Metagenomic analysis of

Mycobacterium tuberculosis in ancient human remains

Dissertation

zur Erlangung des Doktorgrades der Naturwissenschaften (Dr. rer. nat.)
an der Fakultät für Biologie der Ludwig-Maximilians-Universität München

Vorgelegt von

Heidi Yoko Jäger

München, im März 2023

Erstgutachter:	PD Dr. Albert R. Zink
Zweitgutachter:	Prof. Dr. Wolfgang Enard
Tag der Abgabe:	08. März 2023
Tag der mündlichen Prüfung:	26. September 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 08. März 2023

(Heidi Yoko Jäger)

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 LMU Munich.

The research design of this work implemented ethical considerations to the best of the author's knowledge.

おばあちゃんへ

Table of Contents

1 | Summary

Tuberculosis is a globally spread disease found on all inhabited continents and is ranked as one of the top ten causes of deaths worldwide according to the World Health Organization. It is caused by *Mycobacterium tuberculosis* (MTB), a gram-positive opportunistic pathogen that has evolved with humans over many millennia. Today, the emergence of antibiotic-resistant strains and lack of funding for research is heavily impeding the control of this disease, especially in low-income countries. Pathogen genomes, which are reconstructed from ancient DNA (deoxyribonucleic acid) isolated from human remains, can provide important insight into host-pathogen relationships over time and space and could therefore help improve our understanding of the complex evolution of MTB.

The field of paleogenomics is still young but, with the help of advancing sequencing technologies, a substantial amount of genetic data has been generated over the past years, reshaping the evolutionary history of ancient hominins, modern humans, animals, and even causative agents of diseases. However, the success of such studies is highly dependent on the level of sample preservation, as ancient DNA typically displays characteristic damage patterns and high fragmentations with an average length of only around 50 base pairs. In addition, metagenomic sequencing data retrieved from human remains often have high amounts of background DNA resulting from the burial environment or other contaminations. Hence, usually, only minute amounts of endogenous ancient DNA are present in a sample. Finally, the analysis of ancient pathogen DNA in human remains poses an individual challenge, as its detectability is highly dependent on the pathogen's life cycle and disease progression at the time of host death.

Therefore, this thesis first investigates how the sampling process of human remains itself can influence the yield of pathogen DNA by applying an approach sampling multiple locations of a skeletal element for minimal amounts of bone powder to increase the probability of detecting 'hotspots' with highly concentrated pathogen DNA. This metagenomic study presents a first assessment of intra-bone variability of pathogen load and endogenous

human DNA content, as well as a comprehensive analysis of microbial composition of samples across different sampling sites. Finally, it highlights the challenges of obtaining sufficient amounts of ancient pathogen DNA for genome-level analysis while maintaining a balance for minimally destructive sampling.

The DNA extraction using linear polyacrylamide (LPA) for DNA precipitation is a highly efficient and innovative method to recover DNA from environmental and mummified tissue samples. These samples are likely to display high inhibition levels due to the co-extraction of unwanted substances that impede downstream molecular analyses. Comparison with the most commonly used silica-based extraction method shows comparable DNA quality and overall compositions of the LPA-based method, making it a powerful alternative tool for the DNA extraction of complex samples.

Finally, the improved molecular workflow was successfully applied to retrieve high-coverage ancient MTB genomic data from a rib sample of an 18^{th}-century midwife from Vác, Hungary. Here, changes on the skeletal surfaces indicated the presence of the two infectious agents: *Treponema pallidum* and MTB. While the molecular diagnosis of a syphilis infection remained negative, a mixed infection with two distinct strains of the MTB complex was detected. For this, a sophisticated computational pipeline was developed utilizing reference databases of modern and ancient MTB genomic data to ensure the confident detection of mixed strain infections and the phylogenetic analysis on single nucleotide polymorphism-level.

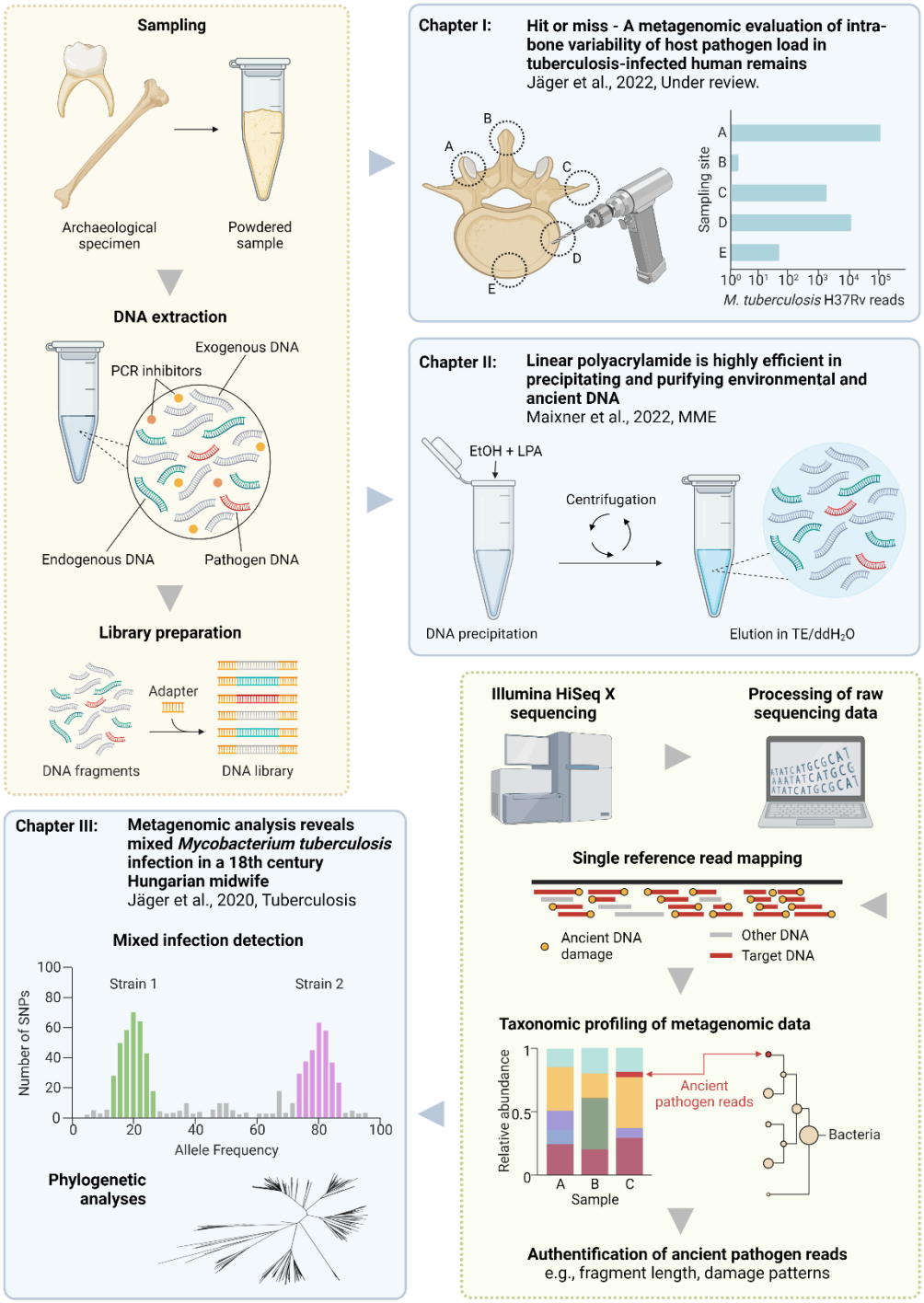

Figure 1: General workflow for ancient *Mycobacterium tuberculosis* DNA detection applied in this thesis. The yellow box indicates laboratory methods, the green box indicates computational methods. Each blue box highlights a chapter of this thesis in form of a graphical summary.

2 | Introduction

2.1 | Ancient pathogen genomics

The origin of human infectious diseases lies deep in the past. Over 12,000 years ago, during the Neolithic revolution, modern humans underwent major transitions adapting their subsistence lifestyle from nomadic hunter-gatherers to settled farmers (Skoglund et al., 2012). This led to higher population density, domestication, agriculture, and trade, which were the main factors leading to exposure to a new set of potential pathogens through zoonotic transmission (Eshed et al., 2004; McKeown, 1988). However, even though substantial historical and archaeological evidence supports epidemic and pandemic events throughout the history of humankind, to this day, the causative agents of these diseases often remain conjectural (Armelagos et al., 1991; Barrett et al., 1998). The emerging field of ancient pathogen genomics combines the interest of research fields such as evolutionary biology, microbiology, archaeology, and anthropology, and aims to illuminate how pathogens have co-evolved and interacted with their hosts over time. For that, ancient DNA (deoxyribose nucleic acid) isolated from ancient specimens is utilized as the principal source of information to decipher the pathogen's genetic history and epidemic patterns on the genome level.

The first step in the analysis of ancient human remains is the anthropological analysis for age and sex estimation and the detection of paleopathological changes, such as bone alternations indicative of chronic infections (Grauer, 2011). Human remains retrieved from burial sites such as leper cemeteries or plague pits can provide first robust indications for the presence of specific infectious agents e.g., *Mycobacterium leprae* or *Yersinia pestis* (Bos et al., 2011; Cole et al., 2022; Spyrou et al., 2016; Tran et al., 2011). Tuberculosis, leprosy, treponemal diseases such as syphilis, and certain fungal infections are infectious diseases that are known to cause characteristic bone lesions in affected individuals. Therefore, the initial paleopathological diagnosis is especially crucial, as the molecular diagnosis highly relies on sampling characteristic bone lesions, where the pathogen is presumed present in higher concentration (Roberts, 2011). However, not all chronic infections leave such

characteristic traces on the skeletal remains (Buikstra & Roberts, 2012). Host susceptibility, immune response, pathology, and the level of disease progression are crucial factors that influence the pathogen load in the host's blood and tissue at the time of death. Furthermore, the burial conditions of the individual greatly influence the abundance and detectability of ancient pathogen DNA in specific sampling locations (Duchêne et al., 2020).

From the 1990s, the accurate detection of pathogens highly relied on polymerase chain reaction (PCR), a molecular technique allowing to target and amplify specific DNA regions of interest and the product of which (i.e., amplicon) was characterized by electrophoresis or Sanger DNA sequencing (Rafi et al., 1994; Salo et al., 1994; Spigelman & Lemma, 1993; Zink et al., 2007). However, this method requires prior knowledge of the potential infectious agent and only targets single loci with amplicon fragment sizes larger than 70 base pairs (bp). Therefore, the analysis is restricted solely to the detection of pathogen DNA and provides limited evolutionary context (Arriaza et al., 1995; Drancourt et al., 1998). With the introduction of next-generation sequencing (NGS) technologies in the early 2000s, cost-effective and rapid high-throughput DNA sequencing was made possible, allowing to parallelly analyze millions of loci from a small fraction of ultra-short DNA fragments (Goodwin et al., 2016). Shotgun sequencing, also known as short-read sequencing, produces short reads between 25 to 500 bp. It is the dominating sequencing method in the field of ancient DNA due to its higher sensitivity and lower error rate in sequencing short DNA fragments compared to newer third-generation sequencing technologies such as Oxford Nanopore Technology or Pacific Biosciences that are specialized in long-read sequencing (Xiao & Zhou, 2020).

The drastic increase in high-resolution data obtained from modern organisms and ancient specimens has simultaneously led to a growing number of reference genomes of various taxa and even individual strains of bacteria and viruses. Thereby, pathogen DNA isolated from ancient specimens can serve to uncover the pathogen's evolutionary history with much higher accuracy and

paleogenetic studies are no longer limited to only identifying a pathogen's presence or absence (Margulies et al., 2005). The first fully reconstructed ancient pathogen genome – *Yersinia pestis* – was published in 2011 by Bos and colleagues and was followed by numerous genome-scale studies that have provided substantial data revising narratives on the global emergence and the long-time prevalence of human pathogens such as *Helicobacter pylori* (Maixner et al., 2016), the hepatitis B virus (Krause-Kyora, Susat, et al., 2018; Patterson Ross et al., 2018), *Treponema pallidum* (Giffin et al., 2020; Majander et al., 2020; Schuenemann, Kumar Lankapalli, et al., 2018), *Yersinia pestis* (Bos et al., 2011; Keller et al., 2019; Spyrou et al., 2016; Spyrou, Keller, et al., 2019; D. M. Wagner et al., 2014), *Plasmodium falciparum* (Marciniak et al., 2016), *Salmonella enterica* (Key et al., 2020; Vågene et al., 2018), the variola virus (Duggan et al., 2016; Mühlemann et al., 2020), *Mycobacterium leprae* (Krause-Kyora, Nutsua, et al., 2018; Schuenemann, Avanzi, et al., 2018), and *Mycobacterium tuberculosis* (Bos et al., 2014; Kay et al., 2015; Vågene et al., 2022).

Thus, ancient pathogen genomics does not only help confirm diagnosis from paleopathology but also helps detect infections in the absence of specific bone changes: historical questions can be answered by working backward into the past, co-infections and co-morbidities can be detected to understand the correlations to other diseases, and finally, past pathogens can be directly compared with current pathogens to make estimations of changes in the genomic makeup over time more accurate.

2.2 | What is ancient DNA?

In 1984, Higuchi and colleagues isolated low molecular weight DNA sequences from a quagga, an extinct member of the horse family which was endemic in South Africa until the late 19th-century (Higuchi et al., 1984). This study marks the first time that ancient DNA was extracted from the remains of a long-dead species and sets the first stone for a new field of research – the potential of which was still unexplored at that time. Now, almost four decades

later, ancient DNA has become an extremely potent tool that has revolutionized our understanding of the history of humankind and has revealed a much more dynamic past than previously anticipated (Gamba et al., 2014; Green et al., 2010; Meyer et al., 2012; Reich et al., 2010; Spyrou, Keller, et al., 2019).

Ancient DNA damage

The success of these studies depends in significant part on the new understanding of ancient DNA. Ancient DNA is characterized by high degradation due to naturally occurring *postmortem* processes. While favorable factors of the environment such as temperature, pH, water, and microbial composition can limit the degree of damage, once the organism's cellular repair mechanisms cease to function, the genome is exposed to intracellular nucleases, microbial intruders, and later to hydrolytic and oxidative processes that threaten the structural integrity of the DNA (**Figure 2A**) (Dabney, Meyer, et al., 2013; Lindahl, 1993). The depurination of the purines guanine and adenine through hydrolysis is the probable cause of fragmentation in ancient DNA. In this process, the N-glycosyl bone between a purine and its sugar is cleaved, resulting in an abasic site. Subsequently, the 3'-aldehydic and 5'-phosphate ends are fragmented through β-elimination, creating a nick of the DNA backbone, and finally slowly breaking down the DNA double helix structure. Moreover, cytosine is deaminated to uracil, which leads to miscoding lesions in ancient DNA as DNA polymerases incorporate adenine across from uracil, and thymine across from adenine, ultimately causing cytosine to thymine (C→T) and guanine to adenine (G→A) substitutions. This cytosine deamination process is faster at apurinic single-stranded sites than within double-stranded parts and accumulates with age. Furthermore, DNA degradation is correlated to the temperature of the environment in which it was conserved over time i.e., low latitudes and high temperatures result in poor DNA preservation, while DNA degradation is delayed under colder/frozen conditions (Hofreiter et al., 2015). Consequently, ancient DNA

is typically present in very low quantities with chemical changes and is highly fragmented with an average length of 30 to 60 bp (Orlando et al., 2021).

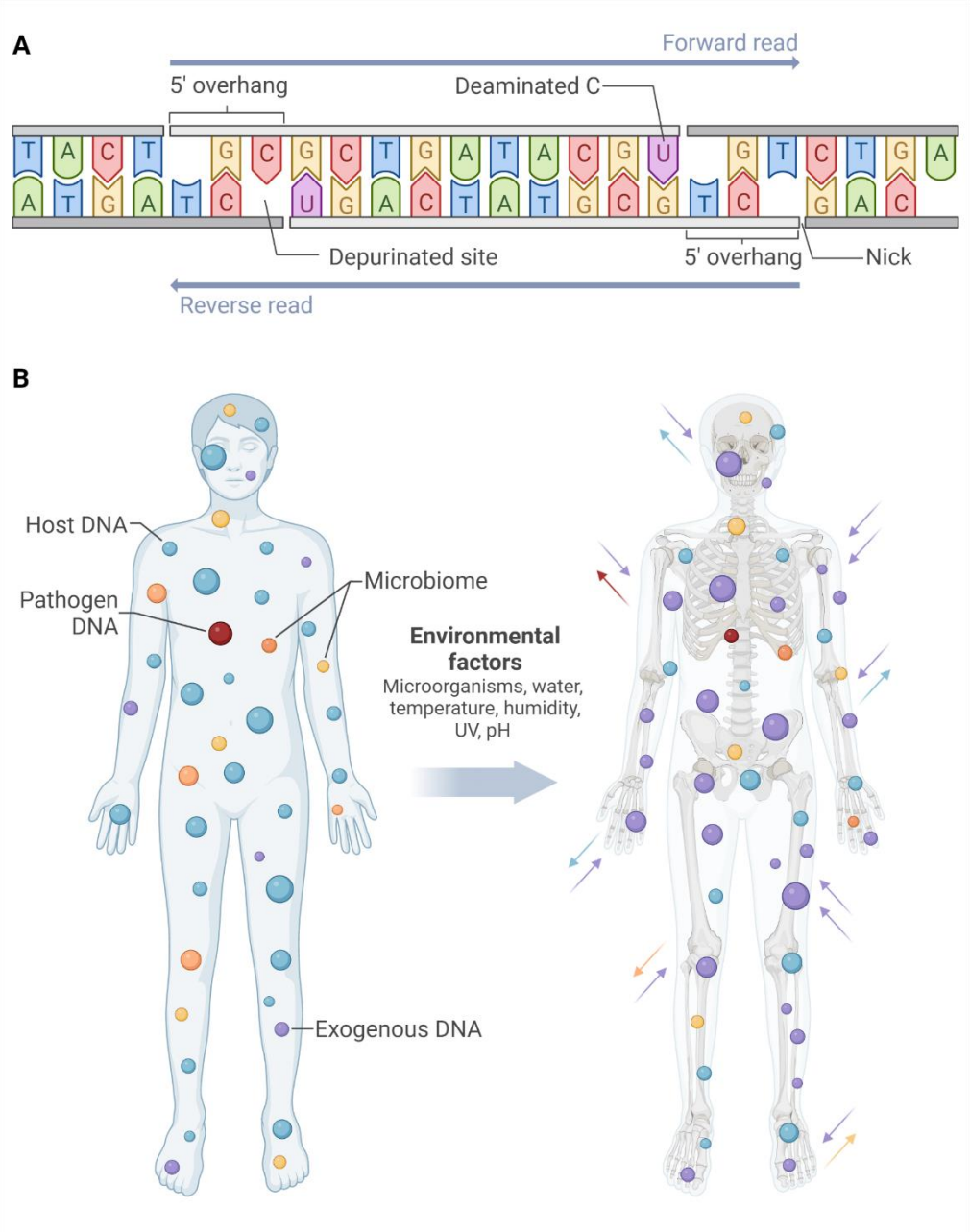

Figure 2: A. Post-mortem DNA damage. Depurination and deamination lead to nicks in double-stranded DNA that cause fragmentations. Adapted from Orlando et al., 2021. **B. Degradation and preservation of ancient DNA over time.** The degradation and loss of endogenous DNA (host and pathogen DNA, microbiome) start immediately at the time of death. Simultaneously, exogenous DNA from the burial environment gradually enters the space, leading to interference with the original signal (short-fragmented, low-quantity DNA versus high molecular weight DNA).

Sources and contamination of ancient DNA

The source of ancient DNA is not only limited to mummified tissue or skeletal elements from human and animal remains. Instead, there is a wide range of materials that have worked successfully, such as dental calculus (Granehäll et al., 2021), hair (Miller et al., 2008), plant remains (Crump et al., 2021; Fordyce et al., 2013), paleofeces (Maixner et al., 2021), sediments (Massilani et al., 2022; Sarhan et al., 2021), and even mollusk shells (der Sarkissian et al., 2017, 2020) or wood (Lendvay et al., 2018; S. Wagner et al., 2018). Depending on the environment in which the specimens were preserved and how they were handled during the excavation and sampling process, various levels of exogenous contamination can be observed. Therefore, archaeological specimens usually contain a mixture of endogenous DNA, i.e., human, host-microbiome, and pathogen DNA, and exogenous DNA i.e., modern human DNA from direct contact with skin, environmental DNA from bio-degrading bacteria or overgrowth of bacteria and fungi in suboptimal storage conditions (Warinner et al., 2017). Additionally, even laboratory reagents such as DNA polymerases or other enzymes can be sources of contamination as they contain traces of bacterial DNA. The entirety of genetic material recovered from an archaeological sample is referred to as the metagenome and allows the genetic analysis of not only one species but the whole diversity in a sample. On average, endogenous ancient DNA comprises 1-10% of the total metagenome. However, the study of ancient diseases is much more complex as ancient pathogen DNA makes up less than 0.5% of the metagenome.

Ancient DNA recovery

Traditionally, for molecular analyses, samples must first be mechanically processed to powder by drilling or using a mixer mill. The sampling process is followed by incubation in buffers with decalcifying, deproteinizing, and delipidating reagents to release the DNA in the sample and hinder it from interacting with any organic or inorganic substances. The extensive fragmentation makes ancient DNA especially challenging to separate from high-molecular-weight DNA and results in the co-extraction of exogenous

contaminations like modern environmental DNA or other small molecules that function as inhibitors in downstream enzymatic reactions (**FIGURE 2B**) (Green et al., 2009; Höss & Pääbo, 1993). Therefore, ancient DNA analyses must be carried out in a dedicated clean laboratory facility following guidelines that ensure a sterile working environment through rigorous decontamination measures (Cooper & Poinar, 2000; Fischer et al., 2016; Fulton & Shapiro, 2019; Knapp et al., 2012).

To address different levels of DNA inhibition and preservation, custom protocols for efficient recovery have been developed (Dabney, Knapp, et al., 2013). For example, including a pre-digestion step or pre-treatment with bleach can improve the removal of easily accessible contaminant DNA and chemical inhibitors, but comes with a trade-off as it can lead to a partial loss of endogenous ancient DNA (Damgaard et al., 2015; Gamba et al., 2016; Korlević et al., 2015). Similarly, the most commonly used DNA purification and precipitation protocols will, to some degree, introduce a bias into the final composition of the extracted DNA. In most methods, DNA is isolated by adsorption to silica particles in chaotropic binding buffer. This reaction can be carried out either in solution (Rohland & Hofreiter, 2007), on spin columns (Dabney, Knapp, et al., 2013), or on suspended silica-coated magnetic beads (Rohland et al., 2018) and is followed by the elution of DNA in low-salt buffer or water. Some protocols additionally suggest removing ancient DNA damage using the USER reagent (New England Biolabs), an enzymatic mix that creates a gap at the location of uracil residues (Rohland et al., 2015). It is to be noted that the usage of this treatment is study-dependent as despite reducing sequencing errors, it also shortens the length of DNA molecules and limits the authentication of ancient DNA.

For NGS, DNA libraries are prepared by adding platform-specific adapters with integrated unique identifying sequences at the ends of the recovered DNA fragments. Thereby, contamination risks are greatly reduced, and multiple samples can be pooled together for sequencing runs. Due to the poor performance of commercially available library preparation kits, new methods

specifically for degraded and damaged DNA have been developed. The most used method is the Meyer and Kircher approach which includes several repair steps to remove or fill in single-stranded overhangs. Thereby, blunt ends are created that facilitate the adapter ligation (Meyer & Kircher, 2010). However, clipping already short DNA fragments, the use of unidirectional adapters, and repeated purification steps before amplification lead to the loss of about half of the unique molecules in the sample (Carøe et al., 2018). Therefore, to increase library complexity, Carøe and colleagues suggest a more advanced single-tube protocol (BEST) that does not require any purification steps and instead uses heat denaturation. Alternatively, DNA molecules can be converted into single-stranded DNA (ssDNA) libraries instead of double-stranded DNA (dsDNA) libraries, which allows to leave the ends of DNA fragments unaltered and hence further increases the DNA recovery yield (Gansauge et al., 2017; Gansauge & Meyer, 2013). Finally, the recently published Santa Cruz Reaction method for ssDNA libraries has great potential to become a leading method in this field, as it is less labor-intensive and more cost-effective in comparison to previous dsDNA and ssDNA library methods (Kapp et al., 2021).

The use of shotgun sequencing to reach sufficient genome coverage of the species of interest is often cost-intensive due to the large amount of background DNA or simply unfeasible due to the limited availability of ancient material. To circumvent this problem, ancient DNA libraries can be enriched for specific DNA sequences before sequencing. These enrichment methods are based on hybridization capture using either DNA or RNA baits that are designed to target specific genomic loci of interest and are either performed on microarrays or in-solution (Fu et al., 2013; Spyrou, Bos, et al., 2019). To put it simply, targeted DNA first hybridizes with the added bait molecules and is subsequently pulled down to separate it from non-target DNA in the library. Thereby, the proportion of DNA fragments with initially very low copy numbers can be increased relative to the amount of background DNA. By applying these enrichment techniques ancient samples that were

previously unsuitable for analysis have been made accessible (Bos et al., 2016; Burbano et al., 2010; Enk et al., 2014).

Computational analysis of ancient DNA

Continuing advances in sequencing technologies have made it economically and technically possible to obtain substantial proportions or even complete nuclear and mitochondrial genome sequences from ancient specimens. When deciding on strategies for the analysis of metagenomic data, the research question is the most relevant factor. Especially when it comes to genomic data retrieved from ancient specimens, special attention is required for downstream computational analysis. For this reason, specialized computational tools have been developed for targeted and nontargeted approaches.

After the production of ancient genomic data, the computational analysis of large-scale datasets remains a crucial step for the successful reconstruction of ancient genomes. Furthermore, the analysis of ancient DNA sequencing data requires rigorous authentication to correctly identify true ancient DNA sequencing reads. While characteristics such as C→T deamination profiles or short fragment length can be useful to determine the authenticity of a sequencing read, they also pose great challenges for genome reconstruction due to unspecific DNA alignment, low coverage, and miscoding lesions that can result in low-confidence genotyping. Thus, in-silico methods specifically tailored for ancient DNA sequencing data are required to ensure ancient DNA authenticity and to maximize the analytical yield of finite archaeological resources.

In case of Illumina shotgun sequencing, raw sequencing data in FASTQ format is first preprocessed by trimming adapter sequences and performing an overall quality assessment, whereby reads that do not meet specific quality requirements are being removed from the dataset. Filtered reads are then mapped against a reference genome of the organism of interest using state-of-the-art short read alignment tools such as Bowtie2 or BWA (Langmead & Salzberg, 2012; Li & Durbin, 2009), followed by various postprocessing steps

to index mapped reads, remove reads with low mapping scores, and remove duplicates. As aforementioned, characteristics such as damage patterns and high fragmentations can be detected using DamageProfiler or mapDamage and are helpful to verify whether sequencing reads are of the truly ancient origin or not (Jónsson et al., 2013; Neukamm et al., 2021). Additionally, the evenness of coverage across the reference genome is also an important factor for ancient DNA authentication and tools like schmutzi or ANGSD allow mitochondrial and human nuclear contamination estimations (Korneliussen et al., 2014; Renaud et al., 2015). Genotyping is performed through variant calling and consensus sequence generation which allows subsequent phylogeny reconstruction commonly using modern reference genomic data. Finally, metagenomic screening for pathogens or other microorganisms has become a routine bioinformatic practice as it is crucial to detect 'invisible' infectious agents without evidence in archaeological records (Warinner et al., 2017; Weyrich et al., 2017).

2.3 | The *Mycobacterium tuberculosis* complex

Mycobacterium tuberculosis (MTB) is the causative agent of tuberculosis (TB), and to date, 23 ancient genomes of this pathogen have been published (Bos et al., 2014; Chan et al., 2013; Jäger et al., 2022; Kay et al., 2015; Sabin et al., 2020; Vågene et al., 2022). The analyses of these genomes revealed valuable insights into past transmission patterns, the time of emergence, and the pathogen's co-evolution with humans over the past millennia.

The global burden of tuberculosis

In 2020, TB was responsible for 1.5 million deaths, making it the second most infectious disease worldwide after COVID-19 (WHO, 2021). It is estimated that 25% of the global population is infected with TB. Among this population, the probability of developing active TB disease is significantly higher for people affected by risk factors such as human immunodeficiency virus (HIV) infections, undernutrition, diabetes, smoking, and alcohol use. However, the leading cause of high TB burden is reduced access to TB diagnosis and

treatment due to poor healthcare system conditions. This is especially true for low- and middle-income countries in South-East Asian (India, Indonesia, Pakistan, Bangladesh), Western Pacific (China, the Philippines), and African (Nigeria, South Africa) regions, which account for over two-thirds of the total global cases. A global TB prevention strategy to combat this severe public health issue was put in place by the World Health Organization in 2014. However, the monitoring and control of this disease are hampered by the emergence of virulent drug-resistant strains and the lack of sustainable funding for the research and development of new vaccines, especially in the most affected countries. Additionally, the progress of ending TB was significantly disrupted by the outbreak of the global COVID-19 pandemic in 2020 leading to the rise of TB deaths, despite declining numbers being reported in the previous year (WHO, 2021). For these reasons, urgent action is needed not only to restore essential TB services and to contain the damage but also to develop better prevention strategies by further deepening our knowledge of this pathogen's origin and history, which has yet to be fully revealed.

The causative agent of tuberculosis

MTB was first described by Robert Koch in 1882 and is a highly specialized, predominantly intracellular pathogen that can actively modulate components of the host's immune response (Koch, 1982; von Both et al., 2018). It is currently hypothesized that MTB emerged as a professional pathogen from an environmental *Mycobacterium* that acquired the ability to persist in free-living protozoa and then gradually adapted to a lifestyle in an intracellular environment (Jang et al., 2008). Another crucial evolutionary step was the development of its unique virulent characteristics and transmission strategy, leading to a highly complex host-pathogen relationship on a cellular level (VanderVen et al., 2016). The transmission of MTB occurs through respiratory droplets produced by coughing or sneezing of the host. While larger droplets fall to the ground almost immediately, smaller droplets evaporate into droplet nuclei and can be airborne for up to six hours. 90% of

new MTB hosts develop a latent infection in which viable MTB survive within infected cells and granuloma but do not grow and instead enter a dormant state (Gengenbacher & Kaufmann, 2012). Hosts with latent infections show no symptoms but can develop active disease if they become more susceptible during their lifetime. In the remaining 10% of the cases, the infected hosts develop active TB disease, where MTB survives and replicates within the granuloma despite evidence of vigorous response of the host immune system, which typically leads to the progressive destruction of lung tissue (pulmonary TB) as well as other areas of the body (Flynn et al., 2011). The life cycle of MTB is fulfilled when it erodes through the lung tissue into the airway and can be transmitted again to another host (Ernst, 2012).

The evolution of the *Mycobacterium tuberculosis* complex

The *Mycobacterium tuberculosis* complex (MTC) comprises the ancestral smooth tubercle bacilli *M. canettii*, nine obligate human pathogenic strains (lineages 1-9), and several animal strains that cause disease in a broad range of domestic and wild mammals such as cattle (*M. bovis*), voles (*M. microti*) and seals (*M. pinnipedii*) (**Figure 4**) (Thomas et al., 2021). In most bacteria, diversification and evolution are mainly driven by horizontal gene transfer. However, due to the highly clonal population structure of the MTC – except for *M. canettii* – with over 99.9% nucleotide sequence identity, this movement of genetic material most likely only played a secondary role in its evolution (Gagneux, 2018). Instead, the driving force for differentiation here is based on large sequence polymorphisms (LSPs) known as regions of difference (RDs) and single nucleotide polymorphisms (SNPs) (Bespiatykh et al., 2021). RDs result from unidirectional unique events such as conservative deletions inherited by all strain descendants and are therefore phylogenetically informative. On the other hand, PE/PPE genes, prophage regions, and regions flanked by insertion sequences are categorized as phylogenetically noninformative RDs. With the development of whole genome sequencing (WGS) technologies, SNPs have also gained great importance as a robust

marker of genetic variation for the phylogenetic analysis of MTC (Achtman, 2008).

Both RDs and SNPs can directly affect the fitness, virulence, and proneness to acquire drug resistance of a strain. In addition to these factors, human lineages 1-9 (L1-9) greatly differ in their geographical distribution (**FIGURE 4**): L1-4 and L7 are viewed as the main MTC lineages and are commonly referred to as *M. tuberculosis sensu stricto*. The East African Indian (EAI) L1 and L7 (Aethiops vetus), which are geographically restricted to the Horn of Africa, derive from the ancestral strain of L2-4 (Nebenzahl-Guimaraes et al., 2016). On the contrary, Beijing L2, Central Asian Strains (CAS) L3, and Euro-American L4 are considered modern strains because of their relatively recent deletion of the MTC-specific deletion 1 region TbD1 (Brosch et al., 2002). Moreover, L2 includes the multidrug-resistant Beijing strains and is especially dominant in East Asia. Together with L4, which is present in all inhabited continents, they are globally the most successful MTC strains. *M. africanum* has been previously referred to as a single entity, however new phylogenomic insights have revealed three separate lineages: L5 and L6 which are restricted to West Africa, and L9 which predominantly occurs in East Africa. (Stucki et al., 2016). It was previously thought that the human-adapted *M. tuberculosis sensu stricto* lineages evolved from *M. bovis*, the causative agent of bovine TB which has a much broader host range. However, *M. africanum* and animal lineages of the MTC are characterized by RDs that are still present in *M. tuberculosis sensu stricto* lineages, which point towards a common predecessor in the past before separating (Brosch et al., 2002). Finally, the recently discovered L8 has only been reported in the African Great Lakes region and represents new evidence supporting an East African origin of the MTC as the ancestral *M. canettii* clade and L8 have both preceded the loss of the cobF genome region involved in the cobalamin/vitamin B12 synthesis which is lacking in all other MTC lineages (Ngabonziza et al., 2020).

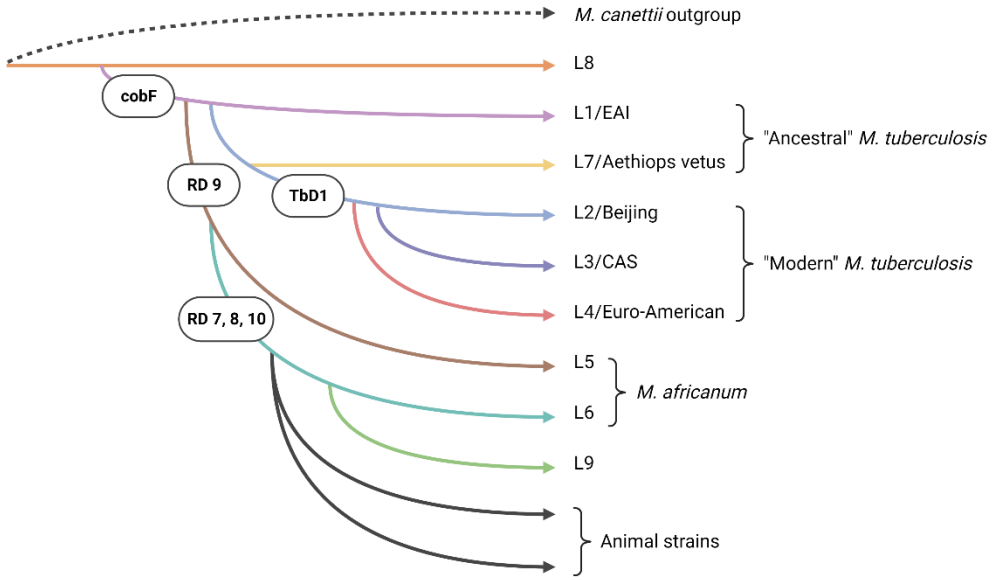

FIGURE 3: EVOLUTIONARY PATHWAY OF THE *MYCOBACTERIUM TUBERCULOSIS* COMPLEX. Boxes highlight key LSPs including RDs and deletion of other genome regions with phylogenetic importance. Based on Brosch et al., 2002.

The origin and date of emergence of TB in humans are being slowly deciphered through the increasing number of well-dated ancient and modern TB genomes from humans as well as animals. The African origin of the MTC is strongly supported by molecular evidence: Foremost, the smooth tubercle bacilli, which is the most closely related living bacteria of the common ancestor of the MTC, is nearly solely found in the Horn of Africa (Gutierrez et al., 2005; Supply et al., 2013). Furthermore, all human-adapted MTC lineages are found on the African continent with a high diversity which decreases with increasing distance from Africa (Comas et al., 2015). This is additionally supported by phylogeographical analyses using whole-genome sequences of human and animal-adapted MTC members (Comas et al., 2013). When it comes to the date of emergence, the current literature suggests two widely different estimates for the most recent common ancestor of the MTC which result from different dating approaches of MTC phylogenies. The study by Comas and colleagues suggests that the MTC emerged through a bottleneck event about 70,000 years ago and co-migrated out of Africa with modern humans. This estimation was calculated using a global dataset of 220 modern MTC strains and universal bacterial substitution rates, excluding

ancient MTC genomes (Comas et al., 2013). On the other hand, other studies have incorporated radiocarbon-dated ancient MTB genomes retrieved from 18th-century Hungarian and Swedish mummies and pre- and peri-contact South American human remains from Peru and Colombia (approx. 500 to 1,000 years old) as calibration points for mutation rate inference and have obtained a much more recent estimation of about 6,000 years before present (Bos et al., 2014; Kay et al., 2015; Sabin et al., 2020; Vågene et al., 2022). All MTB genomes retrieved from South America belong to the *M. pinnipedii* clade, which causes disease in seals, suggesting a zoonotic pinniped to human transmission in the pre-contact era. However, the most recently recovered human-derived ancient MTB strains by Vågene and colleagues are placed at the most basal positions of the *M. pinnipedii* clade, which implies a reverse transmission scenario from human to pinniped (Vågene et al., 2022). Even though all studies relying on ancient DNA calibration have so far agreed on a Neolithic emergence of the MTC, the dating remains controversial as it cannot account for the detection of MTC DNA and TB-specific lipid biomarkers in archaeological material that predates this estimation e.g., 11,000 and 9,000-year-old human remains from Syria and Israel as well as 17,000 years-old bison bones from Wyoming (Baker et al., 2015; Hershkovitz et al., 2008; Lee et al., 2012; Masson et al., 2013; Rothschild et al., 2001). The authenticity of these studies is still debated as they have not been verified by the current state-of-the-art sequencing methods. However, it is also still unclear whether mutation rates obtained from modern clinical samples and relatively recent ancient samples can be deduced to substitution rates over longer periods of time. This is largely because questions on the consistency of mutation rates are yet to be answered as MTC populations do not always display consistent sequence diversity accumulation patterns and the effect of latency periods remains unclear (Comas & Gagneux, 2011; Duchêne et al., 2016). Therefore, to be able to further understand the evolutionary history of the MTC, a robust long-term substitution rate for the MTC needs to be established through the analyses of more ancient MTC genomes.

2.4 | Aims of the thesis

Archaeological records of tuberculosis serve as timestamps to provide a deeper insight into the pathogen's evolution and past epidemic trends. However, the field of ancient pathogen genomics is still young and requires more high-resolution genomic data to answer such fundamental questions. This, in turn, is highly dependent on the improvement of laboratory techniques to overcome the difficult nature of ancient DNA and the development of reliable computational data processing and analysis tools. Therefore, this thesis aims to develop both molecular and computational methodologies to improve the recovery of MTB DNA from human remains and to ultimately perform genome-level metagenomic analyses of ancient pathogens.

For this purpose, in **Chapter I**, I first investigate the intra-bone variability of ancient MTB DNA by applying a comparative metagenomic approach. The study serves as a first assessment to better understand whether the sampling success of pathogen and endogenous human DNA is correlated with the sampling location or microbial composition within or across samples. Moreover, it discusses how future sampling strategies can be improved to maximize the amount of information generated from finite material while also highlighting the challenges of systematic detection and metagenomic analysis of ancient MTB DNA.

In **Chapter II**, a highly purifying DNA precipitation method with linear polyacrylamide is developed for environmental and ancient DNA that efficiently reduces inhibitory effects and recovers very short DNA fragments. A wide variety of sample types including mummified soft tissues, bones, ancient gut content, soil, activated sludge, and animal feces are tested and compared with most widely used silica-based DNA extraction methods to confirm whether this innovative method can provide as a suitable alternative for complex samples.

Finally, **Chapter III** aims to apply the optimized molecular workflow on a rib bone sample of an 18th-century Hungarian midwife. The partially

mummified remains belong to the Vác Mummy Collection which comprises 265 well-documented individuals in different states of natural mummification of which a high proportion display TB infection. The study presents high-coverage ancient MTB genomics data and sheds light on the high prevalence of mixed MTB infections in the pre-industrial and thence pre-antibiotic era with respect to sociohistorical factors.

3 | Results

Chapter I Hit or miss - A metagenomic evaluation of intra-bone variability of host pathogen load in tuberculosis-infected human remains. *Tuberculosis*, 2023

This article is only available at: https://doi.org/10.1016/j.tube.2023.102392

Chapter II Linear polyacrylamide is highly efficient in precipitating and purifying environmental and ancient DNA. *Methods in Ecology and Evolution*, 2022

This article is only available at: https://doi.org/10.1111/2041-210X.13772

Chapter III Metagenomic analysis reveals mixed Mycobacterium tuberculosis infection in a 18th century Hungarian midwife. *Tuberculosis*, 2022

This article is publicly available online at: https://doi.org/10.1016/j.tube.2022.102181

4 | Discussion

Ethical considerations

The cultural, historical, and sometimes even political implications of ancient DNA research have often been insufficiently addressed or wholly neglected in previous paleogenetic studies. This has sparked heavy criticism from outside and within the community of archaeologists, anthropologists, curators, and paleogeneticists but also directly affected Indigenous groups or descendants of the deceased individual in question (Cortezet al., 2021; Katrina G. Claw et al., 2017; Kennett et al., 2017). Based on these events, there has been an engaging discussion on the conduct of ethical ancient DNA research with the aim of setting global guidelines for future study designs. For example, according to Duchêne and colleagues, 236 pathogen genomes from 12,733 individual human remains have been reconstructed by the time of their publication in October 2020, which results in a success rate of only 0.01% (Duchêne et al., 2020). This raises the question of whether the inherent value of reconstructed pathogen genomes outweighs the cost of destructive sampling and furthermore by which parameters this can be measured. As for now, there are no clear answers to these questions and the decision must be made case by case until a collective consensus that meets the needs of all involved parties is established (Alpaslan-Roodenberg et al., 2021; Bardill et al., 2018; J. K. Wagner et al., 2020).

The research conducted as part of this work was designed to the best of the authors' knowledge and to the extent possible, to comply with the ethical guidelines outlined in the studies referenced above.

Challenges of detecting ancient MTB DNA

One aspect on which the ancient DNA research community seems to agree is the prioritization of minimally invasive sampling methods to reduce the irreversible damage to archaeological sources. Hence, for those conducting research with such precious materials it is essential to take this into account while simultaneously maximizing the chances of successful data recovery. Even though new laboratory methods have increased in sensitivity and are continuously decreasing the amount of material needed for ancient DNA

analysis, the irreversible and destructive sampling process must be often repeated multiple times before sufficient genomic data is retrieved (Wilbur et al., 2009). When the research question is focused on the recovery of ancient human DNA, many effective sampling techniques have been established to yield high amounts of endogenous human DNA with little background DNA. For example, the petrous part of the temporal bone has been repeatedly proven to be a highly valuable source of well-conserved endogenous DNA, with ear ossicles and tooth cementum providing suitable alternatives (Damgaard et al., 2015; Hansen et al., 2017; Harney et al., 2021; Pinhasi et al., 2015; Sirak et al., 2020). However, the pathogen load in the body differs depending on the pathogen's life cycle and the individual's infection stage at the time of death, which is challenging to determine post-mortem. Consequently, different sample types such as bone, teeth, and tissue are often needed to successfully detect molecular traces of pathogens in sufficient quantities.

A look at the list of published ancient genomic data reveals a clear bias for specific sample types for molecular analysis of MTB: since TB is primarily a pulmonary disease, mummified soft tissue from the abdominal area is the preferred source (**Table 1**). In case of mummified human bodies, an initial assessment through computed tomography (CT) scans can be valuable to identify calcified nodules in the areas of the lungs that often result from dystrophic calcification after active TB disease (Brown et al., 1994; Sabin et al., 2020). Alternatively, when only skeletal elements are present, vertebrae and ribs have proven to be the most promising candidates due to their anatomical proximity to the lungs (Bos et al., 2014; Bouwman et al., 2012; Jäger et al., 2022; Kay et al., 2015; Vågene et al., 2022). Furthermore, TB is the only disease that displays specific bone changes caused by skeletal TB, a form of extrapulmonary TB that includes collapsed spines (Pott's disease), fused vertebrae with signs of abscesses, and periostitis on ribs. Spinal lesions and periostitis are associated with TB, but they are not specific for it as other diseases such as brucellosis or mycotic infections can manifest similarly (Holloway et al., 2011).

The diagnosis of TB based on paleopathology is limited as skeletal TB occurs in only 2% of active TB cases, and the infected individual does not always live long enough to develop skeletal changes. Because of that, many TB infections in human remains, whether active or latent, remain undetected if solely visually examined. Hence, effective molecular screening methods for pathogen DNA that simultaneously prioritize the preservation of archaeological material are necessary. To facilitate a balance between the maximization of data acquisition and material conservation, it is crucial to increase the perspectives on the molecular preservation of pathogen DNA.

Table 1: Overview of as of date published ancient MTBC genomic data sets.

	Method	Sample type	Age of samples	Number of MTC genome sequence datasets	MTC lineage	Genome coverage
Bouwman et al. (2012)	Targeted enrichment	Rib bone	1,800 CE	1	N/A	N/A
Chan et al. (2013)	Shotgun sequencing	Chest tissue	1,800 CE	1	Lineage 4	32-folds
Bos et al. (2014)	Targeted enrichment	Vertebral body	1,023 - 1,280 CE	3	*M. pinnipedii*	22.7 to 31.4 - folds
Kay et al. (2015)	Targeted enrichment	Rib bone, lung, thorax, abdomen tissue	1,800 CE	14	L4.1.2.1/Haarlem L4.3/LAM L4.10/PGG3	0.4 to 332-folds
Sabin et al. (2020)	Targeted enrichment	Calcified lung nodule	1,800 CE	1	L4.10/PGG3	141-folds
Jäger et al. (2022)	Shotgun sequencing	Rib bone	1,800 CE	1	L4.1.2.1/Haarlem L4.10/PGG3	39.5-folds
Vågene et al. (2022)	Targeted enrichment	Rib bone, vertebral body	1,000 - 1640 CE	3	*M. pinnipedii*	10 to 15-folds

One aspect that has thus far been little understood is the variability of pathogen load within an individual skeletal element of ancient specimens: it has been frequently observed that repeated sampling of bones results in DNA extracts with different pathogen DNA concentrations. Understanding how intra-bone variability can affect the success of recovering ancient pathogen DNA is crucial as often only a limited amount of material is available for molecular analysis.

Our study in **Chapter I** reports an initial assessment of how MTB and endogenous human DNA content varies across different sampling sites within the same skeletal element. The aim was to create full taxonomic profiles of each sampling site and compare MTB and endogenous human DNA contents within and across samples by analyzing normalized shallow shotgun sequencing data (6 million reads). The main criterion for sample selection was that samples were reported as positive for TB-specific DNA markers in peer-reviewed publications. The samples originate from two archaeological sites in Central Europe with contrasting burial histories: two rib bones of two female 18th-century church mummies from Vác, Hungary, and a fused vertebra and a partial ulna of a female individual from a Neolithic site in Halberstadt, Germany (Fletcher, Donoghue, Taylor, et al., 2003; Nicklisch et al., 2012). In

case of the Neolithic German samples, the identical samples as in the original publication by Nicklisch and colleagues were used in our study (Jäger et al., 2023).

The taxonomic assignment of metagenomic shotgun sequencing reads revealed low variation of the microbial composition within samples and between samples of the same archaeological site. The detected phyla of each sample were in line with those typically found in the respective burial environment. For example, Actinobacteria and Proteobacteria belong to the most commonly found phyla in soil samples, and members of Firmicutes are associated with post-mortem bone degradation and have been repeatedly reported in mummified tissues (Janssen, 2006; Javan et al., 2016). This is particularly notable because the level of present-day DNA contamination and ancient DNA preservation in archaeological samples is frequently associated with the level of exposure to harmful environmental influences. However, despite same preservation conditions and highly similar composition of background DNA, discrepancies in MTB and endogenous human DNA content of up to 36- and 63-folds respectively even within distances of as small as 1 cm were detected (Jäger et al., 2023). Since this study only includes a small sample size, observations cannot be generalized. Nevertheless, there are solid indications that ancient pathogen and human DNA is better preserved in certain locations compared to other parts of the bone. Therefore, the detection of 'hotspots' could be by chance, unless future studies with larger sample sizes can provide correlations between anatomical positions and preservation of pathogen and endogenous human DNA.

We highlight how more comprehensive sampling as shown in this study can provide better insight into sample preservation during initial screenings. At first, conducting destructive sampling at multiple locations seemingly contradicts minimally invasive sampling as suggested by the ethical guidelines for ancient DNA research. However, due to high intra-bone variability, the metagenomic detection of the pathogen of interest is not insured if only single sites are sampled for DNA analysis. Initial shallow shotgun sequencing for

screening is common procedure in ancient DNA research and subsequent metagenomic results are crucial to decide whether to proceed with deep shotgun sequencing, a targeted enrichment approach or if the sample should be discarded due to low endogenous DNA content. If discarded, the same material is either resampled or more material (another tissue or bone type) of the same specimen is obtained. Both are labor and time intensive as they require repeated sequencing efforts while the latter, in addition, leads to even greater damage of the specimen (resampling of same material vs sampling of further material).

Furthermore, suggested input volumes for DNA extractions have decreased significantly in more recent methods from up to 500 mg to as low as 50 mg (Dabney & Meyer, 2019; Maixner et al., 2022; Rohland et al., 2018; Rohland & Hofreiter, 2007). Therefore, we believe that initial multiple-sampling for low volumes of bone powder offsets the involved cost of material destruction by increasing the chance to detect positions with high pathogen load and simultaneously decreasing time and labor expenditures.

Discrepancy between PCR-based and metagenomic approaches

Another conspicuous observation to emerge from our comparative metagenomic analysis is that not all samples and sampling sites resulted in a positive TB diagnosis despite evidence for it in previously published results. Especially the Neolithic samples from Germany did not yield any reads for MTB-specific gene markers (i.e., IS6110 and IS8110) in reference-based read mapping as well as MetaPhlAn3 analysis, a computational tool for profiling microbial composition in metagenomic data (Beghini et al., 2021). A supplementary PCR analysis targeting the MTB-specific IS6110 insertion site was performed with identical DNA extracts used for library preparation and sequencing to compare performance of both methods (Jäger et al., 2023). Even though PCR products were not sequenced, nested PCR was performed using internal primers of the IS6110 region to support the specificity of this approach. Only the nested PCR approach resulted in positive results for the Neolithic samples, though these results need to be taken with caution: the

reliability of PCR-based detection of ancient TB has been repeatedly challenged as the insertion sequence-based PCR assay lacks specificity and can lead to false positive results through contamination with environmentally prevalent non-tuberculous Mycobacteria or cross-contamination during laboratory procedures (Campana et al., 2014; Dziadek et al., 2001; Wilbur et al., 2009). Non-tuberculous Mycobacteria (NTM) are environmental bacteria commonly found in water and soil; some NTM are opportunistic pathogens that are genetically similar to the MTC and cause infections in different hosts including humans i.e., *Mycobacterium abscessus* and *Mycobacterium avium* (To et al., 2020). Considering that all Neolithic samples were in fact buried in the soil, contamination could be an explanation for possible false positive results.

A more likely explanation, however, is that low frequency DNA molecules (i.e., endogenous ancient DNA) are less likely to be carried onto the next laboratory step during library preparation compared to high frequency molecules (i.e., environmental microbial DNA). In addition, DNA amplification is a very sensitive process subjected to many variations. For example, increasing number of PCR cycles results in the loss of sequence heterogeneity and can ultimately cause the loss of rare molecules and therefore affect sequencing quality (Adey et al., 2010; Meyer et al., 2008).

Reducing inhibitory effects in complex DNA samples

There has been an ongoing effort to reduce the loss of ancient DNA molecules in downstream processes by developing more efficient protocols that often rely on the modification of already existing methods from related fields such as microbiology, molecular plant sciences, or molecular ecology (Maixner et al., 2022). Commonly used nucleic acid extraction protocols in the field of ancient DNA include reagents to minimize the inhibitory effect of certain substances, such as BSA or dilution. However, both methods can introduce a bias in the composition and complexity of the nucleic acid extract (Hedman & Rådström, 2013; Sidstedt et al., 2020). Furthermore, the addition of chaotropic salts

result in the co-extraction of inhibitory substances which is disadvantageous in the case of complex environmental and archaeological material.

In **Chapter II**, a DNA extraction protocol using salt-free ethanol and linear polyacrylamide (LPA) for DNA precipitation is presented. This method has proven to be highly effective in purifying inhibitory substances from complex environmental and archaeological samples such as activated slug, fen soil, and mummified tissues (Maixner et al., 2022). After lysis and organic purification, DNA precipitation is performed with ice-cold ethanol as it breaks the hydration shell that is formed around DNA through hydrogen bonds of water molecules. Additionally, LPA functions as a neutral carrier which is known to recover even picograms of DNA as short as 20 bps and shows superior precipitation and purification capacities compared to classical DNA precipitation and binding methods using ethanol/acetate, isopropanol, or polyethylene glycol (PEG) (Gaillard & Strauss, 1990; Maixner et al., 2022). The resulting DNA libraries yield similar complexity to those prepared with the commonly used silica-based DNA extraction method developed by Rohland and colleagues if the same lysis buffers are used but display higher performance in recovering ultra-short DNA fragments (Maixner et al., 2022; Rohland et al., 2018). Here, we demonstrate how complex samples that display high inhibition can be tackled by adapting knowledge from other research fields and by developing powerful alternative methods.

Improved extraction method for ancient MTB DNA

In **Chapter III**, a modified version of the DNA extraction method by Maixner and colleagues was applied for the DNA extraction of a bone sample of an 18th-century Hungarian midwife with a suspected TB and syphilis infection (**Figure 4**) (Jäger et al., 2022; Maixner et al., 2022). The initial lysis step was performed with standard 0.5M EDTA buffer and Proteinase K for >48 hours at 56°C instead of overnight at 37°C to ensure demineralization (**Figure 4**, step 2). 2 mL Qiagen PowerBead Pro tubes with 0.7 mm garnets were used for additional mechanical grinding through rotation of the tubes during lysis. Once lysed, samples were subjected to three rounds of deep freezing in liquid nitrogen and subsequent mechanical grinding by vortexing until fully thawed (**Figure 4**, step 3). This step is common practice in MTB DNA isolation as Mycobacteria are known for their extremely thick and lipid-rich cell wall comprised of peptidoglycan, arabinofuran, and mycolic acids (Brennan, 2003).

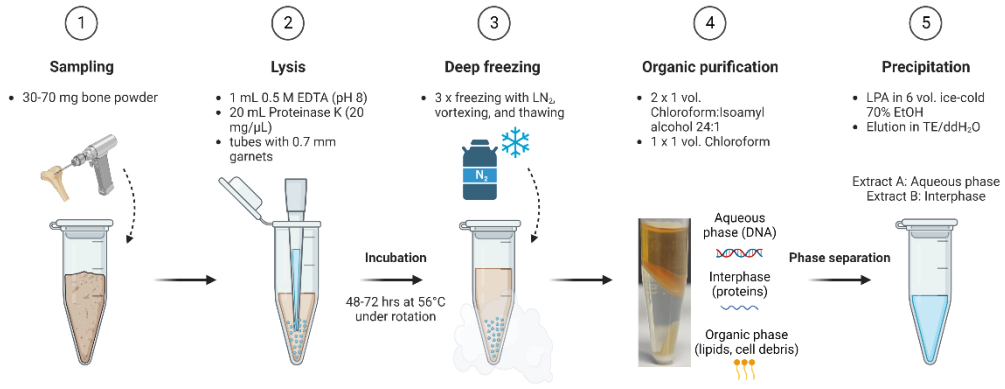

Figure 4: Modified workflow by Maixner et al. as applied in Jäger et al. 2022 for the DNA extraction of a rib bone sample of an 18th-century individual with a suspected TB infection. DNA precipitation of the interphase resulted in a ten times higher coverage.

In fact, mycolic acid profiles can be used as MTB specific lipid biomarkers and serve as a powerful diagnostic tool for ancient TB and leprosy cases in combination with ancient DNA (H. Donoghue et al., 2017). The oldest cases date back to 9,000-year-old human remains from the eastern Mediterranean and a 17,000-year-old bison skeleton, both providing molecular evidence for an

infection with TB (Hershkovitz et al., 2008; Lee et al., 2012). Given this fact, mechanical cell disruption as described above is assumed to be advantageous even for ancient samples to completely breakdown remaining cell walls. The organic purification with chloroform-isoamyl alcohol and chloroform resulted in three separate phases containing DNA in the aqueous upper phase, presumably proteins in the interphase, and lipids and cell debris in the organic lower phase. The interphase was of particular interest, as from our own experience it is usually observed in tissue samples and rarely in bone samples (**Figure 4**, step 4). Therefore, it was isolated in addition to the upper phase and subjected to DNA precipitation with LPA and ethanol (**Figure 4**, step 5).

MTB DNA enrichment effect in the unknown interphase?

The DNA library profile displayed signs of inhibition but nevertheless the successful recovery of short DNA fragments. Surprisingly, shotgun sequencing data of the interphase sample resulted in a ten times higher mean coverage of the *Mycobacterium tuberculosis* H37Rv reference genome compared to the upper phase. The extraction of MTB DNA generally involves a harsh procedure to mechanically and chemically break down the lipid-rich cell wall to access the genetic material (Hosek et al., 2006; Somerville et al., 2005). Therefore, we hypothesize that this 'enrichment effect' could be caused by protein-DNA or lipid-DNA interaction pulling the DNA down instead of remaining in the aqueous phase. This exciting observation could be potentially game-changing for the efficient extraction of MTB DNA from ancient but also clinical samples – however, the results have yet to be reproduced with more samples and the mechanism behind it unraveled.

Clinical application of LPA-based DNA precipitation method

WGS has become the gold standard for MTB genotyping to detect multidrug-resistant and extensively drug-resistant strains and hence providing appropriate care for active TB patients. However, WGS-based detection of MTB usually relies on culture-based methods to retrieve sufficient DNA. This

significantly increases the turnaround time as MTB are slow-growing bacteria with a duplication time of over 18 hours. In addition, repeated culture within-sample can reduce the genetic diversity due to selective growth in culture (Nimmo et al., 2019). Sputum from patients is the most easily accessible and non-invasive sample type for rapid diagnosis. Like ancient DNA, sputum samples are a complex mixture of large proportions of human and oral/nasopharyngeal bacterial DNA but low concentration of MTB DNA. This impedes the nonspecific sequencing process and results in very low pathogen genome coverage and depth, limiting the utility of the data for downstream analysis (McNerney et al., 2017). Targeted enrichment is one way to bypass this problem but comes with relatively high costs, is labor intensive, and requires specific laboratory infrastructure that is not accessible everywhere, especially research facilities in low-income countries. Therefore, the DNA extraction and precipitation method described above (**FIGURE 4**) could potentially provide an alternative, economical approach in the future.

Detection of mixed TB infections

By applying the modified protocol of Maixner and colleagues as illustrated in **FIGURE 5**, we were able to recover high-coverage MTB shotgun sequencing data from the partially mummified remains of Szabina Orlich, an 18th-century midwife from Vác, Hungary. This finding was of particular interest as single nucleotide polymorphism (SNP) analysis revealed a high proportion of heterozygous SNPs, pointing towards a possible mixed infection with two distinct MTB strains.

In general, mixed infections are divided into three different types (**FIGURE 5**). Mixed species infections refer to infections with two or more different species belonging to the same genus simultaneously infecting a single host e.g., *Plasmodium vivax* and *Plasmodium falciparum* (Mayxay et al., 2004) or *Mycobacterium avium* complex and *Mycobacterium abscessus* complex (S. H. Shin et al., 2018). Microevolution and mixed strain infections are both types of polyclonal infections and refer to infections of a single host by two or more genetically distinct strains of the same pathogenic species (Cohen et al., 2011;

McNaughton et al., 2018). In the case of microevolution, the pathogen strain undergoes evolution within the host after infection, leading to minor genetic differences in the resulting progeny. Mixed strain infections on the other hand occur when a single host gets concomitantly or sequentially infected by genetically distinct pathogen strains (Moreno-Molina et al., 2021).

10-20% of TB patients in high burden setting are reported to have mixed infections with two distinct MTB strains (Warren et al., 2004; Zetola et al., 2014). Mixed infections are detected by identifying heterozygous single nucleotide polymorphism (SNP) and are often associated with an underlying immune suppression caused by HIV infection and poor treatment outcomes, especially when different drug resistance patterns between strains are detected (S. S. Shin et al., 2015). However, standard drug susceptibility testing assays are often inadequate due to their low sensitivity and lead to misclassification of cases, delaying effective intervention and posing serious challenges to the control of TB. Whole-genome sequencing enables high-resolution SNP analysis and is therefore the gold standard for the accurate detection of mixed infections in TB patients (Sobkowiak et al., 2018).

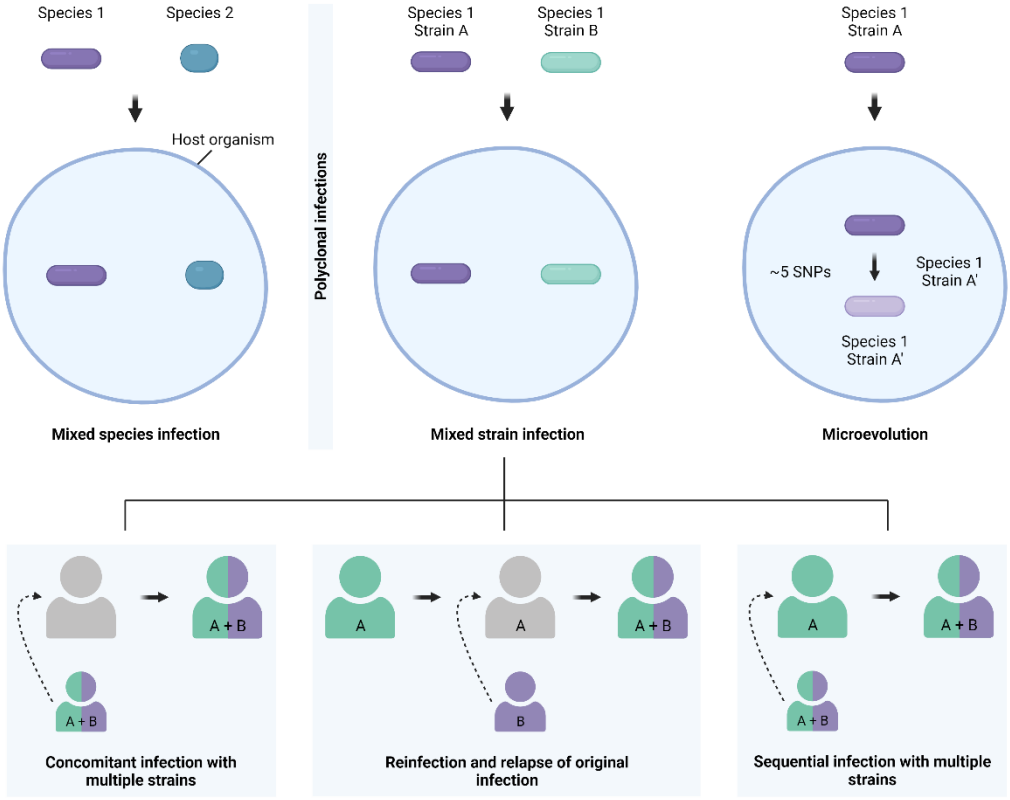

FIGURE 5: DIFFERENT TYPES OF MIXED INFECTIONS. Mixed species infections are caused by infection with two or more pathogens of the same genus. Infections with two pathogens of the same species are referred to as polyclonal infections and include mixed strain infections and microevolution. **Microevolution** occurs when a strain of a pathogen species undergoes small genetical changes within the host. **Mixed strain infections** can occur through various **infection scenarios** (blue boxes) resulting in infection with two or more strains of the same pathogen: grey human icons represent healthy hosts; colored human represent infected hosts; icons dotted arrows represent transmission events.

To avoid misinterpretation and incorrect epidemiological links, appropriate computational pipelines and reference databases have been established for accessible NGS analysis (Byrne et al., 2020; Gabbassov et al., 2021; Napier et al., 2020; Sobkowiak et al., 2018). Hence, the computational pipeline applied in **Chapter III** (**FIGURE 6**) has been developed for metagenomic screening, the confident detection of authentic ancient MTB DNA, and comprehensive genome-wide SNP analysis for phylogeny construction and mixed infection detection. To ensure the authenticity of data and exclude environmental contamination, stringent parameters were applied, and rigorous comparative analyses with publicly available well-documented reference databases of

modern and ancient specimens were taken advantage of. Thereby, we were able to identify a mixed strain infection with the two Euro-American Lineage 4 sublineages L4.1.2.1/Haarlem and L4.10/PGG3. Here, the MTC barcoding reference database developed by Napier and colleagues was of particular importance to identify lineage-specific SNPs as the data of the L4.10/PGG3 strain did not have enough genome coverage for reliable phylogenetic analysis (Napier et al., 2020).

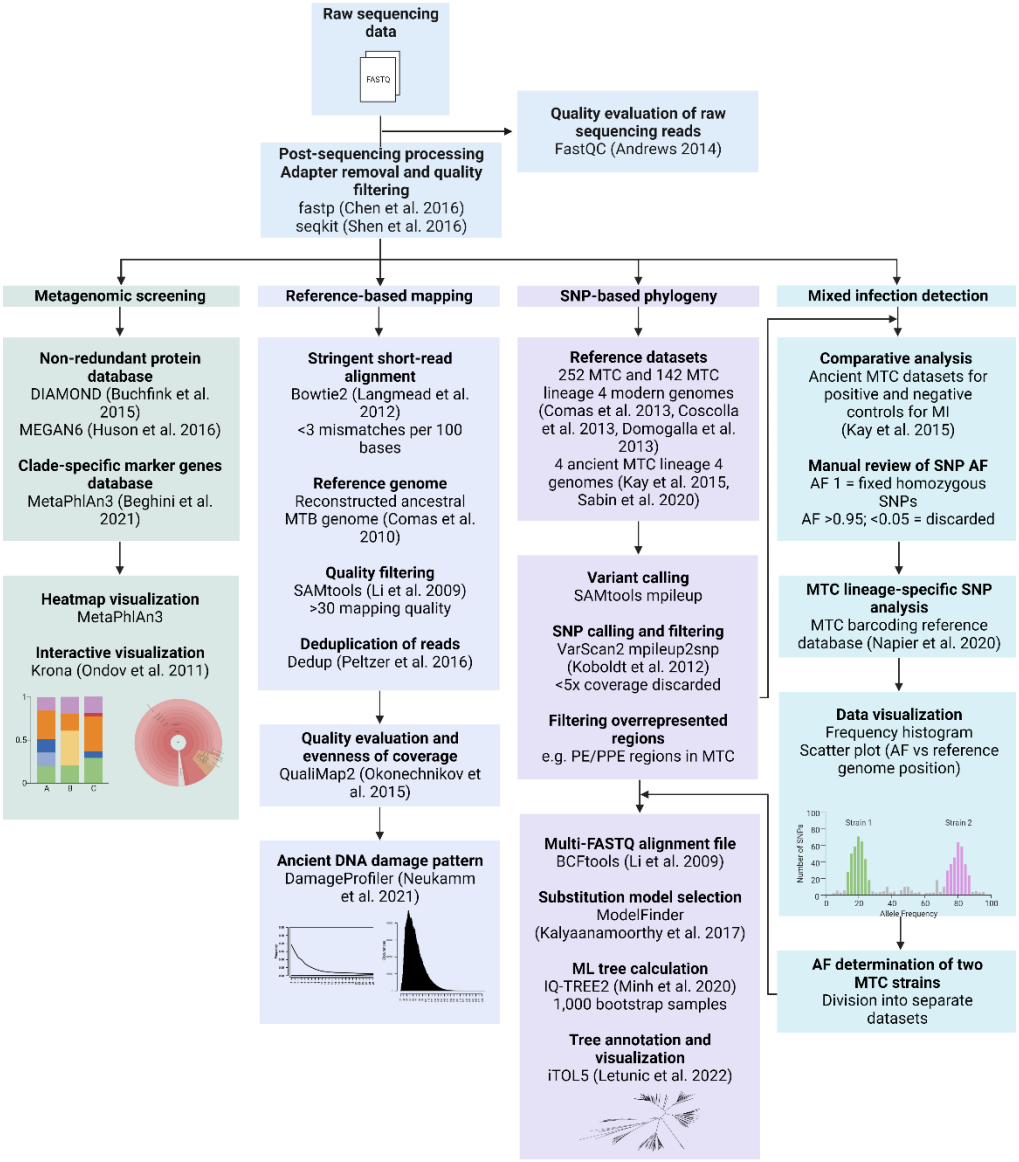

FIGURE 6: COMPUTATIONAL PIPELINE APPLIED IN JÄGER ET AL. (2022) for the analysis of ancient metagenomic shotgun sequencing datasets. First, overall quality of raw sequencing read is evaluated and subjected to trimming of adapter sequences. Reads with low sequencing quality scores are discarded. Quality-passed reads are subjected to metagenomic screening using

various taxonomic classifiers. Mapping against the reference genome is performed using stringent alignment and quality filter parameters. Fragmentation and damage plots are generated for ancient DNA authentication of reads. SNP-based phylogenetic analysis is performed using modern and ancient reference MTC genomes. Variant calling includes SNPs and indels and are subsequently filtered only for SNPs with ≥5x coverage. SNPs in hypervariable positions such as PE/PPE regions are discarded. Allele frequency (AF) distribution is manually examined in filtered variant call format (VCF) files. For confident mixed infection (MI) detection, reference ancient MTB datasets of individuals with MI are used as controls. SNPs with AF of 1 are defined as fixed position for both strains. SNPs with AF >0.95 and <0.05 are discarded due to low confidence level. For additional confirmation MTC-lineage specific SNPs are identified using a reference database for MTC barcoding. AF of each strain is determined through AF frequency histograms and scatter plots displaying AF of each SNP against its respective position on the genome. Datasets are divided according to each strains' AF. A multi-FASTQ alignment file is created including all datasets and the appropriate phylogenetic estimate model is selected for maximum-likelihood (ML) tree calculation. Finally, tree annotation and visualization are managed with an interactive online tool.

Mixed MTB infections in an eighteenth-century Hungarian population

The use of SNP allele frequency histograms is common in metagenomic ancient TB studies to determine whether heterozygous SNPs derive from environmental contamination or are the result of a potential polyclonal infection (Bos et al., 2014; Jäger et al., 2022; Sabin et al., 2020; Vågene et al., 2022). However, thus far, mixed strain infections with two or more distinct MTC strains have only been reported in three studies (Chan et al., 2013; Jäger et al., 2022; Kay et al., 2015): in all three studies, eighteenth-century individuals from the Vác Mummy Collection (VMC) from Hungary have been studied and a total of 15 MTC genome sequence datasets from 9 individuals have been generated (**Table 1**). All 15 MTC genome sequences have been classified as Euro-American Lineage 4 strains, or more precisely, as the generalist sublineages that are known to be highly adaptable and are among the most prevalent sublineages in Europe today (Stucki et al., 2016).

The VMC consists of 265 individuals in various states of natural mummification that were buried between 1731 and 1838 CE in the local Dominican church of Vác, Hungary, and have well-documented archival records revealing occupations and family relationships (H. D. Donoghue et al., 2021). Previous studies have repeatedly confirmed a high proportion of TB infection (>60%) within the local community (H. D. Donoghue et al., 2011; Fletcher, Donoghue, Holton, et al., 2003; Pap et al., 1999; Szikossy I et al., 1997). Here, in **Chapter III**, we were able to further confirm the high prevalence of mixed MTB strain infections in pre-antibiotic eighteenth-century Central Europe, as previously described by Kay and colleagues (Kay et al., 2015). In addition, the study also confirms the long-time prevalence of mixed infection for over one hundred years within the local community of Vác (Jäger et al., 2022).

Archaeological records like the VMC can provide unique insights into the relationship between humans and MTB in the pre-urbanized, pre-antibiotic era. Due to the unique preservation of the individuals isolated from exogenous

influences like microbial contamination from soil or extreme temperatures, the endogenous DNA is mostly well-preserved with relatively low background DNA levels – making them optimal candidates for the successful recovery of ancient MTB DNA. The optimized molecular and computational workflow presented in this dissertation (**FIGURE 1**) will contribute to future studies retrieving further high-coverage MTB genome sequencing data of whole family constellations of this collection. Thereby, we will be able to perform high-resolution SNP analysis to reconstruct epidemiological links between family members or even within the community. This will not only contribute to the better estimation of mutation rates of strains prior to the rise of antibiotic resistances but also help to better understand the influence of infectious diseases on the community and the influence of social responses to a disease that still affect us today.

.

5 | Conclusions and Outlook

The advent of ancient pathogen genomics has led to high expectations within the community to illuminate the evolutionary past of infectious diseases that have plagued humankind for many millennia. However, these expectations are yet to be fulfilled, as continuing effort is required to develop more sensitive methods to recover high-resolution genomic data from finite archaeological records. Moreover, it is crucial that studies of the past adhere to rigorous scientific standards and the conduct of destructive analysis is sufficiently justified. Nevertheless, simultaneously, obtained ancient pathogen genomes have the potential to contribute to biomedical research and perhaps to the development of prevention and treatment strategies.

To this date, all successful metagenomic studies on ancient TB have resulted from either mummified or non-buried individuals highlighting the challenge of recovering high-coverage MTB genomes when contaminated with high levels of background DNA. Whilst evidence based on PCR and biolipid markers for Neolithic TB cases do exist, metagenomic methods remain to increase in sensitivity and resolution to confirm these cases on a genome level and answer the question on the date of emergence of this pathogen and understand its influence on cultures and societies of the past.

The development of new methods plays a crucial part in this young and fast-changing field of research to optimize the balance between the amount of information that can be gained while maintaining the respect towards the inherent value of precious, unrestorable material. Furthermore, the exchange of knowledge between the field of ancient and modern genomics on molecular and computational methods to isolate and genotype MTB DNA is fundamental to overcome global TB burden.

6 | References

Achtman, M. (2008). Evolution, population structure, and phylogeography of genetically monomorphic bacterial pathogens. *Annual Review of Microbiology*, *62*(1), 53–70. https://doi.org/10.1146/annurev.micro.62.081307.162832

Adey, A., Morrison, H. G., (no last name), A., Xun, X., Kitzman, J. O., Turner, E. H., Stackhouse, B., MacKenzie, A. P., Caruccio, N. C., Zhang, X., & Shendure, J. (2010). Rapid, low-input, low-bias construction of shotgun fragment libraries by high-density in vitro transposition. *Genome Biology*, *11*(12), R119. https://doi.org/10.1186/gb-2010-11-12-r119

Alpaslan-Roodenberg, S., Anthony, D., Babiker, H., Bánffy, E., Booth, T., Capone, P., Deshpande-Mukherjee, A., Eisenmann, S., Fehren-Schmitz, L., Frachetti, M., Fujita, R., Frieman, C. J., Fu, Q., Gibbon, V., Haak, W., Hajdinjak, M., Hofmann, K. P., Holguin, B., Inomata, T., ... Zahir, M. (2021). Ethics of DNA research on human remains: Five globally applicable guidelines. *Nature*, *599*(7883), 41–46. https://doi.org/10.1038/s41586-021-04008-x

Armelagos, G. J., Goodman, A. H., & Jacobs, K. H. (1991). The origins of agriculture: population growth during a period of declining health. *Population and Environment 1991 13:1*, *13*(1), 9–22. https://doi.org/10.1007/BF01256568

Arriaza, B. T., Salo, W., Aufderheide, A. C., & Holcomb, T. A. (1995). Pre-Columbian tuberculosis in Northern Chile: Molecular and skeletal evidence. *American Journal of Physical Anthropology*, *98*(1), 37–45. https://doi.org/10.1002/ajpa.1330980104

Baker, O., Lee, O. Y.-C., Wu, H. H. T., Besra, G. S., Minnikin, D. E., Llewellyn, G., Williams, C. M., Maixner, F., O'Sullivan, N., Zink, A., Chamel, B., Khawam, R., Coqueugniot, E., Helmer, D., le Mort, F., Perrin, P., Gourichon, L., Dutailly, B., Pálfi, G., ... Dutour, O. (2015).

Human tuberculosis predates domestication in ancient Syria. *Tuberculosis*, *95*, S4–S12. https://doi.org/10.1016/j.tube.2015.02.001

Bardill, J., Bader, A. C., Garrison, N. A., Bolnick, D. A., Raff, J. A., Walker, A., & Malhi, R. S. (2018). Advancing the ethics of paleogenomics. *Science*, *360*(6387), 384–385. https://doi.org/10.1126/science.aaq1131

Barrett, R., Kuzawa, C. W., McDade, T., & Armelagos, G. J. (1998). Emerging and re-emerging infectious diseases: the third epidemiologic transition. *Annual Review of Anthropology*, *27*(1), 247–271. https://doi.org/10.1146/annurev.anthro.27.1.247

Beghini, F., McIver, L. J., Blanco-Míguez, A., Dubois, L., Asnicar, F., Maharjan, S., Mailyan, A., Manghi, P., Scholz, M., Thomas, A. M., Valles-Colomer, M., Weingart, G., Zhang, Y., Zolfo, M., Huttenhower, C., Franzosa, E. A., & Segata, N. (2021). Integrating taxonomic, functional, and strain-level profiling of diverse microbial communities with bioBakery 3. *ELife*, *10*. https://doi.org/10.7554/eLife.65088

Bespiatykh, D., Bespyatykh, J., Mokrousov, I., & Shitikov, E. (2021). A comprehensive map of *Mycobacterium tuberculosis* complex regions of difference. *MSphere*, *6*(4). https://doi.org/10.1128/mSphere.00535-21

Bos, K. I., Harkins, K. M., Herbig, A., Coscolla, M., Weber, N., Comas, I., Forrest, S. A., Bryant, J. M., Harris, S. R., Schuenemann, V. J., Campbell, T. J., Majander, K., Wilbur, A. K., Guichon, R. A., Wolfe Steadman, D. L., Cook, D. C., Niemann, S., Behr, M. A., Zumarraga, M., ... Krause, J. (2014). Pre-Columbian mycobacterial genomes reveal seals as a source of New World human tuberculosis. *Nature*, *514*(7523), 494–497. https://doi.org/10.1038/nature13591

Bos, K. I., Herbig, A., Sahl, J., Waglechner, N., Fourment, M., Forrest, S. A., Klunk, J., Schuenemann, V. J., Poinar, D., Kuch, M., Golding, G. B., Dutour, O., Keim, P., Wagner, D. M., Holmes, E. C., Krause, J., & Poinar, H. N. (2016). Eighteenth century *Yersinia pestis* genomes reveal

the long-term persistence of an historical plague focus. *ELife*, *5*. https://doi.org/10.7554/eLife.12994

Bos, K. I., Schuenemann, V. J., Golding, G. B., Burbano, H. A., Waglechner, N., Coombes, B. K., McPhee, J. B., Dewitte, S. N., Meyer, M., Schmedes, S., Wood, J., Earn, D. J. D., Herring, D. A., Bauer, P., Poinar, H. N., & Krause, J. (2011). A draft genome of *Yersinia pestis* from victims of the Black Death. *Nature*, *478*(7370), 506–510. https://doi.org/10.1038/nature10549

Bouwman, A. S., Kennedy, S. L., Müller, R., Stephens, R. H., Holst, M., Caffell, A. C., Roberts, C. A., & Brown, T. A. (2012). Genotype of a historic strain of *Mycobacterium tuberculosis*. *Proceedings of the National Academy of Sciences*, *109*(45), 18511–18516. https://doi.org/10.1073/pnas.1209444109

Brennan, P. J. (2003). Structure, function, and biogenesis of the cell wall of *Mycobacterium tuberculosis*. *Tuberculosis*, *83*(1–3), 91–97. https://doi.org/10.1016/S1472-9792(02)00089-6

Brosch, R., Gordon, S. v., Marmiesse, M., Brodin, P., Buchrieser, C., Eiglmeier, K., Garnier, T., Gutierrez, C., Hewinson, G., Kremer, K., Parsons, L. M., Pym, A. S., Samper, S., van Soolingen, D., & Cole, S. T. (2002). A new evolutionary scenario for the *Mycobacterium tuberculosis* complex. *Proceedings of the National Academy of Sciences*, *99*(6), 3684–3689. https://doi.org/10.1073/pnas.052548299

Brown, K., Mund, D. F., Aberle, D. R., Batra, P., & Young, D. A. (1994). Intrathoracic calcifications: radiographic features and differential diagnoses. *RadioGraphics*, *14*(6), 1247–1261. https://doi.org/10.1148/radiographics.14.6.7855339

Buikstra, J., & Roberts, C. (Eds.). (2012). *The global history of paleopathology*. Oxford University Press. https://doi.org/10.1093/acprof:osobl/9780195389807.001.0001

Burbano, H. A., Hodges, E., Green, R. E., Briggs, A. W., Krause, J., Meyer, M., Good, J. M., Maricic, T., Johnson, P. L. F., Xuan, Z., Rooks, M., Bhattacharjee, A., Brizuela, L., Albert, F. W., de la Rasilla, M., Fortea, J., Rosas, A., Lachmann, M., Hannon, G. J., & Pääbo, S. (2010). Targeted investigation of the Neandertal genome by array-based sequence capture. *Science*, *328*(5979), 723–725. https://doi.org/10.1126/science.1188046

Byrne, A. S., Goudreau, A., Bissonnette, N., Shamputa, I. C., & Tahlan, K. (2020). Methods for detecting mycobacterial mixed strain infections – a systematic review. *Frontiers in Genetics*, *11*. https://doi.org/10.3389/fgene.2020.600692

Campana, M. G., Robles García, N., Rühli, F. J., & Tuross, N. (2014). False positives complicate ancient pathogen identifications using high-throughput shotgun sequencing. *BMC Research Notes*, *7*(1), 111. https://doi.org/10.1186/1756-0500-7-111

Carøe, C., Gopalakrishnan, S., Vinner, L., Mak, S. S. T., Sinding, M. H. S., Samaniego, J. A., Wales, N., Sicheritz-Pontén, T., & Gilbert, M. T. P. (2018). Single-tube library preparation for degraded DNA. *Methods in Ecology and Evolution*, *9*(2), 410–419. https://doi.org/10.1111/2041-210X.12871

Chan, J. Z.-M., Sergeant, M. J., Lee, O. Y.-C., Minnikin, D. E., Besra, G. S., Pap, I., Spigelman, M., Donoghue, H. D., & Pallen, M. J. (2013). Metagenomic analysis of tuberculosis in a mummy. *New England Journal of Medicine*, *369*(3), 289–290. https://doi.org/10.1056/NEJMc1302295

Cohen, T., Wilson, D., Wallengren, K., Samuel, E. Y., & Murray, M. (2011). Mixed-strain *Mycobacterium tuberculosis* infections among patients dying in a hospital in KwaZulu-Natal, South Africa. *Journal of Clinical Microbiology*, *49*(1), 385–388. https://doi.org/10.1128/JCM.01378-10

Cole, G., Taylor, G. M., Stewart, G. R., & Dawson-Hobbis, H. (2022). Ancient DNA confirmation of lepromatous leprosy in a skeleton with concurrent osteosarcoma, excavated from the leprosarium of St. Mary Magdalen in Winchester, Hants., UK. *European Journal of Clinical Microbiology & Infectious Diseases*. https://doi.org/10.1007/s10096-022-04494-5

Comas, I., Coscolla, M., Luo, T., Borrell, S., Holt, K. E., Kato-Maeda, M., Parkhill, J., Malla, B., Berg, S., Thwaites, G., Yeboah-Manu, D., Bothamley, G., Mei, J., Wei, L., Bentley, S., Harris, S. R., Niemann, S., Diel, R., Aseffa, A., ... Gagneux, S. (2013). Out-of-Africa migration and Neolithic coexpansion of *Mycobacterium tuberculosis* with modern humans. *Nature Genetics*, *45*(10), 1176–1182. https://doi.org/10.1038/ng.2744

Comas, I., & Gagneux, S. (2011). A role for systems epidemiology in tuberculosis research. *Trends in Microbiology*, *19*(10), 492–500. https://doi.org/10.1016/j.tim.2011.07.002

Comas, I., Hailu, E., Kiros, T., Bekele, S., Mekonnen, W., Gumi, B., Tschopp, R., Ameni, G., Hewinson, R. G., Robertson, B. D., Goig, G. A., Stucki, D., Gagneux, S., Aseffa, A., Young, D., & Berg, S. (2015). Population genomics of *Mycobacterium tuberculosis* in Ethiopia contradicts the Virgin Soil hypothesis for human tuberculosis in Sub-Saharan Africa. *Current Biology*, *25*(24), 3260–3266. https://doi.org/10.1016/j.cub.2015.10.061

Cooper, A., & Poinar, H. N. (2000). Ancient DNA: Do it right or not at all. *Science*, *289*(5482), 1139–1139. https://doi.org/10.1126/science.289.5482.1139b

Cortez, A. D., Bolnick, D. A., Nicholas, G., Bardill, J., & Colwell, C. (2021). An ethical crisis in ancient DNA research: insights from the Chaco

Canyon controversy as a case study. *Journal of Social Archaeology*, *21*(2), 157–178. https://doi.org/10.1177/1469605321991600

Crump, S. E., Fréchette, B., Power, M., Cutler, S., de Wet, G., Raynolds, M. K., Raberg, J. H., Briner, J. P., Thomas, E. K., Sepúlveda, J., Shapiro, B., Bunce, M., & Miller, G. H. (2021). Ancient plant DNA reveals High Arctic greening during the Last Interglacial. *Proceedings of the National Academy of Sciences*, *118*(13). https://doi.org/10.1073/pnas.2019069118

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

Dabney, J., & Meyer, M. (2019). *Extraction of highly degraded DNA from ancient bones and teeth* (pp. 25–29). https://doi.org/10.1007/978-1-4939-9176-1_4

Dabney, J., Meyer, M., & Pääbo, S. (2013). Ancient DNA damage. *Cold Spring Harbor Perspectives in Biology*, *5*(7). https://doi.org/10.1101/CSHPERSPECT.A012567

Damgaard, P. B., Margaryan, A., Schroeder, H., Orlando, L., Willerslev, E., & Allentoft, M. E. (2015). Improving access to endogenous DNA in ancient bones and teeth. *Scientific Reports*, *5*(1), 11184. https://doi.org/10.1038/srep11184

der Sarkissian, C., Möller, P., Hofman, C. A., Ilsøe, P., Rick, T. C., Schiøtte, T., Sørensen, M. V., Dalén, L., & Orlando, L. (2020). Unveiling the ecological applications of ancient DNA from mollusk shells. *Frontiers in Ecology and Evolution*, *8*. https://doi.org/10.3389/fevo.2020.00037

der Sarkissian, C., Pichereau, V., Dupont, C., Ilsøe, P. C., Perrigault, M., Butler, P., Chauvaud, L., Eiríksson, J., Scourse, J., Paillard, C., & Orlando, L. (2017). Ancient DNA analysis identifies marine mollusc shells as new metagenomic archives of the past. *Molecular Ecology Resources*, *17*(5), 835–853. https://doi.org/10.1111/1755-0998.12679

Donoghue, H. D., Páp, I., Szikossy, I., & Spigelman, M. (2011). Detection and characterization of *Mycobacterium tuberculosis* DNA in 18th century Hungarians with pulmonary and extra-pulmonary tuberculosis. In *Yearbook of Mummy Studies* (Vol. 1, pp. 51–56).

Donoghue, H. D., Pap, I., Szikossy, I., & Spigelman, M. (2021). The Vác Mummy Project: investigation of 265 eighteenth-century mummified remains from the TB pandemic era. In *The Handbook of Mummy Studies* (pp. 777–805). Springer Singapore. https://doi.org/https://doi.org/10.1007/978-981-15-3354-9_21

Donoghue, H., Taylor, G., Stewart, G., Lee, O., Wu, H., Besra, G., & Minnikin, D. (2017). Positive diagnosis of ancient leprosy and tuberculosis using ancient DNA and lipid biomarkers. *Diversity*, *9*(4), 46. https://doi.org/10.3390/d9040046

Drancourt, M., Aboudharam, G., Signoli, M., Dutour, O., & Raoult, D. (1998). Detection of 400-year-old *Yersinia pestis* DNA in human dental pulp: An approach to the diagnosis of ancient septicemia. *Proceedings of the National Academy of Sciences*, *95*(21), 12637–12640. https://doi.org/10.1073/pnas.95.21.12637

Duchêne, S., Ho, S. Y. W., Carmichael, A. G., Holmes, E. C., & Poinar, H. (2020). The recovery, interpretation and use of ancient pathogen genomes. *Current Biology*, *30*(19), R1215–R1231. https://doi.org/10.1016/J.CUB.2020.08.081

Duchêne, S., Holt, K. E., Weill, F.-X., le Hello, S., Hawkey, J., Edwards, D. J., Fourment, M., & Holmes, E. C. (2016). Genome-scale rates of

evolutionary change in bacteria. *Microbial Genomics*, *2*(11). https://doi.org/10.1099/mgen.0.000094

Duggan, A. T., Perdomo, M. F., Piombino-Mascali, D., Marciniak, S., Poinar, D., Emery, M. v., Buchmann, J. P., Duchêne, S., Jankauskas, R., Humphreys, M., Golding, G. B., Southon, J., Devault, A., Rouillard, J. M., Sahl, J. W., Dutour, O., Hedman, K., Sajantila, A., Smith, G. L., ... Poinar, H. N. (2016). 17th century variola virus reveals the recent history of smallpox. *Current Biology*, *26*(24), 3407–3412. https://doi.org/https://doi.org/10.1016/j.cub.2016.10.061

Dziadek, J., Sajduda, A., & Boruń, T. M. (2001). Specificity of insertion sequence-based PCR assays for *Mycobacterium tuberculosis* complex. *The International Journal of Tuberculosis and Lung Disease : The Official Journal of the International Union against Tuberculosis and Lung Disease*, *5*(6), 569–574.

Enk, J. M., Devault, A. M., Kuch, M., Murgha, Y. E., Rouillard, J.-M., & Poinar, H. N. (2014). Ancient whole genome enrichment using baits built from modern DNA. *Molecular Biology and Evolution*, *31*(5), 1292–1294. https://doi.org/10.1093/molbev/msu074

Ernst, J. D. (2012). The immunological life cycle of tuberculosis. *Nature Reviews Immunology*, *12*(8), 581–591. https://doi.org/10.1038/nri3259

Eshed, V., Gopher, A., Gage, T. B., & Hershkovitz, I. (2004). Has the transition to agriculture reshaped the demographic structure of prehistoric populations? New evidence from the Levant. *American Journal of Physical Anthropology*, *124*(4), 315–329. https://doi.org/10.1002/ajpa.10332

Fischer, M., Renevey, N., Thür, B., Hoffmann, D., Beer, M., & Hoffmann, B. (2016). Efficacy assessment of nucleic acid decontamination reagents used in molecular diagnostic laboratories. *PLOS ONE*, *11*(7), e0159274. https://doi.org/10.1371/journal.pone.0159274

Fletcher, H. A., Donoghue, H. D., Holton, J., Pap, I., & Spigelman, M. (2003). Widespread occurrence of *Mycobacterium tuberculosis* DNA from 18th-19th century Hungarians. *American Journal of Physical Anthropology*, *120*(2), 144–152. https://doi.org/10.1002/ajpa.10114

Fletcher, H. A., Donoghue, H. D., Taylor, G. M., van der Zanden, A. G., & Spigelman, M. (2003). Molecular analysis of *Mycobacterium tuberculosis* DNA from a family of 18th century Hungarians. *Microbiology*, *149*(1), 143–151. https://doi.org/10.1099/mic.0.25961-0

Flynn, J. L., Chan, J., & Lin, P. L. (2011). Macrophages and control of granulomatous inflammation in tuberculosis. *Mucosal Immunology*, *4*(3), 271–278. https://doi.org/10.1038/mi.2011.14

Fordyce, S. L., Ávila-Arcos, M. C., Rasmussen, M., Cappellini, E., Romero-Navarro, J. A., Wales, N., Alquezar-Planas, D. E., Penfield, S., Brown, T. A., Vielle-Calzada, J.-P., Montiel, R., Jørgensen, T., Odegaard, N., Jacobs, M., Arriaza, B., Higham, T. F. G., Ramsey, C. B., Willerslev, E., & Gilbert, M. T. P. (2013). Deep sequencing of RNA from ancient maize kernels. *PLoS ONE*, *8*(1), e50961. https://doi.org/10.1371/journal.pone.0050961

Fu, Q., Meyer, M., Gao, X., Stenzel, U., Burbano, H. A., Kelso, J., & Pääbo, S. (2013). DNA analysis of an early modern human from Tianyuan Cave, China. *Proceedings of the National Academy of Sciences*, *110*(6), 2223–2227. https://doi.org/10.1073/pnas.1221359110

Fulton, T. L., & Shapiro, B. (2019). *Setting up an ancient DNA laboratory* (pp. 1–13). https://doi.org/10.1007/978-1-4939-9176-1_1

Gabbassov, E., Moreno-Molina, M., Comas, I., Libbrecht, M., & Chindelevitch, L. (2021). SplitStrains, a tool to identify and separate mixed *Mycobacterium tuberculosis* infections from WGS data. *Microbial Genomics*, *7*(6). https://doi.org/10.1099/mgen.0.000607

Gagneux, S. (2018). Ecology and evolution of *Mycobacterium tuberculosis*. *Nature Reviews Microbiology*, *16*(4), 202–213. https://doi.org/10.1038/nrmicro.2018.8

Gaillard, C., & Strauss, F. (1990). Ethanol precipitation of DNA with linear polyacrylamide as carrier. *Nucleic Acids Research*, *18*(2), 378–378. https://doi.org/10.1093/nar/18.2.378

Gamba, C., Hanghøj, K., Gaunitz, C., Alfarhan, A. H., Alquraishi, S. A., Al-Rasheid, K. A. S., Bradley, D. G., & Orlando, L. (2016). Comparing the performance of three ancient DNA extraction methods for high-throughput sequencing. *Molecular Ecology Resources*, *16*(2), 459–469. https://doi.org/10.1111/1755-0998.12470

Gamba, C., Jones, E. R., Teasdale, M. D., McLaughlin, R. L., Gonzalez-Fortes, G., Mattiangeli, V., Domboróczki, L., Kővári, I., Pap, I., Anders, A., Whittle, A., Dani, J., Raczky, P., Higham, T. F. G., Hofreiter, M., Bradley, D. G., & Pinhasi, R. (2014). Genome flux and stasis in a five millennium transect of European prehistory. *Nature Communications*, *5*(1), 5257. https://doi.org/10.1038/ncomms6257

Gansauge, M.-T., Gerber, T., Glocke, I., Korlević, P., Lippik, L., Nagel, S., Riehl, L. M., Schmidt, A., & Meyer, M. (2017). Single-stranded DNA library preparation from highly degraded DNA using *T4* DNA ligase. *Nucleic Acids Research*, gkx033. https://doi.org/10.1093/nar/gkx033

Gansauge, M.-T., & Meyer, M. (2013). Single-stranded DNA library preparation for the sequencing of ancient or damaged DNA. *Nature Protocols*, *8*(4), 737–748. https://doi.org/10.1038/nprot.2013.038

Gengenbacher, M., & Kaufmann, S. H. E. (2012). *Mycobacterium tuberculosis*: Success through dormancy. *FEMS Microbiology Reviews*, *36*(3), 514–532. https://doi.org/10.1111/j.1574-6976.2012.00331.x

Giffin, K., Lankapalli, A. K., Sabin, S., Spyrou, M. A., Posth, C., Kozakaitė, J., Friedrich, R., Miliauskienė, Ž., Jankauskas, R., Herbig, A., & Bos, K. I. (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), 9499. https://doi.org/10.1038/s41598-020-66012-x

Goodwin, S., McPherson, J. D., & McCombie, W. R. (2016). Coming of age: Ten years of next-generation sequencing technologies. *Nature Reviews Genetics*, *17*(6), 333–351. https://doi.org/10.1038/nrg.2016.49

Granehäll, L., Huang, K. D., Tett, A., Manghi, P., Paladin, A., O'Sullivan, N., Rota-Stabelli, O., Segata, N., Zink, A., & Maixner, F. (2021). Metagenomic analysis of ancient dental calculus reveals unexplored diversity of oral archaeal Methanobrevibacter. *Microbiome*, *9*(1), 197. https://doi.org/10.1186/s40168-021-01132-8

Grauer, A. L. (2011). *A companion to paleopathology* (Vol. 23). John Wiley & Sons.

Green, R. E., Briggs, A. W., Krause, J., Prüfer, K., Burbano, H. A., Siebauer, M., Lachmann, M., & Pääbo, S. (2009). The Neandertal genome and ancient DNA authenticity. *The EMBO Journal*, *28*(17), 2494–2502. https://doi.org/10.1038/emboj.2009.222

Green, R. E., Krause, J., Briggs, A. W., Maricic, T., Stenzel, U., Kircher, M., Patterson, N., Li, H., Zhai, W., Fritz, M. H.-Y., Hansen, N. F., Durand, E. Y., Malaspinas, A.-S., Jensen, J. D., Marques-Bonet, T., Alkan, C., Prüfer, K., Meyer, M., Burbano, H. A., ... Pääbo, S. (2010). A draft sequence of the Neandertal genome. *Science*, *328*(5979), 710–722. https://doi.org/10.1126/science.1188021

Gutierrez, M. C., Brisse, S., Brosch, R., Fabre, M., Omaïs, B., Marmiesse, M., Supply, P., & Vincent, V. (2005). Ancient origin and gene mosaicism of

the progenitor of *Mycobacterium tuberculosis*. *PLoS Pathogens*, *1*(1). https://doi.org/10.1371/journal.ppat.0010005

Hansen, H. B., Damgaard, P. B., Margaryan, A., Stenderup, J., Lynnerup, N., Willerslev, E., & Allentoft, M. E. (2017). Comparing ancient DNA preservation in petrous bone and tooth cementum. *PLOS ONE*, *12*(1), e0170940. https://doi.org/10.1371/journal.pone.0170940

Harney, É., Cheronet, O., Fernandes, D. M., Sirak, K., Mah, M., Bernardos, R., Adamski, N., Broomandkhoshbacht, N., Callan, K., Lawson, A. M., Oppenheimer, J., Stewardson, K., Zalzala, F., Anders, A., Candilio, F., Constantinescu, M., Coppa, A., Ciobanu, I., Dani, J., ... Pinhasi, R. (2021). A minimally destructive protocol for DNA extraction from ancient teeth. *Genome Research*, *31*(3), 472–483. https://doi.org/10.1101/gr.267534.120

Hedman, J., & Rådström, P. (2013). *Overcoming inhibition in real-time diagnostic PCR* (pp. 17–48). https://doi.org/10.1007/978-1-60327-353-4_2

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

Higuchi, R., Bowman, B., Freiberger, M., Ryder, O. A., & Wilson, A. C. (1984). DNA sequences from the quagga, an extinct member of the horse family. *Nature*, *312*(5991), 282–284. https://doi.org/10.1038/312282a0

Hofreiter, M., Paijmans, J. L. A., Goodchild, H., Speller, C. F., Barlow, A., Fortes, G. G., Thomas, J. A., Ludwig, A., & Collins, M. J. (2015). The future of ancient DNA: Technical advances and conceptual shifts. *BioEssays*, *37*(3), 284–293. https://doi.org/10.1002/bies.201400160

Holloway, K. L., Henneberg, R. J., de Barros Lopes, M., & Henneberg, M. (2011). Evolution of human tuberculosis: A systematic review and meta-analysis of paleopathological evidence. *HOMO*, *62*(6), 402–458. https://doi.org/10.1016/j.jchb.2011.10.001

Hosek, J., Svastova, P., Moravkova, M., Pavlik, I., & Bartos, M. (2006). Methods of mycobacterial DNA isolation from different biological material: A review. *Veterinarni Medicina*, *51*(5), 180–192.

Höss, M., & Pääbo, S. (1993). DNA extraction from Pleistocene bones by a silica-based purification method. *Nucleic Acids Research*, *21*(16), 3913–3914. https://doi.org/10.1093/nar/21.16.3913

Jäger, H. Y., Atz Zanotelli, D., Maixner, F., Nicklisch, N., Alt, K. W., Meller, H., Pap, I., Szikossy, I., Pálfi, G., & Zink, A. R. (2023). Hit or miss - A metagenomic evaluation of intra-bone variability of host pathogen load in tuberculosis-infected human remains. *Tuberculosis*, *143*, 102392. https://doi.org/10.1016/j.tube.2023.102392

Jäger, H. Y., Maixner, F., Pap, I., Szikossy, I., Pálfi, G., & Zink, A. R. (2022). Metagenomic analysis reveals mixed *Mycobacterium tuberculosis* infection in a 18th century Hungarian midwife. *Tuberculosis*, *137*, 102181. https://doi.org/10.1016/j.tube.2022.102181

Jang, J., Becq, J., Gicquel, B., Deschavanne, P., & Neyrolles, O. (2008). Horizontally acquired genomic islands in the tubercle bacilli. *Trends in Microbiology*, *16*(7), 303–308. https://doi.org/10.1016/j.tim.2008.04.005

Janssen, P. H. (2006). Identifying the dominant soil bacterial taxa in libraries of 16S rRNA and 16S rRNA genes. *Applied and Environmental Microbiology*, *72*(3), 1719–1728. https://doi.org/10.1128/AEM.72.3.1719-1728.2006

Javan, G. T., Finley, S. J., Abidin, Z., & Mulle, J. G. (2016). The thanatomicrobiome: A missing piece of the microbial puzzle of death. *Frontiers in Microbiology*, *7*. https://doi.org/10.3389/fmicb.2016.00225

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

Kapp, J. D., Green, R. E., & Shapiro, B. (2021). A fast and efficient single-stranded genomic library preparation method optimized for ancient DNA. *Journal of Heredity*, *112*(3), 241–249. https://doi.org/10.1093/jhered/esab012

Katrina G. Claw, Dorothy Lippert, Jessica Bardill, Anna Cordova, Keolu Fox, Joseph M. Yracheta, Alyssa C. Bader, Deborah A. Bolnick, Ripan S. Malhi, Kimberly TallBear, & Nanibaa' A. Garrison. (2017). Chaco Canyon dig unearths ethical concerns. *Human Biology*, *89*(3), 177. https://doi.org/10.13110/humanbiology.89.3.01

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

Keller, M., Spyrou, M. A., Scheib, C. L., Neumann, G. U., Kröpelin, A., Haas-Gebhard, B., Päffgen, B., Haberstroh, J., Ribera i Lacomba, A., Raynaud, C., Cessford, C., Durand, R., Stadler, P., Nägele, K., Bates, J. S., Trautmann, B., Inskip, S. A., Peters, J., Robb, J. E., ... Krause, J. (2019). Ancient *Yersinia pestis* genomes from across Western Europe reveal early diversification during the First Pandemic (541–750).

Proceedings of the National Academy of Sciences, *116*(25), 12363–12372. https://doi.org/10.1073/pnas.1820447116

Kennett, D. J., Plog, S., George, R. J., Culleton, B. J., Watson, A. S., Skoglund, P., Rohland, N., Mallick, S., Stewardson, K., Kistler, L., LeBlanc, S. A., Whiteley, P. M., Reich, D., & Perry, G. H. (2017). Archaeogenomic evidence reveals prehistoric matrilineal dynasty. *Nature Communications*, *8*(1), 14115. https://doi.org/10.1038/ncomms14115

Key, F. M., Posth, C., Esquivel-Gomez, L. R., Hübler, R., Spyrou, M. A., Neumann, G. U., Furtwängler, A., Sabin, S., Burri, M., Wissgott, A., Lankapalli, A. K., Vågene, Å. J., Meyer, M., Nagel, S., Tukhbatova, R., Khokhlov, A., Chizhevsky, A., Hansen, S., Belinsky, A. B., ... Krause, J. (2020). Emergence of human-adapted *Salmonella enterica* is linked to the Neolithization process. *Nature Ecology & Evolution*, *4*(3), 324–333. https://doi.org/10.1038/s41559-020-1106-9

Knapp, M., Clarke, A. C., Horsburgh, K. A., & Matisoo-Smith, E. A. (2012). Setting the stage – Building and working in an ancient DNA laboratory. *Annals of Anatomy - Anatomischer Anzeiger*, *194*(1), 3–6. https://doi.org/10.1016/j.aanat.2011.03.008

Koch, R. (1982). The etiology of tuberculosis. *Clinical Infectious Diseases*, *4*(6), 1270–1274. https://doi.org/10.1093/clinids/4.6.1270

Korlević, P., Gerber, T., Gansauge, M.-T., Hajdinjak, M., Nagel, S., Aximu-Petri, A., & Meyer, M. (2015). Reducing microbial and human contamination in DNA extractions from ancient bones and teeth. *BioTechniques*, *59*(2), 87–93. https://doi.org/10.2144/000114320

Korneliussen, T. S., Albrechtsen, A., & Nielsen, R. (2014). ANGSD: Analysis of next generation sequencing data. *BMC Bioinformatics*, *15*(1), 356. https://doi.org/10.1186/s12859-014-0356-4

Krause-Kyora, B., Nutsua, M., Boehme, L., Pierini, F., Pedersen, D. D., Kornell, S. C., Drichel, D., Bonazzi, M., Möbus, L., Tarp, P., Susat, J., Bosse, E., Willburger, B., Schmidt, A. H., Sauter, J., Franke, A., Wittig, M., Caliebe, A., Nothnagel, M., ... Nebel, A. (2018). Ancient DNA study reveals HLA susceptibility locus for leprosy in medieval Europeans. *Nature Communications*, *9*(1). https://doi.org/10.1038/S41467-018-03857-X

Krause-Kyora, B., Susat, J., Key, F. M., Kühnert, D., Bosse, E., Immel, A., Rinne, C., Kornell, S.-C., Yepes, D., Franzenburg, S., Heyne, H. O., Meier, T., Lösch, S., Meller, H., Friederich, S., Nicklisch, N., Alt, K. W., Schreiber, S., Tholey, A., ... Krause, J. (2018). Neolithic and medieval virus genomes reveal complex evolution of hepatitis B. *ELife*, *7*. https://doi.org/10.7554/eLife.36666

Langmead, B., & Salzberg, S. L. (2012). Fast gapped-read alignment with Bowtie 2. *Nature Methods*, *9*(4), 357–359. https://doi.org/10.1038/nmeth.1923

Lee, O. Y.-C., Wu, H. H. T., Donoghue, H. D., Spigelman, M., Greenblatt, C. L., Bull, I. D., Rothschild, B. M., Martin, L. D., Minnikin, D. E., & Besra, G. S. (2012). *Mycobacterium tuberculosis* complex lipid virulence factors preserved in the 17,000-year-old skeleton of an extinct bison, *Bison antiquus*. *PLoS ONE*, *7*(7), e41923. https://doi.org/10.1371/journal.pone.0041923

Lendvay, B., Hartmann, M., Brodbeck, S., Nievergelt, D., Reinig, F., Zoller, S., Parducci, L., Gugerli, F., Büntgen, U., & Sperisen, C. (2018). Improved recovery of ancient DNA from subfossil wood - Application to the world's oldest Late Glacial pine forest. *New Phytologist*, *217*(4), 1737–1748. https://doi.org/10.1111/nph.14935

Li, H., & Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics*, *25*(14), 1754–1760. https://doi.org/10.1093/bioinformatics/btp324

Lindahl, T. (1993). Instability and decay of the primary structure of DNA. *Nature*, *362*(6422), 709–715. https://doi.org/10.1038/362709a0

Maixner, F., Krause-Kyora, B., Turaev, D., Herbig, A., Hoopmann, M. R., Hallows, J. L., Kusebauch, U., Vigl, E. E., Malfertheiner, P., Megraud, F., O'Sullivan, N., Cipollini, G., Coia, V., Samadelli, M., Engstrand, L., Linz, B., Moritz, R. L., Grimm, R., Krause, J., ... Zink, A. (2016). The 5300-year-old *Helicobacter pylori* genome of the Iceman. *Science*, *351*(6269), 162–165. https://doi.org/10.1126/science.aad2545

Maixner, F., Mitterer, C., Jäger, H. Y., Sarhan, M. S., Valverde, G., Lücker, S., Piombino-Mascali, D., Szikossy, I., Molnár, E., Pálfi, G., Pap, I., Cipollini, G., & Zink, A. (2022). Linear polyacrylamide is highly efficient in precipitating and purifying environmental and ancient DNA. *Methods in Ecology and Evolution*, *13*(3), 653–667. https://doi.org/10.1111/2041-210X.13772

Maixner, F., Sarhan, M. S., Huang, K. D., Tett, A., Schoenafinger, A., Zingale, S., Blanco-Míguez, A., Manghi, P., Cemper-Kiesslich, J., Rosendahl, W., Kusebauch, U., Morrone, S. R., Hoopmann, M. R., Rota-Stabelli, O., Rattei, T., Moritz, R. L., Oeggl, K., Segata, N., Zink, A., ... Kowarik, K. (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*(23), 5149-5162.e6. https://doi.org/10.1016/j.cub.2021.09.031

Majander, K., Pfrengle, S., Kocher, A., Neukamm, J., du Plessis, L., Pla-Díaz, M., Arora, N., Akgül, G., Salo, K., Schats, R., Inskip, S., Oinonen, M., Valk, H., Malve, M., Kriiska, A., Onkamo, P., González-Candelas, F., Kühnert, D., Krause, J., & Schuenemann, V. J. (2020). Ancient bacterial

genomes reveal a high diversity of *Treponema pallidum* strains in early modern Europe. *Current Biology*, *30*(19), 3788-3803.e10. https://doi.org/10.1016/J.CUB.2020.07.058

Marciniak, S., Prowse, T. L., Herring, D. A., Klunk, J., Kuch, M., Duggan, A. T., Bondioli, L., Holmes, E. C., & Poinar, H. N. (2016). *Plasmodium falciparum* malaria in 1st–2nd century CE southern Italy. *Current Biology*, *26*(23), R1220–R1222. https://doi.org/10.1016/J.CUB.2016.10.016

Margulies, M., Egholm, M., Altman, W. E., Attiya, S., Bader, J. S., Bemben, L. A., Berka, J., Braverman, M. S., Chen, Y. J., Chen, Z., Dewell, S. B., Du, L., Fierro, J. M., Gomes, X. v., Godwin, B. C., He, W., Helgesen, S., Ho, C. H., Irzyk, G. P., ... Rothberg, J. M. (2005). Genome sequencing in microfabricated high-density picolitre reactors. *Nature*, *437*(7057), 376–380. https://doi.org/10.1038/nature03959

Massilani, D., Morley, M. W., Mentzer, S. M., Aldeias, V., Vernot, B., Miller, C., Stahlschmidt, M., Kozlikin, M. B., Shunkov, M. v., Derevianko, A. P., Conard, N. J., Wurz, S., Henshilwood, C. S., Vasquez, J., Essel, E., Nagel, S., Richter, J., Nickel, B., Roberts, R. G., ... Meyer, M. (2022). Microstratigraphic preservation of ancient faunal and hominin DNA in Pleistocene cave sediments. *Proceedings of the National Academy of Sciences*, *119*(1). https://doi.org/10.1073/pnas.2113666118

Masson, M., Molnár, E., Donoghue, H. D., Besra, G. S., Minnikin, D. E., Wu, H. H. T., Lee, O. Y.-C., Bull, I. D., & Pálfi, G. (2013). Osteological and biomolecular evidence of a 7,000-year-old case of hypertrophic pulmonary osteopathy secondary to tuberculosis from Neolithic Hungary. *PLoS ONE*, *8*(10). https://doi.org/10.1371/journal.pone.0078252

Mayxay, M., Pukrittayakamee, S., Newton, P. N., & White, N. J. (2004). Mixed-species malaria infections in humans. *Trends in Parasitology*, *20*(5), 233–240. https://doi.org/10.1016/j.pt.2004.03.006

McKeown, T. (1988). *The origins of human disease*. Basil Blackwell.

McNaughton, A. L., Sreenu, V. B., Wilkie, G., Gunson, R., Templeton, K., & Leitch, E. C. M. (2018). Prevalence of mixed genotype hepatitis C virus infections in the UK as determined by genotype-specific PCR and deep sequencing. *Journal of Viral Hepatitis*, *25*(5), 524–534. https://doi.org/10.1111/jvh.12849

McNerney, R., Clark, T. G., Campino, S., Rodrigues, C., Dolinger, D., Smith, L., Cabibbe, A. M., Dheda, K., & Schito, M. (2017). Removing the bottleneck in whole genome sequencing of *Mycobacterium tuberculosis* for rapid drug resistance analysis: A call to action. *International Journal of Infectious Diseases*, *56*, 130–135. https://doi.org/10.1016/j.ijid.2016.11.422

Meyer, M., Briggs, A. W., Maricic, T., Höber, B., Höffner, B., Krause, J., Weihmann, A., Pääbo, S., & Hofreiter, M. (2008). From micrograms to picograms: quantitative PCR reduces the material demands of high-throughput sequencing. *Nucleic Acids Research*, *36*(1), e5–e5. https://doi.org/10.1093/nar/gkm1095

Meyer, M., & Kircher, M. (2010). Illumina sequencing library preparation for highly multiplexed target capture and sequencing. *Cold Spring Harbor Protocols*, *2010*(6). https://doi.org/10.1101/pdb.prot5448

Meyer, M., Kircher, M., Gansauge, M., Li, H., Racimo, F., Mallick, S., Schraiber, J., Jay, F., Prüfer, K., de Filippo, C., Sudmant, P. H., Alkan, C., Fu, Q., Do, R., Rohland, N., Tandon, A., Siebauer, Green, R., Bryc, K., ... Pääbo, S. (2012). A high-coverage genome sequence from an archaic Denisovan individual. *Science*, *338*(6104), 222–226. https://doi.org/10.1126/science.1224344

Miller, W., Drautz, D. I., Ratan, A., Pusey, B., Qi, J., Lesk, A. M., Tomsho, L. P., Packard, M. D., Zhao, F., Sher, A., Tikhonov, A., Raney, B., Patterson, N., Lindblad-Toh, K., Lander, E. S., Knight, J. R., Irzyk, G.

P., Fredrikson, K. M., Harkins, T. T., ... Schuster, S. C. (2008). Sequencing the nuclear genome of the extinct woolly mammoth. *Nature*, *456*(7220), 387–390. https://doi.org/10.1038/nature07446

Moreno-Molina, M., Shubladze, N., Khurtsilava, I., Avaliani, Z., Bablishvili, N., Torres-Puente, M., Villamayor, L., Gabrielian, A., Rosenthal, A., Vilaplana, C., Gagneux, S., Kempker, R. R., Vashakidze, S., & Comas, I. (2021). Genomic analyses of *Mycobacterium tuberculosis* from human lung resections reveal a high frequency of polyclonal infections. *Nature Communications*, *12*(1). https://doi.org/10.1038/s41467-021-22705-z

Mühlemann, B., Vinner, L., Margaryan, A., Wilhelmson, H., de la Fuente Castro, C., Allentoft, M. E., de Barros Damgaard, P., Hansen, A. J., Holtsmark Nielsen, S., Strand, L. M., Bill, J., Buzhilova, A., Pushkina, T., Falys, C., Khartanovich, V., Moiseyev, V., Jørkov, M. L. S., Østergaard Sørensen, P., Magnusson, Y., ... Sikora, M. (2020). Diverse variola virus (smallpox) strains were widespread in northern Europe in the Viking Age. *Science*, *369*(6502). https://doi.org/10.1126/science.aaw8977

Napier, G., Campino, S., Merid, Y., Abebe, M., Woldeamanuel, Y., Aseffa, A., Hibberd, M. L., Phelan, J., & Clark, T. G. (2020). Robust barcoding and identification of *Mycobacterium tuberculosis* lineages for epidemiological and clinical studies. *Genome Medicine*, *12*(1). https://doi.org/10.1186/s13073-020-00817-3

Nebenzahl-Guimaraes, H., Yimer, S. A., Holm-Hansen, C., de Beer, J., Brosch, R., & van Soolingen, D. (2016). Genomic characterization of *Mycobacterium tuberculosis* lineage 7 and a proposed name: 'Aethiops vetus.' *Microbial Genomics*, *2*(6). https://doi.org/10.1099/mgen.0.000063

Neukamm, J., Peltzer, A., & Nieselt, K. (2021). DamageProfiler: fast damage pattern calculation for ancient DNA. *Bioinformatics*, *37*(20), 3652–3653. https://doi.org/10.1093/bioinformatics/btab190

Ngabonziza, J. C. S., Loiseau, C., Marceau, M., Jouet, A., Menardo, F., Tzfadia, O., Antoine, R., Niyigena, E. B., Mulders, W., Fissette, K., Diels, M., Gaudin, C., Duthoy, S., Ssengooba, W., André, E., Kaswa, M. K., Habimana, Y. M., Brites, D., Affolabi, D., ... Supply, P. (2020). A sister lineage of the *Mycobacterium tuberculosis* complex discovered in the African Great Lakes region. *Nature Communications*, *11*(1), 2917. https://doi.org/10.1038/s41467-020-16626-6

Nicklisch, N., Maixner, F., Ganslmeier, R., Friederich, S., Dresely, V., Meller, H., Zink, A., & Alt, K. W. (2012). Rib lesions in skeletons from early neolithic sites in Central Germany: on the trail of tuberculosis at the onset of agriculture. *American Journal of Physical Anthropology*, *149*(3), 391–404. https://doi.org/10.1002/ajpa.22137

Nimmo, C., Shaw, L. P., Doyle, R., Williams, R., Brien, K., Burgess, C., Breuer, J., Balloux, F., & Pym, A. S. (2019). Whole genome sequencing *Mycobacterium tuberculosis* directly from sputum identifies more genetic diversity than sequencing from culture. *BMC Genomics*, *20*(1), 389. https://doi.org/10.1186/s12864-019-5782-2

Orlando, L., Allaby, R., Skoglund, P., der Sarkissian, C., Stockhammer, P. W., Ávila-Arcos, M. C., Fu, Q., Krause, J., Willerslev, E., Stone, A. C., & Warinner, C. (2021). Ancient DNA analysis. *Nature Reviews Methods Primers*, *1*(1), 14. https://doi.org/10.1038/s43586-020-00011-0

Pap, I., Józsa, L., Repa, I., Bajzik, G., Lakhani, S., Donoghue, H., & Spigelman, M. (1999). 18-19th century tuberculosis in naturally mummified individuals (Vác, Hungary). In *Tuberculosis: Past and Present* (pp. 421–428). Golden Book Publisher Ltd.

Patterson Ross, Z., Klunk, J., Fornaciari, G., Giuffra, V., Duchêne, S., Duggan, A. T., Poinar, D., Douglas, M. W., Eden, J. S., Holmes, E. C., & Poinar, H. N. (2018). The paradox of HBV evolution as revealed from

a 16th century mummy. *PLOS Pathogens*, *14*(1), e1006750. https://doi.org/10.1371/JOURNAL.PPAT.1006750

Pinhasi, R., Fernandes, D., Sirak, K., Novak, M., Connell, S., Alpaslan-Roodenberg, S., Gerritsen, F., Moiseyev, V., Gromov, A., Raczky, P., Anders, A., Pietrusewsky, M., Rollefson, G., Jovanovic, M., Trinhhoang, H., Bar-Oz, G., Oxenham, M., Matsumura, H., & Hofreiter, M. (2015). Optimal Ancient DNA Yields from the Inner Ear Part of the Human Petrous Bone. *PLOS ONE*, *10*(6), e0129102. https://doi.org/10.1371/journal.pone.0129102

Rafi, A., Spigelman, M., Stanford, J., Lemma, E., Donoghue, H., & Zias, J. (1994). *Mycobacterium leprae* DNA from ancient bone detected by PCR. *The Lancet*, *343*(8909), 1360–1361. https://doi.org/10.1016/S0140-6736(94)92494-5

Reich, D., Green, R. E., Kircher, M., Krause, J., Patterson, N., Durand, E. Y., Viola, B., Briggs, A. W., Stenzel, U., Johnson, P. L. F., Maricic, T., Good, J. M., Marques-Bonet, T., Alkan, C., Fu, Q., Mallick, S., Li, H., Meyer, M., Eichler, E. E., ... Pääbo, S. (2010). Genetic history of an archaic hominin group from Denisova Cave in Siberia. *Nature*, *468*(7327), 1053–1060. https://doi.org/10.1038/nature09710

Renaud, G., Slon, V., Duggan, A. T., & Kelso, J. (2015). Schmutzi: estimation of contamination and endogenous mitochondrial consensus calling for ancient DNA. *Genome Biology*, *16*(1), 224. https://doi.org/10.1186/s13059-015-0776-0

Roberts, C. (2011). The bioarchaeology of leprosy and tuberculosis. In *Social Bioarchaeology* (pp. 252–281). Wiley-Blackwell. https://doi.org/10.1002/9781444390537.ch9

Rohland, N., Glocke, I., Aximu-Petri, A., & Meyer, M. (2018). Extraction of highly degraded DNA from ancient bones, teeth and sediments for high-

throughput sequencing. *Nature Protocols*, *13*(11), 2447–2461. https://doi.org/10.1038/s41596-018-0050-5

Rohland, N., Harney, E., Mallick, S., Nordenfelt, S., & Reich, D. (2015). Partial uracil–DNA–glycosylase treatment for screening of ancient DNA. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *370*(1660), 20130624. https://doi.org/10.1098/rstb.2013.0624

Rohland, N., & Hofreiter, M. (2007). Ancient DNA extraction from bones and teeth. *Nature Protocols*, *2*(7), 1756–1762. https://doi.org/10.1038/nprot.2007.247

Rothschild, B. M., Martin, L. D., Lev, G., Bercovier, H., Bar-Gal, G. K., Greenblatt, C., Donoghue, H., Spigelman, M., & Brittain, D. (2001). *Mycobacterium tuberculosis* complex DNA from an extinct bison dated 17,000 years before the present. *Clinical Infectious Diseases*, *33*(3), 305–311. https://doi.org/10.1086/321886

Sabin, S., Herbig, A., Vågene, Å. J., Ahlström, T., Bozovic, G., Arcini, C., Kühnert, D., & Bos, K. I. (2020). A seventeenth-century *Mycobacterium tuberculosis* genome supports a Neolithic emergence of the *Mycobacterium tuberculosis* complex. *Genome Biology*, *21*(1). https://doi.org/10.1186/s13059-020-02112-1

Salo, W. L., Aufderheide, A. C., Buikstra, J., & Holcomb, T. A. (1994). Identification of *Mycobacterium tuberculosis* DNA in a pre-Columbian Peruvian mummy. *Proceedings of the National Academy of Sciences*, *91*(6), 2091–2094. https://doi.org/10.1073/pnas.91.6.2091

Sarhan, M. S., Lehmkuhl, A., Straub, R., Tett, A., Wieland, G., Francken, M., Zink, A., & Maixner, F. (2021). Ancient DNA diffuses from human bones to cave stones. *IScience*, *24*(12), 103397. https://doi.org/10.1016/j.isci.2021.103397

Schuenemann, V. J., Avanzi, C., Krause-Kyora, B., Seitz, A., Herbig, A., Inskip, S., Bonazzi, M., Reiter, E., Urban, C., Dangvard Pedersen, D., Taylor, G. M., Singh, P., Stewart, G. R., Velemínský, P., Likovsky, J., Marcsik, A., Molnár, E., Pálfi, G., Mariotti, V., ... Krause, J. (2018). Ancient genomes reveal a high diversity of *Mycobacterium leprae* in medieval Europe. *PLOS Pathogens*, *14*(5). https://doi.org/10.1371/journal.ppat.1006997

Schuenemann, V. J., Kumar Lankapalli, A., Barquera, R., Nelson, E. A., Iraíz Hernández, D., Acuña Alonzo, V., Bos, K. I., Márquez Morfín, L., Herbig, A., & Krause, J. (2018). Historic *Treponema pallidum* genomes from Colonial Mexico retrieved from archaeological remains. *PLOS Neglected Tropical Diseases*, *12*(6), e0006447. https://doi.org/10.1371/journal.pntd.0006447

Shin, S. H., Jhun, B. W., Kim, S.-Y., Choe, J., Jeon, K., Huh, H. J., Ki, C.-S., Lee, N. Y., Shin, S. J., Daley, C. L., & Koh, W.-J. (2018). Nontuberculous mycobacterial lung diseases caused by mixed infection with *Mycobacterium avium* complex and *Mycobacterium abscessus* complex. *Antimicrobial Agents and Chemotherapy*, *62*(10). https://doi.org/10.1128/AAC.01105-18

Shin, S. S., Modongo, C., Ncube, R., Sepako, E., Klausner, J. D., & Zetola, N. M. (2015). Advanced immune suppression is associated with increased prevalence of mixed-strain *Mycobacterium tuberculosis* infections among persons at high risk for drug-resistant tuberculosis in Botswana. *Journal of Infectious Diseases*, *211*(3), 347–351. https://doi.org/10.1093/infdis/jiu421

Sidstedt, M., Rådström, P., & Hedman, J. (2020). PCR inhibition in qPCR, dPCR and MPS—mechanisms and solutions. *Analytical and Bioanalytical Chemistry*, *412*(9), 2009–2023. https://doi.org/10.1007/s00216-020-02490-2

Sirak, K., Fernandes, D., Cheronet, O., Harney, E., Mah, M., Mallick, S., Rohland, N., Adamski, N., Broomandkhoshbacht, N., Callan, K., Candilio, F., Lawson, A. M., Mandl, K., Oppenheimer, J., Stewardson, K., Zalzala, F., Anders, A., Bartík, J., Coppa, A., ... Pinhasi, R. (2020). Human auditory ossicles as an alternative optimal source of ancient DNA. *Genome Research*, *30*(3), 427–436. https://doi.org/10.1101/gr.260141.119

Skoglund, P., Malmström, H., Raghavan, M., Storå, J., Hall, P., Willerslev, E., Gilbert, M. T. P., Götherström, A., & Jakobsson, M. (2012). Origins and genetic legacy of Neolithic farmers and hunter-gatherers in Europe. *Science*, *336*(6080), 466–469. https://doi.org/10.1126/science.1216304

Sobkowiak, B., Glynn, J. R., Houben, R. M. G. J., Mallard, K., Phelan, J. E., Guerra-Assunção, J. A., Banda, L., Mzembe, T., Viveiros, M., McNerney, R., Parkhill, J., Crampin, A. C., & Clark, T. G. (2018). Identifying mixed *Mycobacterium tuberculosis* infections from whole genome sequence data. *BMC Genomics*, *19*(1). https://doi.org/10.1186/s12864-018-4988-z

Somerville, W., Thibert, L., Schwartzman, K., & Behr, M. A. (2005). Extraction of *Mycobacterium tuberculosis* DNA: a question of containment. *Journal of Clinical Microbiology*, *43*(6), 2996–2997. https://doi.org/10.1128/JCM.43.6.2996-2997.2005

Spigelman, M., & Lemma, E. (1993). The use of the polymerase chain reaction (PCR) to detect *Mycobacterium tuberculosis* in ancient skeletons. *International Journal of Osteoarchaeology*, *3*(2), 137–143. https://doi.org/10.1002/oa.1390030211

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

Spyrou, M. A., Keller, M., Tukhbatova, R. I., Scheib, C. L., Nelson, E. A., Andrades Valtueña, A., Neumann, G. U., Walker, D., Alterauge, A., Carty, N., Cessford, C., Fetz, H., Gourvennec, M., Hartle, R., Henderson, M., von Heyking, K., Inskip, S. A., Kacki, S., Key, F. M., ... Krause, J. (2019). Phylogeography of the second plague pandemic revealed through analysis of historical *Yersinia pestis* genomes. *Nature Communications*, *10*(1). https://doi.org/10.1038/s41467-019-12154-0

Spyrou, M. A., Tukhbatova, R. I., Feldman, M., Drath, J., Kacki, S., Beltrán De Heredia, J., Arnold, S., Sitdikov, A. G., Castex, D., Wahl, J., Gazimzyanov, I. R., Nurgaliev, D. K., Herbig, A., Bos, K. I., & Krause, J. (2016). Historical *Y. pestis* genomes reveal the European Black Death as the source of ancient and modern plague pandemics. *Cell Host and Microbe*, *19*(6), 874–881. https://doi.org/10.1016/J.CHOM.2016.05.012

Stucki, D., Brites, D., Jeljeli, L., Coscolla, M., Liu, Q., Trauner, A., Fenner, L., Rutaihwa, L., Borrell, S., Luo, T., Gao, Q., Kato-Maeda, M., Ballif, M., Egger, M., Macedo, R., Mardassi, H., Moreno, M., Vilanova, G. T., Fyfe, J., ... Gagneux, S. (2016). *Mycobacterium tuberculosis* lineage 4 comprises globally distributed and geographically restricted sublineages. *Nature Genetics*, *48*(12), 1535–1543. https://doi.org/10.1038/ng.3704

Supply, P., Marceau, M., Mangenot, S., Roche, D., Rouanet, C., Khanna, V., Majlessi, L., Criscuolo, A., Tap, J., Pawlik, A., Fiette, L., Orgeur, M., Fabre, M., Parmentier, C., Frigui, W., Simeone, R., Boritsch, E. C., Debrie, A.-S., Willery, E., ... Brosch, R. (2013). Genomic analysis of smooth tubercle bacilli provides insights into ancestry and pathoadaptation of *Mycobacterium tuberculosis*. *Nature Genetics*, *45*(2), 172–179. https://doi.org/10.1038/ng.2517

Szikossy I, Bernert Z, & Pap I. (1997). Anthropological investigation of the 18th-19th century ossuary of the Dominican Church at Vác, Hungary. *Acta Biologica Szegediensis*, *42*, 145–150.

Thomas, J., Balseiro, A., Gortázar, C., & Risalde, M. A. (2021). Diagnosis of tuberculosis in wildlife: a systematic review. *Veterinary Research*, *52*(1), 31. https://doi.org/10.1186/s13567-020-00881-y

To, K., Cao, R., Yegiazaryan, A., Owens, J., & Venketaraman, V. (2020). General overview of nontuberculous Mycobacteria opportunistic pathogens: *Mycobacterium avium* and *Mycobacterium abscessus*. *Journal of Clinical Medicine*, *9*(8), 2541. https://doi.org/10.3390/jcm9082541

Tran, T.-N.-N., Signoli, M., Fozzati, L., Aboudharam, G., Raoult, D., & Drancourt, M. (2011). High throughput, multiplexed pathogen detection authenticates plague waves in Medieval Venice, Italy. *PLoS ONE*, *6*(3), e16735. https://doi.org/10.1371/journal.pone.0016735

Vågene, Å. J., Herbig, A., Campana, M. G., Robles García, N. M., Warinner, C., Sabin, S., Spyrou, M. A., Andrades Valtueña, A., Huson, D., Tuross, N., Bos, K. I., & Krause, J. (2018). *Salmonella enterica* genomes from victims of a major sixteenth-century epidemic in Mexico. *Nature Ecology & Evolution*, *2*(3), 520–528. https://doi.org/10.1038/s41559-017-0446-6

Vågene, Å. J., Honap, T. P., Harkins, K. M., Rosenberg, M. S., Giffin, K., Cárdenas-Arroyo, F., Leguizamón, L. P., Arnett, J., Buikstra, J. E., Herbig, A., Krause, J., Stone, A. C., & Bos, K. I. (2022). Geographically dispersed zoonotic tuberculosis in pre-contact South American human populations. *Nature Communications*, *13*(1), 1195. https://doi.org/10.1038/s41467-022-28562-8

VanderVen, B. C., Huang, L., Rohde, K. H., & Russell, D. G. (2016). The minimal unit of infection: *Mycobacterium tuberculosis* in the macrophage. *Microbiology Spectrum*, *4*(6). https://doi.org/10.1128/microbiolspec.TBTB2-0025-2016

von Both, U., Berk, M., Agapow, P.-M., Wright, J. D., Git, A., Hamilton, M. S., Goldgof, G., Siddiqui, N., Bellos, E., Wright, V. J., Coin, L. J., Newton, S. M., & Levin, M. (2018). *Mycobacterium tuberculosis* exploits

a molecular off switch of the immune system for intracellular survival. *Scientific Reports*, *8*(1), 661. https://doi.org/10.1038/s41598-017-18528-y

Wagner, D. M., Klunk, J., Harbeck, M., Devault, A., Waglechner, N., Sahl, J. W., Enk, J., Birdsell, D. N., Kuch, M., Lumibao, C., Poinar, D., Pearson, T., Fourment, M., Golding, B., Riehm, J. M., Earn, D. J. D., DeWitte, S., Rouillard, J. M., Grupe, G., ... Poinar, H. (2014). *Yersinia pestis* and the plague of Justinian 541-543 AD: a genomic analysis. *The Lancet Infectious Diseases*, *14*(4), 319–326. https://doi.org/10.1016/S1473-3099(13)70323-2

Wagner, J. K., Colwell, C., Claw, K. G., Stone, A. C., Bolnick, D. A., Hawks, J., Brothers, K. B., & Garrison, N. A. (2020). Fostering responsible research on ancient DNA. *The American Journal of Human Genetics*, *107*(2), 183–195. https://doi.org/10.1016/J.AJHG.2020.06.017

Wagner, S., Lagane, F., Seguin-Orlando, A., Schubert, M., Leroy, T., Guichoux, E., Chancerel, E., Bech-Hebelstrup, I., Bernard, V., Billard, C., Billaud, Y., Bolliger, M., Croutsch, C., Čufar, K., Eynaud, F., Heussner, K. U., Köninger, J., Langenegger, F., Leroy, F., ... Orlando, L. (2018). High-throughput DNA sequencing of ancient wood. *Molecular Ecology*, *27*(5), 1138–1154. https://doi.org/10.1111/mec.14514

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

Warren, R. M., Victor, T. C., Streicher, E. M., Richardson, M., Beyers, N., van Pittius, N. C. G., & van Helden, P. D. (2004). Patients with active tuberculosis often have different strains in the same sputum specimen. *American Journal of Respiratory and Critical Care Medicine*, *169*(5), 610–614. https://doi.org/10.1164/rccm.200305-714OC

Weyrich, L. S., Duchene, S., Soubrier, J., Arriola, L., Llamas, B., Breen, J., Morris, A. G., Alt, K. W., Caramelli, D., Dresely, V., Farrell, M., Farrer, A. G., Francken, M., Gully, N., Haak, W., Hardy, K., Harvati, K., Held, P., Holmes, E. C., ... Cooper, A. (2017). Neanderthal behaviour, diet, and disease inferred from ancient DNA in dental calculus. *Nature*, *544*(7650), 357–361. https://doi.org/10.1038/nature21674

WHO. (2021). *Global tuberculosis report 2021*. https://www.who.int/publications/i/item/9789240037021

Wilbur, A. K., Bouwman, A. S., Stone, A. C., Roberts, C. A., Pfister, L.-A., Buikstra, J. E., & Brown, T. A. (2009). Deficiencies and challenges in the study of ancient tuberculosis DNA. *Journal of Archaeological Science*, *36*(9), 1990–1997. https://doi.org/10.1016/j.jas.2009.05.020

Xiao, T., & Zhou, W. (2020). The third generation sequencing: the advanced approach to genetic diseases. *Translational Pediatrics*, *9*(2), 163–173. https://doi.org/10.21037/tp.2020.03.06

Zetola, N. M., Shin, S. S., Tumedi, K. A., Moeti, K., Ncube, R., Nicol, M., Collman, R. G., Klausner, J. D., & Modongo, C. (2014). Mixed *Mycobacterium tuberculosis* complex infections and false-negative results for rifampin resistance by GeneXpert MTB/RIF Aae associated with poor clinical outcomes. *Journal of Clinical Microbiology*, *52*(7), 2422–2429. https://doi.org/10.1128/JCM.02489-13

Zink, A. R., Molnár, E., Motamedi, N., Pálfy, G., Marcsik, A., & Nerlich, A. G. (2007). Molecular history of tuberculosis from ancient mummies and skeletons. *International Journal of Osteoarchaeology*, *17*(4). https://doi.org/10.1002/oa.909

7 | Acknowledgement

First of all, I would like to thank Dr. Albert Zink, the head of the Eurac Research Institute for Mummy Studies and my formal supervisor at LMU, for allowing me to work at the institute for the past four years. He introduced me to this institute in 2017 when I was working as his teaching assistant. Little did I know that I would soon embark on this PhD journey, especially considering that I could not even point at Bolzano on a map at that time. Thank you, Albert, for always creating a positive atmosphere and making everyone feel at home.

I am deeply grateful for having crossed paths with my supervisor Dr. Frank Maixner. You have sincerely been nothing but the most excellent mentor to me throughout the past years. Since the beginning, you have always given me the space and time to simply explore and grow. Whenever I needed advice, I have always felt listened to, and you never failed to help me find new inspiration. Frank, your trust, support, encouragement, and guidance are the foundation of my success, and I thank you for that. I will always think back to our hour-long talks about everything and nothing with a smile on my face.

I also want to express my gratitude to all my colleagues and friends at the Institute for Mummy Studies. I have learned so much from each one of you and I feel honored to have worked in such a unique team. Thank you, Renate, Manu, Stefania, Christina, Lena, Alex, Sabry, Alice, Vale, Giovanna, Myriam, Guido, and Marco for your constant support, your pep talks, and most importantly our coffee and lunch breaks.

My special thanks go to Christina Wurst, who has supported me since the first time I came to the institute as a master student in 2017. She was a fantastic mentor back then and will always be my number one emergency contact for anything lab-related.

Lastly, I am sending a big, warm hug to each of my loved ones. To the ones that have been there since the very beginning, the ones that have stuck with me, and those I was fortunate enough to meet along the way. It has indeed been **something** that we have been through, not just inside academia. Your support, patience, love, and of course, all the cups of coffee and home-cooked meals were my driving force to keep going – you make it all worth it.

Tausend Dank, mille grazie, a thousand thanks...

an Tim und die Familie Reska

a Mattia e alla famiglia Secchi

a Pesco e alla famiglia Barbieri

a Nico e Giulia

to Eric, Tessa and Storm

an Franzi und Clarisse

an Nazza, Lukas und Michi

...and finally, of course **Atlas**

ママとパパ、聖治と孝治、いつも支えてくれてありがとうございます。

Cari Carlotta, Mara e Francesco, voglio ringraziarvi di cuore per il vostro continuo sostegno, le parole di incoraggiamento, la pazienza, tutti i pasti, per avermi dato un posto dove sentirmi al sicuro e a casa e soprattutto per avermi fatto mangiare abbastanza insalata. Grazie, grazie e grazie!

♥

8 | Appendices

8.1 | Statement of own contributions

Chapter I **Jäger, H.Y.**, Atz Zanotelli, D., Maixner F., Nicklisch N., Alt, K.W., Meller, H., ... Zink, A.R. (2023). Hit or miss - A metagenomic evaluation of intra-bone variability of host pathogen load in tuberculosis-infected human remains. *Tuberculosis*, *143*, 102392.

I conceptualized the idea and designed the methodology, managed the project and co-supervised, conducted the molecular investigations (sampling, DNA extraction, library, and sequencing preparation), performed bioinformatical analyses of the sequencing data, designed and created the figures, and led the writing of the manuscript.

Chapter II Maixner, F., Mitterer, C., **Jäger, H. Y.**, Sarhan, M. S., Valverde, G., Lücker, S., ... Zink, A. (2022). Linear polyacrylamide is highly efficient in precipitating and purifying environmental and ancient DNA. *Methods in Ecology and Evolution*, 13(3), 653–667.

I conducted significant parts of the molecular investigation (sampling, DNA extraction, library, and sequencing preparation), the bioinformatical analyses of the sequencing data, and contributed to the writing of the manuscript.

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

I conceptualized the idea and designed the methodology, managed the project, conducted the molecular investigation (sampling, DNA extraction, library, and sequencing preparation), developed the bioinformatical pipeline, analyzed the sequencing data, designed, and created the figures, and led the writing of the manuscript.

8.2 | List of peer-reviewed publications

Jäger, H.Y., Atz Zanotelli, D., Maixner F., Nicklisch N., Alt, K.W., Meller, H., ... Zink, A.R. (2023). Hit or miss - A metagenomic evaluation of intra-bone variability of host pathogen load in tuberculosis-infected human remains. *Tuberculosis*, *143*, 102392. https://doi.org/10.1016/j.tube.2023.102392

Lee, O. Y.-C., Wu, H. H. T., Besra, G. S., Minnikin, D. E., **Jäger, H. Y.**, Maixner, F., ... Pálfi, G. (2023). Sensitive lipid biomarker detection for tuberculosis in late Neanderthal skeletons from Subalyuk Cave, Hungary. *Tuberculosis*, *143*, 102420. https://doi.org/10.1016/j.tube.2023.102420

Zink, A., Maixner, F., **Jäger, H. Y**., Szikossy, I., Pálfi, G., & Pap, I. (2023). Tuberculosis in mummies – new findings, perspectives, and limitations. *Tuberculosis*, *143*, 102371. https://doi.org/10.1016/j.tube.2023.102371

Spekker, O., Váradi, O. A., Szekeres, A., **Jäger, H. Y.**, Zink, A., Berner, M., ... Tihanyi, B. (2022). A rare case of calvarial tuberculosis from the Avar Age (8th century CE) cemetery of Kaba–Bitózug (Hajdú-Bihar County, Hungary) – Pathogenesis and differential diagnostic aspects. *Tuberculosis*, 135, 102226. doi: 10.1016/j.tube.2022.102226

Dominelli, N., **Jäger, H. Y.**, Langer, A., Brachmann, A., & Heermann, R. (2022). High-throughput sequencing analysis reveals genomic similarity in phenotypic heterogeneous *Photorhabdus luminescens* cell populations. Annals of Microbiology, 72(1), 20. doi: 10.1186/s13213-022-01677-5

Jäger, H. Y., Maixner, F., Pap, I., Szikossy, I., Pálfi, G., & Zink, A. R. (2022). Metagenomic analysis reveals mixed *Mycobacterium tuberculosis* infection in a 18th century Hungarian midwife. *Tuberculosis*, 102181. doi: 10.1016/j.tube.2022.102181

Maixner, F., Mitterer, C., **Jäger, H. Y.**, Sarhan, M. S., Valverde, G., Lücker, S., ... Zink, A. (2022). Linear polyacrylamide is highly efficient in precipitating

and purifying environmental and ancient DNA. *Methods in Ecology and Evolution*, 13(3), 653–667. doi: 10.1111/2041-210X.13772

Loufouma Mbouaka, A., Gamble, M., Wurst, C., **Jäger, H. Y.**, Maixner, F., Zink, A., ... Binder, M. (2021). The elusive parasite: comparing macroscopic, immunological, and genomic approaches to identifying malaria in human skeletal remains from Sayala, Egypt (third to sixth centuries AD). *Archaeological and Anthropological Sciences*, 13(7), 115. doi: 10.1007/s12520-021-01350-z

Jäger, H. Y., Han, C. S., & Dingemanse, N. J. (2019). Social experiences shape behavioral individuality and within-individual stability. *Behavioral Ecology*, 30(4), 1012–1019. doi: 10.1093/beheco/arz042

Han, C. S., **Jäger, H. Y.**, & Dingemanse, N. J. (2016). Individuality in nutritional preferences: a multi-level approach in field crickets. *Scientific Reports*, 6(1), 29071. doi: 10.1038/srep29071

8.3 | Congress contributions

Jäger, H.Y., Maixner, F., Zink, A. (2022) Chasing ancient pathogens: A look into the ancient tuberculosis research of the Vác Mummy Collection. *10th World Congress on Mummy Studies, Bolzano (Italy).* Poster presentation.

Jäger, H.Y., Maixner, F., Zink, A. (2022) A leap in the dark – How intra-bone variability of pathogen load affects metagenomic detection in ancient *Mycobacterium tuberculosis* DNA in skeletal human remains. *ICEPT-3 TB evolution meeting, Szeged (Hungary).* Talk, invited speaker.

Zink, A., **Jäger, H.Y.**, Valverde, G., Pap, I., Szikossy, I., Pálfi, G., Maixner, F. (2022). Tuberculosis in mummies – New findings, perspectives, and limitations. *ICEPT-3 TB evolution meeting, Szeged (Hungary).* Talk, contributing author.

Maixner, F., **Jäger, H.Y.**, Sarhan, M.S., Valverde, G., Zink, A. (2022) *Mycobacterium tuberculosis* DNA extraction from mummified and skeletal human remains using linear polyacrylamide (LPA). *ICEPT-3 TB evolution meeting, Szeged (Hungary).* Talk, contributing author.

Spekker, O., Váradi O.A., Szekeres, A., **Jäger, H.Y.**, Zink, A., Berner, M., ... Tihanyi, B. (2022) The first reported archaeological case with calvarial tuberculosis from the present-day territory of Hungary. *ICEPT-3 TB evolution meeting, Szeged (Hungary).* Talk, contributing author.

Lee, O., Wu, H., Minnikin, D, Besra, G.S., Llewellyn, G., **Jäger H.Y.**, ... Pálfi, G. (2022) Sensitive detection of mycobacterial cell wall lipid biomarkers in Neanderthal skeletons from Subalyuk, Hungary. *ICEPT-3 TB evolution meeting, Szeged (Hungary).* Talk, contributing author.

Váradi, O.A., Spekker, O., **Jäger, H.Y.**, Maixner, F., Zink, A., Rakk, D., ... Minnikin D. (2022) The development, verification and objectives of a mycocerosic acid based HPLC-HRMS method for TB diagnostics in

paleopathology. *ICEPT-3 TB evolution meeting, Szeged (Hungary).* Talk, contributing author.

Minnikin D., Lee, O., Wu, H., Llwellyn, G., Williams, C., **Jäger, H.Y.**, ... Zink, A. (2020) Lipid biomarkers for tuberculosis are present in Neanderthal skeletal remains from Subalyuk, Hungary. *26th Annual Meeting of the European Association of Archaeologists, Budapest (Hungary)* 2020. Talk, contributing author.

Jäger, H.Y., Maixner, F., Zink, A. (2019) Design and evaluation of a targeted enrichment assay capturing *Mycobacterium tuberculosis* complex DNA in ancient human remains. *4th International Conference on Clinical Metagenomics, Genève (Switzerland).* Talk, speaker.

Jäger, H.Y., Maixner, F., Zink, A. (2019) Design and evaluation of a hybridization capture assay targeting *Mycobacterium tuberculosis* complex DNA in ancient human remains. *Health and Life in Ancient Egypt – Mummies in Focus, Budapest (Hungary).* Talk, invited speaker.

Jäger, H.Y., Bergmeier, F.S., Haszprunar, G., Jörger, K.M. (2017) The first *Pruvotina* (Mollusca, Solenogastres) from the Pacific - a lost loner in the depths of the Sea of Okhotsk? *4th International Congress on Invertebrate Morphology, Moscow (Russia).* Poster presentation.