Liu K-C, Lay C-J, et al. Bacteremia caused by *Mycobacterium wolinskyi*. *Emerg Infect Dis*. 2008;14(11):1818.
44. Adékambi T, Stein A, Carvajal J, Raoult D, Drancourt M. Description of *Mycobacterium conceptionense* sp. nov., a *Mycobacterium fortuitum* group organism isolated from a posttraumatic osteitis inflammation. *J Clin Microbiol*. 2006;44(4):1268–73.
45. Kazda J. *Mycobacterium sphagni* sp. nov. *Int J Syst Evol Microbiol*. 1980;30(1):77–81.
46. Lee SA, Raad II, Adachi JA, Han XY. Catheter-related bloodstream infection caused by *Mycobacterium brumae*. *J Clin Microbiol*. 2004;42(11):5429–31.
47. Dahl JL, Gatlin W III, Tran PM, Sheik CS. *Mycobacterium nivoides* sp. nov. isolated from a peat bog. *Int J Syst Evol Microbiol*. 2021;71(3). <https://www.microbiologyresearch.org/content/journal/ijsem/10.1099/ijsem.0.004438>.
48. Luis BAL, Díaz-Lomelí P, Gómez-Albarrán LP, Martínez-Gamboa A, Ponce-de-León A. *Mycobacterium obuense* bacteremia in a patient with pneumonia. *Emerg Infect Dis*. 2019;25(5):1015.
49. Marie I, Heliot P, Roussel F, Herve F, Muir J, Levesque H. Fatal *Mycobacterium peregrinum* pneumonia in refractory polymyositis treated with infliximab. *Rheumatology*. 2005;44(9):1201–2.
50. Adékambi T, Foucault C, La Scola B, Drancourt M. Report of two fatal cases of *Mycobacterium mucogenicum* central nervous system infection in immunocompetent patients. *J Clin Microbiol*. 2006;44(3):837–40.
51. Xu C, Wu W, Pan H, Hui T, Wu Q, Zhou Z, et al. *Mycobacterium Agri* skin infection in a previously healthy patient: a case study. *Infection Drug Resistance*. 2021;14:2965.
52. Vuorio R, Andersson MA, Rainey FA, Kroppenstedt RM, Kämpfer P, Busse H-J, et al. A new rapidly growing mycobacterial species, *Mycobacterium murale* sp. nov., isolated from the indoor walls of a children's day care Centre. *Int J Syst Evol Microbiol*. 1999;49(1):25–35.
53. Balcázar JL, Planas M, Pintado J. *Mycobacterium hippocampi* sp. nov., a rapidly growing scotochromogenic species isolated from a seahorse with tail rot. *Curr Microbiol*. 2014;69(3):329–33.
54. Tsukamura M. *Mycobacterium parafortuitum*: a new species. *Microbiology*. 1966;42(1):7–12.
55. Tsukamura M, Mizuno S, Toyama H. *Mycobacterium pulveris* sp. nov., a nonphotochromogenic mycobacterium with an intermediate growth rate. *Int J Syst Evol Microbiol*. 1983;33(4):811–5.
56. Kondo A, Mori K, Iwata J, Tamura M, Yamamoto T, Nakao Y, et al. Caseous necrotic granuloma in the pituitary stalk due to nontuberculous mycobacteria (*Mycobacterium tokaiense*) infection. *Neurol Med Chir*. 2006;46(2):80–3.
57. Singh J, Antony SJ. Prosthetic joint infection due to *Mycobacterium morioakaense* in an immunocompetent patient after a total knee replacement. In: *Baylor University medical center proceedings*: Taylor & Francis; 2020. p. 97–9.
58. Tsukamura M, Mizuno S, Tsukamura S. Numerical analysis of rapidly growing, scotochromogenic mycobacteria, including *Mycobacterium obuense* sp. nov., nom. Rev., *Mycobacterium rhodesiae* sp. nov., nom. Rev., *Mycobacterium aichiense* sp. nov., nom. Rev., *Mycobacterium chubuense* sp. nov., nom. Rev., and *Mycobacterium tokaiense* sp. nov., nom. Rev. *Int J Syst Evol Microbiol*. 1981;31(3):263–75.
59. Lévy-Frédault V, Prome J-C, Grandry J, Boisvert H, David HL. *Mycobacterium fallax* sp. nov. *Int J Syst Evol Microbiol*. 1983;33(2):336–43.
60. Lee C-H, You H-L, Wang J-W, Tang Y-F, Liu J-W. Prosthetic joint infection caused by *Mycobacterium alvei* in an elderly patient. *J Clin Microbiol*. 2011;49(8):3096–8.
61. Ausina V, Luquin M, Barcelo MG, Lanéelle M, Lévy-Frédault V, Belda F, et al. *Mycobacterium alvei* sp. nov. *Int J Syst Evol Microbiol*. 1992;42(4):529–35.
62. Huth RG, Brown-Elliott BA, Wallace RJ Jr. *Mycobacterium mageritense* pulmonary disease in patient with compromised immune system. *Emerg Infect Dis*. 2011;17(3):556.
63. Pettit AC, Jahangir AA, Wright PW. *Mycobacterium doricum* osteomyelitis and soft tissue infection. *Emerg Infect Dis*. 2011;17(11):2075.
64. Ballester F, Pujol I, Alcaide F, Pizarro I, Simó JM, Joven J, et al. First human isolate of *Mycobacterium poriferiae* in the sputum of a patient with chronic bronchitis. *J Clin Microbiol*. 2011;49(8):3107–8.
65. Casal M, Rey C. *Mycobacterium gadium* sp. nov. a new species of rapid-growing scotochromogenic mycobacteria. *Tubercle*. 1974;55(4):299–308.
66. Trujillo ME, Velazquez E, Kroppenstedt RM, Schumann P, Rivas R, Mateos PF, et al. *Mycobacterium psychrotolerans* sp. nov., isolated from pond water near a uranium mine. *Int J Syst Evol Microbiol*. 2004;54(5):1459–63.
67. Ballester F, Alcaide F, Pujol I, Hernández-Flix S, Simó JM, Joven J, et al. First human isolate of *Mycobacterium madagascariense* in the sputum of a patient with tracheobronchitis. *Clin Chem Lab Med*. 2013;51(2):e35–6.
68. Kirschner P, Teske A, Schröder K-H, Kroppenstedt R, Wolters J, Böttger E. *Mycobacterium confluentis* sp. nov. *Int J Syst Evol Microbiol*. 1992;42(2):257–62.
69. Adékambi T, Berger P, Raoult D, Drancourt M. *rpoB* gene sequence-based characterization of emerging non-tuberculous mycobacteria with descriptions of *Mycobacterium bolletii* sp. nov., *Mycobacterium phocaicum* sp. nov. and *Mycobacterium aubagnense* sp. nov. *Int J Syst Evol Microbiol*. 2006;56(1):133–43.
70. Adékambi T. *Mycobacterium mucogenicum* group infections: a review. *Clin Microbiol Infect*. 2009;15(10):911–8.
71. Shojaei H, Hashemi A, Heidarieh P, Hosseini N, Naser AD. Chronic pulmonary disease due to *Mycobacterium monacense* infection: the first case from Iran. *Ann Lab Med*. 2012;32(1):87–90.
72. Tortoli E, Baruzzo S, Hejdra Y, Klenk H-P, Lauria S, Mariottini A, et al. *Mycobacterium insubricum* sp. nov. *Int J Syst Evol Microbiol*. 2009;59(6):1518–23.
73. Zhang Y, Zhang J, Fang C, Pang H, Fan J. *Mycobacterium litorale* sp. nov., a rapidly growing mycobacterium from soil. *Int J Syst Evol Microbiol*. 2012;62(Pt_5):1204–7.
74. Zhang D-F, Chen X, Zhang X-M, Zhi X-Y, Yao J-C, Jiang Y, et al. *Mycobacterium sediminis* sp. nov. and *Mycobacterium arabicense* sp. nov., two rapidly growing members of the genus *Mycobacterium*. *Int J Syst Evol Microbiol*. 2013;63(Pt_11):4081–6.
75. Kim B-J, Kim J-M, Kim B-R, Lee S-Y, Kim G, Jang Y-H, et al. *Mycobacterium anyangense* sp. nov., a rapidly growing species isolated from blood of Korean native cattle, Hanwoo (*Bos taurus coreanae*). *Int J Syst Evol Microbiol*. 2015;65(Pt_7):2277–85.
76. Tran PM, Dahl JL. *Mycobacterium sarraceniae* sp. nov. and *Mycobacterium helvum* sp. nov., isolated from the pitcher plant *Sarracenia purpurea*. *Int J Syst Evol Microbiol*. 2016;66(11):4480–5.
77. Gordon RE, Smith MM. Rapidly growing, acid fast bacteria I: species' descriptions of *Mycobacterium phlei* Lehmann and Neumann and *Mycobacterium smegmatis* (Trevisan) Lehmann and Neumann. *J Bacteriol*. 1953;66(1):41–8.
78. Hormisch D, Brost I, Kohring G-W, Giffhorn F, Kroppenstedt R, Stackebrandt E, et al. *Mycobacterium fluoranthenorans* sp. nov., a fluoranthene and aflatoxin B1 degrading bacterium from contaminated soil of a former coal gas plant. *Syst Appl Microbiol*. 2004;27(6):653–60.
79. Go JR, Wengenack NL, Abu Saleh OM, Corsini Campioli C, Deml SM, Wilson JW. *Mycobacterium septicum*: a 6-year clinical experience from a tertiary hospital and reference laboratory. *J Clin Microbiol*. 2020;58(12):e02091–20.
80. Schinsky MF, McNeil MM, Whitney AM, Steigerwalt AG, Lasker BA, Floyd MM, et al. *Mycobacterium septicum* sp. nov., a new rapidly growing species associated with catheter-related bacteraemia. *Int J Syst Evol Microbiol*. 2000;50(2):575–81.
81. Tsukamura M, van der Meulen HJ, Grabow WO. Numerical taxonomy of rapidly growing, scotochromogenic mycobacteria of the *Mycobacterium parafortuitum* complex: *Mycobacterium austroafricanum* sp. nov. and *Mycobacterium diernhoferi* sp. nov., nom. Rev. *Int J Syst Evol Microbiol*. 1983;33(3):460–9.
82. Konjek J, Souled S, Guerardel Y, Trivelli X, Bernut A, Kremer L, et al. *Mycobacterium lutetiense* sp. nov., *Mycobacterium montmartrense* sp. nov. and *Mycobacterium arcueilense* sp. nov., members of a novel group of non-pigmented rapidly growing mycobacteria

- recovered from a water distribution system. *Int J Syst Evol Microbiol*. 2016;66(9):3694–702.
83. Awadh H, Mansour M, Shorman M. Bacteremia with an unusual pathogen: *Mycobacterium neoaurum*. *Case Rep Infect Dis*. 2016;2016:5167874.
 84. Heckman GA, Hawkins C, Morris A, Burrows LL, Bergeron C. Rapidly progressive dementia due to *Mycobacterium neoaurum* meningoen- cephalitis. *Emerg Infect Dis*. 2004;10(5):924–7.
 85. Gharbi R, Khanna V, Frigui W, Mhenni B, Brosch R, Mardassi H. Phe- notypic and genomic hallmarks of a novel, potentially pathogenic rapidly growing *Mycobacterium* species related to the *Mycobacterium fortuitum* complex. *Sci Rep*. 2021;11(1):13011.
 86. Levasseur A, Asmar S, Robert C, Drancourt M. Draft genome sequence of *Mycobacterium houstonense* strain ATCC 49403T. *Genome Announc*. 2016;4(3):e00443–16.
 87. Biet F, Boschirolu ML. Non-tuberculous mycobacterial infections of veterinary relevance. *Res Vet Sci*. 2014;97:S69–77.
 88. Weitzman I, Osadczy D, Corrado M, Karp D. *Mycobacterium thermore- sistibile*: a new pathogen for humans. *J Clin Microbiol*. 1981;14(5):593–5.
 89. Ruan J, Li X-Y, Chen H. *Mycobacterium chubuense* hand infection. *IDCases*. 2020;20:e00742.
 90. Tanaka S, Hoshino Y, Sakagami T, Fukano H, Matsui Y, Hiranuma O. Pathogenicity of *Mycobacterium phlei*, a non-pathogenic nontu- berculous mycobacterium in an immunocompetent host carrying anti-interferon gamma autoantibodies: a case report. *BMC Infect Dis*. 2019;19(1):1–6.
 91. Deinhardt-Emmer S, Höring S, Mura C, Hillemann D, Hermann B, Sachse S, et al. First time isolation of *Mycobacterium hassiacum* from a respiratory sample. *Clin Med Insights Circul Respir Pulmon Med*. 2018;12:1179548417747529.
 92. Salzer HJF, Chitechi B, Hillemann D, Mandl M, Paar C, Mitterhumer M, et al. Nontuberculous mycobacterial pulmonary disease from *Mycobacterium hassiacum*, Austria. *Emerg Infect Dis*. 2020;26(11):2776–8.
 93. Sood S, Yadav A, Shrivastava R. *Mycobacterium aurum* is unable to survive *Mycobacterium tuberculosis* latency associated stress condi- tions: implications as non-suitable model organism. *Indian J Microbiol*. 2016;56(2):198–204.
 94. Tsukamura M. *Mycobacterium chitae*: a new species. *Japanese J Micro- biol*. 1967;11(1):43–7. <https://doi.org/10.1111/j.1348-0421.1967.tb00319.x>.
 95. Liu B, Zheng D, Jin Q, Chen L, Yang J. VFDB 2019: a comparative pathog- enomic platform with an interactive web interface. *Nucleic Acids Res*. 2019;47(D1):D687–92.
 96. Rivera-Calzada A, Famelis N, Llorca O, Geibel S. Type VII secretion systems: structure, functions and transport models. *Nat Rev Microbiol*. 2021;19(9):567–84.
 97. Be NA, Lamichhane G, Grosset J, Tyagi S, Cheng Q-J, Kim KS, et al. Murine model to study the invasion and survival of *Mycobac- terium tuberculosis* in the central nervous system. *J Infect Dis*. 2008;198(10):1520–8.
 98. van Leeuwen LM, Boot M, Kuijl C, Picavet DI, van Stempvoort G, van der Pol SM, et al. *Mycobacteria* employ two different mechanisms to cross the blood–brain barrier. *Cell Microbiol*. 2018;20(9):e12858.
 99. Sala A, Bordes P, Genevaux P. Multiple toxin-antitoxin systems in *Mycobacterium tuberculosis*. *Toxins*. 2014;6(3):1002–20.
 100. Keren I, Minami S, Rubin E, Lewis K. Characterization and transcrip- tome analysis of *Mycobacterium tuberculosis* persisters. *MBio*. 2011;2(3):e00100–11.
 101. Cater JC, Redmond WB. *Mycobacterial* phages isolated from stool specimens of patients with pulmonary disease. *Am Rev Respir Dis*. 1963;87(5):726–9.
 102. Huh H, Wong S, Jean JS, Slavcev R. Bacteriophage interactions with mammalian tissue: therapeutic applications. *Adv Drug Deliv Rev*. 2019;145:4–17.
 103. Glickman C, Kammlade SM, Hasan NA, Epperson LE, Davidson RM, Strong M. Characterization of integrated prophages within diverse species of clinical nontuberculous mycobacteria. *Virology*. 2020;17(1):1–13.
 104. Sassi M, Gouret P, Chabrol O, Pontarotti P, Drancourt M. *Mycobacte- riophage*-driven diversification of *Mycobacterium abscessus*. *Biol Direct*. 2014;9(1):1–15.
 105. Scheidegger S. Mittelalterliche Quecksilbervergiftungen. In: Dhom G, ed. *Verhandlungen der Deutschen Gesellschaft für Pathologie*. Stutt- gart: Gustav Fisher Verlag. 1977;61:374.
 106. Wujastyk D. Histories of mercury in medicine across Asia and beyond. *Asiatische Studien-Études Asiatiques*. 2015;69(4):819–30.
 107. Bachour N. Healing with mercury: the uses of mercury in Ara- bic medical literature. *Asiatische Studien-Études Asiatiques*. 2015;69(4):831–66.
 108. O'Shea JG. 'Two minutes with venus, two years with mercury'- mercury as an antisiphilic chemotherapeutic agent. *J R Soc Med*. 1990;83(6):392–5.
 109. Norn S, Permin H, Kruse E, Kruse PR. Mercury—a major agent in the history of medicine and alchemy. *Dansk medicinhistorisk arbog*. 2008;36:21–40.
 110. Fornaciari A, Chericoni S, Stefanelli F, Fornaciari G, Giuffra V. Renais- sance mercurial therapy in the mummies of saint Domenico Maggiore in Naples: a palaeopathological and palaeotoxicological approach. *Archaeol Anthropol Sci*. 2022;14(3):1–13.
 111. Gad SC, Pham T. Mercury. In: Wexler P, editor. *Encyclopedia of toxicol- ogy* (third edition). Oxford: Academic; 2014. p. 207–10.
 112. Park J-D, Zheng W. Human exposure and health effects of inorganic and elemental mercury. *J Prev Med Publ Health = Yebang Uihakhoe chi*. 2012;45(6):344–52.
 113. Friberg L, Mottet NK. Accumulation of methylmercury and inorganic mercury in the brain. *Biol Trace Elem Res*. 1989;21(1):201–6.
 114. Liu J, Cui J, Wei X, Li W, Liu C, Li X, et al. Investigation on selenium and mercury interactions and the distribution patterns in mice organs with LA-ICP-MS imaging. *Anal Chim Acta*. 2021;1182:338941.
 115. LaFond RE, Lukehart SA. Biological basis for syphilis. *Clin Microbiol Rev*. 2006;19(1):29–49.
 116. Park K-H, Lee MS, Hong IK, Sung J-Y, Choi S-H, Park SO, et al. Bone involvement in secondary syphilis: a case report and systematic review of the literature. *Sex Transm Dis*. 2014;41(9):532–7.
 117. Çakmak SK, Tamer E, Karadağ AS, Waugh M. Syphilis: a great imitator. *Clin Dermatol*. 2019;37(3):182–91.
 118. Schütz J. Ueber Bacillen bei syphilis. *DMW-Deutsche Medizinische Wochenschrift*. 1885;11(19):320–1.
 119. Lustgarten S. The bacillus of syphilis. *Lancet*. 1885;125(3214):609–10.
 120. Bos KI, Jäger G, Schuenemann VJ, Vågøene ÅJ, Spyrou MA, Herbig A, et al. Parallel detection of ancient pathogens via array-based DNA capture. *Philos Transact Royal Soc B: Biol Sci*. 2015;370(1660):20130375.
 121. Ratnatunga CN, Lutzky VP, Kupz A, Doolan DL, Reid DW, Field M, et al. The rise of non-tuberculosis mycobacterial lung disease. *Front Immunol*. 2020;11. <https://www.frontiersin.org/articles/10.3389/fimmu.2020.00303/full>.
 122. Ahmed I, Tiberi S, Farooqi J, Jabeen K, Yeboah-Manu D, Migliori GB, et al. Non-tuberculous mycobacterial infections—a neglected and emerging problem. *Int J Infect Dis*. 2020;92:546–50.
 123. Desai AN, Hurtado R. Nontuberculous mycobacterial infections. *JAMA*. 2021;325(15):1574.
 124. Fellows Yates JA, Velsko IM, Aron F, Posth C, Hofman CA, Austin RM, et al. The evolution and changing ecology of the African hominid oral microbiome. *Proc Natl Acad Sci*. 2021;118(20):e2021655118.
 125. Maixner F, Mitterer C, Jäger HY, Sarhan MS, Valverde G, Lückner S, et al. Linear polyacrylamide is highly efficient in precipitating and purifying environmental and ancient DNA. *Methods Ecol Evol*. 2021;n/a(n/a). <https://besjournals.onlinelibrary.wiley.com/doi/full/10.1111/2041-210X.13772>.
 126. Meyer M, Kircher M. Illumina sequencing library preparation for highly multiplexed target capture and sequencing. *Cold Spring Harb Protoc*. 2010;2010(6):pdb prot5448.
 127. Chen S, Zhou Y, Chen Y, Gu J. Fastp: an ultra-fast all-in-one FASTQ pre- processor. *Bioinformatics*. 2018;34(17):i884–90.
 128. Shen W, Le S, Li Y, Hu F. SeqKit: a cross-platform and ultrafast toolkit for FASTA/Q file manipulation. *PLoS One*. 2016;11(10):e0163962.
 129. Rosenbloom KR, Armstrong J, Barber GP, Casper J, Clawson H, Diekhans M, et al. The UCSC genome browser database: 2015 update. *Nucleic Acids Res*. 2015;43(Database issue):D670–81.
 130. Andrews RM, Kubacka I, Chinnery PF, Lightowlers RN, Turnbull DM, Howell N. Reanalysis and revision of the Cambridge reference sequence for human mitochondrial DNA. *Nat Genet*. 1999;23(2):147.

131. Li H, Durbin R. Fast and accurate long-read alignment with Burrows-wheeler transform. *Bioinformatics*. 2010;26(5):589–95.
132. Okonechnikov K, Conesa A, Garcia-Alcalde F. Qualimap 2: advanced multi-sample quality control for high-throughput sequencing data. *Bioinformatics*. 2016;32(2):292–4.
133. Renaud G, Slon V, Duggan AT, Kelso J. Schmutzi: estimation of contamination and endogenous mitochondrial consensus calling for ancient DNA. *Genome Biol*. 2015;16:224.
134. Weissensteiner H, Pacher D, Kloss-Brandstatter A, Forer L, Specht G, Bandelt HJ, et al. HaploGrep 2: mitochondrial haplogroup classification in the era of high-throughput sequencing. *Nucleic Acids Res*. 2016;44(W1):W58–63.
135. Skoglund P, Storå J, Götherström A, Jakobsson M. Accurate sex identification of ancient human remains using DNA shotgun sequencing. *J Archaeol Sci*. 2013;40:4477–82.
136. Buchfink B, Reuter K, Drost H-G. Sensitive protein alignments at tree-of-life scale using DIAMOND. *Nat Methods*. 2021;18(4):366–8.
137. Coordinators NR. Database resources of the national center for biotechnology information. *Nucleic Acids Res*. 2014;42(D1):D7–D17.
138. Wang Z-P, Xing H-L, Dong L, Zhang H-Y, Han C-Y, Wang X-C, et al. Egg cell-specific promoter-controlled CRISPR/Cas9 efficiently generates homozygous mutants for multiple target genes in *Arabidopsis* in a single generation. *Genome Biol*. 2015;16(1):144.
139. Nurk S, Meleshko D, Korobeynikov A, Pevzner PA. metaSPAdes: a new versatile metagenomic assembler. *Genome Res*. 2017;27(5):824–34.
140. Kang DD, Li F, Kirton E, Thomas A, Egan R, An H, et al. MetaBAT 2: an adaptive binning algorithm for robust and efficient genome reconstruction from metagenome assemblies. *PeerJ*. 2019;7:e7359.
141. Wu YW, Singer SW. Recovering individual genomes from metagenomes using MaxBin 2.0. *Curr Protoc*. 2021;1(5):e128.
142. Alneberg J, Bjarnason BS, de Bruijn I, Schirmer M, Quick J, Ijaz UZ, et al. Binning metagenomic contigs by coverage and composition. *Nat Methods*. 2014;11(11):1144–6.
143. Parks DH, Imelfort M, Skennerton CT, Hugenholtz P, Tyson GW. CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. *Genome Res*. 2015;25(7):1043–55.
144. Zhou C, Xu Q, He S, Ye W, Cao R, Wang P, et al. GTDB: an integrated resource for glycosyltransferase sequences and annotations. *Database (Oxford)*. 2020;2020. <https://academic.oup.com/database/article/doi/10.1093/database/baaa047/5857526>.
145. Chaumeil PA, Mussig AJ, Hugenholtz P, Parks DH. GTDB-Tk: a toolkit to classify genomes with the genome taxonomy database. *Bioinformatics*. 2019. <https://academic.oup.com/bioinformatics/article/36/6/1925/5626182>.
146. Swindell SR, Plasterer TN, Seqman. In: *Sequence data analysis guidebook*. Springer; 1997. p. 75–89.
147. Asnicar F, Thomas AM, Beghini F, Mengoni C, Manara S, Manghi P, et al. Precise phylogenetic analysis of microbial isolates and genomes from metagenomes using PhyloPhlAn 3.0. *Nat Commun*. 2020;11(1):2500.
148. Katoh K, Standley DM. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol Biol Evol*. 2013;30(4):772–80.
149. Capella-Gutiérrez S, Silla-Martínez JM, Gabaldón T. trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics*. 2009;25(15):1972–3.
150. Price MN, Dehal PS, Arkin AP. FastTree 2—approximately maximum-likelihood trees for large alignments. *PLoS One*. 2010;5(3):e9490.
151. Stamatakis A. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics*. 2014;30(9):1312–3.
152. Langmead B, Salzberg SL. Fast gapped-read alignment with bowtie 2. *Nat Methods*. 2012;9:357.
153. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. The sequence alignment/map format and SAMtools. *Bioinformatics*. 2009;25(16):2078–9.
154. Seemann T. Prokka: rapid prokaryotic genome annotation. *Bioinformatics*. 2014;30(14):2068–9.
155. Hyatt D, Chen G-L, LoCascio PF, Land ML, Larimer FW, Hauser LJ. Prodigal: prokaryotic gene recognition and translation initiation site identification. *BMC Bioinformatics*. 2010;11(1):119.
156. Lagesen K, Hallin P, Rødland EA, Stærfeldt H-H, Rognes T, Ussery DW. RNAMmer: consistent and rapid annotation of ribosomal RNA genes. *Nucleic Acids Res*. 2007;35(9):3100–8.
157. Gilchrist CL, Chooi Y-H. Clinker & clustermap. Js: automatic generation of gene cluster comparison figures. *Bioinformatics*. 2021;37(16):2473–5.
158. Emms DM, Kelly S. OrthoFinder: phylogenetic orthology inference for comparative genomics. *Genome Biol*. 2019;20(1):1–14.
159. Xie Y, Wei Y, Shen Y, Li X, Zhou H, Tai C, et al. TADB 2.0: an updated database of bacterial type II toxin–antitoxin loci. *Nucleic Acids Res*. 2017;46(D1):D749–53.
160. Arndt D, Grant JR, Marcu A, Sajed T, Pon A, Liang Y, et al. PHASTER: a better, faster version of the PHAST phage search tool. *Nucleic Acids Res*. 2016;44(W1):W16–21.
161. Nayfach S, Camargo AP, Schulz F, Eloë-Fadrosch E, Roux S, Kyrpides NC. CheckV assesses the quality and completeness of metagenome-assembled viral genomes. *Nat Biotechnol*. 2021;39(5):578–85.
162. Ruiz-Perez CA, Conrad RE, Konstantinidis KT. MicrobeAnnotator: a user-friendly, comprehensive functional annotation pipeline for microbial genomes. *BMC Bioinformatics*. 2021;22(1):1–16.
163. Steinegger M, Söding J. MMseqs2 enables sensitive protein sequence searching for the analysis of massive data sets. *Nat Biotechnol*. 2017;35(11):1026–8.
164. Roux S, Páez-Espino D, Chen I-MA, Palaniappan K, Ratner A, Chu K, et al. IMG/VR v3: an integrated ecological and evolutionary framework for interrogating genomes of uncultivated viruses. *Nucleic Acids Res*. 2021;49(D1):D764–75.
165. Sequencing data of the mummified Lady from Barfüsser church from Basel, Switzerland. ENA accession PRJEB44723. <https://www.ebi.ac.uk/ena/browser/view/PRJEB44723> (2022)

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions



Paper III: Supplementary Information

(This page intentionally left blank)

Supplementary figures

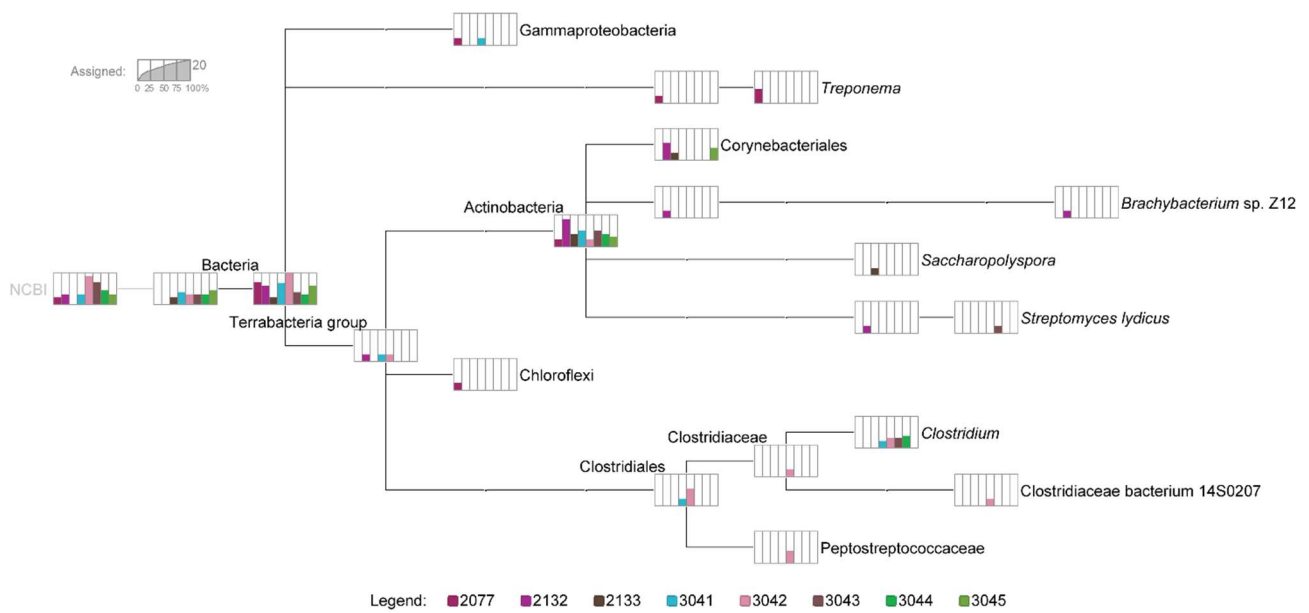


Figure S1: NCBI taxonomy-based cladogram of the metagenomic reads mapped to the reference genome of *Treponema pallidum* subsp. *pallidum* (NZ_CP010561.1). The shotgun metagenomic reads of all samples were initially mapped against the indexed reference genome using Burrows-Wheeler Aligner (BWA) with lenient parameters “bwa aln -n 0.01 -l 16”. After filtering the reads with a mapping quality < 30, the reads were exported into FASTA format. Reads were aligned against the NCBI-nt database, using the basic local alignment search tools (BLAST) with the options “blastn” and “--word-size 7”. The resulting blast tables were used for lowest common ancestor (LCA) assignment applying the tool blast2rma of MEGAN. The final output was visualized with the absolute read counts, using MEGAN v6.21.16.

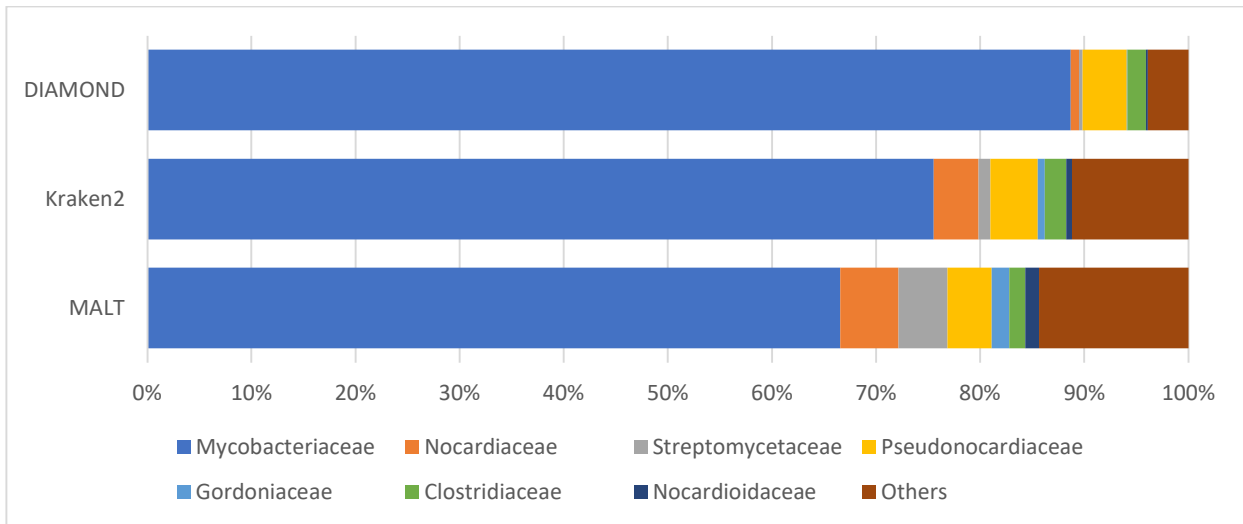


Figure S2: Top bacterial families in the brain tissue as inferred by different taxonomic classifiers. DIAMOND search was performed against the NCBI non-redundant protein database (NCBI-nr), then the reads were assigned to their lowest common ancestor (LCA) using MEGAN [1]. While Kraken2 search was performed against Kraken standard database and the family-level abundances were estimated using BRACKEN [2]. MEGAN alignment tool (MALT) was ran against database of ~16,600 bacterial representative genomes [3] and MEGAN LCA was used similar to DIAMOND.

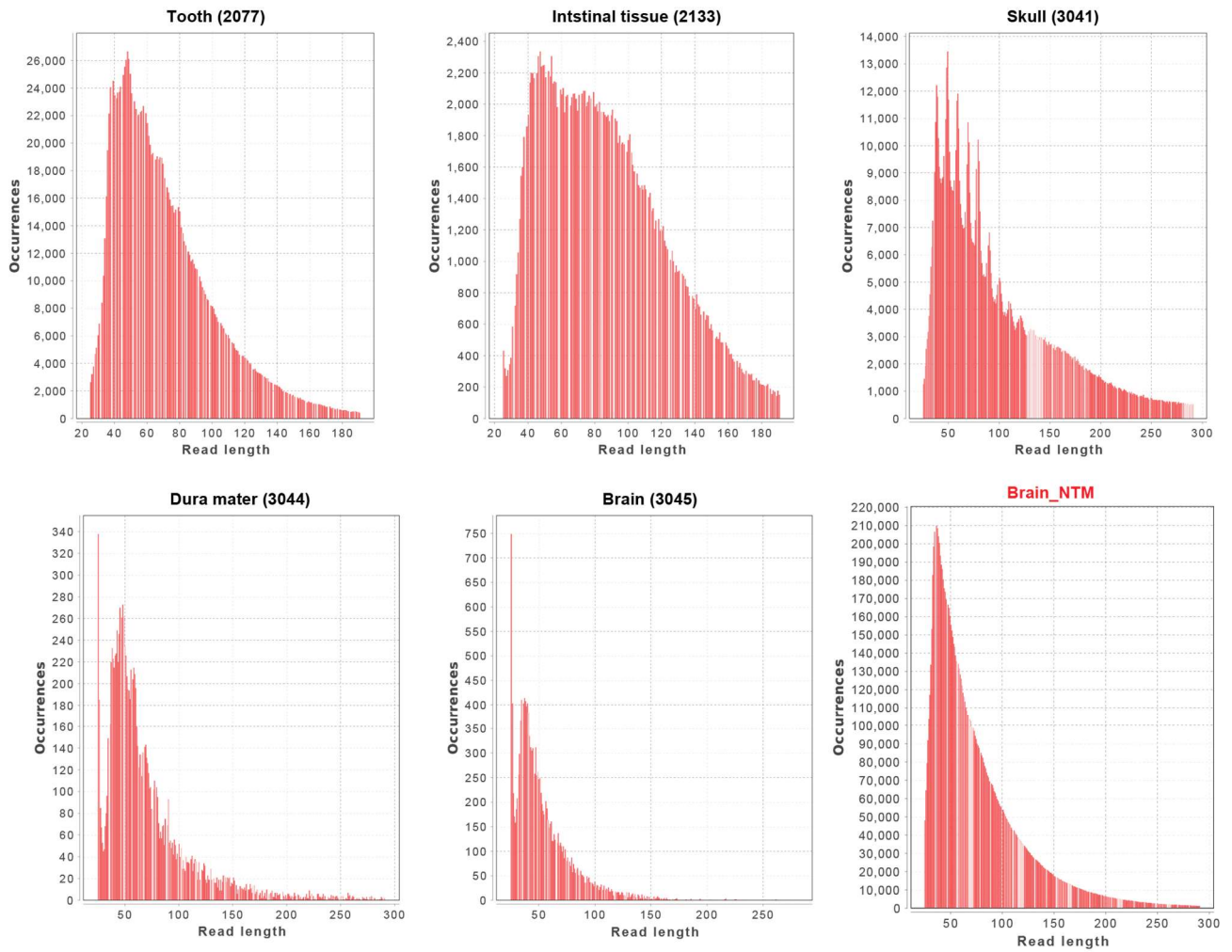


Figure S3: Read lengths distribution of human DNA of different tissues as well as the brain NTM. For the different body tissues (Tooth, Intestinal tissues, Skull, Dura mater, and brain), the metagenomic reads were mapped against the human reference genome (hg19), then the read lengths were calculated using MapDamage2.0 [4]. While for the brain_NTM, the brain metagenomic reads were mapped against the assembled genome of the brain_NTM, then the read lengths were calculated using DamageProfiler [5]. For further information on the samples' origin, please refer to **Additional file 1: Table S1**.

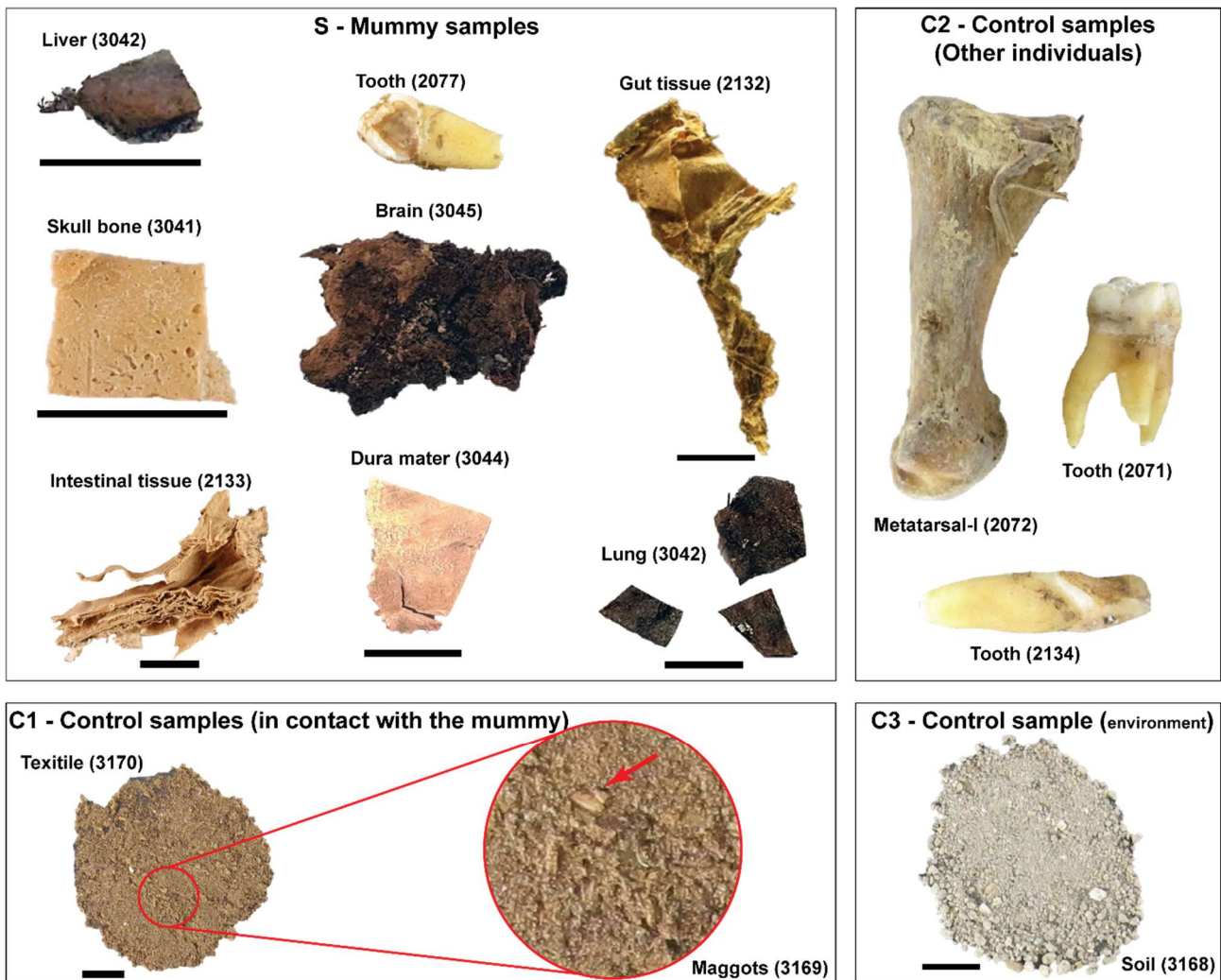


Figure S4: Description of the analyzed samples. S, the mummy samples are those collected from the mummy itself and assigned to particular tissues based on morphological and histological analyses [6], while the C1 samples, are those which were collected from underneath the mummy and were in between the body and the cloths (i.e., the textile and the maggots). C2, the samples collected from the skeletons which were found in the other coffin in the same grave. C3, a soil sample collected from the soil layer that was covering the mummy, most likely was added after the reburial during the 18th century. Please refer to **Additional file 1: Table S1** for further details. The black bars refer to scales of 10 mm.

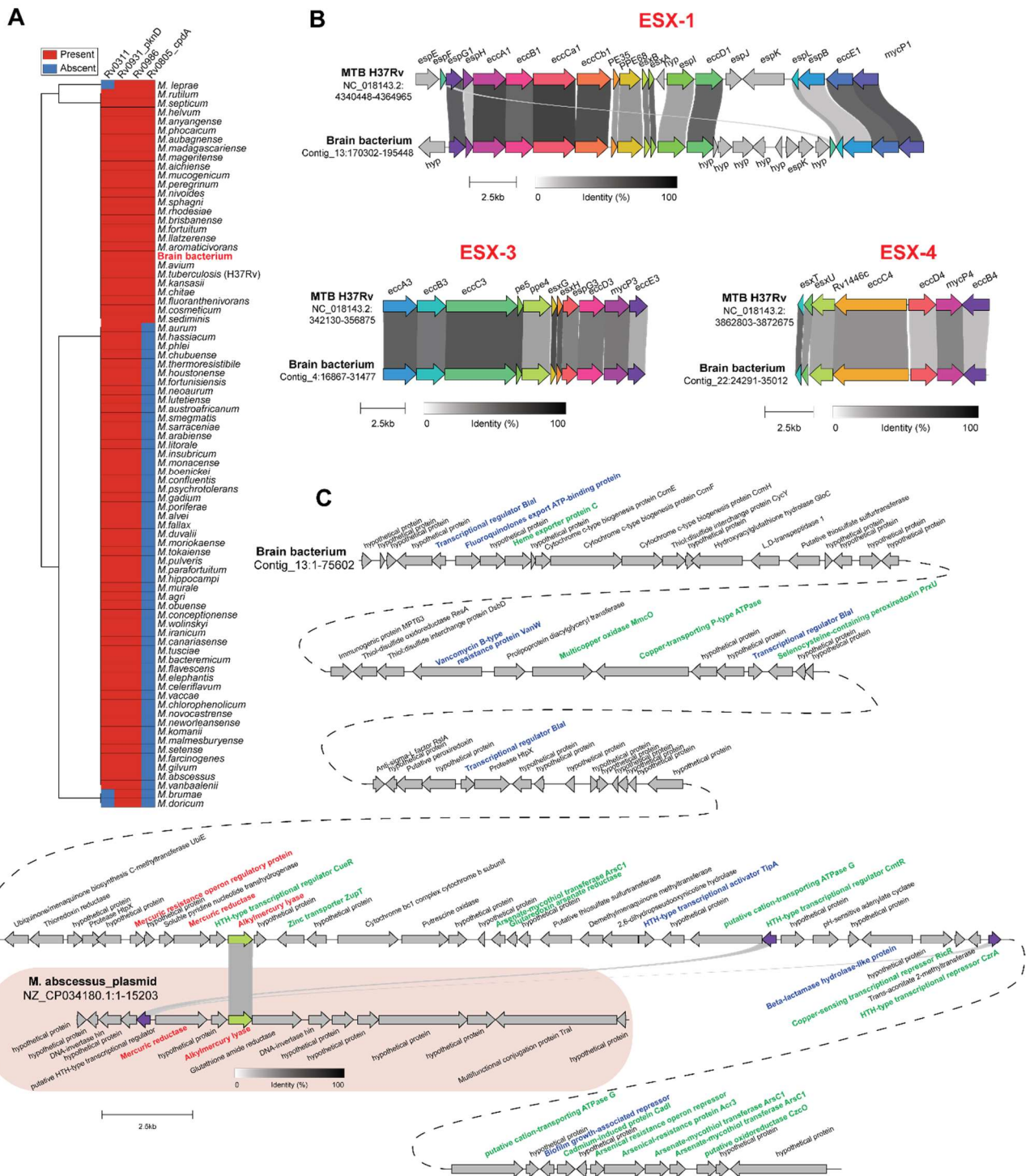


Figure S5: Virulence genes of the brain NTM (A) Heatmap showing the presence/absence of genes involved in crossing Blood-Brain Barrier (BBB) and brain invasion. **(B)** Gene cluster comparison of type VII secretion systems in *Mycobacterium tuberculosis* and ACB brain bacterium. **(C)** Synteny map of the mercury resistance operon of *Mycobacterium abscessus* compared with the mercury resistance genes (in red) of the ACB brain NTM. The map also shows the neighboring heavy metal- (in green) and antibiotics (in blue) resistance genes.

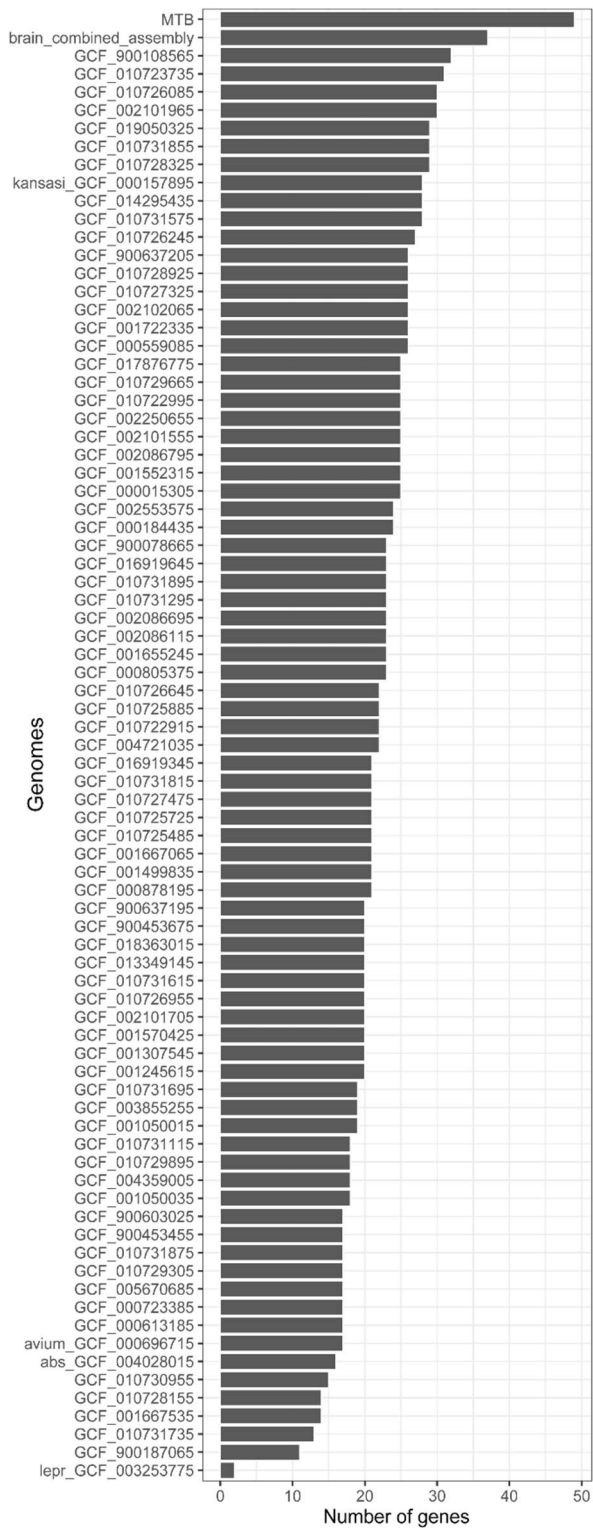


Figure S6: Toxin/Antitoxin (TA) Systems in the analyzed mycobacterial genomes. The horizontal bars refer to the sum of the toxin/antitoxin proteins found in the analyzed genomes.

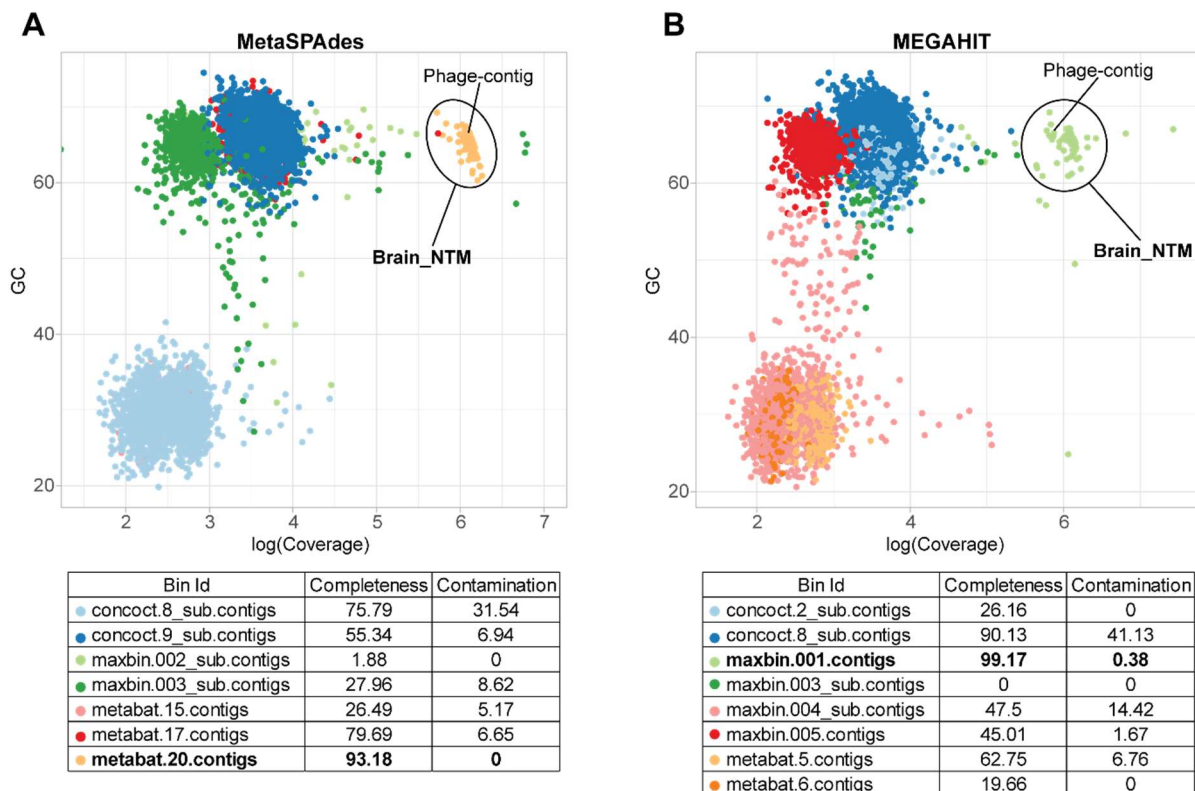


Figure S7: Metagenomic binning of Anna Catharina Bischoff's (ACB) brain sample. The upper biplots show the clustering of the contigs assembled by MetaSPAdes (**A**) and MEGAHIT (**B**). The colors refer to the different bins as resulted from DAS_Tool, which combined the output of three different bidders (Metabat2, MaxBin2, and CONCOCT). The circled contigs refer to the brain_NTM, and the black line inside the circle refers to the contig_38 (which contains a phage genome). The lower panel shows the completeness and the contamination of the resulting bins as estimated by CheckM. The bins in bold fonts are the brain_NTM.

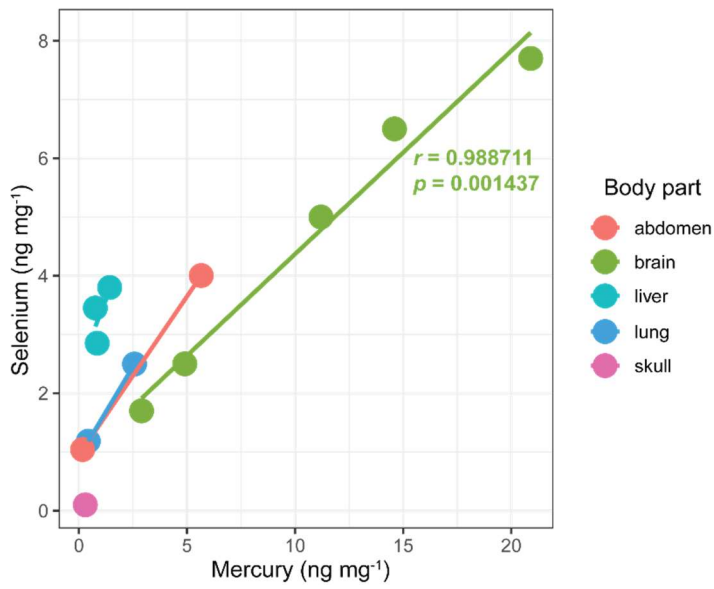


Figure S8: Correlation analysis between concentrations of mercury (Hg) and Selenium (Se) in different body parts.

Supplementary references

1. Huson DH, Beier S, Flade I, Gorska A, El-Hadidi M, Mitra S, Ruscheweyh HJ, Tappu R: **MEGAN Community Edition - Interactive Exploration and Analysis of Large-Scale Microbiome Sequencing Data**. *PLoS Comput Biol* 2016, **12**(6):e1004957.
2. Lu J, Rincon N, Wood DE, Breitwieser FP, Pockrandt C, Langmead B, Salzberg SL, Steinegger M: **Metagenome analysis using the Kraken software suite**. *Nature protocols* 2022:1-25.
3. Herbig A, Maixner F, Bos KI, Zink A, Krause J, Huson DH: **MALT: Fast alignment and analysis of metagenomic DNA sequence data applied to the Tyrolean Iceman**. *bioRxiv* 2016:050559.
4. Jónsson H, Ginolhac A, Schubert M, Johnson P, Orlando L: **mapDamage2.0: fast approximate Bayesian estimates of ancient DNA damage parameters**. *Bioinformatics* 2013, **29**(13):1682-1684.
5. Neukamm J, Peltzer A, Nieselt K: **DamageProfiler: Fast damage pattern calculation for ancient DNA**. *Bioinformatics* 2021.
6. Hotz G, Opitz-Belakhal C: **Anna Catharina Bischoff. Die Mumie aus der Barfüsserkirche**. Basel, Switzerland: Christoph Merian Verlag; 2021.

Discussion

(This page intentionally left blank)

Discussion

Throughout the study, we had access to different types of ancient biological materials, preserved under different geoclimatic conditions, which we aimed to analyze for different purposes. The main aim of this study was to improve the current analytical methods of ancient DNA, either *in situ* (during the sampling from the site), *in vitro* (during the processing the lab), and/or even *in silico* (during the data analysis). This should cover most of the aspects of the ancient DNA field and is expected to lead to results that have not been yet achieved.

DNA preservation versus isolation

Considering the preservation conditions, in the first study, the analyzed materials were human skeletal bones surrounded by calcite stone deposits, preserved in Wimsener water caves under constant water dripping at constant temperature of 7-8 °C. While in the second study, we analyzed paleofeces from the Hallstatt salt mines in Austria, where the extraordinary preservation of the analyzed ancient biomolecules, either proteins or nucleic acids, was mainly attributed to the combination effect of constant temperature of 8 °C throughout the year and the high salt concentration inside the mines. In the third study, the samples were taken from the mummy of the Anna Catharina Bischoff (ACB), which was buried in the Barfüsserkirche in Basel, Switzerland, where the hygroscopic effect of the salt surrounding the coffin and the high concentration of mercury inside the mummy prevented the decomposition of the mummy and aided in the preservation of the DNA. However, such unusual presence of chemical elements might have aided to preserve the ancient materials for centuries and even for millennia, however they represent a big challenge or DNA extraction, since they interfere with downstream enzymatic processes. Therefore, we used a newly developed method for ancient DNA extraction, developed in our labs using linear polyacrylamide (LPA) for precipitating DNA after enzymatic digestion of soft tissues or demineralization of bones (Maixner et al., 2022). Such a method was proved in comparison with other methods dedicated for ancient DNA extraction, to be more efficient in obtaining ultra-small DNA fragments (as small as 20 bp), which is one of the inherent features of ancient DNA.

While it remains to assess ancient DNA damage levels, we need to consider two main type of damages that occurs to DNA post-deposition. First, the physical degradation of DNA which is linked to its half-life under different physical conditions (Dabney et al., 2013). Second, the

chemical changes in the DNA like terminal cytosine-to-thymine deamination (C-to-T). In our study, and under such unusual preservation conditions, all samples displayed the first type of damage, as they showed median fragment length of 50-75 nt. However, they were showing low chemical damage (i.e., C-to-T) considering their age (except for the Wimsener cave samples). Interestingly, within the Barfüsser mummy samples, we observed negative correlation between the levels of mercury concentrations and levels of DNA deamination, which demonstrates intra-mummy variability of ancient DNA damage levels (Sarhan et al., 2023). Unfortunately, such kind of analysis was not feasible to be performed systematically in all samples, however it represents an indication on the mercury and high-salt effect on ancient DNA preservation.

Sampling the unsampled

Sampling ancient materials for DNA analysis is often faced with many obstacles that hinder the progress of ancient DNA search. Those obstacles are related to ethical considerations, museums' internal rules, and distortion of the archeological objects, since sampling often involves physical damages, e.g., drilling skulls to access the petrous bone, grinding teeth, or cutting pieces of soft tissues. In 2021, a consortium of researchers (including archaeologists, anthropologists, museum curators, and geneticists) published global guidelines for ethics of DNA research on human remains (Alpaslan-Roodenberg et al., 2021). The third guideline stated, "*Researchers must minimize damage to human remains*". In our first study (Sarhan et al., 2021), our aim initially was to molecularly identify the sex of the individual and the mitochondrial haplogroup using some skeletal remains. But since the bones were laid in the cave and surrounded with calcite stone layers, accumulated over the years as a result of continuous water dripping from the cave ceiling, we hypothesized that there might have been a kind of migration of DNA molecules from the bones to the stones. To test this hypothesis, we sampled additional spots from the stone parts, and exposed them to the same molecular and shotgun analyses of the bones. The results showed that the stone parts contained ancient DNA fragments belonging to the same individual, which proved our initial hypothesis on the migration of the DNA from the bones to the stones, under the effect of water dripping. This was also applicable for the microbial DNA, as we found similar microbial DNA profiles in both phases. Simply, we can use the analogy of a chromatographic system, where we can consider the bone/stone complex as the stationary phase and the water dripping as the mobile phase. So, the DNA molecules under such a dynamic system migrated towards the direction of the gravity.

The same concept of sampling the surroundings of archeological objects can be extended for other similar cases. For example, we analyzed the defrosted water from the Iceman mummy (Ötzi) as a proxy for the whole-body microbiome, and additionally retrieved his complete mitochondrial genome (Data not published). Another example, in the third study where we analyzed the cloths and insects from the coffin of the mummy of Anna Catharina Bischoff (ACB) and were able to identify her mitochondrial haplogroup. Similarly, Pedersen et al. reported retrieval of ancient human DNA from the cement that glues head lice nits to the hair of 1500–2000 years old Argentinian mummies (Pedersen et al., 2022). The percentages of human DNA they retrieved from the nit cement was comparable to the human DNA they retrieved from a tooth.

Our findings bring different benefits to the field of ancient DNA, since they bring into attention potential additional sources of ancient DNA, which can help in the future to avoid the widely used “inevitable” destructive sampling of ancient materials. This also might convince stakeholder to allow research on more precious materials, like mummies or other valuable museum showcases.

Feces wealth for human health

One example for the unusual ancient samples that are already established in the field is the ancient stools (also known as paleofeces or coprolites) that are rarely found in archeological sites. Paleofeces are very particular since they represent a snapshot in the persons past. In this study, we analyzed different paleofeces from the Hallstatt salt mines of Austria, which is considered to be the oldest salt mines in the world, as there are records for continuous mining activities for 7000 years (Spinei, 2010). The radiocarbon dating indicated that the samples belong to the Bronze Age, the Iron Age, and to the Baroque period. Therefore, it was very intriguing to deeply analyze those samples, using multidisciplinary approaches to gain insights into the gut microbiome structure, dietary habits, and health status of the people lived in such a particular place throughout time (Maixner et al., 2021).

Recent studies on human gut microbiome showed that lifestyle and dietary habits can influence the microbiome structure (Rampelli et al., 2015b; Shanahan et al., 2021). On the global level, human gut microbiome was suggested to be divided into Westernized- and non-Westernized microbiomes (Pasolli et al., 2019; Tett et al., 2019). The Westernized life style and urbanization was linked with loss of microbial diversity in the gut due to industrialization, urbanization and high fat diets, while the non-Westernized microbiome was

linked to higher microbial diversity and high fibrous diets (Pasolli et al., 2019; Segata, 2015; Tett et al., 2019). Compared with modern human gut microbiomes, the structural and functional analysis of the microbiome of Hallstatt paleofeces displayed resemblance to the non-Westernized microbiomes. This was even noticeable in the more recent sample of the Baroque times, which indicates that the Westernization effect and loss of diversity happened at a later timepoint. For sure this is just an indication now, and we would need to include more samples from similar era in Europe, i.e., 18th, 19th, and early 20th centuries, to confirm that and to determine exactly at which time point such a shift happened, and how people dietary habits changed with the major sociological, economical, and political changes.

In addition to the microbiome analysis, we attempted to infer the dietary components, that are among the main drivers of the microbiome composition, by analyzing the plant and animal DNA content of the feces. We developed a computational pipeline to compare the shotgun metagenomic sequences against different genomic, organellar, and curated marker-genes databases, and on the top of that, we employed microscopic and proteomic analyses to gain further support for the metagenomic results. The dietary molecular analysis results (metagenomic and metaproteomic) did not match totally with the microscopic results, and this is likely because of the databases biases and limitations, and the strict selection and filtering criteria we set to minimize as much as possible the false positives. For example, to infer the plant diet, we compared our sequences to NCBI-nr, chloroplast genome, mitochondrial genome, *rbcL/matK*, and plant inter-transcribed spacers (ITS). In the meantime, the complete genome sequences of many plants, particularly those not of economic importance, have not yet been achieved, due to several reasons (Michael and VanBuren, 2020; Twyford, 2018). For instance, if we have a plant with marker gene/chloroplast genome available, but the complete nuclear and mitochondrial genomes are not yet available, this plant will likely be excluded based on the current selection scheme, which was based on a majority rule. The same picture is applicable to the protein data since it targets only the coding sequences. In the future, this heuristic selection criteria will need to be based on weighted scores, considering whether the target sequences are available in the used database or not. Additionally, it would need also to be validated with control samples of modern human stool samples (Maixner, 2019), preferably with similar lifestyle and microbiome structures.

From genome reconstruction to inferring function

In one of the analyzed Iron Age paleofeces samples, we noticed high prevalence of two ascomycetous fungal species, i.e., *Penicillium roqueforti* and *Saccharomyces cerevisiae*. Their abundance was high enough, so we opted to reconstruct their genomes by aligning the metagenomic reads to their reference genomes, which resulted in reconstruction of near-complete genomes of both species. Then, different questions were imposed - Why were they there? Were they part of the diet? Were they just environmental contaminant that cumulated after deposition? In standard situations, transcriptomic analysis would be the most suitable approach to follow. For example, if it is possible to identify transcripts that are known to be specific for certain fermentation fate, this would give a direct explanation for their presence in the feces. However, this is not feasible yet with the current knowledge on the ancient RNA extraction and handling in the lab. There are only few studies which succeeded to recover ancient RNA molecules, specifically eukaryotic microRNA (Keller et al., 2017; Smith et al., 2019) and RNA viruses (Smith et al., 2014). Therefore, we had to choose different approach to infer the function of those two fungi without having transcriptomic data. But since both species are of high economic importance in food and beverage fermentation, there were plenty of genomes available representing different functional fates. We opted to use combination of phylogenetic and population genetic analyses, in addition to detection of function-specific marker genes to understand the potential functions of the reconstructed ancient genomes. In case of the *Penicillium roqueforti*, the ancient genome showed clear clustering to cheese making strains, being basal to the non-Roquefort cheese clade, not to the environmental strains. The same indication was obtained from the admixture analysis. While, in case of the *Saccharomyces cerevisiae*, the ancient genomes clustered basal to one of the beer clades, and additionally contained the RTM1 gene which is a specific marker gene for beer fermenting yeasts.

The approach we used to infer the potential functions of ancient fungal genomes is new to the field, since so far and to the best of our knowledge there is only one ancient fungal genome reported, i.e., the basidiomycetous *Malassezia restricta*, from an 18th century blood-stained paper sample (De-Dios et al., 2020). With our two ancient fungal genomes, we report the first ancient ascomycetous genomes and the increase the total count of ancient fungal genomes to three. Additionally, the results of this study extended the horizon of the ancient DNA research to the food processing techniques since we are presenting the very first molecular genomic evidence for ancient food fermentation. It would be interesting to understand more

about the process itself, like how they performed the fermentation, and at what temperature, and whether the miners exploited the constant temperature of the cave to store the cheese and to age the beer or not. To answer those questions, more efforts would be needed combining archeological, anthropological, and molecular data to come up with a solid picture on how these processes were carried out. Considering that there is a growing interest now in analyzing paleofeces for different purposes, these questions are expected to be addressed during the coming years.

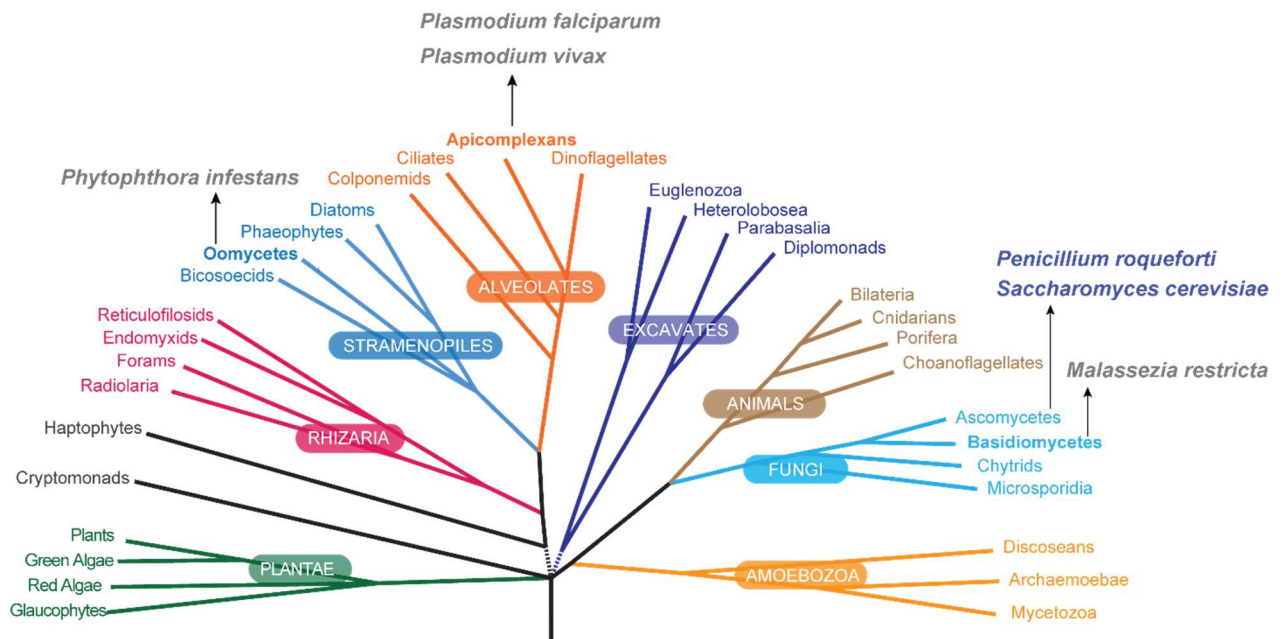


Figure 4: Tree of eukaryotes – Adapted from (Keeling, 2019). A schematic tree summarizing current data on the relationships between major lineages of eukaryotic domain. The black arrows refer to the reported ancient microbial species (in grey font) with available mitochondrial or nuclear genomes as of 2020. The species in dark blue font are the fungal genomes reported in our study.

Going beyond the current pool of pathogens

In the ancient DNA field, most studies used mapping-to-reference approach to detect and to reconstruct ancient genomes. This approach only addresses the known microbes with available reference genomes and neglect the vast majority of the metagenome (i.e., the microbial dark matter), including not-yet-described taxa. In addition, such an approach could be biased with the reference genome used for the alignment process. For instance, if the ancient genome had gone through major genomic deletions or rearrangement, it will not be possible to detect such a change using the reference-based approaches. Likewise, if we are dealing with novel organisms that have not been yet described.

In the third study of this thesis (Sarhan et al., 2023), we analyzed different samples from the 18th century mummy of Anna Catharina Bischoff (ACB) – A mummy that was found in 1975 in during the renovation of Franciscan church in Basel, Switzerland. The reason behind her mumification was highly attributed to the high mercury concentrations all-over her body, particularly in the brain (Hotz and Opitz-Belakhal, 2021). The computed tomography (CT) scanning of the mummy showed some bone lesions that might be indicating to syphilis, or even tuberculosis or Paget’s disease (Tzankov and Bircher, 2021). Hence, our aim was to molecularly assess these assumptions, using ancient DNA techniques. Therefore, we analyzed different samples from different body parts (i.e., teeth, skull, brain, dura matter, lungs, liver, intestines, and stool), in addition to other samples from the coffin she was found in, as control samples (i.e., cloths/textile, maggots, and soil).

To assess the previous assumptions, we initially used a targeted approach for the data analysis, i.e., mapping-to-reference approach, which failed to prove presence of either *Treponema pallidum* (syphilis-causing) or *Mycobacterium tuberculosis* (tuberculosis-causing). Therefore, we decided to follow an explorative untargeted approach for the data analysis, by employing *de novo* assembly. We used different metagenomic assemblers which use different assembly algorithms (i.e., MEGAHIT and MetaSPAdes) to perform the metagenomic assembly, by which we were able to retrieve two high-quality genomes (>90% completeness and <5% contamination), one from each assembler. We used the whole genome sequences of both genomes to assign taxonomy, which resulted in reliable taxonomy to the family-level, namely to the Mycobacteriaceae, with no clear assignment to lower taxonomic levels. Mycobacteriaceae members are generally regarded as environmental microbes (Honda et al., 2018), despite of being containing common typical pathogens like *Mycobacterium tuberculosis* and *Mycobacterium leprae* (Natarajan et al., 2020; Röltgen et al., 2020). Other members of the family are known as non-tuberculous mycobacteria (NTM). They are characterized by being infectious to immunocompromised patients, although being ubiquitous in the different environmental niches like water and soil (Pennington et al., 2021; Tan and Kasperbauer, 2021). Therefore, it was crucial for us to deeply characterize the genomes we recovered to understand whether they are environmental contaminants, or they are endogenous to the brain of the mummy. Deep analysis of the reconstructed genome revealed the following: i) containment of different virulence gene clusters similar to *Mycobacterium tuberculosis* complex; ii) potential ability to cross the blood brain barrier (BBB); iii) presence of a phage genome which might have increased the virulence of this bacterium; iv) the

genome contains variety of gene systems that enable survival in harsh conditions of heavy metals and antibiotics.

In modern clinical diagnostics, bacterial infections can be diagnosed in a variety methods, e.g., nucleic acid detection from blood cultures, direct detection of pathogen blood samples, direct metagenomics and transcriptomics, and rapid antibiotic susceptibility testing (AST) methods (Peri et al., 2021). However, most of those tools are not applicable for ancient pathogens, due to some of the inherent features of the field, e.g., cultivation is almost impossible because of the cellular degradation of hosts and pathogens, RNA protocols are not yet established, and scarcity of analyzed materials and blood stream accessibility. Therefore, majority of the reported genomes of ancient pathogens are those of historical significance and with enough modern genomic information (e.g., *Yersinia pestis*), which makes it feasible for researchers to use mapping-to-reference approach to detect them (Spyrou et al., 2019). Evidently, the case we analyzed in this study demonstrate one of the major limitations of relying solely on mapping-to-reference approach to diagnose pathogens, in general, and ancient pathogens in particular, as we could not assign the symptoms to any known pathogen, although we analyzed samples from all over the body. Using *de novo* metagenomic assembly facilitated identification of a novel undescribed pathogen, which is likely to be the reason behind the clinical symptoms of the mummy. Moreover, the approach we followed for combining the output, of two different assembly algorithms, resulted in a higher quality genome even with modern genomic standards (**Figure 3c, Table 1**).

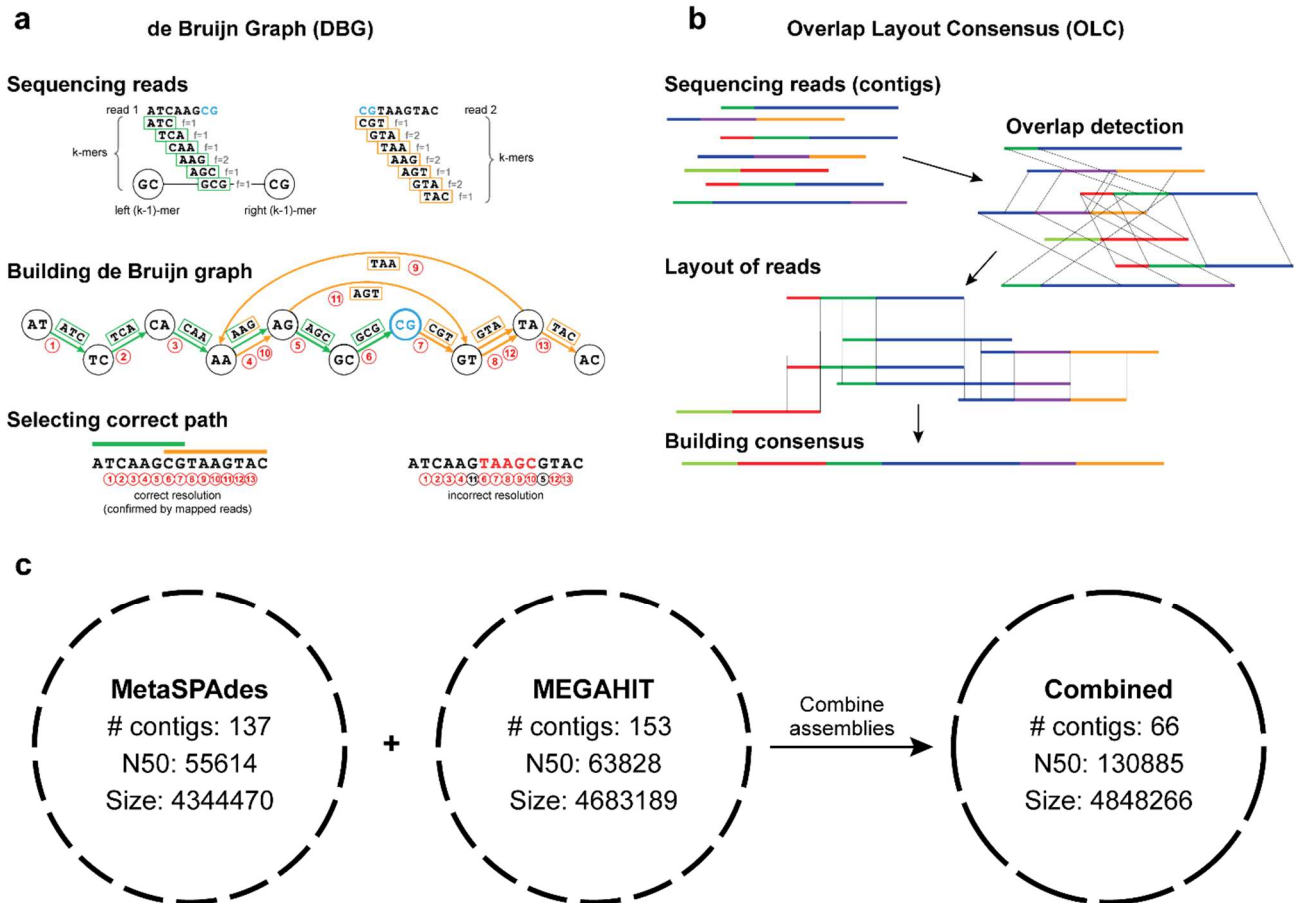


Figure 5: *De novo* assembly approaches. a, de Bruijn graph assembly (DBG). b, Overlap layout consensus (OLC). The scheme of DBG and OLC are adapted from Vollmers et al. (Vollmers et al., 2017). c, combining the outputs of MEGAHIT and MetaSPAdes assemblers which are using DBG algorithm for assembly. The numbers refer to the values obtained from the paper III in the results sections (Sarhan et al., 2023).

It is worth to mention that there is a noticeable increase of ancient DNA studies which successfully implemented *de novo* metagenomic assembly to reconstruct ancient microbial genomes. One of the early uses of *de novo* assembly in ancient DNA field was to assemble the genome of a medieval *Mycobacterium leprae* (Schuenemann et al., 2013). Recently, *de novo* assembly was used to analyze ancient oral microbiota which facilitated the discovery of not-yet-described archaeal diversities that are not anymore present in the contemporary populations, suggesting possible shift in the human oral microbiota over the last millennia (Granehall et al., 2021). It was also used on a large-scale to analyze ancient paleofeces to reconstruct and discover undescribed human gut microbes (Wibowo et al., 2021). It is important to highlight that our study represent a proof-of-concept for describing previously unknown ancient pathogen, using *de novo* metagenomic assembly.

***De novo* assembly as affected by ancient DNA damages**

Since ancient DNA can be damaged and modified, one question would impose itself is how far *de novo* assembly can be influenced by ancient DNA damaged sites? And whether these damaged sites can be integrated within the resulted contigs resulting in chimeric sequences. In the seminal paper of using *de novo* assembly to reconstruct genomes of ancient gut microbiota, the authors treated the extracted ancient DNA with uracil DNA glycosylase (UDG), an enzyme that cleaves DNA at uracil-containing sites (Wibowo et al., 2021). This step was done to assess whether repairing ancient DNA damages can result in different microbial profiles and improve genome assembly. They reported similar microbial profiles between UDG-treated and UDG-untreated libraries, however the assembly quality measures were less in the UDG-treated libraries, due to average shorter fragment lengths.

Apart from that, *de novo* metagenomic assemblers, e.g., MEGAHIT and MetaSPAdes, in their first step start to build dictionaries of all k-mers present in one sample (hash table), then they discard all singleton k-mers. Exclusion of singleton k-mers before building the assembly graph will likely minimize integration of ancient DNA damaged sites in further steps, provided that libraries are void of duplicates. Therefore, performing PCR-deduplication step before starting the assembly process is highly recommended particularly for highly damaged samples.

Do we really need *de novo* assembly?

As demonstrated earlier, *de novo* assembly can reveal unknown pathogen, but how about if we have particular focus on one pathogen and would like to understand it in an evolutionary context? Currently, our ability to reconstruct ancient microbial genomes using mapping-to-reference methods, is limited with the availability of representative reference genomes. Therefore, we need to differentiate between two types of pathogens: First, clonal pathogens which exhibit clonal populations with limited genetic diversity, like *Mycobacterium tuberculosis*, which is an intracellular pathogen that depends mainly on transmission from one individual to another. These kinds of pathogens usually maintain similar genome size and entertain very homogenous population since the main source of genetic variation in the intrinsic mutation rate. Second, the recombinant pathogens, like *Mycobacterium avium* complex or *Helicobacter pylori*, which have diverse genome sizes and can display genetic recombination which leads to heterogenous population structure. These recombinant

pathogens get in contact with environmental milieu where they can be transformed or transduced by external DNA sources which increase their genetic survivability.

The mapping-to-reference approach can be still enough for analysis of clonal pathogens, but for recombinant pathogens, it is incapable of unveiling the major genomic rearrangement events, like transductions or big deletions. Therefore, *de novo* assembly would have more value in those cases.

Opportunities and challenges

Application of *de novo* metagenomic assembly in ancient DNA field would have great implications on understanding the human microbiome evolution and how human habits influenced their gut and oral flora (Granehall et al., 2021; Wibowo et al., 2021). The *de novo* assembly not only will enable identification of novel extinct and previously undescribed pathogens, but also can identify new genomic features of known genomes, that might have been excised/deleted during earlier times. This might seem to be a speculation, but similar findings were reported in *de novo* modern human genome assembly (Wong et al., 2018).

It is important however not to assume the applicability of *de novo* assembly in all ancient DNA cases, since the method has also some inherent limitations similar to modern DNA (Yang et al., 2021): i) it can create chimeras by assembling sequences from different genomes; ii) it needs high memory usage; iii) it cannot handle highly repetitive sequences; and iv) it can be influenced with uneven sequencing depth.

References

- Achtman, M., and Wagner, M.J.N.r.m. (2008). Microbial diversity and the genetic nature of microbial species. *6*, 431-440.
- Almeida, A., Mitchell, A.L., Boland, M., Forster, S.C., Gloor, G.B., Tarkowska, A., Lawley, T.D., and Finn, R.D. (2019). A new genomic blueprint of the human gut microbiota. *Nature* *568*, 499-504.
- Alpaslan-Roodenberg, S., Anthony, D., Babiker, H., Bánffy, E., Booth, T., Capone, P., Deshpande-Mukherjee, A., Eisenmann, S., Fehren-Schmitz, L., and Frachetti, M. (2021). Ethics of DNA research on human remains: Five globally applicable guidelines. *Nature* *599*, 41-46 %! Ethics of DNA research on human remains: Five globally applicable guidelines %@ 1476-4687.
- Anastasiou, E., Papathanasiou, A., Schepartz, L.A., and Mitchell, P.D. (2018). Infectious disease in the ancient Aegean: intestinal parasitic worms in the Neolithic to Roman period inhabitants of Kea, Greece. *Journal of Archaeological Science: Reports* *17*, 860-864.
- Aufderheide, A.C., Salo, W., Madden, M., Streitz, J., Buikstra, J., Guhl, F., Arriaza, B., Renier, C., Wittmers Jr, L.E., and Fornaciari, G. (2004). A 9,000-year record of Chagas' disease. *Proceedings of the National Academy of Sciences* *101*, 2034-2039.
- Bos, K.I., Jäger, G., Schuenemann, V.J., Vågene, Å.J., Spyrou, M.A., Herbig, A., Nieselt, K., and Krause, J. (2015). Parallel detection of ancient pathogens via array-based DNA capture. *Philosophical Transactions of the Royal Society B: Biological Sciences* *370*, 20130375.
- Bowers, R.M., Kyrpides, N.C., Stepanauskas, R., Harmon-Smith, M., Doud, D., Reddy, T.B.K., Schulz, F., Jarett, J., Rivers, A.R., Elie-Fadrosh, E.A., *et al.* (2017). Minimum information about a single amplified genome (MISAG) and a metagenome-assembled genome (MIMAG) of bacteria and archaea. *Nat Biotechnol* *35*, 725-731 %727 2017/2008/2009 %2018 Aug 2018 %! Minimum information about a single amplified genome (MISAG) and a metagenome-assembled genome (MIMAG) of bacteria and archaea %@ 1546-1696 (Electronic) 1087-0156 (Linking).
- Brown, C.T., Hug, L.A., Thomas, B.C., Sharon, I., Castelle, C.J., Singh, A., Wilkins, M.J., Wrighton, K.C., Williams, K.H., and Banfield, J.F. (2015). Unusual biology across a group comprising more than 15% of domain Bacteria. *Nature* *523*, 208-211.
- Cann, R.L., Stoneking, M., and Wilson, A.C. (1987). Mitochondrial DNA and human evolution. *Nature* *325*, 31-36.
- Carpenter, M.L., Buenrostro, J.D., Valdiosera, C., Schroeder, H., Allentoft, M.E., Sikora, M., Rasmussen, M., Gravel, S., Guillén, S., and Nekhrizov, G. (2013). Pulling out the 1%: whole-genome capture for the targeted enrichment of ancient DNA sequencing libraries. *The American Journal of Human Genetics* *93*, 852-864.
- Dabney, J., Meyer, M., and Pääbo, S. (2013). Ancient DNA damage. *Cold Spring Harbor perspectives in biology* *5*, a012567.
- De-Dios, T., van Dorp, L., Charlier, P., Morfopoulou, S., Lizano, E., Bon, C., Le Bitouzé, C., Alvarez-Estape, M., Marquès-Bonet, T., and Balloux, F. (2020). Metagenomic analysis of a blood stain from the French revolutionary Jean-Paul Marat (1743–1793). *Infection, Genetics and Evolution* *80*, 104209 %! Metagenomic analysis of a blood stain from the French revolutionary Jean-Paul Marat (101743–101793) %@ 101567-101348.
- De-Dios, T., van Dorp, L., Gelabert, P., Carøe, C., Sandoval-Velasco, M., Fregel, R., Escosa, R., Aranda, C., Huijben, S., and Balloux, F. (2019). Genetic affinities of an eradicated European *Plasmodium falciparum* strain. *Microbial genomics* *5*.
- Drancourt, M., Aboudharam, G., Signoli, M., Dutour, O., and Raoult, D.J.P.o.t.N.A.o.S. (1998). Detection of 400-year-old *Yersinia pestis* DNA in human dental pulp: an approach to the diagnosis of ancient septicemia. *95*, 12637-12640.
- Duchêne, S., Ho, S.Y., Carmichael, A.G., Holmes, E.C., and Poinar, H. (2020). The recovery, interpretation and use of ancient pathogen genomes. *Current Biology* *30*, R1215-R1231.
- Gelabert, P., Sandoval-Velasco, M., Olalde, I., Fregel, R., Rieux, A., Escosa, R., Aranda, C., Paaijmans, K., Mueller, I., and Gilbert, M.T.P. (2016). Mitochondrial DNA from the eradicated European

Plasmodium vivax and P. falciparum from 70-year-old slides from the Ebro Delta in Spain. *Proceedings of the National Academy of Sciences* 113, 11495-11500.

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

Ginolhac, A., Rasmussen, M., Gilbert, M.T.P., Willerslev, E., and Orlando, L. (2011). mapDamage: testing for damage patterns in ancient DNA sequences. *Bioinformatics* 27, 2153-2155.

Gitschier, J.J.P.G. (2008). Imagine: an interview with svante pääbo. 4, e1000035.

GraneHäll, L., Huang, K.D., Tett, A., Manghi, P., Paladin, A., O'Sullivan, N., Rota-Stabelli, O., Segata, N., Zink, A., and Maixner, F. (2021). Metagenomic analysis of ancient dental calculus reveals unexplored diversity of oral archaeal Methanobrevibacter. *Microbiome* 9, 1-18.

Green, E.J., and Speller, C.F. (2017). Novel substrates as sources of ancient DNA: prospects and hurdles. *Genes* 8, 180 %! Novel substrates as sources of ancient DNA: prospects and hurdles %@ 2073-4425.

Green, R.E., Krause, J., Briggs, A.W., Maricic, T., Stenzel, U., Kircher, M., Patterson, N., Li, H., Zhai, W., and Fritz, M.H.-Y.J.s. (2010). A draft sequence of the Neandertal genome. 328, 710-722.

Hagelberg, E., Hofreiter, M., and Keyser, C. (2015). Ancient DNA: the first three decades (The Royal Society), pp. 20130371 %! Ancient DNA: the first three decades %@ 20130962-20138436.

Hansen, H.B., Damgaard, P.B., Margaryan, A., Stenderup, J., Lynnerup, N., Willerslev, E., and Allentoft, M.E. (2017). Comparing ancient DNA preservation in petrous bone and tooth cementum. *PLoS one* 12, e0170940.

Hawass, Z., Gad, Y.Z., Ismail, S., Khairat, R., Fathalla, D., Hasan, N., Ahmed, A., Elleithy, H., Ball, M., and Gaballah, F. (2010). Ancestry and pathology in King Tutankhamun's family. *Jama* 303, 638-647.

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

Honda, J.R., Viridi, R., and Chan, E.D. (2018). Global environmental nontuberculous mycobacteria and their contemporaneous man-made and natural niches. *Frontiers in microbiology* 9, 2029.

Hotz, G., and Opitz-Belakhal, C. (2021). Anna Catharina Bischoff. Die Mumie aus der Barfüsserkirche (Basel, Switzerland: Christoph Merian Verlag).

Hugenholtz, P.J.G.b. (2002). Exploring prokaryotic diversity in the genomic era. 3, 1-8.

Jäger, H.Y., Maixner, F., Pap, I., Szikossy, I., Pálfi, G., and Zink, A.R. (2022). Metagenomic analysis reveals mixed Mycobacterium tuberculosis infection in a 18th century Hungarian midwife. *Tuberculosis*, 102181.

Keeling, P.J. (2019). Combining morphology, behaviour and genomics to understand the evolution and ecology of microbial eukaryotes. *Philosophical Transactions of the Royal Society B* 374, 20190085 %! Combining morphology, behaviour and genomics to understand the evolution and ecology of microbial eukaryotes %@ 20190962-20198436.

Keller, A., Kreis, S., Leidinger, P., Maixner, F., Ludwig, N., Backes, C., Galata, V., Guerriero, G., Fehlmann, T., and Franke, A. (2017). miRNAs in ancient tissue specimens of the Tyrolean Iceman. *Molecular Biology and Evolution* 34, 793-801 %! miRNAs in ancient tissue specimens of the Tyrolean Iceman %@ 0737-4038.

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

Knapp, M., and Hofreiter, M. (2010). Next generation sequencing of ancient DNA: requirements, strategies and perspectives. *Genes* 1, 227-243 %! Next generation sequencing of ancient DNA: requirements, strategies and perspectives %@ 2073-4425.

Kocher, A., Papac, L., Barquera, R., Key, F.M., Spyrou, M.A., Hübner, R., Rohrlach, A.B., Aron, F., Stahl, R., and Wissgott, A. (2021). Ten millennia of hepatitis B virus evolution. *Science* 374, 182-188.

Lasken, R.S.J.N.R.M. (2012). Genomic sequencing of uncultured microorganisms from single cells. *10*, 631-640.

Lewis, W.H., Tahon, G., Geesink, P., Sousa, D.Z., and Ettema, T.J.J.N.R.M. (2021). Innovations to culturing the uncultured microbial majority. *19*, 225-240.

MacKenzie, L., Speller, C.F., Holst, M., Keefe, K., and Radini, A. (2021). Dental calculus in the industrial age: human dental calculus in the Post-Medieval period, a case study from industrial Manchester. *Quaternary International*.

Maixner, F. (2019). Molecular Reconstruction of the Diet in Human Stool Samples. *Msystems* *4*, e00634-00619.

Maixner, F., Krause-Kyora, B., Turaev, D., Herbig, A., Hoopmann, M.R., Hallows, J.L., Kusebauch, U., Vigl, E.E., Malferteiner, P., Megraud, F., *et al.* (2016). The 5300-year-old *Helicobacter pylori* genome of the Iceman. *Science* *351*, 162-165 %167 2016/2001/2009 %2018 Jan 2018 %! The 5300-year-old *Helicobacter pylori* genome of the Iceman %@ 1095-9203 (Electronic) 0036-8075 (Linking).

Maixner, F., Mitterer, C., Jäger, H.Y., Sarhan, M.S., Valverde, G., Lücker, S., Piombino-Mascalì, D., Szikossy, I., Molnár, E., and Pálfi, G. (2022). Linear polyacrylamide is highly efficient in precipitating and purifying environmental and ancient DNA. *Methods in Ecology and Evolution* *13*, 653-667 %! Linear polyacrylamide is highly efficient in precipitating and purifying environmental and ancient DNA %@ 2041-2210X.

Maixner, F., Sarhan, M.S., Huang, K.D., Tett, A., Schoenafinger, A., Zingale, S., Blanco-Míguez, A., Manghi, P., Cemper-Kiesslich, J., Rosendahl, W., *et al.* (2021). Hallstatt miners consumed blue cheese and beer during the Iron Age and retained a non-Westernized gut microbiome until the Baroque period. *Current Biology* *31*, 5149-5162.e5146.

Maixner, F., Thomma, A., Cipollini, G., Widder, S., Rattei, T., and Zink, A. (2014). Metagenomic analysis reveals presence of *Treponema denticola* in a tissue biopsy of the Iceman. *PLoS One* *9*, e99994.

Maixner, F., Turaev, D., Cazenave-Gassiot, A., Janko, M., Krause-Kyora, B., Hoopmann, M.R., Kusebauch, U., Sartain, M., Guerriero, G., and O'Sullivan, N. (2018). The Iceman's last meal consisted of fat, wild meat, and cereals. *Current Biology* *28*, 2348-2355. e2349.

Majander, K., Pfrengle, S., Kocher, A., Neukamm, J., du Plessis, L., Pla-Díaz, M., Arora, N., Akgül, G., Salo, K., and Schats, R. (2020). Ancient bacterial genomes reveal a high diversity of *Treponema pallidum* strains in early modern Europe. *Current Biology* *30*, 3788-3803. e3710.

Marciniak, S., Prowse, T.L., Herring, D.A., Klunk, J., Kuch, M., Duggan, A.T., Bondioli, L., Holmes, E.C., and Poinar, H.N. (2016). *Plasmodium falciparum* malaria in 1st–2nd century CE southern Italy. *Current Biology* *26*, R1220-R1222.

Margaryan, A., Hansen, H.B., Rasmussen, S., Sikora, M., Moiseyev, V., Khoklov, A., Epimakhov, A., Yepiskoposyan, L., Kriiska, A., Varul, L.J.E., *et al.* (2018). Ancient pathogen DNA in human teeth and petrous bones. *8*, 3534-3542.

Martin, M.D., Cappellini, E., Samaniego, J.A., Zepeda, M.L., Campos, P.F., Seguin-Orlando, A., Wales, N., Orlando, L., Ho, S.Y., and Dietrich, F.S. (2013). Reconstructing genome evolution in historic samples of the Irish potato famine pathogen. *Nature communications* *4*, 1-7.

Michael, T.P., and VanBuren, R. (2020). Building near-complete plant genomes. *Current Opinion in Plant Biology* *54*, 26-33.

Mühlemann, B., Jones, T.C., Damgaard, P.d.B., Allentoft, M.E., Shevnina, I., Logvin, A., Usmanova, E., Panyushkina, I.P., Boldgiv, B., and Bazartseren, T. (2018). Ancient hepatitis B viruses from the Bronze Age to the Medieval period. *Nature* *557*, 418-423.

Natarajan, A., Beena, P., Devnikar, A.V., and Mali, S. (2020). A systemic review on tuberculosis. *Indian Journal of Tuberculosis* *67*, 295-311.

Nerlich, A.G., Haas, C.J., Zink, A., Szeimies, U., and Hagedorn, H.G. (1997). Molecular evidence for tuberculosis in an ancient Egyptian mummy. *The Lancet* *350*, 1404.

Neukamm, J., Peltzer, A., and Nieselt, K. (2021). DamageProfiler: Fast damage pattern calculation for ancient DNA. *Bioinformatics* 37, 3652-3653 %! DamageProfiler: Fast damage pattern calculation for ancient DNA %@ 1367-4803.

Noonan, J.P., Coop, G., Kudaravalli, S., Smith, D., Krause, J., Alessi, J., Chen, F., Platt, D., Paabo, S., and Pritchard, J.K.J.s. (2006). Sequencing and analysis of Neanderthal genomic DNA. 314, 1113-1118.

Orlando, L., Allaby, R., Skoglund, P., Der Sarkissian, C., Stockhammer, P.W., Ávila-Arcos, M.C., Fu, Q., Krause, J., Willerslev, E., and Stone, A.C. (2021). Ancient DNA analysis. *Nature Reviews Methods Primers* 1, 1-26.

Pääbo, S. (1985). Molecular cloning of Ancient Egyptian mummy DNA. *Nature* 314, 644-645 %648 1985/1904/1901 %! Molecular cloning of Ancient Egyptian mummy DNA %@ 1476-4687.

Pääbo, S., Higuchi, R.G., and Wilson, A.C. (1989). Ancient DNA and the polymerase chain reaction: the emerging field of molecular archaeology (Minireview). *The Journal of biological chemistry* 264, 9709-9712.

Pallen, M.J.J.I.J.o.S., and Microbiology, E. (2021). The status Candidatus for uncultured taxa of Bacteria and Archaea: SWOT analysis. 71.

Parks, D.H., Imelfort, M., Skennerton, C.T., Hugenholtz, P., and Tyson, G.W. (2015). CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. *Genome Res* 25, 1043-1055 %1047 2015/1005/1016 %1048 Jul %! CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes %@ 1549-5469 (Electronic) 1088-9051 (Linking).

Pasolli, E., Asnicar, F., Manara, S., Zolfo, M., Karcher, N., Armanini, F., Beghini, F., Manghi, P., Tett, A., Ghensi, P., *et al.* (2019). Extensive Unexplored Human Microbiome Diversity Revealed by Over 150,000 Genomes from Metagenomes Spanning Age, Geography, and Lifestyle. *Cell* 176, 649-662 e620 %647 2019/2001/2022 %2018 Jan 2024 %! Extensive Unexplored Human Microbiome Diversity Revealed by Over 2150,2000 Genomes from Metagenomes Spanning Age, Geography, and Lifestyle %@ 1097-4172 (Electronic) 0092-8674 (Linking).

Pedersen, M.W., Antunes, C., De Cahsan, B., Moreno-Mayar, J.V., Sikora, M., Vinner, L., Mann, D., Klimov, P.B., Black, S., and Michieli, C.T. (2022). Ancient Human Genomes and Environmental DNA from the Cement Attaching 2,000-Year-Old Head Lice Nits. *Molecular biology and evolution* 39, msab351.

Pennington, K.M., Vu, A., Challener, D., Rivera, C.G., Shweta, F., Zeuli, J.D., and Temesgen, Z. (2021). Approach to the diagnosis and treatment of non-tuberculous mycobacterial disease. *Journal of Clinical Tuberculosis and Other Mycobacterial Diseases* 24, 100244.

Peri, A.M., Stewart, A., Hume, A., Irwin, A., and Harris, P.N. (2021). New microbiological techniques for the diagnosis of bacterial infections and sepsis in ICU including point of care. *Current Infectious Disease Reports* 23, 1-11.

Pinhasi, R., Fernandes, D., Sirak, K., Novak, M., Connell, S., Alpaslan-Roodenberg, S., Gerritsen, F., Moiseyev, V., Gromov, A., Raczky, P., *et al.* (2015). Optimal Ancient DNA Yields from the Inner Ear Part of the Human Petrous Bone. *PLOS ONE* 10, e0129102.

Poulakakis, N., Parmakelis, A., Lymberakis, P., Mylonas, M., Zouros, E., Reese, D.S., Glaberman, S., and Caccone, A.J.B.I. (2006). Ancient DNA forces reconsideration of evolutionary history of Mediterranean pygmy elephantids. 2, 451-454.

Prentice, M.B., Gilbert, T., and Cooper, A.J.T.L.I.D. (2004). Was the Black Death caused by *Yersinia pestis*? 4, 72.

Rampelli, S., Schnorr, S.L., Consolandi, C., Turrioni, S., Severgnini, M., Peano, C., Brigidi, P., Crittenden, A.N., Henry, A.G., and Candela, M. (2015a). Metagenome sequencing of the Hadza hunter-gatherer gut microbiota. *Current Biology* 25, 1682-1693 %! Metagenome sequencing of the Hadza hunter-gatherer gut microbiota %@ 0960-9822.

Rampelli, S., Schnorr, S.L., Consolandi, C., Turrioni, S., Severgnini, M., Peano, C., Brigidi, P., Crittenden, A.N., Henry, A.G., and Candela, M. (2015b). Metagenome sequencing of the Hadza hunter-gatherer gut microbiota. *Current Biology* 25, 1682-1693.

Reich, D., Green, R.E., Kircher, M., Krause, J., Patterson, N., Durand, E.Y., Viola, B., Briggs, A.W., Stenzel, U., Johnson, P.L., *et al.* (2010). Genetic history of an archaic hominin group from Denisova Cave in Siberia. *Nature* 468, 1053-1060 %1057 2010/1012/1024 %1058 Dec 1023 %! Genetic history of an archaic hominin group from Denisova Cave in Siberia %@ 1476-4687 (Electronic) 0028-0836 (Linking).

Renaud, G., Slon, V., Duggan, A.T., and Kelso, J. (2015). Schmutzi: estimation of contamination and endogenous mitochondrial consensus calling for ancient DNA. *Genome biology* 16, 1-18.

Röltgen, K., Pluschke, G., Spencer, J.S., Brennan, P.J., and Avanzi, C. (2020). The immunology of other mycobacteria: *M. ulcerans*, *M. leprae*. Paper presented at: Seminars in immunopathology (Springer).

Sabin, S., Herbig, A., Vågene, Å.J., Ahlström, T., Bozovic, G., Arcini, C., Kühnert, D., and Bos, K.I. (2020). A seventeenth-century *Mycobacterium tuberculosis* genome supports a Neolithic emergence of the *Mycobacterium tuberculosis* complex. *Genome biology* 21, 1-24 %! A seventeenth-century *Mycobacterium tuberculosis* genome supports a Neolithic emergence of the *Mycobacterium tuberculosis* complex %@ 1474-1760X.

Saheb Kashaf, S., Almeida, A., Segre, J.A., and Finn, R.D. (2021). Recovering prokaryotic genomes from host-associated, short-read shotgun metagenomic sequencing data. *Nature Protocols* 16, 2520-2541.

Sarhan, M.S., Hamza, M.A., Youssef, H.H., Patz, S., Becker, M., ElSawey, H., Nemr, R., Daanaa, H.-S.A., Mourad, E.F., and Morsi, A.T.J.J.o.A.R. (2019). Culturomics of the plant prokaryotic microbiome and the dawn of plant-based culture media—a review. *19*, 15-27.

Sarhan, M.S., Lehmkuhl, A., Straub, R., Tett, A., Wieland, G., Francken, M., Zink, A., and Maixner, F. (2021). Ancient DNA diffuses from human bones to cave stones. *iScience* 24, 103397 %103397 102022/103301/103307 %103398 Dec 103317 %! Ancient DNA diffuses from human bones to cave stones %@ 102589-100042 (Electronic) 102589-100042 (Linking).

Sarhan, M.S., Wurst, C., Tzankov, A., Bircher, A.J., Wittig, H., Briellmann, T., Augsburger, M., Hotz, G., Zink, A., and Maixner, F.J.B.b. (2023). A nontuberculous mycobacterium could solve the mystery of the lady from the Franciscan church in Basel, Switzerland. *21*, 1-16.

Schuenemann, V.J., Singh, P., Mendum, T.A., Krause-Kyora, B., Jäger, G., Bos, K.I., Herbig, A., Economou, C., Benjak, A., and Busso, P. (2013). Genome-wide comparison of medieval and modern *Mycobacterium leprae*. *Science* 341, 179-183.

Segata, N. (2015). Gut microbiome: westernization and the disappearance of intestinal diversity. *Current Biology* 25, R611-R613.

Shanahan, F., Ghosh, T.S., and O'Toole, P.W. (2021). The healthy microbiome—what is the definition of a healthy gut microbiome? *Gastroenterology* 160, 483-494.

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* 111, 2229-2234.

Smith, O., Clapham, A., Rose, P., Liu, Y., Wang, J., and Allaby, R.G. (2014). A complete ancient RNA genome: identification, reconstruction and evolutionary history of archaeological Barley Stripe Mosaic Virus. *Scientific reports* 4, 1-6 %! A complete ancient RNA genome: identification, reconstruction and evolutionary history of archaeological Barley Stripe Mosaic Virus %@ 2045-2322.

Smith, O., Dunshea, G., Sinding, M.-H.S., Fedorov, S., Germonpre, M., Bocherens, H., and Gilbert, M.T.P. (2019). Ancient RNA from Late Pleistocene permafrost and historical canids shows tissue-specific transcriptome survival. *PLoS biology* 17, e3000166.

Spinei, V. (2010). Salz–Reich. 7000 Jahre Hallstatt, Hrsg. Anton Kern, Kerstin Kowarik, Andreas W. Rausch, Hans Reschreiter. *Arheologia Moldovei* 33, 344-346 %! Salz–Reich. 7000 Jahre Hallstatt, Hrsg. Anton Kern, Kerstin Kowarik, Andreas W. Rausch, Hans Reschreiter %@ 2501-5893.

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

Susat, J., Lübke, H., Immel, A., Brinker, U., Macāne, A., Meadows, J., Steer, B., Tholey, A., Zagorska, I., Gerhards, G., *et al.* (2021). A 5,000-year-old hunter-gatherer already plagued by *Yersinia pestis*. *Cell Reports* 35, 109278.

Tan, S., and Kasperbauer, S. (2021). Nontuberculous mycobacteria. Paper presented at: Seminars in Respiratory and Critical Care Medicine (Thieme Medical Publishers, Inc.).

Tett, A., Huang, K.D., Asnicar, F., Fehlner-Peach, H., Pasolli, E., Karcher, N., Armanini, F., Manghi, P., Bonham, K., and Zolfo, M. (2019). The *Prevotella copri* complex comprises four distinct clades underrepresented in westernized populations. *Cell host & microbe* 26, 666-679. e667.

Tett, A., Pasolli, E., Masetti, G., Ercolini, D., and Segata, N. (2021). *Prevotella* diversity, niches and interactions with the human host. *Nature Reviews Microbiology* 19, 585-599 %! *Prevotella* diversity, niches and interactions with the human host %@ 1740-1534.

Twyford, A.D. (2018). The road to 10,000 plant genomes. *Nature Plants* 4, 312-313 %! The road to 310,000 plant genomes %@ 2055-0278.

Tzankov, A., and Bircher, A.J. (2021). Mikroskopische Untersuchung. In Anna Catharina Bischoff Die Mumie aus der Barfüsserkirche (Rekonstruktion einer Basler Frauenbiografie des 18 Jahrhunderts), G. Hotz, and C. Opitz-Belakhal, eds. (Basel, Switzerland: Christoph Merian Verlag %! Mikroskopische Untersuchung %@ 978-3-85616-959-6).

van der Kuyl, A.C., Kuiken, C.L., Dekker, J.T., Perizonius, W.R., and Goudsmit, J. (1995). Nuclear counterparts of the cytoplasmic mitochondrial 12S rRNA gene: a problem of ancient DNA and molecular phylogenies. *Journal of Molecular Evolution* 40, 652-657.

Van Dorp, L., Gelabert, P., Rieux, A., De Manuel, M., De-Dios, T., Gopalakrishnan, S., Carøe, C., Sandoval-Velasco, M., Fregel, R., and Olalde, I. (2020). *Plasmodium vivax* malaria viewed through the lens of an eradicated European strain. *Molecular biology and evolution* 37, 773-785 %! *Plasmodium vivax* malaria viewed through the lens of an eradicated European strain %@ 0737-4038.

Vollmers, J., Wiegand, S., and Kaster, A.-K. (2017). Comparing and evaluating metagenome assembly tools from a microbiologist's perspective-not only size matters! *PloS one* 12, e0169662.

Warinner, C., Speller, C., and Collins, M.J. (2015). A new era in palaeomicrobiology: prospects for ancient dental calculus as a long-term record of the human oral microbiome. *Philosophical Transactions of the Royal Society B: Biological Sciences* 370, 20130376 %! A new era in palaeomicrobiology: prospects for ancient dental calculus as a long-term record of the human oral microbiome %@ 20130962-20138436.

Weyrich, L.S., Duchene, S., Soubrier, J., Arriola, L., Llamas, B., Breen, J., Morris, A.G., Alt, K.W., Caramelli, D., and Dresely, V. (2017). Neanderthal behaviour, diet, and disease inferred from ancient DNA in dental calculus. *Nature* 544, 357-361.

Wibowo, M.C., Yang, Z., Borry, M., Hübner, A., Huang, K.D., Tierney, B.T., Zimmerman, S., Barajas-Olmos, F., Contreras-Cubas, C., and García-Ortiz, H. (2021). Reconstruction of ancient microbial genomes from the human gut. *Nature* 594, 234-239.

Wong, K.H., Levy-Sakin, M., and Kwok, P.-Y. (2018). De novo human genome assemblies reveal spectrum of alternative haplotypes in diverse populations. *Nature communications* 9, 1-9.

Yang, C., Chowdhury, D., Zhang, Z., Cheung, W.K., Lu, A., Bian, Z., and Zhang, L. (2021). A review of computational tools for generating metagenome-assembled genomes from metagenomic sequencing data. *Computational and Structural Biotechnology Journal* 19, 6301-6314.

Yoshida, K., Schuenemann, V.J., Cano, L.M., Pais, M., Mishra, B., Sharma, R., Lanz, C., Martin, F.N., Kamoun, S., and Krause, J. (2013). The rise and fall of the *Phytophthora infestans* lineage that triggered the Irish potato famine. *elife* 2, e00731.

Zink, A.R., Sola, C., Reischl, U., Grabner, W., Rastogi, N., Wolf, H., and Nerlich, A.G. (2003). Characterization of *Mycobacterium tuberculosis* complex DNAs from Egyptian mummies by spoligotyping. *Journal of clinical microbiology* 41, 359-367.

Acknowledgements

First and foremost, I would like to express my deepest gratitude to my supervisors, Dr. Frank Maixner and PD. Dr. Albert Zink, for their unlimited support, guidance, and encouragement throughout my PhD journey. Your expertise, patience, flexibility, and dedication have been invaluable in shaping my research and helping me grow as a researcher.

I would also like to extend my appreciation to my colleagues at Eurac Research, whose camaraderie and collaboration have been instrumental in creating a stimulating and enjoyable research environment. In particular, I would like to thank my PhD colleagues Lena Granehall, Heidi Yoko Jäger, Christina Wurst, Stefania Zingale, and Alexandra Mussauer for their invaluable feedback on my work and for the many fruitful discussions and debates we have had. My thanks are also extended to my officemates Marco Samadelli, Alice Paladin, and Guido Valverde, and also Valentina Coia and Myriam Croze for enlightening me with ideas and thoughts that pushed my thinking beyond my direct specialization.

I would like to acknowledge Giovanna Cipollini, Renate Cassar, and Manuela Milan for their tireless support in handling the logistical and bureaucratic aspects of my PhD experience. Your efficient and professional work has made my life as a researcher much easier.

I am grateful to the intern students who have joined the group throughout the past three years, especially Emilio Rudbeck, Lisa Meiseleder, Jasmin Niederkofler, and Mia Heidenreich. Your enthusiasm and fresh perspectives have enriched my research experience and provided much-needed support and fun.

Lastly, I would like to express my heartfelt gratitude to my family for their unwavering love, support, and encouragement throughout this journey. Without your belief in me, this humble achievement would not have been possible.

I would like also to thank my dear friends (members of حارة النباشيين) Elhoussein Mourad, Mohamed Abdelfadil, Mahmoud Mabrouk, and Ammar Abdalrahem for their thoughtful discussions and feedback which have been invaluable to me, and I am grateful for their friendship.

I'm also grateful to our co-authors from the National History Museum of Vienna, Life Science Compute Cluster (LiSC) of University of Vienna, University of Tübingen, Center of Integrative Biology (CIBIO) of University of Trento, the Natural History Museum of Basel, University hospital of Basel, Institute for Systems Biology – Seattle, and University of Innsbruck, for their fruitful collaboration.