

Ancient DNA investigation into infectious diseases of medieval Europe

Dissertation

in fulfillment of the requirements for the degree “*Dr. rer. nat.*”
of the Faculty of Mathematics and Natural Sciences at Kiel University

submitted by

Joanna Helena Bonczarowska

Kiel, 2022

First examiner: Prof. Dr. Manuela Dittmar
Second examiner: Prof. Dr. Almut Nebel

Date of oral examination: 28.11.2022

Approved for print on: _____

Approved by: _____
Prof. Dr. Frank Kempken, Dean

TABLE OF CONTENTS

Summary.....	3
Zusammenfassung.....	5
Introduction.....	7
Aims of the dissertation.....	23
Material & Methods.....	37
PART I: PATHOGEN-DRIVEN EVOLUTION OF HUMAN IMMUNITY.....	41
Chapter 1: Ancient DNA study in medieval Europeans shows an association between HLA-DRB1*03 and paratyphoid fever (published).....	43
Chapter 2: Human Leukocyte Antigens (HLA) and leprosy - a genetic association study in medieval Europeans.....	53
PART II: PATHOGEN EVOLUTION.....	105
Chapter 3: Pathogen genomics study of an early medieval community in Germany reveals extensive co-infections (published).....	107
Chapter 4: <i>Yersinia pestis</i> strains from Latvia show depletion of the <i>pla</i> virulence gene at the end of the second plague pandemic (published).....	125
Chapter 5: <i>Yersinia pestis</i> responsible for ancient plague outbreaks lacked zonula occludens toxin (Zot) present in strains of the modern pandemic (submitted).....	137
Chapter 6: Genomic and phylogenetic analysis of medieval <i>Yersinia pestis</i>	161
Discussion.....	179
Declaration.....	203
Acknowledgements.....	205

Supplementary material can be found online as specified in each chapter.

Summary

From wide-spread epidemics to sporadic infections, transmissible diseases have been a major burden to human populations around the globe, posing an immense selective pressure on our genomes. Many infectious diseases responsible for epidemic outbreaks in the Middle Ages are virtually absent in Europe today. This hampers the study of genetic factors associated with infection susceptibility in Europeans. Moreover, the constant evolution of pathogens, exacerbated by the ongoing climate change, poses a major threat to human health. It is therefore of high importance to detect and analyze genomic evolutionary changes related to pathogen virulence that may contribute to disease outbreaks or disappearance. In my dissertation, I applied ancient genomics to investigate the pathogen diversity and infectious disease burden in medieval Europe as well as to identify Human Leukocyte Antigens (HLA) alleles involved in predisposition to leprosy and paratyphoid fever in medieval Europeans. To this end, I extracted DNA from over 300 medieval human remains, performed an osteological analysis of a subset of individuals and used various bioinformatic tools to analyze pathogen and human genomic data from over 500 samples. Chapter 1 presents the first study on the genetic predisposition to *Salmonella* infection in Europeans and the only association analysis on paratyphoid fever C, which resulted in identification of the HLA-DRB1*03 as a risk factor in the European population (published in *Frontiers in Immunology*). I investigated complete HLA profiles of 300 medieval Danish and German individuals in a case-control association study on leprosy in chapter 2. My research resulted in confirmation of two associations previously made for non-Europeans and identification of three potentially novel variants. The systematic screening of the early medieval Lauchheim settlement in chapter 3 allowed me to estimate the infection prevalence at the time, providing the first evidence of the high infectious disease burden (crude period prevalence = 31.4%) in the Early Middle Ages. Seven genomes belonging to four pathogen species (*Mycobacterium leprae* (*M. leprae*), variola virus (VARV), parvovirus B19 and hepatitis B virus (HBV)) were reconstructed and analyzed. The results shed light onto the adaptation of VARV to the human host, spread of *M. leprae* in the Middle Ages and spatio-temporal changes in the distribution of HBV genotypes (published in *Genome Biology*). The genomic analysis of medieval *Yersinia pestis* (*Y. pestis*) strains presented in chapter 4 showed the depletion of the virulence-associated *pla* gene exclusively in isolates dating to the end of the Medieval pandemic. This change might have contributed to the disappearance of plague

from Europe in the 18th cent. AD (published in Scientific Reports). In the following chapter 5, I performed a comparative analysis between ancient, medieval and modern *Y. pestis* genomes, which demonstrated that only a subset of modern strains carried a chromosomal phage encoding the zonula occludens toxin– a known virulence factor. The phage-positive strains were associated with a wide geographic spread and a large proportion of human infections, including the Modern pandemic, which could suggest that the phage acquisition enhanced the transmission of these strains. Moreover, I detected homologous sequences in several *Enterobacteria*, which may point to ongoing evolutionary changes towards increased virulence. In chapter 6, I provided evidence that plague victims were not exclusively buried in mass graves in medieval Denmark, but also at parish cemeteries (4-7%). The findings concerning the infectious disease epidemiology in the Middle Ages as well as both the pathogen and the host genetic factors involved in the disease progression contribute to a better understanding of the host-pathogen coevolution and the genetic implications of past epidemics on our health today.

Zusammenfassung

Von weitreichenden Epidemien bis hin zu sporadischen Infektionen - übertragbare Krankheiten stellten schon immer eine große Belastung für die Menschen auf der ganzen Welt dar und übten zudem einen immensen Selektionsdruck auf unser Genom aus. Viele Infektionskrankheiten, die noch im Mittelalter epidemische Ausmaße erreichten, gibt es heute in Europa praktisch nicht mehr. Dies erschwert die Erforschung der genetischen Suszeptibilitätsfaktoren von Europäern für diese Infektionskrankheiten. Darüber hinaus stellt die kontinuierliche Anpassung von Krankheitserregern, welche durch den Klimawandel noch verschärft wird, eine große Gefahr für die menschliche Gesundheit dar. Es ist daher von großer Relevanz, die genetischen Veränderungen, die mit der Virulenz von Pathogenen oder generell dem Ausbruch oder Verschwinden von Krankheiten in Zusammenhang stehen, zu erkennen und zu analysieren. In meiner Dissertation habe ich daher genomische Datensätze von Menschen aus dem Mittelalter genutzt, um die Erregervielfalt und die Belastung durch Infektionskrankheiten im mittelalterlichen Europa zu untersuchen. Ein weiteres Ziel war es, das menschliche Leukozytenantigensystem (HLA) der mittelalterlichen Individuen zu untersuchen, um darin mögliche genetische Prädispositionen für Lepra und Paratyphus zu identifizieren. Hierfür habe ich DNA aus über 300 mittelalterlichen menschlichen Überresten extrahiert, führte bei einer Untergruppe der Individuen eine osteologische Analyse durch und verwendete verschiedene bioinformatische Verfahren zur Analyse von Erreger- und Humangenomdaten aus insgesamt über 500 Proben. In Kapitel 1 wird die erste Studie zur genetischen Prädisposition für Salmonelleninfektionen bei Europäern und die bislang einzige Assoziationsanalyse zu Paratyphus C vorgestellt, in der HLA-DRB1*03 als Risikofaktor in der europäischen Bevölkerung identifiziert wurde (veröffentlicht in *Frontiers in Immunology*). In Kapitel 2 untersuchte ich im Rahmen einer Fall-Kontroll-Assoziationsstudie zu Lepra vollständige HLA-Profile von 300 mittelalterlichen dänischen und deutschen Individuen. Durch meine Forschung konnten zwei zuvor für Nichteuropäer publizierte Assoziationen bestätigt und drei neue potenziell assoziierte HLA-Allele detektiert werden. Ein systematisches Pathogenscreening der Individuen aus der frühmittelalterlichen Siedlung Lauchheim (Kapitel 3) ermöglichte es mir, die damalige Infektionsprävalenz abzuschätzen, welches den ersten Nachweis für eine hohe Belastung durch Infektionskrankheiten (rohe Periodenprävalenz = 31,4 %) im Frühmittelalter darstellt. Insgesamt habe ich sieben Genome von vier Erregern

(*Mycobacterium leprae* (*M. leprae*), Variola-Virus (VARV), Parvovirus B19 und Hepatitis-B-Virus (HBV)) rekonstruiert und analysiert. Die Ergebnisse geben Aufschluss über die Anpassung von VARV an den menschlichen Wirt, die Ausbreitung von *M. leprae* im Mittelalter und die räumlich-zeitlichen Veränderungen in der Verteilung der HBV-Genotypen (veröffentlicht in Genome Biology). Die in Kapitel 4 vorgestellte Genomanalyse mittelalterlicher *Yersinia pestis* (*Y. pestis*)-Stämme hat gezeigt, dass das Virulenz-assoziierte *pla*-Gen in Isolaten aus dem Ende der mittelalterlichen Pandemie abgenommen hat. Diese Veränderung könnte zum Verschwinden der Pest aus Europa im 18. Jh. n. Chr. beigetragen haben (veröffentlicht in Scientific Reports). In Kapitel 5 führte ich eine vergleichende Analyse zwischen prähistorischen, mittelalterlichen und modernen *Y. pestis*-Genomen durch. Die Analysen zeigten, dass ein heutiges *Y. pestis*-Biovar im Chromosom Phagen-DNA enthält, welches unter anderem für das Zonula occludens-Toxin, einen bekannten Virulenzfaktor, kodiert. Die Phagen-positiven Stämme zeichnen sich durch eine weite geografische Verbreitung aus und sind für einen großen Anteil der modernen menschlichen Pestinfektionen, einschließlich der letzten Pandemie im 19. Jh., verantwortlich. Dies deutet darauf hin, dass die Aufnahme der Phagen-DNA eine leichtere Übertragung dieser Stämme begünstigte. Darüber hinaus habe ich homologe Sequenzen zu mehreren *Enterobacteriaceae* entdeckt, welches ein Hinweis für eine fortlaufende Veränderung zugunsten einer erhöhten Virulenz ist. In Kapitel 6 zeige ich, dass Pestopfer im mittelalterlichen Dänemark häufig auch regulär auf Pfarrfriedhöfen (4-7 %) beigesetzt wurden und nicht, wie bislang allgemein angenommen, ausschließlich in Massengräbern. Die Erkenntnisse über die Epidemiologie der Infektionskrankheiten im Mittelalter sowie über die am Krankheitsverlauf beteiligten genetischen Faktoren von Erreger und Wirt tragen zu einem besseren Verständnis der Koevolution von Wirt und Erreger bei und zeigen die Auswirkungen vergangener Epidemien auf unsere heutige Gesundheit auf.

Introduction

Infectious disease research – insights from ancient genomics

Pathogens are thought to have accompanied human kind since its origin. Until recently, our knowledge on infectious disease outbreaks in the past was based on written and pictorial historical sources (e.g., Fox & Karp, 1988), paleopathological analyses (Ortner, 2003) and presence of mass burials (e.g., Haller et al., 2021; Willmot et al., 2020) which were utilized in difficult times of raging epidemics. However, due to the inherent personal, cultural and linguistic biases in the descriptions and depictions of past people and events (McCullagh, 2002), such sources may represent a speculative evidence of a specific illness. Osteological analysis, although can help in this regard, does not remain without limitations. Firstly, many infectious diseases do not manifest in the skeleton. Secondly, diseases which leave traces of infection on bone often do it in a nonspecific manner and despite evidence of a systemic infection, in the form of periostitis or osteomyelitis, a differential diagnosis is impossible (Ortner, 2003). Matters are further complicated in instances where the skeletal assemblage is fragmentary and disease-specific patterns of lesions cannot be discerned. Moreover, the application of modern criteria for diagnosis assumes that the pathogenesis and clinical manifestation of the disease has not changed over time. In the 1980s, emergence of the ancient DNA (aDNA) field (Higuchi et al., 1984) eventually allowed pathogen detection in archaeological skeletal remains based on short genomic fragments, using PCR technology (Spigelman & Lemma, 1993). It thus became possible to make a diagnosis of an infection using molecular evidence. With the advent of Next Generation Sequencing (NGS) (Margulies et al., 2005), the amount of genomic data available for analysis drastically increased, enabling reconstruction and analysis of complete pathogen genomes (Bos et al., 2011). Over the last decade, the aDNA field has been a valuable source of knowledge on the emergence, virulence and spread of infectious diseases, providing us with insights into the evolutionary history of pathogenic microorganisms and the never-ending war between humans and their pathogens.

Infectious diseases in medieval Europe

There is an abundance of historical and archaeological evidence depicting the centuries-long struggles of medieval Europeans with infectious diseases (Fig. 1). The “speckled monster”, nowadays known as smallpox, is thought to have arrived in Europe sometime around the 5-6th

cent. AD. Fortunately, the disease caused by variola virus has been eradicated. In the past, however, it claimed millions of lives (Riedel, 2005). It was estimated that smallpox killed approx. 400 000 people annually in 18th-century Europe and blinded around one third of the survivors (Barquet, 1997). Deadly epidemics in the Middle Ages were also caused by paratyphoid fever – following an infection with *Salmonella enterica* Paratyphi C (Haller et al., 2021) (Fig. 2) – and measles (Retief & Cilliers, 2010).

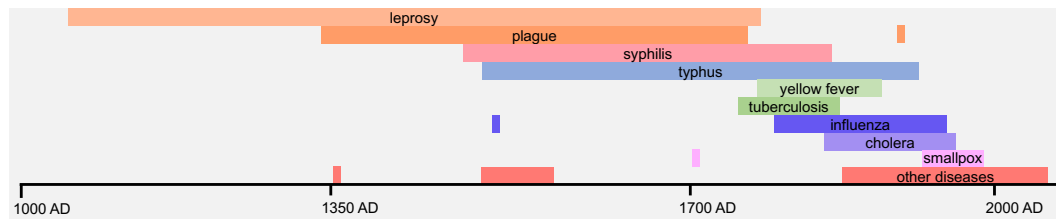


Figure 1. Timeline of the European outbreaks of infectious diseases since the High Middle Ages based on historical and molecular data. The “other diseases” category includes paratyphoid and typhoid fever, sweating sickness, malaria, encephalitis lethargica, variant Creutzfeldt-Jakob disease, q-fever and COVID19.

The most notorious infectious disease of the Middle Ages was, however, the plague. The word “plague” still inspires terror today. *Y. pestis* was the causative agent of at least three major pandemics in human history (the Justinianic plague (6th-8th cent. AD), the Medieval plague (14th-18th cent. AD) and the Modern plague (19th-20th cent. AD)), with the Black Death being the most infamous period. Black Death started the Medieval pandemic and lasted for about six years during which approximately 30-50% of the European population died from plague (Poos, 1991; Scott & Duncan, 2001; Wood et al., 2003). As the pathogen is an extremely fast killer, the plague pandemic left behind a number of mass graves, which have been a valuable source of knowledge about the Medieval plague and its epidemiology (e.g. Grainger et al., 2008; Hartle, 2017; Pfizenmaier, 2016; Stoakley, 2019; Willmot et al., 2020). Another infectious disease that reached large-scale epidemic proportions in medieval Europe was leprosy, caused by *Mycobacterium leprae* (*M. leprae*). Although this disease is not deadly *per se*, it is disfiguring and highly debilitating, which often led to social ostracism of the infected individuals and their isolation in infirmaries called leprosaria. It was estimated that by the late Middle Ages, there were thousands of such institutions across Europe (Raffcliffe, 2006). Both

plague and leprosy, although vastly different in their pathogenesis (acute vs. chronic infection), persisted in Europe for centuries until their disappearance in the late Middle Ages.



Figure 2. Two mass burials containing victims of a 14th cent. AD paratyphoid fever epidemic in Lübeck, Germany (Haller et al., 2021).

Our knowledge about medieval epidemics is mostly based on historical sources and the exact pathogens behind the outbreaks are often unknown. Although aDNA studies helped to identify the causative agent in several instances, e.g., *Y. pestis* as the pathogen responsible for the Justinianic pandemic and the Black Death (Bos et al., 2011; Wagner et al., 2014), most aDNA findings represent isolated cases of infection. Furthermore, when looking at the timeline of European infectious disease outbreaks (Fig. 1), a large gap in knowledge regarding the pathogen landscape in the Middle Ages becomes apparent. Up to now, no systematic pathogen screening of skeletal remains buried in common cemeteries has been performed. This kind of analysis would improve our understanding of the infectious disease burden in medieval Europeans.

Plague and evolution of *Yersinia pestis*

Y. pestis evolved from the relatively benign pathogen *Y. pseudotuberculosis* approx. 7000 years ago (Susat et al., 2021) to become one of the most dangerous infectious agents known to humankind (Tian & Zheng, 2014). Plague is a zoonotic disease with rodents being the main host of the pathogen (Gage & Kosoy, 2005). Currently, sylvatic plague reservoirs are present nearly globally with the exception of Australia and Europe (Dennis & Mead, 2006). The oriental rat flea (*Xenopsylla cheopis*) is the most common vector for the disease and the

classical transmission mode (rat > flea > human) was first described in 1898 (Simond et al., 1998). It was, however, proposed that lice and human fleas could also contribute to the spread of *Y. pestis* (Dean et al., 2018). Moreover, plague can be transmitted between humans via respiratory droplets (Cleri et al., 1997; Dennis & Mead, 2006). The infection mode affects the type of targeted tissues, leading to differences in the disease pathophysiology. Three clinical forms of plague can be distinguished: bubonic, pneumonic and septicemic. The bubonic plague occurs following a bite of an infected flea. After an incubation period of approx. 2-8 days, the bacteria infect the lymphatic system by entering and proliferating within lymph nodes, leading to the development of “buboes” (painfully swollen lymph nodes) accompanied by flu-like symptoms (Butler, 1983; Dennis, 2005). If left untreated, the mortality of bubonic plague reaches approx. 60% (Dennis, 2005). Inhalation of the pathogen via the respiratory route results in a severe infection of the lungs, which can also occur following the spread of the bacterium from the blood into the lung tissue (secondary pneumonic plague). Primary infection of the lungs is acute and rapid in its progression. The symptoms include fever, chest pain, cough and hemoptysis. Unless a treatment is applied within the first 20 hours, death may occur in around three days after infection (mortality rate reaches 100%) (Butler, 1983; Prentice & Rahalison, 2007). Infection of the bloodstream (septicemic plague) can develop as a complication of the bubonic and/or the pneumonic form. It can, however, also be acquired after a flea and/or an animal bite (primary septicemic plague). The flu-like symptoms and, in some cases, gastrointestinal disturbances are non-specific. Thus, diagnosis of plague is usually delayed, leaving patients with very low chances of survival (Butler, 1983).

Y. pestis gained its remarkably high virulence by several genomic alterations. These include acquisition of two virulence-associated plasmids (pPCP1 and pMT1) as well as pseudogenization and loss of several chromosomal genes, which greatly contributed to transmissibility of the pathogen. The pMT1 plasmid carries an important for *Y. pestis* virulence gene (*ymt*), which encodes *Yersinia* murine toxin (Hinnebush et al., 2002), and the *caf* operon, encoding an immuno-protective capsule (Galyov et al., 1990). Together with the *ymt* acquisition, inactivation of four chromosomal genes (*rcaA*, *pde2*, *pde3* and *ureD*) allowed the bacterium to cause the flea-borne plague and expand the range of mammalian hosts (Hinnebush et al., 2002). Additionally, *pla* surface protein (*pla*) encoded on the pPCP1 plasmid promotes the dissemination of the bacteria in the lymphatic nodes and their extensive growth in the lower airways (Banerjee et al., 2020; Sodeinde et al., 1992). Shared among pathogenic *Yersinia* is

also the pCD1 plasmid, which carries genes encoding *Yersinia* outer membrane proteins (Yops) responsible for phagocytosis inhibition, downregulation of cytokine production and cell death induction (Grabowski et al., 2017). Analysis of prehistorical Eurasian *Y. pestis* genomes dating to the Late Neolithic and Early Bronze Age periods (LNBA) revealed that the early *Y. pestis* strains (N=28, dating to approx. 3350-507 BCE) did not possess the full capacity of flea colonization. These strains lacked the *ymt* gene as well as the ability to produce biofilm – an important factor for flea-borne transmission. LNBA genomes also carried active forms of genes responsible for the fast killing of the fleas (Rascovan et al., 2019; Rasmussen et al., 2015; Valtueña et al., 2017; 2022; Susat et al., 2021; Yu et al., 2020). In the 2nd millennium BCE, the first *Y. pestis* strains fully capable of vector transmission (*ymt*-positive) emerged (RT5 and I2470, approx. 1918-1231 BCE) (Spyrou et al., 2018; Valtueña et al., 2022). It was suggested that prior to the *ymt* acquisition, the pathogen spread mainly via the aerosol and oral routes with the vector-borne transmission being limited or absent (Valtueña et al., 2017; Slavin & Sebbane, 2022). Based on these findings, the flea-adapted LNBA strains and the strains that were not adapted to flea colonization temporally co-existed until they became extinct. A similar fate was met by the *Y. pestis* lineages responsible for the Justinianic pandemic and the post-Black Death outbreaks in the Middle Ages. Currently, the phylogenetic tree of *Y. pestis* constitutes 5 branches (0-4). Branch 0 contains the LNBA (N=30) (Rascovan et al., 2019; Rasmussen et al., 2015; Valtueña et al., 2017; 2022; Spyrou et al., 2018; Susat et al., 2021; Yu et al., 2020) and Justinianic plague strains (N=8) (Damgaard et al., 2018; Feldman et al., 2016; Keller et al., 2019) as well as modern strains from the territories of China, Mongolia and the former Soviet Union. The phylogenetic branches 1-4 were formed through rapid population expansion and can be found across the globe (Cui et al., 2013). All strains responsible for the Medieval and Modern plague pandemics cluster within branch 1, which exhibits the widest geographical distribution.

The Medieval plague pandemic is thought to have started in Asia in 1346 AD, from where it was brought to Europe via marine trade, reaching Sicily in October 1347 (Spyrou et al., 2016). Plague then rapidly spread across the continent, arriving in northern Europe by 1348 (Bos et al., 2011). The first six years of the pandemic were particularly deadly and this period is known as the Black Death (1346-1353 AD). It was estimated that between 30 to 50% of the European population died following an infection during that time (Poos, 1991; Scott & Duncan, 2001; Wood et al., 2003). After 1353 AD, sporadic plague outbreaks persisted for another 400 years.

It was previously proposed that the post-Black death plague epidemics were a result of multiple waves of *Y. pestis*' introduction to Europe from Asia (Haensch et al., 2010). However, phylogenetic analyses of medieval *Y. pestis* genomes (N=48) revealed a low diversity of the strains (Bos et al., 2011; Spyrou et al., 2016; 2019; 2022), suggesting that after a single introduction of the pathogen to Europe, *Y. pestis* established reservoirs in local rodent populations. Thus, although we have no knowledge or evidence of potential plague reservoirs in medieval Europe, the post-Black death outbreaks were likely a consequence of environmental spill-overs of plague from rodents to humans. The *Y. pestis* phylogeny also shows that two separate lineages of *Y. pestis* were responsible for the Black Death and the subsequent outbreaks. Whilst the Black Death migrated to Asia and eventually became the source of the Modern pandemic, the post-Black death lineage went extinct (Spyrou et al., 2016).

Plague disappeared from Europe in the 18th cent. AD for reasons which are still debated. Several possible explanations were previously proposed, including improvement of hygienic standards and living conditions, change of the host from the black rat to the brown rat and disruption of the ecological niche of local reservoirs by industrialization (Appleby, 1980; Carmichael, 2014). Although Europe is currently free of plague, *Y. pestis* reservoirs are still present in Asia, Africa and both Americas (Dennis & Mead, 2006). In the mid 19th cent. AD, the Modern plague pandemic broke out in China from where it quickly spread around the globe, killing approx. 15 million people (Frith, 2012). Since then, sporadic outbreaks of plague have still occurred (e.g. Dutt et al., 1994; Leslie et al., 2011; Nguyen et al., 2018), claiming more victims despite the existence of antibiotics against *Y. pestis* infection. The emergence of antibiotic-resistant strains provides a reason for concern. The development of various resistance mechanisms (Lei & Kumar, 2022) reflects the constant evolution of the pathogen and reminds us that plague remains a serious threat to human health.

Leprosy and the spread of *Mycobacterium leprae*

Mentions of a leprosy-like disease date far back into antiquity with the earliest generally accepted evidence dating to 2000 BCE (Robbins et al., 2009). Leprosy is a chronic infectious disease caused by *M. leprae*. As temperatures below 37°C are optimal for the pathogens' growth, the infection mainly affects the skin and superficial structures, such as mucous membranes, eyes, testes and peripheral nerves. *M. leprae* exhibits a high affinity to Schwann

cells (SC) in the neural system (Cole et al., 2001) as well as keratinocytes and macrophages in the skin (Cogen et al., 2012). In the SC, mycobacterial infection leads to chronic inflammation, demyelination, fibrosis and eventually axonal death, resulting in sensory loss and disability (Britton, 1998; Rambukkana et al., 2002). Numb areas are susceptible to injuries and secondary infection, usually leading to deformation and in some cases even loss of extremities. Another source of visible signs of illness are skin lesions (single or numerous) - most often found on the limbs. Moreover, chronic rhinitis can cause the nose to collapse (saddle nose) (Andersen & Manchester, 1992). Leprosy is a lifelong disease; the average incubation period is estimated to vary from 3 to 5 years (Pinheiro et al., 2011), with a minimum of a few weeks (Montestruc & Berdonneau, 1954). It can, however, reach even 30 years (Pinheiro et al., 2011). As leprosy progresses over decades, it eventually affects the skeletal system causing various types of lesions (Ortner, 2003).

Approximately 90% of people infected with *M. leprae* do not develop clinical symptoms (Joyce, 2012; Sales et al., 2011). Different degrees of the immunological response to *M. leprae* infection lead to a variety of clinical manifestations and thus – the spectral concept of leprosy. Several classification systems for leprosy exist which assign the disease type based on the severity of the symptoms and the host's immune response. A commonly used system was proposed by Ridley and Jopling in 1966 in which the classification is based on the cell-mediated response to the infection (Ridley & Jopling, 1966). In this classification, five leprosy forms are recognized: tuberculoid-tuberculoid (TT), borderline-tuberculoid (BT), borderline-borderline (BB), borderline-lepromatous (BL) and lepromatous-lepromatous (LL). The mild end of the spectrum (TT) is characterized by a relatively effective immune system and containment of the infection in granulomas. In contrast, LL is associated with widespread lesions, absence of granulomas, relatively high contagiousness and poor response from the immune system of the affected individual. In the borderline forms, the immunological response to the infection is incomplete and patients can experience shifts along the leprosy spectrum (Walker & Lockwood, 2007).

Skeletal manifestation of leprosy is mainly attributed to LL. Characteristic disease lesions include pathological changes to the facial skeleton (*facies leprosa*, rhino-maxillary syndrome) focused around the maxilla and the nasal bones (Møller-Christensen, 1974; 1978) (Fig. 3), as well as atrophy of the phalanges in hands and feet and nonspecific periosteal reactions in tibia

and fibula (Møller-Christensen, 1961; Ortner, 2003). In severe cases, muscle paralysis leads to immobility and subsequently results in atrophy and osteoporosis (Andersen & Manchester, 1992; Ortner, 2003). In TT, skeletal lesions are more likely to be a result of injuries and secondary infections and any deformities of the facial skeleton are unilateral (Manchester, 2002). The presence of lesions typical of leprosy have been highly beneficial for the study of this disease in past populations, including its epidemiology and society's attitude towards the disabled (i.e. leprosaria and burial practices). It can also improve our understanding of the disease by providing valuable insights into the mechanisms of *M. leprae* infection in the absence of antibiotic treatment.



Figure 3. Cranium of an adult male buried in a cemetery adjacent to the St. Jørgen leprosarium in Odense, Denmark. Lesions pathognomonic for LL can be observed on the maxilla and nasal bones. These include widening of the nasal aperture, resorption of the anterior nasal spine, destruction of the nasal bones and the alveolar process of the maxilla as well as antemortem loss of anterior dentition (photograph: B. Krause-Kyora).

The exact way of leprosy transmission is still unknown. However, it is generally accepted that the respiratory route is the most probable (Noordeen, 1994; Cole et al., 2001). Direct contact between lesions of affected patients and skin erosions of healthy individuals can possibly result in an infection (Job et al., 1999). Transmission via blood (Ghorpade, 2002) or breast milk (Pedley, 1968) cannot be excluded. Furthermore, nine-banded armadillos (*Dasypus novemcinctus*) are a known reservoir of the leprosy bacillus (Balamayooran et al., 2015), which has been directly implicated in the transmission of the disease to humans (Truman et al., 2011; Sharma et al., 2015). *M. leprae* has also been identified in red squirrels (*Sciurus vulgaris*) (Avanzi et al., 2016) and apes (Donham & Leininger, 1977; Hockings et al., 2021). Moreover, there is evidence suggesting that environmental reservoirs of mycobacteria, such as soil (Tió-Coma et al., 2019), water (Matsuoka et al., 1999) and free-living amoebas (Lahiri & Krahenbuhl, 2008) may also contribute to leprosy transmission.

Whole genome sequencing and comparative genomics have provided insight into the biology and evolution of *M. leprae*, contributing to our understanding of its geno- and phenotypical behavior (e.g. Monot et al., 2005; 2009; Schuenemann et al., 2018). The size of the *M. leprae* genome is much smaller relative to other mycobacteria species (Cole et al., 2001; Brosch et al., 2000). Moreover, the pathogen carries the highest number of pseudogenes among bacteria (27%) (Brosch et al., 2000; Cole et al., 2001; Liu et al., 2004) and only 49.5% of the sequences are gene-coding. It is therefore thought that the bacterium must have gone through a massive gene decay since the divergence from its last common ancestor with *M. tuberculosis* (Cole et al., 2001), approx. 66 million years ago (Monot et al., 2009; Schuenemann et al., 2013). Reductive genome evolution is known to occur in organisms that transition from a free-living environment to obligate intracellular parasitism (Cole et al., 2001) and although the precise mechanisms of *M. leprae* pseudogenization are still debated, the event was estimated to have occurred around 20 million years ago (Han et al., 2009).

The genome of *M. leprae* exhibits greatly limited variability, with all strains analyzed to date being over 99% identical (Schuenemann et al., 2013). Up to now, 42 whole genomes of ancient *M. leprae* have been analyzed (Fotakis et al., 2020; Krause-Kyora et al., 2018; Schuenemann et al., 2013; 2018; Morozova et al., 2020; Neukamm et al., 2020; Pfrengle et al., 2021). All except one (Abusir1630, 342-117 BCE, Egypt (Neukamm et al., 2020)) were isolated from medieval and early modern Europe with the dating of the genomes spanning approx. 1300 years

(5th – 20th cent. AD). Such a high genome stability can be advantageous in the field of phylogeography, as it can provide insight into the spread of pathogens, migration of their hosts and disease evolution. Indeed, the phylogenetic tree of *M. leprae* has been shown to reflect ancient human migrations. At the current phylogenetic resolution, six main branches (0-5) can be distinguished, with branch 2 being further divided into sub-branches 2E, 2H and 2F (Avanzi et al., 2020; Benjak et al., 2018; Schuenemann et al., 2013; 2018). The majority of medieval *M. leprae* strains belong to branch 3, including the early medieval genomes (before the 10th cent. AD) (N=2, EDI006 and GC96). Interestingly, so do the modern *M. leprae* strains isolated from red squirrels in the United Kingdom, which points to a possible contribution of squirrels to the spread of leprosy in the Middle Ages (Avanzi et al., 2016; Inskip et al., 2017), as squirrel meat and fur were common trade materials at the time (Veale, 1966). It is thought that branch 3 strains were brought from Europe to the Americas in the 15th cent. AD (Mark, 2017) infecting both humans and nine-banded armadillos (Balamayooran et al., 2015; Truman et al., 2011; Sharma et al., 2015). Despite the general prevalence of branch 3 in medieval Europe, *M. leprae* genomes dating to the Late Middle Ages (after the 10th cent. AD) represent almost all major phylogenetic branches, which are nowadays found in Asia (branches 0 and 1), Africa (branches 2 and 4) and the Americas (branches 3 and 4). Furthermore, previous studies revealed a high diversity at leprosarium-associated cemeteries (Krause-Kyora et al., 2018; Pfrengle et al., 2021; Schuenemann et al., 2013). These findings suggest that leprosaria could have housed not only locals, but also foreign patients and pilgrims (Pfrengle et al., 2021). Such a high diversity of *M. leprae* strains observed in medieval Europe could be explained by two different models proposed by Schuenemann and colleagues (Schuenemann et al., 2018). Model 1 assumes that leprosy originated in western Eurasia or possibly in Europe from where it spread across the globe. On the other hand, model 2 infers introduction of the pathogen to Europe from different regions before and during the Middle Ages. The extensive trade networks and urban development in Europe in that period support both models. Whilst prevalence of the ancestral branch 0 in East Asia and the Middle East adds verisimilitude to model 2, the majority of the analyzed-to-date medieval strains are ancestral to modern isolates favoring model 1 (Schuenemann et al., 2018). As the origin of leprosy, its spread, and the reasons for its disappearance from Europe still remain elusive, it is imperative to identify and genotype more archaeological *M. leprae* isolates from various locations in order to understand the spatio-temporal changes in disease patterns.

Pathogen-driven selection on the Human Leukocyte Antigen (HLA) region

Our immune system constantly protects us from any potentially harmful foreign particles (antigens). It comprises structural and chemical barriers, such as skin and mucous membranes, as well as the innate and adaptive immunity. Whilst the innate immune response is rapid and non-specific, the adaptive response is tailored to specific antigens and essential in cases when the innate response is insufficient. Characteristic for the adaptive immunity is the immunological memory – the ability of the immune system to react more rapidly and effectively upon reinfection (Marshall et al., 2018). The Human Leukocyte Antigen (HLA) system (or human Major Histocompatibility Complex, MHC) is involved in both the innate and adaptive immunity. It is a large gene complex (3 Mb) located on chromosome 6, which encodes cell-surface proteins responsible for regulation of the immune response. The HLA proteins contribute to self and non-self-recognition as well as presentation of foreign antigens to the effector cells, playing a key role in infection, autoimmunity and transplantation medicine. Thus, there is a genetic component to infectious disease susceptibility (infection, progression, manifestation and mortality). The HLA can be divided into two major groups: class I (A, B, C, E, F and G) and class II (DR, DQ, DP, DM and DO) (Fig. 4). Class I HLA are present on all nucleated cells and present endogenous antigens to cytotoxic T cells. Class II HLA, on the other hand, are expressed only on the surface of antigen presenting cells, such as macrophages, dendritic cells and B cells, and are responsible for presentation of exogenous antigens to helper T cells. (Fig. 4). Each class recognizes distinct peptides, which bind to the binding groove, allowing for the presentation of a wide range of antigens (Janeway et al., 2001).

The HLA complex is the most polymorphic locus in the human genome with 33490 known HLA alleles, as of March 2022 (Robinson et al., 2020). The high polymorphism of the HLA genes requires a systematic nomenclature (Fig. 5). The first-field (first two digits after the locus name, e.g., HLA-A*01) groups alleles of similar sequence and function. The second field (digits after the first “:”, e.g., HLA-A*01:01) indicates differences in the sequence (both within and outside of the binding domain) that affect the encoded molecule. The following fields of resolution indicate differences in nucleotide sequence which do not affect the coded protein (e.g., HLA-A*01:01:01) (Marsh et al., 2010).

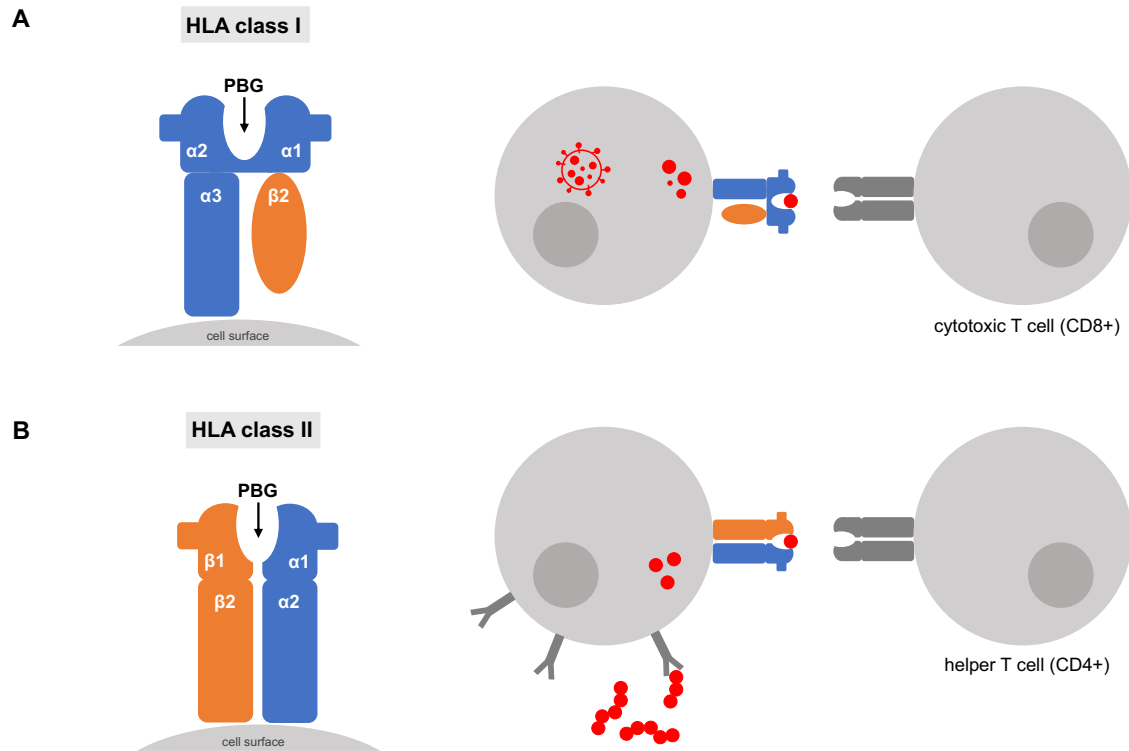


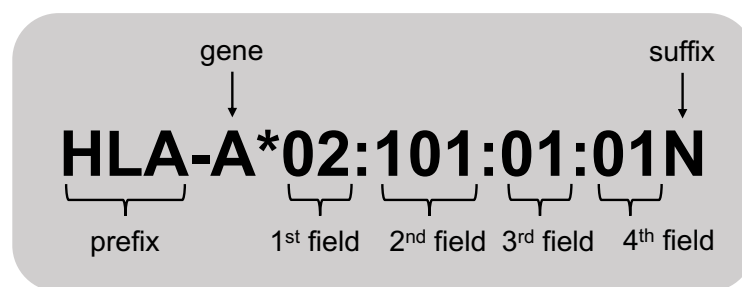
Figure 4. HLA class I (A) and class II (B) molecules (after Janeway et al., 2001).

A - Class I molecule consists of the alpha chain with three domains forming the peptide-binding groove (PBG), and the beta-2 microglobulin. They are present on all nucleated cells and present endogenous antigens to cytotoxic T cells (CD8+).

B- Class II molecule consists of one alpha and one beta chains with two domains each – both contribute to the PBG structure. They are expressed on the surface of antigen presenting cells and present exogenous antigens to T helper cells (CD4+).

The extremely high level of HLA polymorphism is thought to be a result of balancing selection, which is likely a consequence of the immense selective pressure of pathogens on the human genome. The influence of pathogens on the HLA diversity is supported by several lines of evidence. Firstly, the variation of the binding groove sequences is considerably higher relative to other loci (Hughes & Nei, 1998; Bitarello et al., 2016). Secondly, HLA polymorphism has been previously correlated with pathogen diversity (Prugnolle et al., 2005). Thirdly, previous computational and experimental studies point to the influence of pathogen pressure on HLA variability (Özer & Lenz, 2021; Penn et al., 2002). Furthermore, HLA alleles are linked with susceptibility and resistance to infectious diseases including acquired immunodeficiency

syndrome (AIDS), hepatitis B and C, tuberculosis, leprosy, typhoid fever, leishmaniasis, malaria and schistosomiasis (Blackwell et al., 2009; Sanchez-Mazas, 2020). Such associations are discovered by performing a case-control analysis, in which the allele frequencies are compared between two groups: cases (individuals diagnosed with the disease under study) and controls (individuals without the disease). Interestingly, there is an overlap between HLA alleles linked to susceptibility/resistance to infections and chronic inflammatory and autoimmune conditions. For instance, a genome-wide association study (GWAS) of leprosy in the Chinese (Zhang et al., 2009) identified a susceptibility variant in the HLA-DRB1 locus, which was also linked to Crohn's disease (Ahmad et al., 2006) suggesting that shared immune signaling pathways could regulate both diseases, which highlights the pleiotropy of HLA alleles (Schurr & Gros, 2009). It has been proposed that past epidemics of infectious diseases affected the HLA diversity, leading to a change in the frequency of variants associated with autoimmune and inflammatory conditions today (Krause-Kyora et al., 2018).



LEGEND
prefix – marks the HLA region of the chromosome
gene – particular HLA locus
1st field – allele group
2nd field – specific HLA protein
3rd field – indicates synonymous substitution within the coding region
4th field – indicates synonymous substitution within the non-coding region
suffix – denotes changes in expression
"-", "*" and ":" – separators

Figure 5. Nomenclature of the HLA system (after Marsh et al., 2010).

As Europe is free of many infectious diseases that used to be widespread in the past, for instance leprosy and paratyphoid fever, up to now, all modern association studies have been carried out in non-European populations. Therefore, not much is known about the genetic susceptibility to

various infectious diseases in European populations. It also remains largely unexplored to what extent past epidemics have affected the present-day HLA diversity and possibly the predisposition to chronic inflammation and autoimmunity. With aDNA, case-control association studies of diseases that in Europe are considered to be long gone became possible (Immel et al., 2021; Krause-Kyora et al., 2018). Such analyses can shed light onto possible reasons for the disappearance of certain infectious diseases, such as leprosy, and provide us with a better understanding of the environmental factors behind the balancing selection that has been acting on the HLA locus.

Aims of the dissertation

Part I: pathogen-driven evolution of human immunity

Susceptibility to infection and clinical manifestation of infectious diseases are influenced by several factors, including the genetic contribution of the host. Frequencies of the disease-associated variants may differ between populations due to their distinct genetic and demographic histories. In association studies, the HLA system is most often linked to predisposition to infectious diseases across diverse populations. Interestingly, HLA alleles are also associated with various chronic inflammatory conditions, which highlights the pleiotropy of the HLA complex. As many infectious diseases that are endemic in the Americas, Africa and Asia, such as leprosy and typhoid fever, are nowadays virtually absent in Europe, the majority of association studies have been conducted in present-day non-European populations. This leaves a gap in knowledge on the genetic predisposition to these infections in Europeans.

- ❖ In the first part of my thesis, I applied aDNA methodology to investigate the association between the HLA complex and predisposition to infections with *S. enterica* Paratyphi C (paratyphoid fever) as well as *M. leprae* (leprosy) in medieval Europeans. My main aims included:
 - a) testing of the association between HLA-DRB1*03 and *S. enterica* infection, reported for South Asians (chapter 1);
 - b) testing of the association between HLA-DRB1*15:01 and *M. leprae* infection, reported for medieval Europeans (chapter 2);
 - c) identification of possibly novel variants linked to leprosy predisposition and replication of the associations previously described in present-day non-Europeans (chapter 2);
 - d) evaluation of the influence of past epidemics on HLA variability and predisposition to autoimmune diseases and chronic inflammation today.

Part II: pathogen evolution

Infectious diseases are thought to have been an every-day threat to the people of medieval Europe. Most of our knowledge regarding this topic is speculative. The advent of aDNA

methodology greatly aided the study of infectious diseases in past populations by allowing identification of the pathogen. The majority of aDNA analyses focused on the analysis of mass burials and isolated cases. Although these studies succeeded in identifying specific pathogens responsible for deadly disease outbreaks and provided important insights into their evolution, they represent a rather small window into the pathogen landscape of medieval Europe. Moreover, the scarcity of data from the early medieval period leaves both the pathogen diversity and the disease prevalence at the time unexplored.

- ❖ The first systematic pathogen screening of early medieval burials, located within the borders of a German settlement, was performed (chapter 3) with the aims of
 - a) assessing the pathogen landscape at the time;
 - b) evaluating the infectious disease burden of the community.

One of the most devastating pandemics in the human history was the Medieval plague caused by *Y. pestis*. It started with the Black Death (1346-1353 AD) after which sporadic outbreaks persisted in Europe for another 400 years, vanishing from the continent in the 18th cent. AD. Although genomic and phylogenetic analyses of medieval *Y. pestis* genomes provided us with hypotheses concerning the source of the post-Black Death epidemics, possible reasons for plague disappearance are speculative. Analysis of a larger number of *Y. pestis* isolates from various locations is needed to expand our knowledge on the diversity of the pathogen in medieval Europe. Moreover, as *Y. pestis* remains a major threat to human health, a better understanding of the bacterium's evolution since the Middle Ages is of great interest in times of emerging and re-emerging infectious diseases.

- ❖ Metagenomic samples isolated from skeletal remains buried at medieval mass graves in Latvia (chapter 4) and cemeteries in Denmark and Germany (chapter 6) were screened for the presence of *Y. pestis*' DNA with the aim to:
 - a) determine whether the mass graves were a result of a plague outbreak;
 - b) assess the prevalence of plague in medieval cemeteries;
 - c) investigate the geographic and temporal diversity of medieval *Y. pestis* strains;
 - d) perform a comparative analysis between the ancient and medieval as well as modern *Y. pestis* genomes (chapter 4 and 5) with the aim of detecting genomic differences that could affect the pathogen's virulence.

References

- Ahmad, T., Marshall, S., & Jewell, D. (2006). Genetics of inflammatory bowel disease: The role of the HLA complex. *World Journal of Gastroenterology*, 12(23), 3628. <https://doi.org/10.3748/wjg.v12.i23.3628>
- Andersen, J. G., & Manchester, K. (1992). The rhinomaxillary syndrome in leprosy: A clinical, radiological and palaeopathological study. *International Journal of Osteoarchaeology*. <https://doi.org/10.1002/oa.1390020204>
- Appleby, A. B. (1980). The Disappearance of Plague: A Continuing Puzzle. *The Economic History Review*, 33(2), 161–173. <https://doi.org/10.1111/j.1468-0289.1980.tb01821.x>
- Aufderheide, A., & Rodriguez-Martin, C. (1998). *The Cambridge encyclopedia of human palaeopathology*. Cambridge University Press.
- Avanzi, C., Del-Pozo, J., Benjak, A., Stevenson, K., Simpson, V. R., Busso, P., McLuckie, J., Loiseau, C., Lawton, C., Schoening, J., Shaw, D. J., Piton, J., Vera-Cabrera, L., Velarde-Felix, J. S., McDermott, F., Gordon, S. v., Cole, S. T., & Meredith, A. L. (2016). Red squirrels in the British Isles are infected with leprosy bacilli. *Science*, 354(6313), 744–747. <https://doi.org/10.1126/science.aah3783>
- Avanzi, C., Singh, P., Truman, R. W., & Suffys, P. N. (2020). Molecular epidemiology of leprosy: An update. *Infection, Genetics and Evolution*, 86, 104581. <https://doi.org/10.1016/j.meegid.2020.104581>
- Balamayooran, G., Pena, M., Sharma, R., & Truman, R. W. (2015). The armadillo as an animal model and reservoir host for Mycobacterium leprae. *Clinics in Dermatology*, 33(1), 108–115. <https://doi.org/10.1016/j.clindermatol.2014.07.001>
- Banerjee, S. K., Crane, S. D., & Pechous, R. D. (2020). A Dual Role for the Plasminogen Activator Protease During the Preinflammatory Phase of Primary Pneumonic Plague. *The Journal of Infectious Diseases*, 222(3), 407–416. <https://doi.org/10.1093/infdis/jiaa094>
- Barquet, N. (1997). Smallpox: The Triumph over the Most Terrible of the Ministers of Death. *Annals of Internal Medicine*, 127(8), 635–642. https://doi.org/10.7326/0003-4819-127-8_Part_1-199710150-00010
- Benjak, A., Avanzi, C., Singh, P., Loiseau, C., Girma, S., Busso, P., Fontes, A. N. B., Miyamoto, Y., Namisato, M., Bobosha, K., Salgado, C. G., da Silva, M. B., Bouth, R. C., Frade, M. A. C., Filho, F. B., Barreto, J. G., Nery, J. A. C., Bühner-Sékula, S., Lupien, A., ... Cole, S. T. (2018). Phylogenomics and antimicrobial resistance of the leprosy bacillus Mycobacterium leprae. *Nature Communications*, 9, 352. <https://doi.org/10.1038/s41467-017-02576-z>
- Bitarello, B. D., Francisco, R. dos S., & Meyer, D. (2016). Heterogeneity of dN/dS ratios at the classical HLA class I genes over divergence time and across the allelic phylogeny. *Journal of Molecular Evolution*, 82(1), 38–50. <https://doi.org/10.1007/s00239-015-9713-9>

Blackwell, J. M., Jamieson, S. E., & Burgner, D. (2009). HLA and Infectious Diseases. *Clinical Microbiology Reviews*, 22(2), 370–385. <https://doi.org/10.1128/CMR.00048-08>

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, 506-510. <https://doi.org/10.1038/nature10549>

Britton, W. (1998). The management of leprosy reversal reactions. *Leprosy Review*, 69(3), 225–234. <https://doi.org/10.5935/0305-7518.19980024>

Brosch, R., Gordon, S. v, Eiglmeier, K., Garnier, T., & Cole, S. T. (2000). Comparative genomics of the leprosy and tubercle bacilli. *Research in Microbiology*, 151(2), 135–142. [https://doi.org/10.1016/S0923-2508\(00\)00117-0](https://doi.org/10.1016/S0923-2508(00)00117-0)

Butler, T. (1983). *Plague and Other Yersinia Infections*. Plenum Medical Book Company.

Cardona-Castro, N., Cortés, E., Beltrán, C., Romero, M., Badel-Mogollón, J. E., & Bedoya, G. (2015). Human Genetic Ancestral Composition Correlates with the Origin of *Mycobacterium leprae* Strains in a Leprosy Endemic Population. *PLOS Neglected Tropical Diseases*, 9(9), e0004045. <https://doi.org/10.1371/journal.pntd.0004045>

Carmichael, A. (2014). *Pandemic Disease in the Medieval World: Rethinking the Black Death* (M. Green, Ed.). ARC Medieval Press.

Cleri, D. J., Vernaleo, J. R., Lombardi, L. J., Rabbat, M. S., Mathew, A., Marton, R., & Reyelt, M. C. (1997). Plague pneumonia disease caused by *Yersinia pestis*. *Semin Respir Infect*, 12(1), 12–23.

Cogen, A. L., Walker, S. L., Roberts, C. H., Hagge, D. A., Neupane, K. D., Khadge, S., & Lockwood, D. N. J. (2012). Human Beta-Defensin 3 Is Up-Regulated in Cutaneous Leprosy Type 1 Reactions. *PLoS Neglected Tropical Diseases*, 6(11), e1869. <https://doi.org/10.1371/journal.pntd.0001869>

Cole, S. T., Eiglmeier, K., Parkhill, J., James, K. D., Thomson, N. R., Wheeler, P. R., Honoré, N., Garnier, T., Churcher, C., Harris, D., Mungall, K., Basham, D., Brown, D., Chillingworth, T., Connor, R., Davies, R. M., Devlin, K., Duthoy, S., Feltwell, T., ... Barrell, B. G. (2001). Massive gene decay in the leprosy bacillus. *Nature*, 409, 1007-1011. <https://doi.org/10.1038/35059006>

Cui, Y., Yu, C., Yan, Y., Li, D., Li, Y., Jombart, T., Weinert, L. A., Wang, Z., Guo, Z., Xu, L., Zhang, Y., Zheng, H., Qin, N., Xiao, X., Wu, M., Wang, X., Zhou, D., Qi, Z., Du, Z., ... Yang, R. (2013). Historical variations in mutation rate in an epidemic pathogen, *Yersinia pestis*. *Proceedings of the National Academy of Sciences*, 110(2). <https://doi.org/10.1073/pnas.1205750110>

Damgaard, P. de B., Marchi, N., Rasmussen, S., Peyrot, M., Renaud, G., Korneliussen, T., Moreno-Mayar, J. V., Pedersen, M. W., Goldberg, A., Usmanova, E., Baimukhanov, N., Loman, V., Hedeager, L., Pedersen, A. G., Nielsen, K., Afanasiev, G., Akmatov, K., Aldashev,

- A., Alpaslan, A., ... Willerslev, E. (2018). 137 ancient human genomes from across the Eurasian steppes. *Nature*, 557(7705), 369–374. <https://doi.org/10.1038/s41586-018-0094-2>
- Dean, K. R., Krauer, F., Walløe, L., Lingjærde, O. C., Bramanti, B., Stenseth, N. Chr., & Schmid, B. v. (2018). Human ectoparasites and the spread of plague in Europe during the Second Pandemic. *Proceedings of the National Academy of Sciences*, 115(6), 1304–1309. <https://doi.org/10.1073/pnas.1715640115>
- Dennis, D. (2005). Plague as a biological weapon. In I. Fong & K. Alibek (Eds.), *Bioterrorism and Infectious Disease: A New Dilemma for the 21st Century* (pp. 37–70). Springer.
- Dennis, D. T., & Mead, P. S. (2006). Plague. In *Tropical Infectious Diseases*. Elsevier. <https://doi.org/10.1016/B978-0-443-06668-9.50047-8>
- Donham, K. J., & Leininger, J. R. (1977). Spontaneous Leprosy-Like Disease in a Chimpanzee. *Journal of Infectious Diseases*, 136(1), 132–136. <https://doi.org/10.1093/infdis/136.1.132>
- Dutt, A., Akhtar, R., & McVeigh, M. (1994). Surat plague of 1994 re-examined. *Southeast Asian Journal of Tropical Medicine and Public Health*, 37(4), 755–760.
- Feldman, M., Harbeck, M., Keller, M., Spyrou, M. A., Rott, A., Trautmann, B., Scholz, H. C., Pääfgen, B., Peters, J., McCormick, M., Bos, K., Herbig, A., & Krause, J. (2016). A High-Coverage *Yersinia pestis* Genome from a Sixth-Century Justinianic Plague Victim. *Molecular Biology and Evolution*, 33(11), 2911–2923. <https://doi.org/10.1093/molbev/msw170>
- Fotakis, A. K., Denham, S. D., Mackie, M., Orbegozo, M. I., Mylopotamitaki, D., Gopalakrishnan, S., Sicheritz-Pontén, T., Olsen, J. v., Cappellini, E., Zhang, G., Christophersen, A., Gilbert, M. T. P., & Vågene, Å. J. (2020). Multi-omic detection of *Mycobacterium leprae* in archaeological human dental calculus. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 375(1812), 20190584. <https://doi.org/10.1098/rstb.2019.0584>
- Fox, D., & Karp, D. (1988). Images of Plague: Infectious Disease in the Visual Arts. In E. Fee & D. Fox (Eds.), *AIDS: The Burdens of History* (pp. 173–190). University of California Press.
- Frith, J. (2012). The History of Plague – Part 1. The Three Great Pandemics. *Journal of Military and Veterans' Health*, 20(2), 11–16.
- Gage, K. L., & Kosoy, M. Y. (2005). Natural History of Plague: Perspectives from More than a Century of Research. *Annual Review of Entomology*, 50(1), 505–528. <https://doi.org/10.1146/annurev.ento.50.071803.130337>
- Galyov, E. E., Smirnov, O. Yu., Karlishchev, A. V., Volkovoy, K. I., Denesyuk, A. I., Nazimov, I. V., Rubtsov, K. S., Abramov, V. M., Dalvadyanz, S. M., & Zav'yalov, V. P. (1990). Nucleotide sequence of the *Yersinia pestis* gene encoding F1 antigen and the primary structure of the protein. *FEBS Letters*, 277(1–2), 230–232. [https://doi.org/10.1016/0014-5793\(90\)80852-A](https://doi.org/10.1016/0014-5793(90)80852-A)

Ghorpade, A. (2002). Inoculation (tattoo) leprosy: a report of 31 cases. *Journal of the European Academy of Dermatology and Venereology*, 16(5), 494–499. <https://doi.org/10.1046/j.1468-3083.2002.00548.x>

Grabowski, B., Schmidt, M. A., & Rüter, C. (2017). Immunomodulatory Yersinia outer proteins (Yops)—useful tools for bacteria and humans alike. *Virulence*, 8(7), 1124–1147. <https://doi.org/10.1080/21505594.2017.1303588>

Grainger, I., Hawkins, D., Cowal, L., & Mikulski, R. (2008). *The Black Death Cemetery, East Smithfield*. Museum of London Archaeology Service.

Haensch, S., Bianucci, R., Signoli, M., Rajerison, M., Schultz, M., Kacki, S., Vermunt, M., Weston, D. A., Hurst, D., Achtman, M., Carniel, E., & Bramanti, B. (2010). Distinct clones of *Yersinia pestis* caused the black death. *PLoS Pathogens*, 6(10), e1001134. <https://doi.org/10.1371/journal.ppat.1001134>

Haller, M., Callan, K., Susat, J., Flux, A. L., Immel, A., Franke, A., Herbig, A., Krause, J., Kupczok, A., Fouquet, G., Hummel, S., Rieger, D., Nebel, A., & Krause-Kyora, B. (2021). Mass burial genomics reveals outbreak of enteric paratyphoid fever in the Late Medieval trade city Lübeck. *IScience*, 24(5), 102429. <https://doi.org/10.1016/j.isci.2021.102419>

Han, X. Y., Sizer, K. C., Thompson, E. J., Kabanja, J., Li, J., Hu, P., Gómez-Valero, L., & Silva, F. J. (2009). Comparative Sequence Analysis of *Mycobacterium leprae* and the New Leprosy-Causing *Mycobacterium lepromatosis*. *Journal of Bacteriology*, 191(19), 6067–6074. <https://doi.org/10.1128/JB.00762-09>

Hartle, R. (2017). *The New Churchyard: From Moorfields Marsh to Bethlem Burial Ground, Brokers Row and Liverpool Street (MOLA/Crossrail Archaeology 10)*. Museum of London Archaeology.

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>

Hinnebusch, B. J., Rudolph, A. E., Cherepanov, P., Dixon, J. E., Schwan, T. G., & Forsberg, Å. (2002). Role of Yersinia Murine Toxin in Survival of *Yersinia pestis* in the Midgut of the Flea Vector. *Science*, 296(5568), 733–735. <https://doi.org/10.1126/science.1069972>

Hockings, K. J., Mubemba, B., Avanzi, C., Pleh, K., Dux, A., Bersacola, E., Bessa, J., Ramon, M., Metzger, S., Patrono, L. v., Jaffe, J. E., Benjak, A., Bonneaud, C., Busso, P., Couacy-Hymann, E., Gado, M., Gagneux, S., Johnson, R. C., Kodio, M., ... Leendertz, F. H. (2021). Leprosy in wild chimpanzees. *Nature*, 598(7882), 652–656. <https://doi.org/10.1038/s41586-021-03968-4>

Hughes, A. L., & Nei, M. (1988). Pattern of nucleotide substitution at major histocompatibility complex class I loci reveals overdominant selection. *Nature*, 335(6186), 167–170. <https://doi.org/10.1038/335167a0>

Immel, A., Key, F. M., Szolek, A., Barquera, R., Robinson, M. K., Harrison, G. F., Palmer, W. H., Spyrou, M. A., Susat, J., Krause-Kyora, B., Bos, K. I., Forrest, S., Hernández-Zaragoza, D.

- I., Sauter, J., Solloch, U., Schmidt, A. H., Schuenemann, V. J., Reiter, E., Kairies, M. S., ... Krause, J. (2021). Analysis of Genomic DNA from Medieval Plague Victims Suggests Long-Term Effect of *Yersinia pestis* on Human Immunity Genes. *Molecular Biology and Evolution*, 38(10), 4059–4076. <https://doi.org/10.1093/molbev/msab147>
- Inskip, S., Taylor, G. M., Anderson, S., & Stewart, G. (2017). Leprosy in pre-Norman Suffolk, UK: biomolecular and geochemical analysis of the woman from Hoxne. *Journal of Medical Microbiology*, 66(11), 1640–1649. <https://doi.org/10.1099/jmm.0.000606>
- Janeway, C., Travers Jr, P., Walport, M., & Shlomchik, M. (2001). *Immunobiology: The immune system in health and disease* (5th edition). Garland Science.
- Job, C., Jayakumar, J., & Aschhoff, M. (1999). “Large numbers” of *Mycobacterium leprae* are discharged from the intact skin of lepromatous patients; a preliminary report. *International Journal of Leprosy and Other Mycobacterial Diseases*, 67(2), 164–167.
- Joyce, M. P. (2012). Historic aspects of human susceptibility to leprosy and the risk of conjugal transmission. *Memórias Do Instituto Oswaldo Cruz*, 107(suppl 1), 17–21. <https://doi.org/10.1590/S0074-02762012000900004>
- Keller, M., Spyrou, M. A., Scheib, C. L., Neumann, G. U., Kröpelin, A., Haas-Gebhard, B., Pääfgen, 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>
- 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, 1569. <https://doi.org/10.1038/s41467-018-03857-x>
- Lahiri, R., & Krahenbuhl, J. (2008). The role of free-living pathogenic amoeba in the transmission of leprosy: a proof of principle. *Leprosy Reviews*, 79(4), 401–409.
- Lei, C., & Kumar, S. (2022). *Yersinia pestis* antibiotic resistance: a systematic review. *Osong Public Health and Research Perspectives*, 13(1), 24–36. <https://doi.org/10.24171/j.phrp.2021.0288>
- Leslie, T., Whitehouse, C. A., Yingst, S., Baldwin, C., Kakar, F., Mofleh, J., Hami, A. S., Mustafa, L., Omar, F., Ayazi, E., Rossi, C., Noormal, B., Ziar, N., & Kakar, R. (2011). Outbreak of gastroenteritis caused by *Yersinia pestis* in Afghanistan. *Epidemiology and Infection*, 139(5), 728–735. <https://doi.org/10.1017/S0950268810001792>
- Liu, Y., Harrison, P. M., Kunin, V., & Gerstein, M. (2004). Comprehensive analysis of pseudogenes in prokaryotes: widespread gene decay and failure of putative horizontally transferred genes. *Genome Biology*, 5(9), R64. <https://doi.org/10.1186/gb-2004-5-9-r64>

Manchester, K. (2002). Infective Bone Changes of Leprosy. In C. Roberts, M. Lewis, & K. Manchester (Eds.), *The Past and Present of Leprosy: archaeological, historical, and clinical approaches*. (pp. 69–72). Archaeopress: BAR International Series.

Margulies, M., Egholm, M., Altman, W. E., Attiya, S., Bader, J. S., Bembien, 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>

Mark, S. (2017). Early Human Migrations (ca. 13,000 Years Ago) or Postcontact Europeans for the Earliest Spread of *Mycobacterium leprae* and *Mycobacterium lepromatosis* to the Americas. *Interdisciplinary Perspectives on Infectious Diseases*, 2017, 1–8. <https://doi.org/10.1155/2017/6491606>

Marsh, S. G. E., Albert, E. D., Bodmer, W. F., Bontrop, R. E., Dupont, B., Erlich, H. A., Fernández-Viña, M., Geraghty, D. E., Holdsworth, R., Hurley, C. K., Lau, M., Lee, K. W., Mach, B., Maiers, M., Mayr, W. R., Müller, C. R., Parham, P., Petersdorf, E. W., Sasazuki, T., ... Trowsdale, J. (2010). Nomenclature for factors of the HLA system, 2010. *Tissue Antigens*, 75(4), 291–455. <https://doi.org/10.1111/j.1399-0039.2010.01466.x>

Marshall, J. S., Warrington, R., Watson, W., & Kim, H. L. (2018). An introduction to immunology and immunopathology. *Allergy, Asthma & Clinical Immunology*, 14(S2), 49. <https://doi.org/10.1186/s13223-018-0278-1>

Matsuoka, M., Izumi, S., Budiawan, T., Nakata, N., & Saeki, K. (1999). *Mycobacterium leprae* DNA in daily using water as a possible source of leprosy infection. *Indian Journal of Leprosy*, 71(1), 61–67.

Mccullagh, C. B. (2000). Bias in Historical Description, Interpretation, and Explanation. *History and Theory*, 39(1), 39–66. <https://doi.org/10.1111/0018-2656.00112>

Møller-Christensen, V. (1961). *Bone Changes in Leprosy*. Munksgaard.

Møller-Christensen, V. (1974). Changes in the anterior nasal spine and the alveolar process of the maxillae in leprosy a clinical examination. *International Journal of Leprosy and Other Mycobacterial Diseases*, 42(4), 431–435.

Møller-Christensen, V. (1978). *Leprosy Changes of the Skull*. Odense University Press.

Monot, M., Honoré, N., Garnier, T., Araoz, R., Coppée, J.-Y., Lacroix, C., Sow, S., Spencer, J. S., Truman, R. W., Williams, D. L., Gelber, R., Virmond, M., Flageul, B., Cho, S.-N., Ji, B., Paniz-Mondolfi, A., Convit, J., Young, S., Fine, P. E., ... Cole, S. T. (2005). On the Origin of Leprosy. *Science*, 308(5724), 1040–1042. <https://doi.org/10.1126/science/1109759>

Monot, M., Honoré, N., Garnier, T., Zidane, N., Sherafi, D., Paniz-Mondolfi, A., Matsuoka, M., Taylor, G. M., Donoghue, H. D., Bouwman, A., Mays, S., Watson, C., Lockwood, D., Khamesipour, A., Dowlati, Y., Jianping, S., Rea, T. H., Vera-Cabrera, L., Stefani, M. M., ... Cole, S. T. (2009). Comparative genomic and phylogeographic analysis of *Mycobacterium leprae*. *Nature Genetics*, 41, 1282–1289. <https://doi.org/10.1038/ng.477>

- Montestruc, E., & Berdonneau, R. (1954). New cases of leprosy in infants in Martinique. *Bulletin de La Societe de Pathologie Exotique*, 47(6), 781–783.
- Morozova, I., Kasianov, A., Bruskin, S., Neukamm, J., Molak, M., Batieva, E., Pudło, A., Rühli, F. J., & Schuenemann, V. J. (2020). New ancient Eastern European *Yersinia pestis* genomes illuminate the dispersal of plague in Europe. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 375(1812), 20190569. <https://doi.org/10.1098/rstb.2019.0569>
- Mühlemann, B., Margaryan, A., de Barros Damgaard, P., Allentoft, M. E., Vinner, L., Hansen, A. J., Weber, A., Bazaliiskii, V. I., Molak, M., Arneborg, J., Bogdanowicz, W., Falys, C., Sablin, M., Smrčka, V., Sten, S., Tashbaeva, K., Lynnerup, N., Sikora, M., Smith, D. J., ... Jones, T. C. (2018). Ancient human parvovirus B19 in Eurasia reveals its long-term association with humans. *Proceedings of the National Academy of Sciences of the United States of America*, 115(29), 7557–7562. <https://doi.org/10.1073/pnas.1804921115>
- Neukamm, J., Pfrengle, S., Molak, M., Seitz, A., Francken, M., Eppenberger, P., Avanzi, C., Reiter, E., Urban, C., Welte, B., Stockhammer, P. W., Teßmann, B., Herbig, A., Harvati, K., Nieselt, K., Krause, J., & Schuenemann, V. J. (2020). 2000-year-old pathogen genomes reconstructed from metagenomic analysis of Egyptian mummified individuals. *BMC Biology*, 18(1), 108. <https://doi.org/10.1186/s12915-020-00839-8>
- Nguyen, V. K., Parra-Rojas, C., & Hernandez-Vargas, E. A. (2018). The 2017 plague outbreak in Madagascar: Data descriptions and epidemic modelling. *Epidemics*, 25, 20–25. <https://doi.org/10.1016/j.epidem.2018.05.001>
- Noordeen, S. (1994). Epidemiology of Leprosy. In *Mycobacteria. hapman & Hall Medical Microbiology Series*. Springer.
- Ortner, D. J. (2003). Identification of Pathological Conditions in Human Skeletal Remains. In *Identification of Pathological Conditions in Human Skeletal Remains*. <https://doi.org/10.1016/B978-0-12-528628-2.X5037-6>
- Özer, O., & Lenz, T. L. (2021). Unique Pathogen Peptidomes Facilitate Pathogen-Specific Selection and Specialization of MHC Alleles. *Molecular Biology and Evolution*, 38(10), 4376–4387. <https://doi.org/10.1093/molbev/msab176>
- Pedley, J. (1968). Presence of *M. leprae* in the nipple secretion and lumina of the hypertrophied mammary gland. In a case of gynaecomastia associated with active and untreated lepromatous leprosy. *Leprosy Reviews*, 39(2), 67–70.
- Penn, D. J., Damjanovich, K., & Potts, W. K. (2002). MHC heterozygosity confers a selective advantage against multiple-strain infections. *Proceedings of the National Academy of Sciences*, 99(17), 11260–11264. <https://doi.org/10.1073/pnas.162006499>
- Pfizenmaier, S. (2016). *Charterhouse Square: Black Death Cemetery and Carthusian Monastery, Meat Market and Suburb (MOLA/Crossrail Archaeology 7)*. Museum of London Archaeology.

- Pfrenkle, S., Neukamm, J., Guellil, M., Keller, M., Molak, M., Avanzi, C., Kushniarevich, A., Montes, N., Neumann, G. U., Reiter, E., Tukhbatova, R. I., Berezina, N. Y., Buzhilova, A. P., Korobov, D. S., Suppersberger Hamre, S., Matos, V. M. J., Ferreira, M. T., González-Garrido, L., Wasterlain, S. N., ... Schuenemann, V. J. (2021). *Mycobacterium leprae* diversity and population dynamics in medieval Europe from novel ancient genomes. *BMC Biology*, 19(1), 220. <https://doi.org/10.1186/s12915-021-01120-2>
- Pinheiro, R. O., de Souza Salles, J., Sarno, E. N., & Sampaio, E. P. (2011). *Mycobacterium leprae*-host-cell interactions and genetic determinants in leprosy: An overview. In *Future Microbiology*. <https://doi.org/10.2217/fmb.10.173>
- Poos, L. (1991). *A rural society after the Black Death: Essex, 1350–1525*. Cambridge University Press.
- Prentice, M. B., & Rahalison, L. (2007). Plague. *The Lancet*, 369(9568), 1196–1207. [https://doi.org/10.1016/S0140-6736\(07\)60566-2](https://doi.org/10.1016/S0140-6736(07)60566-2)
- Prugnotte, F., Manica, A., Charpentier, M., Guégan, J. F., Guernier, V., & Balloux, F. (2005). Pathogen-Driven Selection and Worldwide HLA Class I Diversity. *Current Biology*, 15(11), 1022–1027. <https://doi.org/10.1016/j.cub.2005.04.050>
- Raffcliffe, C. (2006). *Leprosy in Medieval England*. Boydell Press.
- Rambukkana, A., Zanazzi, G., Tapinos, N., & Salzer, J. L. (2002). Contact-Dependent Demyelination by *Mycobacterium leprae* in the Absence of Immune Cells. *Science*, 296(5569), 927–931. <https://doi.org/10.1126/science.1067631>
- Rascovan, N., Sjögren, K.-G., Kristiansen, K., Nielsen, R., Willerslev, E., Desnues, C., & Rasmussen, S. (2019). Emergence and Spread of Basal Lineages of *Yersinia pestis* during the Neolithic Decline. *Cell*, 176(1–2). <https://doi.org/10.1016/j.cell.2018.11.005>
- Rasmussen, S., Allentoft, M., Nielsen, K., Orlando, L., Sikora, M., Sjögren, K., Pedersne, A., Schubert, M., Vam Dam, A., Kapel, C., Nielsen, H., Btunak, S., Avetisyan, P., Epimakhov, A., Khalyapin, M., Guni, A., Kriiska, A., Lasak, I., Metspalu, M., ... Willerslev, E. (2015). Early divergent strains of *Yersinia pestis* in Eurasia 5,000 years ago. *Cell*, 163(3), 571–582. <https://doi.org/10.1016/j.cell.2015.10.009>
- Retief, F., & Cilliers, L. (2010). Measles in antiquity and the Middle Ages. *South African Medical Journal*, 100(4), 216. <https://doi.org/10.7196/SAMJ.3504>
- Ridley, D. S., & Jopling, W. H. (1966). Classification of Leprosy According to Immunity. *International Journal of Leprosy*.
- Riedel, S. (2005). Edward Jenner and the History of Smallpox and Vaccination. *Baylor University Medical Center Proceedings*, 18(1), 21–25. <https://doi.org/10.1080/08998280.2005.11928028>
- Robbins, G., Tripathy, V. M., Misra, V. N., Mohanty, R. K., Shinde, V. S., Gray, K. M., & Schug, M. D. (2009). Ancient Skeletal Evidence for Leprosy in India (2000 B.C.). *PLoS ONE*, 4(5), e5669. <https://doi.org/10.1371/journal.pone.0005669>

Robinson, J., Barker, D. J., Georgiou, X., Cooper, M. A., Flicek, P., & Marsh, S. G. E. (2020). IPD-IMGT/HLA Database. *Nucleic Acids Research*, 48(D1), D948–D955. <https://doi.org/10.1093/nar/gkz950>

Sales, A. M., Ponce de Leon, A., Düppre, N. C., Hacker, M. A., Nery, J. A. C., Sarno, E. N., & Penna, M. L. F. (2011). Leprosy among Patient Contacts: A Multilevel Study of Risk Factors. *PLoS Neglected Tropical Diseases*, 5(3), e1013. <https://doi.org/10.1371/journal.pntd.0001013>

Sanchez-Mazas, A. (2020). A review of HLA allele and SNP associations with highly prevalent infectious diseases in human populations. *Swiss Medical Weekly*. <https://doi.org/10.4414/smww.2020.20214>

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), e1006997. <https://doi.org/10.1371/journal.ppat.1006997>

Schuenemann, V. J., Singh, P., Mendum, T. A., Krause-Kyora, B., Jäger, G., Bos, K. I., Herbig, A., Economou, C., Benjak, A., Busso, P., Nebel, A., Boldsen, J. L., Kjellström, A., Wu, H., Stewart, G. R., Taylor, G. M., Bauer, P., Lee, O. Y. C., Wu, H. H. T., ... Krause, J. (2013). Genome-wide comparison of medieval and modern *Mycobacterium leprae*. *Science*, 341(6142), 179–183. <https://doi.org/10.1126/science.1238286>

Schurr, E., & Gros, P. (2009). A Common Genetic Fingerprint in Leprosy and Crohn's Disease? *New England Journal of Medicine*, 361(27), 2666–2668. <https://doi.org/10.1056/NEJMe0910690>

Scott, S., & Duncan, C. J. (2001). *Biology of Plagues: Evidence fro Historical populations*. Cambridge University Press.

Sharma, R., Singh, P., Loughry, W. J., Lockhart, J. M., Inman, W. B., Duthie, M. S., Pena, M. T., Marcos, L. A., Scollard, D. M., Cole, S. T., & Truman, R. W. (2015). Zoonotic Leprosy in the Southeastern United States. *Emerging Infectious Diseases*, 21(12). <https://doi.org/10.3201/eid2112.150501>

Simond, M., Godley, M., & Pierre, D. (1998). Paul-Louis Simond and his discovery of plague transmission by rat fleas: a centenary. *Journal of the Royal Society of Medicine*, 91, 101–104.

Slavin, P., & Sebbane, F. (2022). Emergence and spread of ancestral *Yersinia pestis* in Late-Neolithic and Bronze-Age Eurasia, ca . 5,000 to 2,500 y B.P. *Proceedings of the National Academy of Sciences*, 119(21). <https://doi.org/10.1073/pnas.2204044119>

Sodeinde, O. A., Subrahmanyam, Y. V. B. K., Stark, K., Quan, T., Bao, Y., & Goguen, J. D. (1992). A Surface Protease and the Invasive Character of Plague. *Science*, 258(5084), 1004–1007. <https://doi.org/10.1126/science.1439793>

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., 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., Gourvenec, 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, 4470. <https://doi.org/10.1038/s41467-019-12154-0>

Spyrou, M. A., Musralina, L., Gneccchi Ruscone, G. A., Kocher, A., Borbone, P.-G., Khartanovich, V. I., Buzhilova, A., Djansugurova, L., Bos, K. I., Kühnert, D., Haak, W., Slavin, P., & Krause, J. (2022). The source of the Black Death in fourteenth-century central Eurasia. *Nature*, 606(7915), 718–724. <https://doi.org/10.1038/s41586-022-04800-3>

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>

Spyrou, M. A., Tukhbatova, R. I., Wang, C.-C., Valtueña, A. A., Lankapalli, A. K., Kondrashin, V. v., Tsybin, V. A., Khokhlov, A., Kühnert, D., Herbig, A., Bos, K. I., & Krause, J. (2018). Analysis of 3800-year-old *Yersinia pestis* genomes suggests Bronze Age origin for bubonic plague. *Nature Communications*, 9(1), 2234. <https://doi.org/10.1038/s41467-018-04550-9>

Stoakley, M. (2019). ‘Great fears of the sickness here in the City ... God preserve us all ...’. A Plague Burial Ground in Leith, 1645: An Archaeological Excavation at St Mary’s (Leith) RC Primary School, Leith Links, Edinburgh. *Scottish Archaeological Internet Reports*, 86, 1–82.

Susat, J., Lübke, H., Immel, A., Brinker, U., Macăne, A., Meadows, J., Steer, B., Tholey, A., Zagorska, I., Gerhards, G., Schmöcke, U., Kalniņš, M., Franke, A., Pētersone-Gordina, E., Teßman, B., Törv, M., Schreiber, S., Andree, C., Bērziņš, V., ... Krause-Kyora, B. (2021). A 5,000-year-old hunter-gatherer already plagued by *Yersinia pestis*. *Cell Reports*, 35(13), 109278. <https://doi.org/10.1016/j.celrep.2021.109278>

Tian, D., & Zheng, T. (2014). Comparison and Analysis of Biological Agent Category Lists Based On Biosafety and Biodefense. *PLoS ONE*, 9(6), e101163. <https://doi.org/10.1371/journal.pone.0101163>

Tió-Coma, M., Wijnands, T., Pierneef, L., Schilling, A. K., Alam, K., Roy, J. C., Faber, W. R., Menke, H., Pieters, T., Stevenson, K., Richardus, J. H., & Geluk, A. (2019). Detection of *Mycobacterium leprae* DNA in soil: multiple needles in the haystack. *Scientific Reports*, 9(1), 3165. <https://doi.org/10.1038/s41598-019-39746-6>

Truman, R. W., Singh, P., Sharma, R., Busso, P., Rougemont, J., Paniz-Mondolfi, A., Kapopoulou, A., Brisse, S., Scollard, D. M., Gillis, T. P., & Cole, S. T. (2011). Probable

Zoonotic Leprosy in the Southern United States. *New England Journal of Medicine*, 364(17), 1626–1633. <https://doi.org/10.1056/NEJMoA1010536>

Valtueña, A., Mittnik, A., Key, F. M., Haak, W., Allmäe, R., Belinskij, A., Daubaras, M., Feldman, M., Jankauskas, R., Janković, I., Massy, K., Novak, M., Pfrengle, S., Reinhold, S., Šlaus, M., Spyrou, M. A., Szécsényi-Nagy, A., Törv, M., Hansen, S., ... Krause, J. (2017). The Stone Age Plague and Its Persistence in Eurasia. *Current Biology*, 27(23), 3683–3691.e8. <https://doi.org/10.1016/j.cub.2017.10.025>

Valtueña, A., Neumann, G. U., Spyrou, M. A., Musralina, L., Aron, F., Beisenov, A., Belinskiy, A. B., Bos, K. I., Buzhilova, A., Conrad, M., Djansugurova, L. B., Dobeš, M., Ernée, M., Fernández-Eraso, J., Frohlich, B., Furmanek, M., Hałuszko, A., Hansen, S., Harney, É., ... Herbig, A. (2022). Stone Age *Yersinia pestis* genomes shed light on the early evolution, diversity, and ecology of plague. *Proceedings of the National Academy of Sciences*, 119(17), e2116722119. <https://doi.org/10.1073/pnas.2116722119>

Veale, E. (1966). *The English Fur Trade in the Later Middle Ages*. Clarendon Press.

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](https://doi.org/10.1016/S1473-3099(13)70323-2)

Walker, S. L., & Lockwood, D. N. J. (2007). Leprosy. *Clinics in Dermatology*, 25(2), 165–172. <https://doi.org/10.1016/j.clindermatol.2006.05.012>

Willmott, H., Townend, P., Swales, D. M., Poinar, H., Eaton, K., & Klunk, J. (2020). A Black Death mass grave at Thornton Abbey: the discovery and examination of a fourteenth-century rural catastrophe. *Antiquity*, 94(373), 179–196. <https://doi.org/10.15184/aqy.2019.213>

Wood, J. W., Ferrell, R. J., & Dewitte-Avina, S. N. (2003). The Temporal Dynamics of the Fourteenth-Century Black Death: New Evidence from English Ecclesiastical Records. *Human Biology*, 75(4), 427–448. <https://doi.org/10.1353/hub.2003.0067>

Yu, H., Spyrou, M. A., Karapetian, M., Shnaider, S., Radzevičiūtė, R., Nägele, K., Neumann, G. U., Penske, S., Zech, J., Lucas, M., LeRoux, P., Roberts, P., Pavlenok, G., Buzhilova, A., Posth, C., Jeong, C., & Krause, J. (2020). Paleolithic to Bronze Age Siberians Reveal Connections with First Americans and across Eurasia. *Cell*, 181(6), 1232–1245.e20. <https://doi.org/10.1016/j.cell.2020.04.037>

Zhang, F., Liu, H., Chen, S., Wang, C., Zhu, C., Zhang, L., Chu, T., Liu, D., Yan, X., & Liu, J. (2009). Evidence for an association of HLA-DRB1*15 and DRB1*09 with leprosy and the impact of DRB1*09 on disease onset in a Chinese Han population. *BMC Medical Genetics*, 10(1), 133. <https://doi.org/10.1186/1471-2350-10-133>

Material & Methods

An overview of the material and methods used in this dissertation can be seen in Figure 1. More detailed descriptions can be found in the Material and Methods sections of the corresponding chapters.

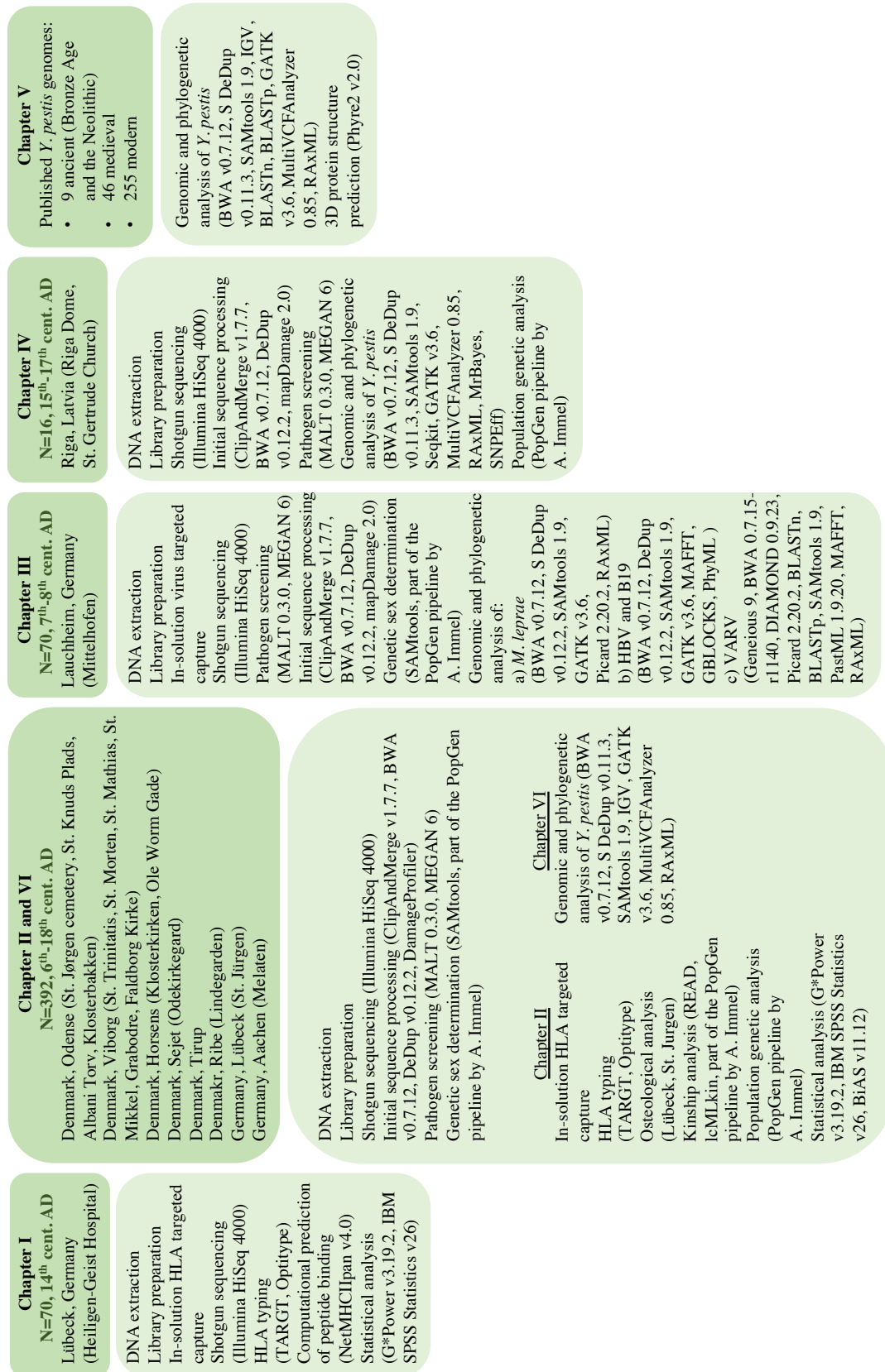


Figure 1. Material (dark green) and methods (light green) implemented in all six chapters of this dissertation.

PART I: PATHOGEN-DRIVEN EVOLUTION OF HUMAN IMMUNITY

Chapter 1: Ancient DNA study in medieval Europeans shows an association between HLA-DRB1*03 and paratyphoid fever

Frontiers in Immunology 2021

doi: 10.3389/fimmu.2021.691475

Supplementary material for this chapter can be found here:

<https://www.frontiersin.org/articles/10.3389/fimmu.2021.691475/full#supplementary-material>



Ancient DNA Study in Medieval Europeans Shows an Association Between HLA-DRB1*03 and Paratyphoid Fever

Magdalena Haller^{1†}, Joanna H. Bonczarowska^{1†}, Dirk Rieger², Tobias L. Lenz^{3,4}, Almut Nebel¹ and Ben Krause-Kyora^{1*}

¹ Institute of Clinical Molecular Biology, Kiel University, Kiel, Germany, ² Department of Archaeology, Hanseatic City of Lübeck Historic Monuments Protection Authority, Lübeck, Germany, ³ Research Group for Evolutionary Immunogenomics, Max Planck Institute for Evolutionary Biology, Plön, Germany, ⁴ Research Unit for Evolutionary Immunogenomics, Department of Biology, Universität Hamburg, Hamburg, Germany

OPEN ACCESS

Edited by:

Pedro Jose Esteves,
Centro de Investigação em
Biodiversidade e Recursos Genéticos
(CIBIO-InBIO), Portugal

Reviewed by:

Jean Roudier,
Aix Marseille Université, France
Javier Santander,
Memorial University of Newfoundland,
Canada

*Correspondence:

Ben Krause-Kyora
b.krause-kyora@ikmb.uni-kiel.de

[†]These authors have contributed
equally to this work and share
first authorship

Specialty section:

This article was submitted to
Comparative Immunology,
a section of the journal
Frontiers in Immunology

Received: 06 April 2021

Accepted: 28 June 2021

Published: 15 July 2021

Citation:

Haller M, Bonczarowska JH,
Rieger D, Lenz TL, Nebel A
and Krause-Kyora B (2021)
Ancient DNA Study in Medieval
Europeans Shows an Association
Between HLA-DRB1*03
and Paratyphoid Fever.
Front. Immunol. 12:691475.
doi: 10.3389/fimmu.2021.691475

Outbreaks of infectious diseases repeatedly affected medieval Europe, leaving behind a large number of dead often inhumed in mass graves. Human remains interred in two burial pits from 14th century CE Germany exhibited molecular evidence of *Salmonella enterica* Paratyphi C (S. Paratyphi C) infection. The pathogen is responsible for paratyphoid fever, which was likely the cause of death for the buried individuals. This finding presented the unique opportunity to conduct a paratyphoid fever association study in a European population. We focused on HLA-DRB1*03:01 that is a known risk allele for enteric fever in present-day South Asians. We generated HLA profiles for 29 medieval S. Paratyphi C cases and 24 contemporaneous controls and compared these to a modern German population. The frequency of the risk allele was higher in the medieval cases (29.6%) compared to the contemporaneous controls (13%; $p = 0.189$), albeit not significantly so, possibly because of small sample sizes. Indeed, in comparison with the modern controls ($n = 39,689$; 10.2%; $p = 0.005$) the frequency difference became statistically significant. This comparison also suggested a slight decrease in the allele's prevalence between the medieval and modern controls. Up to now, this is the first study on the genetic predisposition to *Salmonella* infection in Europeans and the only association analysis on paratyphoid fever C. Functional investigation using computational binding prediction between HLA variants and S. Paratyphi and S. Typhi peptides supported a reduced recognition capacity of bacterial proteins by DRB1*03:01 relative to other common DRB1 variants. This pattern could potentially explain the disease association. Our results suggest a slightly reduced predisposition to paratyphoid fever in modern Europeans. The causative allele, however, is still common today, which can be explained by a trade-off, as DRB1*03:01 is protective against infectious respiratory diseases such as severe respiratory syndrome (SARS). It is thus possible that the allele also provided resistance to corona-like viruses in the past.

Keywords: *Salmonella enterica* Paratyphi C, enteric fever, human leukocyte antigen, trade-off, antigen binding prediction

INTRODUCTION

In Europe, infectious diseases were particularly common in the Middle Ages (1–6), occasionally reaching immense proportions and wreaking havoc across the whole continent. For instance, the infamous Black Death and subsequent plague outbreaks left behind millions of dead who, due to the scale of the pandemic, were often inhumed in mass burials (7). Studying past contagions is not only of interest from the historical perspective, but also important for our understanding of current and future epidemics. Excavation on the grounds of the medieval Heiligen-Geist-Hospital (HGH) in Lübeck, northern Germany, revealed the presence of two mass graves and two smaller burial pits (8–10). Molecular analysis of the human remains (1370 – 1400 CE) indicated that individuals buried in the pits were infected with *Salmonella enterica* subsp. *enterica* serovar Paratyphi C at the time of death (11). The pathogen is a causative agent of paratyphoid fever, which is a life-threatening illness transmitted through contaminated food and water (12). Nowadays *S. Paratyphi C* is virtually absent in Europe, however, it can be assumed that the pathogen's prevalence used to be higher in the past. Discovery of *S. Paratyphi C* in medieval Germany (11) and Norway (13) suggests a change in the geographic distribution of *Salmonella enterica* species over time. It is thus possible that paratyphoid fever used to have a substantial impact on the health of Europeans at that time (11). As the pathogen's genome is thought to have been stable for millennia (13), genetic factors of the host could have contributed to the disappearance of this infectious disease in modern Europe.

Apart from *S. Paratyphi C*, paratyphoid fever can be caused by two other *Salmonella* serovars (*S. Paratyphi A* and *B*). Clinically similar typhoid fever is caused by *Salmonella Typhi* and together with *S. Paratyphi* infections, the diseases are known under the umbrella term of enteric fever. Several genetic host factors have been previously associated with susceptibility to the disease (14–17). Among those factors is the polymorphic human leukocyte antigen (HLA) region. The high level of HLA variation together with a number of associations between HLA alleles and a variety of infectious diseases [e. g. (18–24)] suggests that pathogens have acted as a powerful selective pressure on the human genome as a result of host-pathogen co-evolution (25–27). Regarding enteric fever, the HLA alleles DRB1*03:01 and DRB1*04:05 were identified as risk and protective factors, respectively, in present-day Vietnamese and Nepalese (16, 17). However, these associations were only shown for *S. Typhi* and *S. Paratyphi A* infections; as of yet, there are no studies exploring the genetic predisposition to *S. Paratyphi C* infections in modern populations. Nowadays the disease is very rare in Europe and usually a result of migrations from other regions of the world (12, 28, 29). Thus, analysis of susceptibility to paratyphoid fever in European populations is only possible with an ancient DNA (aDNA) approach. In this study, we generated HLA genotype data for 53 skeletal remains from the HGH site and in particular examined the frequency of the DRB1*03:01 allele to assess whether variation at the DRB1 locus predisposed medieval Europeans to enteric infection.

MATERIALS AND METHODS

Materials

The skeletal material analyzed in this study was excavated from four mass burials (contexts 4528, 4529, 4562 and 4571) located near the Heiligen-Geist-Hospital ("Hospital of the Holy Ghost", HGH) in Lübeck, northern Germany (8–10). The graves were dug between 1340 and 1370 CE (11).

DNA Extraction and Library Preparation

Seventy individuals from the HGH site were analyzed in this study [Tables S1, Table S3 (11) for details on samples]. We sampled almost exclusively teeth and used the whole specimens for aDNA extraction. The DNA was isolated with the BioRobot® EZ1 Advanced applying a custom-made version of the Large-Volume-Protocol and the EZ1 DNA Investigator® Kit (Qiagen) as described previously (11, 30). The DNA extracts were then converted into indexed partial Uracil-DNA Glycosylase (UDG)-treated libraries. The libraries were sequenced on an Illumina HiSeq 4000 platform. To reliably exclude contamination, all procedures were conducted in clean rooms specifically dedicated to ancient DNA analysis, and negative controls were included in both the extraction and library preparation steps. Furthermore, short tandem repeat profiles at seven loci were generated as an additional authentication criterion and to avoid double sampling Table S2 (11).

HLA Targeted Capture and Typing

In-solution enrichment of the HLA regions was successful for 53 of the 70 samples analyzed (excavation context 4528, $n = 13$; 4529, $n = 11$; 4562, $n = 6$; 4571, $n = 23$). Genotyping was performed using the semi-automated HLA-typing pipeline TARGT (Targeted Analysis of sequencing Reads for GenoTyping), which was designed for the analysis of low-coverage sequences as found in aDNA extracts (31). The pipeline automatically identifies and sorts target-specific reads from sequence data. Subsequently, the sorted sequences were manually analyzed to determine the HLA alleles (HLA calling) in individual samples (31). In this study, the manual genotyping was done by three researchers independently, thus increasing the reliability of the results. In addition, we applied the algorithm OptiType (32), which is only available for class I genes, to verify class I allele calls obtained with the TARGT pipeline. The OptiType software (32) was used with default parameters. Only those samples that had alleles concordantly called by both methods were included in the results (Table S1).

Computational Prediction of Peptide Binding by HLA

Full proteomes of *Salmonella enterica* Paratyphi C RKS4594 (accession NC_012125.1, France) and *S. enterica* Typhi CT18 (accession NC_003198.1, Vietnam) were used to predict the binding of 15mer peptides (967,163 and 1,201,560 possible 15mers from *S. Paratyphi* and *S. Typhi*, respectively) for all DRB1 variants with an allele frequency >1% in the present-day German population (AlleleFrequencies.net Database (AFND): Germany, Population 8, $n = 39,689$; 33). Peptide binding

prediction was performed with the established algorithm NetMHCIIpan v4.0, using the default threshold for ‘strong binding’ (33). For each HLA-DRB1 variant, the total number of bound peptides overall as well as the total number of bacterial proteins ‘recognized’ (at least one peptide of a given protein predicted to be bound by the given HLA-DRB1 variant) were obtained.

Statistical Analysis

Fisher’s exact test and odds ratio calculations were performed for the DRB1*03 allele using IBM SPSS Statistics (Version 26) predictive analytics software. Power calculations on required sample sizes were calculated with the software G*Power, v3.1.9.2 (34). For DRB1*03:01, the power calculation showed that an association analysis was appropriate with our sample size. Assuming an OR of 4 (17), an allocation ratio of 10 and an allele frequency of 10.2% in modern controls (35), 28 alleles were required in cases to obtain an *a priori* power of 75% at a significance level of 0.05.

RESULTS

In this study, we generated HLA genotype data from the skeletal remains of 53 individuals excavated from the HGH mass burial in Lübeck. The site consists of two mass graves (archaeological contexts 4528 and 4529) and two burial pits (4562 and 4571, **Figure 1**). Of the 70 metagenomic samples analyzed, 24 from 4528/4529 and 29 from 4562/4571 yielded HLA data after in-solution DNA enrichment. Genotyping of the HLA alleles was performed for the class I (HLA-A, -B and -C) and class II (HLA-DQB1 and -DRB1) loci using the TARGT pipeline which is well suited to the analysis of low-coverage sequences (31) (**Table S1**). Despite such a dedicated pipeline, the highly degraded nature of aDNA (36) poses a severe challenge to HLA typing so that allele calling at the preferred 2nd-field (4-digit) resolution is not always possible. Nevertheless, even at 1st-field level of resolution, HLA

typing can be highly informative, as it distinguishes among major functional groups (‘supertypes’) of classical HLA alleles. For the same reason, it was also not always possible to call alleles at both haplotypes, so that the total number of called alleles was lower than 2N. The investigation here was performed following the strategy outlined for an association study with aDNA data (20). As the individuals buried in the pits (4562/4571) had been infected with *S. Paratyphi* C (11), they were considered cases in the association test (with 27 successfully called HLA-DRB1 alleles across the 29 individuals). The individuals buried in the mass graves (4528/4529) were used as contemporaneous controls (with 23 called HLA-DRB1 alleles across the 24 individuals), based on the fact that no molecular traces of the focal pathogen were observed in the remains (11). In a secondary analysis, allele frequencies in the historical cases were compared with a large cohort of modern individuals (AFND: Germany, Population 8, *n* = 39,689; 33).

As DRB1*04:05 is protective against enteric fever, its frequency would be expected to be lower relative to the modern controls (0.45%, Germany, AFND Population 8, *n* = 39,689; 33). Thus, DRB1*04:05 was not investigated due to the case sample being too small for a meaningful statistical analysis based on the power calculation. On the contrary, for the DRB1*03:01 risk allele, an association study was feasible making it the main focus of this investigation. Out of the 11 individuals with a DRB1*03 allele call at 1st-field resolution, calls at 2nd-field resolution were possible for four individuals who consistently showed DRB1*03:01. Among the known alleles of the DRB1*03 allele group, DRB1*03:01 is by far the most common in the modern German population (freq = 10.2%), with the next frequent allele being two orders of magnitude less frequent (DRB1*03:02, freq = 0.1%). It is therefore highly likely that the other seven individuals with a DRB1*03 call also carried the DRB1*03:01 allele, justifying an association analysis based on the DRB1*03 calls. Moreover, four individuals were possibly homozygous for the DRB1*03 type, given the available read depth and lack of other DRB1 sequences. However, a conservative approach was used here and these individuals were regarded as



FIGURE 1 | Heiligen-Geist-Hospital (“Holy-Ghost-Hospital”) archaeological context 4562 in which *S. Paratyphi* C-positive individuals were found.

heterozygotes. It is thus likely that the obtained frequency is underestimated. The frequency of the DRB1*03 allele in the cases (29.6%) was considerably higher in comparison to the contemporaneous controls (13%, $p = 0.189$), albeit it did not reach statistical significance, likely owing to the small sample sizes for the medieval groups. In line with this suggestive association result, the comparison of medieval cases with modern individuals yielded a frequency difference that was highly significant (10.2%; $p = 0.005$, OR = 3.61) (**Table S1**). Interestingly, all four genotypes with a possible 2nd field allele call of DRB1*03:01 also showed the alleles DQB1*02:01 or DQB1*02 at the neighboring HLA-DQB1 locus (**Table S1**), which could be an indication of genetic linkage between DRB1*03:01 and DQB1*02:01, a common haplotype in present-day populations (37).

Given the role of HLA molecules in peptide presentation for antigen-specific immunity, the binding of *S. Paratyphi* and *S. Typhi* peptides to common HLA-DRB1 variants was also explored, including the risk variant DRB1*03:01. Using computational peptide binding prediction, no particular difference in the total number of bound bacterial peptides was observed between the risk variant and the other DRB1 variants (**Figure 2A**). However, the number of ‘recognized’ bacterial proteins (defined as proteins of which at least one peptide was bound by a given HLA variant) was among the lowest for DRB1*03:01 compared to the other HLA variants (for both *S. Paratyphi* C and *S. Typhi*; **Figure 2B**). This peculiarity in peptide binding of DRB1*03:01 became particularly evident when the number of proteins that were uniquely missed by each DRB1 variant (i.e., were predicted to be recognized by all other DRB1 variants but not the given one) were quantified. DRB1*03:01 missed 85 *S. Paratyphi* proteins that were recognized by all other DRB1 variants, more than thrice the number compared to the other DRB1 variants (median = 26 ± 16 SD, **Figure 2C**). A similar effect was observed for *S. Typhi* (**Figure 2C**).

DISCUSSION

In this study, variation at the HLA-DRB1 locus was investigated in the context of genetic susceptibility to *S. Paratyphi* C infection in human remains from Germany dating to the 14th century CE. The DRB1*03:01 allele has been associated with an increased risk of enteric fever in modern Vietnamese and Nepalese populations (in this case caused by *S. Typhi* and *S. Paratyphi* A only) (16, 17). Up to date, no association study with an *S. Paratyphi* C infection has been performed in present-day populations, as it is relatively rare (38). However, there is evidence suggesting a high prevalence of this infectious disease in the Middle Ages (11, 13, 39).

HLA typing was successfully performed for 53 individuals excavated from four archaeological contexts: two mass graves ($n = 24$) and two smaller pits ($n = 29$). Individuals buried in the pits were found to exhibit genetic traces of infection with *S. Paratyphi* C, likely representing the cause of death. The larger mass graves were stratigraphically older and no pathogen DNA

was detected in the remains inhumed there (11). It has been previously hypothesized that individuals in the mass graves were victims of plague (9, 40). However, no evidence has been found to support these claims (11). In addition, to the best of our knowledge, no link between HLA variation and plague susceptibility has been reported in human association studies. Thus, even if plague was responsible for the death of the individuals in the mass graves, the disease was not expected to influence the frequencies of the examined allele. Based on these assumptions, the individuals from the mass graves were treated as contemporaneous controls for the medieval paratyphoid fever cases buried in the small pits. The cases were also compared to a representative modern German population from the AlleleFrequencies.net Database (35). As genetic continuity between the HGH Lübeck population and present-day northern Europeans was shown previously (11), it allows for such an analysis to be carried out. Furthermore, kinship was not detected among the medieval individuals (11). Our results indicate that the frequency of DRB1*03 was significantly higher in the medieval paratyphoid fever cases in comparison to modern controls ($p = 0.005$). The obtained odds ratio of 3.6 is comparable to that in previous reports for enteric fever in modern South Asian populations (16, 17). The DRB1*03 frequency in the cases (29.6%) was also substantially higher relative to medieval controls (13%). Although calls at 2nd-field resolution (i.e., DRB1*03:01) were only possible for four individuals, all DRB1*03 calls were interpreted as DRB1*03:01, since this allele is by far the most common within the DRB1*03 group in Germans today (35).

This is the first report of an allelic association with an *S. Paratyphi* C infection. However, due to the relatively small number of specimens available for the present study, an inherent limitation of aDNA studies, this finding should be confirmed in a larger sample comprised of Far Eastern populations in which the disease is still common today (12). Nevertheless, as paratyphoid fever is no longer endemic in Europe, aDNA analysis is the only approach that allows us to explore whether DRB1*03:01 is involved in genetic predisposition to an *S. Paratyphi* C infection in Europeans.

The functional follow-up analysis, using a computational algorithm for prediction of binding between HLA variants and bacterial peptides, revealed an interesting pattern for the HLA-DRB1 risk variant that could potentially explain its association with susceptibility. The variant DRB1*03:01 was predicted to miss recognition (i.e., fail to bind at least one peptide) of the largest number of bacterial proteins that were recognized by all other common DRB1 variants. However, it should be noted that the number of missed proteins was small compared to the total number of proteins in the bacterial proteomes. It thus remains to be investigated whether this peculiar binding pattern of DRB1*03:01 would significantly affect susceptibility of its carriers.

The decreased frequency of the HLA risk allele in the modern population (10.2%) relative to the medieval controls (13%) might indicate slight negative selection. The limited available sample size for ancient remains for both cases and contemporaneous controls

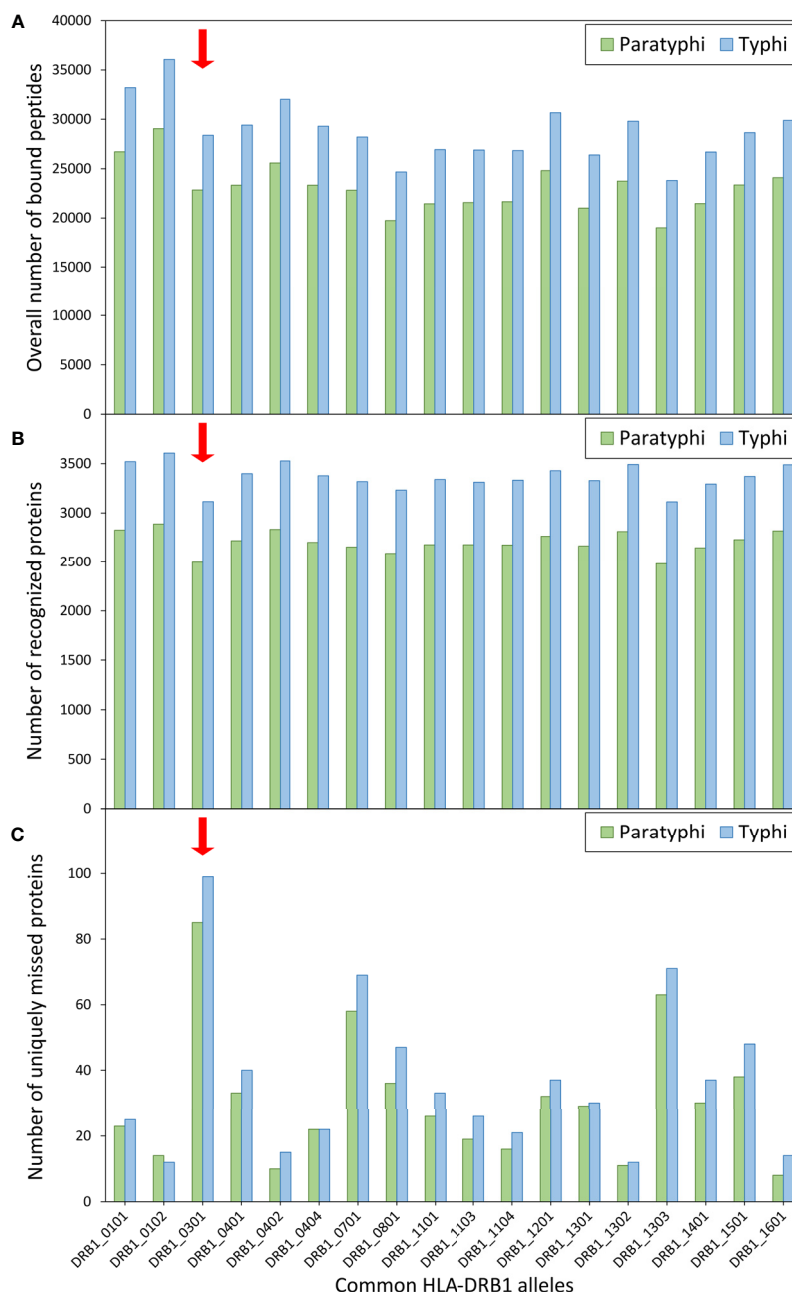


FIGURE 2 | Computationally predicted binding of common HLA-DRB1 variants to bacterial peptides. Binding prediction was performed for all HLA-DRB1 variants with allele frequency >1% in present-day Germany (n=18), using the full proteomes of both *Salmonella enterica* Paratyphi C (green bars) and *S. Typhi* (blue bars). For each DRB1 variant, **(A)** the overall number of bound peptides, **(B)** the total number of 'recognized' proteins (at least one peptide bound by a given DRB1 variant), and **(C)** the number of uniquely 'missed' proteins (recognized by all other variants) are shown.

prevents any strong conclusion about the potential selective effect of this association. However, in view of the negative effect of DRB1*03:01, it is tempting to speculate about a trade-off that might explain as to why this allele is still present at such a high frequency today. Interestingly, the allele was shown to be protective against severe acute respiratory syndrome (SARS) infection in the modern Chinese population (41–43). Although currently there is no

data confirming this association for present-day Europeans, it is possible that ancient epidemics of paratyphoid fever affected the level of genetic resistance against SARS, predisposing present-day Europeans to an infection with the causative coronavirus. There is further evidence linking DRB1*03:01 with various respiratory diseases such as resistance to tuberculosis (44). It is therefore conceivable that this allele was protective against corona-like

viruses or other infectious agents in the past. In addition, it has been suggested that DRB1*03:01 provides protection against allergy to animal-derived proteins (dairy and eggs) (45, 46). Recently, we discovered that the allele was absent in a Neolithic population of 23 individuals (47). This observation is striking when considered in the context of dietary changes that were introduced with the Neolithic agricultural transition. All these factors and the notable presence of DRB1*03:01 in the current gene pool indicate that this HLA allele has been subject to balancing selection. DRB1*03:01 is a risk factor for type I diabetes, autoimmune hepatitis, multiple sclerosis, celiac disease, sarcoidosis, Grave's disease, systemic lupus erythematosus and Sjögren's syndrome in modern populations across the globe (48–55). Due to the relatively low number of pathogens in the modern environment, negative effects of this HLA allele might now predominate, contributing to the increasing prevalence of inflammatory and autoimmune diseases today.

DATA AVAILABILITY STATEMENT

The datasets generated for this study can be found in the European Nucleotide Archive under the accession number PRJEB42927.

ETHICS STATEMENT

Ethical review and approval were not required for the study on human participants in accordance with the local legislation and institutional requirements.

REFERENCES

- Boldsen JL. Leprosy and Mortality in the Medieval Danish Village of Tirup. *Am J Phys Anthropol* (2005) 126:159–68. doi: 10.1002/ajpa.20085
- Dangvard Pedersen D, Milner GR, Kolmos HJ, Boldsen JL. Tuberculosis in Medieval and Early Modern Denmark: A Paleoepidemiological Perspective. *Int J Paleopathol* (2019) 27:101–8. doi: 10.1016/j.ijpp.2018.11.003
- Hopkins DR. *Greatest Killer: Smallpox in History*. Chicago: University of Chicago Press (2002). 415 p.
- Mühlemann B, Vinner L, Margaryan A, Wilhelmson H, de La Fuente Castro C, Allentoft ME, et al. Diverse Variola Virus (Smallpox) Strains Were Widespread in Northern Europe in the Viking Age. *Science* (2020) 369:eaaw8977. doi: 10.1126/science.aaw8977
- Schuenemann VJ, Avanzi C, Krause-Kyora B, Seitz A, Herbig A, Inskip S, et al. Ancient Genomes Reveal a High Diversity of Mycobacterium Lepae in Medieval Europe. *PLoS Pathog* (2018) 14:e1006997. doi: 10.1371/journal.ppat.1006997
- Spyrou MA, Keller M, Tukhbatova RI, Scheib CL, Nelson EA, Andrades Valtueña A, et al. Phylogeography of the Second Plague Pandemic Revealed Through Analysis of Historical Yersinia Pestis Genomes. *Nat Commun* (2019) 10:4470. doi: 10.1038/s41467-019-12154-0
- Willmott H, Townend P, Swales DM, Poinar H, Eaton K, Klunk J. A Black Death Mass Grave at Thornton Abbey: The Discovery and Examination of a Fourteenth-Century Rural Catastrophe. *Antiquity* (2020) 94:179–96. doi: 10.15184/aqy.2019.213
- Lüttger SA. "Archäologische Untersuchungen der Massenbestattungen am Heiligen-Geist-Hospital zu Lübeck: Auswertung der Befunde und Funde". In: M Gläser, editor. *Lübecker Schriften zu Archäologie und Kulturgeschichte* 26. Bonn: Rudolf Habelt (2002). p. 139–243.
- Prechel M. "Anthropologische Untersuchungen der Skelettreste aus einem Pestmassengrab am Heiligen-Geist-Hospital zu Lübeck". In: GP Fehring,

AUTHOR CONTRIBUTIONS

BK-K conceived and designed the study. MH generated aDNA data. MH, JB, and TL analyzed the data. MH, JB, TL, AN, and BK-K interpreted the results. DR provided data and interpretation on the archaeological material. JB, MH, TL, AN, and BK-K wrote the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

This study was supported by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) through the projects 2901391021 (CRC 1266), 437857095 and Germany's Excellence Strategy EXC2167-390884018. JB was funded by the International Max Planck Research School for Evolutionary Biology.

ACKNOWLEDGMENTS

We thank Edma Loku for her help with the genotype calling.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fimmu.2021.691475/full#supplementary-material>

- editor. *Lübecker Schriften zu Archäologie und Kulturgeschichte* 24. Bonn: Rudolf Habelt (1996). p. 323–39.
- Prechel M. "Eine Lübecker Population von 1350. Krankheiten und Mangelerscheinungen: Archäologische Untersuchungen auf dem Lübecker Stadthügel: Befunde und Funde". In: M Gläser, editor. *Lübecker Schriften zu Archäologie und Kulturgeschichte* 26. Bonn: Rudolf Habelt (2002). p. 245–86.
- Haller M, Callan K, Susat J, Flux AL, Immel A, Franke A, et al. Mass Burial Genomics Reveals Outbreak of Enteric Paratyphoid Fever in the Late Medieval Trade City Lübeck. *iScience* (2021) 24:102419. doi: 10.1016/j.isci.2021.102419
- Crump JA, Mintz ED. Global Trends in Typhoid and Paratyphoid Fever. *Clin Infect Dis* (2010) 50:241–6. doi: 10.1086/649541
- Zhou Z, Lundström I, Tran-Dien A, Duchêne S, Alikhan N-F, Sergeant MJ, et al. Pan-Genome Analysis of Ancient and Modern Salmonella Enterica Demonstrates Genomic Stability of the Invasive Para C Lineage for Millennia. *Curr Biol* (2018) 28:2420–28.e10. doi: 10.1016/j.cub.2018.05.058
- Ali S, Vollaard AM, Widjaja S, Surjadi C, van de Vosse E, van Dissel JT. PARK2/PACRG Polymorphisms and Susceptibility to Typhoid and Paratyphoid Fever. *Clin Exp Immunol* (2006) 144:425–31. doi: 10.1111/j.1365-2249.2006.03087.x
- Alvarez MI, Glover LC, Luo P, Wang L, Theusch E, Oehlers SH, et al. Human Genetic Variation in VAC14 Regulates Salmonella Invasion and Typhoid Fever Through Modulation of Cholesterol. *Proc Natl Acad Sci USA* (2017) 114:E7746–55. doi: 10.1073/pnas.1706070114
- Dunstan SJ, Stephens HA, Blackwell JM, Duc CM, Lanh MN, Dudbridge F, et al. Genes of the Class II and Class III Major Histocompatibility Complex Are Associated With Typhoid Fever in Vietnam. *J Infect Dis* (2001) 183:261–8. doi: 10.1086/317940
- Dunstan SJ, Hue NT, Han B, Li Z, Tram TT, Sim KS, et al. Variation at HLA-DRB1 is Associated With Resistance to Enteric Fever. *Nat Genet* (2014) 46:1333–6. doi: 10.1038/ng.3143

18. Asti M, Martinetti M, Zavaglia C, Cuccia MC, Gusberty L, Tinelli C, et al. Human Leukocyte Antigen Class II and III Alleles and Severity of Hepatitis C Virus-Related Chronic Liver Disease. *Hepatology* (1999) 29:1272–9. doi: 10.1002/hep.510290445
19. Hu J, Li L, Pang L, Chen Y, Yang L, Liu C, et al. HLA-DRB1*1501 and HLA-DQB1*0301 Alleles are Positively Associated With HPV16 Infection-Related Kazakh Esophageal Squamous Cell Carcinoma in Xinjiang China. *Cancer Immunol Immunother* (2012) 61:2135–41. doi: 10.1007/s00262-012-1281-x
20. Krause-Kyora B, Nutsua M, Boehme L, Pierini F, Pedersen DD, Kornell S-C, et al. Ancient DNA Study Reveals HLA Susceptibility Locus for Leprosy in Medieval Europeans. *Nat Commun* (2018) 9:1569. doi: 10.1038/s41467-018-03857-x
21. Mbarek H, Ochi H, Urabe Y, Kumar V, Kubo M, Hosono N, et al. A Genome-Wide Association Study of Chronic Hepatitis B Identified Novel Risk Locus in a Japanese Population. *Hum Mol Genet* (2011) 20:3884–92. doi: 10.1093/hmg/ddr301
22. Sveinbjornsson G, Gudbjartsson DF, Halldorsson BV, Kristinsson KG, Gottfredsson M, Barrett JC, et al. HLA Class II Sequence Variants Influence Tuberculosis Risk in Populations of European Ancestry. *Nat Genet* (2016) 48:318–22. doi: 10.1038/ng.3498
23. Zhang F, Liu H, Chen S, Wang C, Zhu C, Zhang L, et al. Evidence for an Association of HLA-DRB1*15 and DRB1*09 With Leprosy and the Impact of DRB1*09 on Disease Onset in a Chinese Han Population. *BMC Med Genet* (2009) 10:133. doi: 10.1186/1471-2350-10-133
24. Arora J, Pierini F, McLaren PJ, Carrington M, Fellay J, Lenz TL. HLA Heterozygote Advantage Against HIV-1 Is Driven by Quantitative and Qualitative Differences in HLA Allele-Specific Peptide Presentation. *Mol Biol Evol* (2020) 37:639–50. doi: 10.1093/molbev/msz249
25. Field Y, Boyle EA, Telis N, Gao Z, Gaulton KJ, Golan D, et al. Detection of Human Adaptation During the Past 2000 Years. *Science* (2016) 354:760–4. doi: 10.1126/science.aag0776
26. Penn DJ, Damjanovich K, Potts WK. MHC Heterozygosity Confers a Selective Advantage Against Multiple-Strain Infections. *Proc Natl Acad Sci USA* (2002) 99:11260–4. doi: 10.1073/pnas.162006499
27. Fumagalli M, Sironi M, Pozzoli U, Ferrer-Admetlla A, Ferrer-Admetlla A, Pattini L, et al. Signatures of Environmental Genetic Adaptation Pinpoint Pathogens as the Main Selective Pressure Through Human Evolution. *PLoS Genet* (2011) 7:e1002355. doi: 10.1371/journal.pgen.1002355
28. Achtman M, Wain J, Weill F-X, Nair S, Zhou Z, Sangal V, et al. Multilocus Sequence Typing as a Replacement for Serotyping in *Salmonella* Enterica. *PLoS Pathog* (2012) 8:e1002776. doi: 10.1371/journal.ppat.1002776
29. Uzzau S, Brown DJ, Wallis T, Rubino S, Leori G, Bernard S, et al. Host Adapted Serotypes of *Salmonella* Enterica. *Epidemiol Infect* (2000) 125:229–55. doi: 10.1017/S0950268899004379
30. Fehren-Schmitz L, Reindel M, Cagigao ET, Hummel S. Pre-Columbian Population Dynamics in Coastal Southern Peru: A Diachronic Investigation of mtDNA Patterns in the Palpa Region by Ancient DNA Analysis. *Am J Phys Anthropol* (2010) 141:208–21. doi: 10.1002/ajpa.21135
31. Pierini F, Nutsua M, Böhm L, Özer O, Bonczarowska J, Susat J, et al. Targeted Analysis of Polymorphic Loci From Low-Coverage Shotgun Sequence Data Allows Accurate Genotyping of HLA Genes in Historical Human Populations. *Sci Rep* (2020) 10:7339. doi: 10.1038/s41598-020-64312-w
32. Szolek A, Schubert B, Mohr C, Sturm M, Feldhahn M, Kohlbacher O. OptiType: Precision HLA Typing From Next-Generation Sequencing Data. *Bioinformatics* (2014) 30:3310–6. doi: 10.1093/bioinformatics/btu548
33. Reynisson B, Alvarez B, Paul S, Peters B, Nielsen M. NetMHCpan-4.1 and NetMHCIPan-4.0: Improved Predictions of MHC Antigen Presentation by Concurrent Motif Deconvolution and Integration of MS MHC Eluted Ligand Data. *Nucleic Acids Res* (2020) 48:W449–54. doi: 10.1093/nar/gkaa379
34. Faul F, Erdfelder E, Buchner A, Lang A-G. Statistical Power Analyses Using G*Power 3.1: Tests for Correlation and Regression Analyses. *Behav Res Methods* (2009) 41:1149–60. doi: 10.3758/BRM.41.4.1149
35. Gonzalez-Galarza FF, McCabe A, Santos EJ, Jones J, Takeshita L, Ortega-Rivera ND, et al. Allele Frequency Net Database (AFND) 2020 Update: Gold-Standard Data Classification, Open Access Genotype Data and New Query Tools. *Nucleic Acids Res* (2020) 48:D783–8. doi: 10.1093/nar/gkz1029
36. Dabney J, Meyer M, Pääbo S. Ancient DNA damage. *Cold Spring Harb Perspect Biol* (2013) 5:a012567. doi: 10.1101/cshperspect.a012567
37. Zhang R, Duffy BF, Lange V, Eby CS, Liu C. Association Between the HLA-DRB1*03:01-DQB1*02:01 Haplotype and PF4/heparin Antibodies. *Blood Adv* (2019) 3:3136–42. doi: 10.1182/bloodadvances.2019000311
38. Appiah GD, Hughes MJ, Chatham-Stephens K. “Typhoid & Paratyphoid Fever”. In: *CDC Yellow Book 2020: Health Information for International Travel*. New York: Oxford University Press (2020).
39. Vågene ÅJ, Herbig A, Campana MG, Robles García NM, Warinner C, Sabin S, et al. *Salmonella* Enterica Genomes From Victims of a Major Sixteenth-Century Epidemic in Mexico. *Nat Ecol Evol* (2018) 2:520–8. doi: 10.1038/s41559-017-0446-6
40. Koppmann K. Die Chroniken der niedersächsischen Städte, Lübeck, Band 2. Leipzig: S. Hirzel (1899).
41. Wang S-F, Chen K-H, Chen M, Li W-Y, Chen Y-J, Tsao C-H, et al. Human-Leukocyte Antigen Class I Cw 1502 and Class II DR 0301 Genotypes are Associated With Resistance to Severe Acute Respiratory Syndrome (SARS) Infection. *Viral Immunol* (2011) 24:421–6. doi: 10.1089/vim.2011.0024
42. Ng MH, Lau K-M, Li L, Cheng S-H, Chan WY, Hui PK, et al. Association of Human-Leukocyte-Antigen Class I (B*0703) and Class II (DRB1*0301) Genotypes With Susceptibility and Resistance to the Development of Severe Acute Respiratory Syndrome. *J Infect Dis* (2004) 190:515–8. doi: 10.1086/421523
43. Sanchez-Mazas A. HLA Studies in the Context of Coronavirus Outbreaks. *Swiss Med Wkly* (2020) 150:w20248. doi: 10.4414/sm.w.2020.20248
44. Malkova A, Starshinova A, Zinchenko Y, Basantsova N, Mayevskaya V, Yablonskiy P, et al. The Opposite Effect of Human Leukocyte Antigen Genotypes in Sarcoidosis and Tuberculosis: A Narrative Review of the Literature. *ERJ Open Res* (2020) 6:00155–2020. doi: 10.1183/23120541.00155-2020
45. Dimitrov I, Doytchinova I. Associations Between Milk and Egg Allergens and the HLA-DRB1/DQ Polymorphism: A Bioinformatics Approach. *Int Arch Allergy Immunol* (2016) 169:33–9. doi: 10.1159/00044172
46. Ogródowczyk AM, Dimitrov I, Wróblewska B. Two Faces of Milk Proteins Peptides With Both Allergenic and Multidimensional Health Beneficial Impact- Integrated. In Vitro/In Silico Approach. *Foods* (2021) 10:163. doi: 10.3390/foods10010163
47. Immel A, Pierini F, Rinne C, Meadows J, Barquera R, Szolek A, et al. Genome-Wide Study of a Neolithic Wartberg Grave Community Reveals Distinct HLA Variation and Hunter-Gatherer Ancestry. *Commun Biol* (2021) 4:113. doi: 10.1038/s42003-020-01627-4
48. Chen QY, Huang W, She JX, Baxter F, Volpe R, McLaren NK. HLA-DRB1*08, DRB1*03/DRB3*0101, and DRB3*0202 are Susceptibility Genes for Graves’ Disease in North American Caucasians, Whereas DRB1*07 is Protective. *J Clin Endocrinol Metab* (1999) 84:3182–6. doi: 10.1210/jcem.84.9.5991
49. Cruz-Tapias P, Pérez-Fernández OM, Rojas-Villarraga A, Rodríguez-Rodríguez A, Arango M-T, Anaya J-M. Shared HLA Class II in Six Autoimmune Diseases in Latin America: A Meta-Analysis. *Autoimmune Dis* (2012) 2012:569728. doi: 10.1155/2012/569728
50. Fernando MM, Stevens CR, Sabeti PC, Walsh EC, McWhinnie AJ, Shah A, et al. Identification of Two Independent Risk Factors for Lupus Within the MHC in United Kingdom Families. *PLoS Genet* (2007) 3:e192. doi: 10.1371/journal.pgen.0030192
51. Kwon OJ, Karni A, Israel S, Brautbar C, Amar A, Meiner Z, et al. HLA Class II Susceptibility to Multiple Sclerosis Among Ashkenazi and non-Ashkenazi Jews. *Arch Neurol* (1999) 56:555–60. doi: 10.1001/archneur.56.5.555
52. Morris DL, Taylor KE, Fernando MM, Nititham J, Alarcón-Riquelme ME, Barcellos LF, et al. Unraveling Multiple MHC Gene Associations With Systemic Lupus Erythematosus: Model Choice Indicates a Role for HLA Alleles and Non-HLA Genes in Europeans. *Am J Hum Genet* (2012) 91:778–93. doi: 10.1016/j.ajhg.2012.08.026
53. Nakken B, Jonsson R, Brokstad KA, Omholt K, Nerland AH, Haga HJ, et al. Associations of MHC Class II Alleles in Norwegian Primary Sjögren’s

- Syndrome Patients: Implications for Development of Autoantibodies to the Ro52 Autoantigen. *Scand J Immunol* (2001) 54:428–33. doi: 10.1046/j.1365-3083.2001.00993.x
54. Shankarkumar U, Ghosh K, Badakere SS, Mohanty D. HLA-DRB1*03 and DQB1*0302 Associations in a Subset of Patients Severely Affected With Systemic Lupus Erythematosus From Western India. *Ann Rheum Dis* (2003) 62:92–3. doi: 10.1136/ard.62.1.92
 55. van Gerven NM, de Boer YS, Zwiers A, Verwer BJ, Drenth JP, Van Hoek B, et al. HLA-DRB1*03:01 and HLA-DRB1*04:01 Modify the Presentation and Outcome in Autoimmune Hepatitis Type-1. *Genes Immun* (2015) 16:247–52. doi: 10.1038/gene.2014.82

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2021 Haller, Bonczarowska, Rieger, Lenz, Nebel and Krause-Kyora. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Chapter 2: Human Leukocyte Antigens (HLA) and leprosy – a genetic association study in medieval Europeans

This chapter was prepared as a hybrid of results in press and a dissertation chapter. Examination of the skeletal collection from Lübeck St. Jürgen is published as a chapter (in German) in Rieger, D., Fuchs, K., Bonczarowska, J. H. & Krause-Kyora, B. *pauperibus leprosis in (...) Sanctum Georgium – eine Stiftung der Kaufleute des Lübecker Gründungsviertels*. In: Rieger, D. & Schneider, M. (ed.) *Ausgrabungen im Lübecker Gründungsviertel II. Archäoparasitologie, Handelsgeschichte, Paläopathologie und Anthropologie*. Lübeck 2022.

Supplementary material can be found here: <https://doi.org/10.57892/100-4>

Human Leukocyte Antigens (HLA) and leprosy - a genetic association study in medieval Europeans

Abstract

Leprosy is an infectious disease which mostly affects superficial tissues, such as peripheral nerves, skin and eyes. It is a chronic, highly disfiguring and debilitating illness caused by *Mycobacterium leprae* (*M. leprae*). Europe struggled with leprosy in the Middle Ages, as thousands of institutions were built for the purpose of isolating leprosy-affected individuals from society. Nowadays, the disease is absent in Europe and genetic factors have been considered as a major contributing factor to leprosy disappearance. As leprosy has virtually vanished from the continent, not much is known about the genetic predisposition to *M. leprae* infection in Europeans. In this study, an exploratory case-control association analysis between Human Leukocyte Antigen (HLA) alleles and leprosy was performed in a large sample of individuals whose skeletal remains were excavated from 17 medieval cemeteries across Denmark (N=15) and Germany (N=2). The “case vs. control” status was assessed based on the presence of skeletal lesions pathognomonic for leprosy and/or molecular evidence of *M. leprae* infection. The analyzed dataset comprised HLA profiles of 300 individuals (99 cases and 201 controls). Five HLA alleles were found to be nominally associated with leprosy susceptibility. HLA-A*23 ($p=0.001$, OR=16.9), A*26 ($p=0.037$, OR=3.2) and B*38 ($p=0.009$, OR=12.6) were identified as risk factors, whilst HLA-C*06 ($p=0.044$, OR=0.4) and DRB1*13 ($p=0.022$, OR=0.48) were associated with leprosy protection. These findings support the genetic contribution to leprosy susceptibility in Europeans. Interestingly, when comparing HLA frequencies between the medieval controls and a modern German population, a temporal change was noted for several alleles. These variants were previously associated with common infectious diseases as well as autoimmune conditions in modern non-European and European populations. The pattern of the associations suggests an evolutionary trade-off between susceptibility to infectious diseases and autoimmunity, which likely reflects the influence of ancient epidemics on the HLA diversity in present-day Europe.

Introduction

Leprosy (Hansen's disease) is a chronic infectious disease caused by *Mycobacterium leprae* (*M. leprae*). Infection with this bacterium usually occurs via the respiratory route (Noordeen, 1994; Cole et al., 2001) and can affect individuals at any stage of life (de Vries & Perry, 1985). Leprosy mainly influences the skin and superficial structures, such as mucous membranes, eyes, testis and peripheral nerves (Aufderheide and Rodriguez-Martin, 1998). Over time, chronic rhinitis can cause the nose to collapse (saddle nose) and chronic inflammation of the ocular nerve may lead to blindness (Andersen and Manchester, 1992). Neurologically, leprosy also manifests as sensory loss in hands and feet (Britton, 1998; Rambukkana et al., 2002). Such numb areas of the skin are susceptible to injuries and secondary infection, which can result in deformation and complete loss of extremities, making leprosy a highly debilitating disease.

In Europe, leprosy has been virtually absent since the 16th cent. AD, with only a few cases being reported in Norway until the 19th century (Vogelsang, 1965; Stone et al., 2009). Although nowadays Europeans are free of leprosy, the disease was very common on the continent in the Middle Ages, with the prevalence peak between the 11th to 14th cent. AD (Boldsen and Møllerup, 2006; Schuenemann et al., 2013; Taylor et al., 2013). In the medieval period, disability was feared and often thought to be a punishment for sins. This led to ostracism and consequent isolation of leprosy-affected individuals from society. For this purpose, specialized institutions called leprosaria were established (Mazzi, 1992; Raffcliffe, 2006). As leprosy is a lifelong disease, leprosaria became a new home for the sick who were often eventually buried at the associated cemeteries. Although leprosy *per se* is not considered as a deadly disease (Bryceson and Pfaltzgraff, 1990), complications, such as amyloidosis, can occur and result in death (Klioze, 2000). Coupled with hormonal disturbances (Gunawan et al., 2020; Khanna et al., 2014; Saporta & Yuksel, 1994) and social ostracism, the disease likely hindered the reproductive success of leprosy-affected individuals, possibly leading to negative selection of leprosy-associated genetic risk factors in Europeans.

Several genetic association studies have identified genetic variants conferring susceptibility to or protection against the disease (e.g., Ali et al., 2003; Alvarado-Navarro et al., 2008; Malhotra et al.,

2005; Sousa et al., 2012; Wang et al., 2016) for populations in which it is endemic today, such as India, China, Brazil and Mexico. Many associations were reported for variants in the Human Leukocyte Antigen (HLA) region (Tab. S1). These variants are the most often studied polymorphisms with regard to infectious disease susceptibility due to the important role of HLA molecules in antigen recognition and presentation (Blackwell et al., 2009). As Europe has been free of the disease for centuries, all modern association studies were carried out in non-European populations. Therefore, not much is known about the genetic predisposition to leprosy and *M. leprae*-driven selection effects in present-day Europeans. However, analysis of medieval European HLA alleles could help answer the questions regarding the leprosy disappearance from Europe and provide a better understanding of the human-pathogen coevolution.

So far, only one aDNA leprosy case-control association study on medieval Europeans has been performed, on individuals whose skeletal remains were excavated from various cemeteries across Denmark and Germany (Krause-Kyora et al., 2018). Identification of the leprosy-affected individuals (cases) was possible due to the presence of skeletal lesions characteristic for severe leprosy (lepromatous leprosy, LL) (Møller-Christensen 1952; 1974; 1978) and molecular evidence of an *M. leprae* infection. Moreover, these individuals were buried near a well-known medieval leprosarium (St. Jørgen) in Odense (Boldsen, 2001). Another institution recognized for housing leprosy-affected individuals in the Middle Ages was the Melaten leprosarium, located in Aachen, Germany (1230-1550 AD) (Prescher & Wagner, 2016).

In the present study, individuals from cemeteries associated with these two leprosoria were considered as potential cases. Conversely, skeletal remains from 14 medieval parish graveyards were treated as potential controls (healthy individuals). The nature of individuals buried at an additional archaeological site included in this study – a cemetery adjacent to the St. Jürgen hospital in Lübeck (1270-1550 AD, Germany) – was unclear. Historical sources suggested the possibility that the hospital could have been used to house people suffering from leprosy (Schulze, 1993). Thus, the case vs control status of individuals from Lübeck had to be established based on osteological and molecular analysis of the skeletal remains prior to the association analysis.

The aim of this study was to explore the association between HLA variants and leprosy in the European population on a large sample of 392 medieval Europeans (Danish and German). The objectives included:

- 1) obtaining the material for the analysis:
 - DNA extraction and library preparation
 - Next Generation Sequencing: untargeted sequencing of the samples for metagenomic analyses (pathogen screening and population genetic analysis)
- 2) determination of leprosy cases and controls based on the:
 - screening of the metagenomic samples for the presence of *M. leprae* DNA (all samples)
 - information provided about skeletal lesions (Melaten in Aachen and St. Jørgen in Odense)
 - osteological analysis (St. Jørgen in Lübeck)
- 3) confirmation of the genetic continuity across medieval and modern Europeans via principal component analysis
- 4) HLA targeted capture and sequencing
- 5) HLA allele typing and frequency calculation
- 6) statistical association analyses

Material

The analyzed sample comprised 526 bones and teeth belonging to 392 individuals (Tab. S2) from 17 medieval cemeteries across Denmark and Germany (Tab. 1, Supplementary Data 1) that were provided by Prof. Jesper Boldsen (Institute of Forensic Medicine at the University of Southern Denmark; Danish samples), Dr Dirk Rieger (Archaeology and Monument Preservation Lübeck; Lübeck samples) and Prof. Andreas Prescher (Uniklinik RWTH Aachen; Melaten samples). The archaeological dates of the sites reported in this chapter were obtained from written sources (Prescher & Wagner, 2016; Siegfried, 2018) and personal communication (Boldsen, 2018).

Table 1. Number of examined individuals per archaeological site.

Site	Location	Dating (year AD)	N of individuals
Denmark			
St. Jørgen cemetery	Odense	1270-1550	68
St. Knuds Plads		1150-1550	2
Albani Torv		1150-1550	8
Klosterbakken		1150-1550	1
St. Trinitatis	Viborg	1150-1550	15
St. Morten		1250-1529	4
St. Mathias		1150-1550	3
St. Mikkel		1100-1529	12
Grabodre		1150-1550	3
Faldborg Kirke		1200-1600	4
Klosterkirken	Horsens	1600-1800	52
Ole Worm Gade		EMA-1480	26
Odekirkegard	Sejet	1150-1550	32
Tirup	Tirup	1150-1350	31
Lindegarden	Ribe	1050-1738	28
Germany			
St. Jürgen	Lübeck	1270-1550	34
Melaten	Aachen	1230-1550	69
Total	-	EMA-1738	392

EMA – Early Middle Ages (6th – 10th cent. AD)

Methods

Osteological analysis of skeletal remains from Lübeck St. Jürgen

The skeletal collection from Lübeck comprised 1160 commingled skeletal elements. Due to the highly challenging nature of the collection, a new protocol was established to record all skeletal elements. All bone fragments were subsequently macroscopically analyzed. Both written and photographic documentation of the remains were prepared. Osteological analysis included sex and age assessment, where possible, as well as estimation of the minimum number of individuals present. All observed pathological lesions were described.

Osteological sex was estimated for available crania, mandibles and pelvic bones, which are commonly used for sex assessment (Klepinger, 2006; White & Folkens, 2005). For each element, several morphological features (total N=24) were assessed based on previously published methods (Ascadi & Nemeskeri, 1970; Buikstra & Ubelaker, 1994; Novotny, 1975; Phenice, 1969; Piontek, 1999) and scored according to a custom protocol (Tab. S3). The average score was then calculated for each element. In instances where postcranial elements were highly robust or extremely gracile, sex estimation was made based on the general robustness of the bone. As sexual dimorphism fully develops only after puberty, sex was not predicted for subadult individuals.

Age-at-death of subadult individuals was estimated based on the patterns of dentition development (Anderson, 1976; Moorrees, 1963; Ubelaker, 1989), epiphyseal fusion (Knussmann & Martin, 1988; Scheuer & Black, 2004) and bone growth (Gindhart, 1973; Stloukal & Hanakova, 1978). For adult individuals, cranial suture closure (Meindl & Lovejoy, 1985; Piontek, 1999), tooth wear (Lovejoy, 1985; Miles, 1963) as well as morphology of the pubic symphysis (Suchey-Brooks, 1990; Todd, 1920) and auricular surface of the ilium (Lovejoy, 1985) were examined in the context of age assessment.

One of the main goals of the osteological examination was to assess whether the individuals buried at the cemetery associated with the St. Jürgen hospital in Lübeck suffered from leprosy. A differential diagnosis of leprosy based on such an analysis is mainly possible due to characteristic

lesions that the disease leaves on the skeleton at the advanced stage of illness. The first comprehensive description of the skeletal manifestations of leprosy was a result of palaeopathological studies carried out on the remains buried on the grounds of the St. Jørgen hospital in Naestved (Møller-Christensen 1952; 1974; 1978). Based on the work of Møller-Christensen and his successors (e.g., Andersen & Manchester, 1992; Boldsen & Freund, 2006; Ortner & Putschar, 1981), it has been widely accepted that a combination of characteristic facial skeleton deformities (*facies leprosa*, rhinomaxillary syndrome) together with lesions of the extremities (atrophy of phalanges in hands and feet, periostitis of tibia and/or fibula) is considered pathognomonic of leprosy. Skeletal remains from Lübeck St. Jürgen were screened for the presence of such lesions.

DNA extraction, library preparation and sequencing

All DNA samples were extracted and processed in a dedicated ancient DNA facility at the University of Kiel following the guidelines on contamination control in ancient DNA (Knapp et al., 2012; Pilli et al., 2013; Yang & Watt, 2005), according to a previously published protocol for the non-UGD treated samples (Krause-Kyora et al., 2018). Shotgun sequencing was performed on the Illumina HiSeq 6000 (2x100) platform of the Institute of Clinical Molecular Biology (IKMB) in Kiel.

Read preprocessing, mapping and authenticity evaluation

Adapter sequences were removed and paired-end reads were merged with ClipAndMerge v1.7.7. (Peltzer et al., 2016). Shotgun sequence data was mapped to the human genome (build hg19) using BWA v0.7.12 (Li & Durbin, 2009) with a reduced mapping stringency parameter “-n 0.01” to account for mismatches in aDNA. Duplicated reads were removed with DeDup v0.12.2 (Peltzer et al., 2016).

To confirm the ancient origin of the sequences, terminal damage of the reads (C to T substitutions) was assessed with DamageProfiler (Neukamm et al., 2021). After the validation, the first two positions from the 5’ and 3’-ends of the reads were trimmed.

Pathogen screening

All 526 clipped and merged metagenomic datasets were screened for the presence of *M. leprae* DNA with the Megan Alignment Tool 0.3.0 (MALT) (Herbig et al., 2016) (SemiGlobal alignment mode, identity threshold = 90%), using a custom database containing bacterial genomes available at the NCBI platform (24.01.2019). Output alignments were inspected visually in MEGAN 6 (Huson et al., 2007)

Determination of the “case vs control” status

The “case” status of individuals from St. Jørgen in Odense was confirmed in a previous study by Krause-Kyora and colleagues (Krause-Kyora et al., 2018). The initial assumption that all individuals buried at the Melaten cemetery (Aachen) were indeed suffering from leprosy was further tested based on the presence of 1) skeletal lesions strongly suggestive of leprosy, such as *facies leprosa* and atrophy of the phalanges, and 2) molecular evidence of *M. leprae* infection. Only individuals fulfilling at least one of these two criteria were considered as “strong cases” and included in the association analysis. All remaining individuals were excluded due to their uncertain disease status.

Osteological analysis and pathogen screening were used as means of confirming the “case vs. control” status for the individuals buried near the St. Jürgen hospital in Lübeck.

In the Middle Ages, leprosy-affected individuals were also buried at parish cemeteries among the general population (Boldsen, 2008; Boldsen et al., 2013). Thus, all individuals from the remaining 14 sites were screened for the presence of *M. leprae* DNA in the metagenomic samples. Unfortunately, no information on osteological lesions for these individuals was available. Thus, those exhibiting molecular evidence of infection were considered as cases and the remaining individuals were used as controls.

Sex determination and kinship analysis

Genetic sex of the analyzed individuals was determined based on the ratio of sequences aligning to the X and Y chromosomes compared to the autosomes (Fu et al., 2016).

Kinship analysis was performed using two softwares: READ (Monroy Kuhn et al., 2018) and lcMLkin (Lipatov et al., 2015). The former infers kinship based on the proportion of non-matching autosomal genotypes. LcMLkin, on the other hand, estimates identity by descent (IBD) probabilities based on genotype likelihoods. Due to the relatively poor quantity of data, characteristic for aDNA, presence of relatedness was recognized only in instances where kinship was supported by two lines of evidence (indicated by both softwares) and the number of SNPs considered in the estimation did not fall below 1000.

Population genetic analysis

Pseudo-haploid genotypes were generated for a pool of 1,233,013 SNPs, commonly used for population genetic analyses, with SequenceTools 1.2.2 (Haak et al., 2015; Lazaridis et al., 2014; Mathieson et al., 2015). The minimal threshold of single nucleotide polymorphisms (SNPs) used for assessment of genetic relationships between humans varies between aDNA studies, ranging from 10,000 (Immel et al., 2021) to 30,000 SNPs (Mathieson et al., 2015). Here, an intermediate threshold of 20,000 SNPs was set and samples with at least 20,000 typed SNPs were included in the principal component analysis (PCA). In total, genotype data for 182 medieval individuals was merged with a dataset of previously published 1138 individuals from 66 contemporary West Eurasian populations (Patterson et al., 2012) using *mergeit* from the EIGENSOFT package. The PCA was performed with *smartpca* and the 'lsqproject' option (Patterson et al., 2006).

HLA targeted capture and typing

One sample per individual was chosen for in-solution targeted enrichment of the HLA exonic regions. A sample was considered to be a good candidate for the capture when the endogenous DNA content in the shotgun data exceeded one percent and the proportion of duplicated reads

mapping to the human genome (hg19) was below 40 percent. Based on these criteria, enrichment was performed for 324 of the 392 samples.

Individuals for which the HLA capture was successful (N = 323) were then typed with the semi-automated HLA-typing pipeline TARGT (Targeted Analysis of sequencing Reads for GenoTyping), designed for the analysis of low-coverage sequences typical of aDNA (Pierini et al., 2020). Five classical HLA loci were typed: class I (HLA-A, B and C) as well as class II (HLA-DRB1 and DQB1). These genes were chosen based on previous associations with leprosy (Tab. S1). Class I HLA alleles were also called using the OptiType software (Szolek et al., 2014) with default parameters (available only for class I HLA alleles). Results obtained with OptiType were used in instances where the call made with TARGT was ambiguous (two possible options). In such cases, the allele supported by OptiType was chosen. When OptiType did not support either of the two TARGT options, the allele was not called.

Reproducibility of the manual typing using TARGT was tested on a randomly chosen set of samples. Assuming a good overall agreement level of 85% and confidence level of 95% ($\pm 7.5\%$), the sample size for the reproducibility test was calculated with BiAS v 11.12 (Mace, 1962). Consequently, 86 alleles were typed by two independent observers of different experience levels and the overall agreement [%] was calculated.

HLA allele frequency calculation and statistical analysis

HLA alleles were successfully typed for 271 individuals for which HLA capture data was obtained in this study. In addition, HLA calls for 68 individuals from Odense reported in a previous publication were included (Tab. S7 in Pierini et al., 2020). The sample sets from this and the previous analysis overlapped for 47 individuals out of which HLA calls were available for 32 samples in both studies (Tab. S4). The alleles were called using the same typing pipeline TARGT. They were, however, based on sequences from independent DNA extracts. For the overlapping individuals (N=47), the calls were merged in a way that would increase the resolution of the calls. In instances of discrepancy, the allele supported by OptiType results generated in this study was preferred (class I HLA). For class II alleles, the call of a higher resolution was selected. Moreover,

heterozygous calls were chosen over homozygous calls. For the remaining 21 individuals from the previously published dataset (Pierini et al., 2020), HLA alleles were called for 13 individuals. In the end, merging of the two datasets resulted in increasing the number of cases from Odense from 44 (successfully typed in this study) to 73 (Tab. S5). Thus, the final dataset for the HLA allele frequency calculations and association analysis consisted of 300 individuals (99 cases (Tab. S5 and S6) and 201 controls (Tab. S7)).

Exploratory statistical analysis was performed with the Fisher's exact test or Chi² test. The association analysis and calculation of the odds ratios (ORs) were done using IBM SPSS Statistics (version 26). Statistical power was calculated with G*Power v3.1.9.2 (Faul et al., 2009). Due to the limited sample size, correction for multiple testing was not applied. Thus, the associations reported here are based on nominally significant p-values.

Results

Sequencing data from 526 samples belonging to 392 individuals was obtained and analyzed in this study (Tab. S2). Skeletal remains of these individuals were uncovered from 17 medieval cemeteries located in today's Denmark (N=15) and Germany (N=2) (Tab. 1). Based on the assessment of the terminal damage of the reads (C to T substitutions) (Fig. S1), 12 individuals were excluded from further analysis. They are marked as “contaminated” in Tab. S2.

Osteological and molecular analysis of the remains from Lübeck St. Jürgen

In total, 1160 skeletal elements, excavated from eight primary burial contexts and two overlapping burial pits discovered near the St. Jürgen hospital in Lübeck, were identified as human bones.

The minimum number of individuals (MNI) present was estimated based on the number of specific bones, taking into consideration age-at-death and sex, where possible. At least 12 subadult individuals were present in the collective burial (Tab. S8), including an unborn child, two infants (0-3 years old.), one young child (3-6 y.o.), two older children (6-12 y.o.) and four adolescents

(12-17 y.o.). The MNI for adults (MNI=37) was estimated based on the number of left tibiae (Tab. S9). Assignment of adult individuals into age categories was more problematic relative to subadults. Almost 90% of adult bones could only be classified as “adult” and no specific age range was estimated. Osteological sex was assessed for 80 skeletal elements (Tab. S10).

Various pathologies related to metabolic deficiencies, hemorrhage, strenuous physical activity, trauma and inflammation were recorded on the bones (Fig. 1). Periosteal reactions on the tibiae and fibulae were noted for 33 bones. This kind of changes, when found in isolation, is highly nonspecific and could be a result of infections other than leprosy (e.g., tuberculosis, treponematosi s or glanders) and trauma followed by a secondary infection. Neither *facies leprosa* nor atrophy of distal phalanges (typical of leprosy) were observed. Therefore, osteological examination of skeletal remains buried near the St. Jürgen hospital in Lübeck did not confirm the previous hypothesis that these individuals suffered from leprosy.

In the next step of the analysis, the St. Jürgen collection was sampled for molecular analysis. In total, 43 samples were taken from 22 petrous bones, 19 teeth and two foot bones. The foot bones were analyzed specifically for pathogen screening due to visible pathological lesions indicative of a severe infection (Fig. 1C). DNA was successfully isolated from 40 samples belonging to 34 individuals (26 adults and eight subadults) (Tab. 2, S3). Genetic data allowed to pair cranial fragments with corresponding mandibles for three individuals (4, 9 and 16 in Tab. S2), which was impossible based on osteological analysis due to the fragmentary and commingled nature of the burials. One of the samples was excluded from analysis based on possible contamination.

Osteological sex was estimated for 19 out of the 33 individuals for which DNA was available (57.6%) (Tab. 2). The results were in agreement with the genetic sex for 11 individuals (57.9%). Among the remaining eight individuals where discrepancies were noted, five individuals (62.5%) were represented only by highly fragmentary remains, which is likely the cause for the elevated misclassification rate.



Figure 1. Examples of pathological lesions observed on skeletal remains from Lübeck St. Jürgen (photographs: K. Fuchs).

A - Roof of the right orbit. The arrow points to small openings in the external layer of the bone (*cribra orbitalia*), which are associated with iron deficiency.

B - Bones of the right lower limb: femur (left) with tibia and fibula (right) - posterior view. The arrows point to the visible bending of the shafts caused by vitamin D deficiency (rickets).

C - Tarsal and metatarsal bones of the left foot - superior view. Fused joints (ankylosis) are marked with white arrows. Observed ankylosis and deformation of tarsal bones indicate an ongoing inflammation process, likely due to septic arthritis.

D - Ventral surface of the 3rd and 4th ribs (left). Hemorrhage-associated new bone deposits are marked with arrows.

E - Right clavicle - inferior view. Marked with the arrow is an oval lesion at the attachment site of the costoclavicular ligament. Presence of such changes indicates a significant and chronic mechanical load on the pectoral girdle.

F - Distal end of the humerus (right) - antero-lateral view. The arrows point to the eburnation of the *capitulum humeri* (upper) and margin lipping of the articular surface for the radius (lower) reflective of the advanced wear of the elbow joint due to osteoarthritis.

Table 2. Characteristics of the samples taken for molecular analysis (Lübeck St. Jürgen skeletal collection).

Individual ID	Sample type	Osteological sex	Genetic sex	Age-at-death
1	T	F	F	18-25
	PB			
2	T	M	M	30-35
3	T	-	M	24-35
4	T	-	M	>55
	PB			
5	T	-	M	6-9
6	T	PF	F	25-35
7	T	PF	M	35-55
8	T	-	M	>35
	foot			
	foot			
9	T	PM	M	16-18
	PB			
10	PB	-	F	>21
11	PB	PF	M	>21
12	PB	-	F	<21
13	PB	F	F	25-35
14	PB	F	M	>21
15	PB	-	M	40-58
16	PB	-	M	>21
	T			
17	PB	-	M	4-5
18	PB	PF	F	>21
19	PB	F	F	40-50
20	PB	F	M	38-58
21	PB	PF	M	>21
22	PB	-	F	1-3
23	PB	PF	F	>21
24	T	PF	F	35-55
25	T	PM	M	>50
26	T	-	M	2.5-3.5
27	T	-	F	7-8

Continuation of Table 2. Characteristics of the samples taken for molecular analysis (Lübeck St. Jürgen skeletal collection).

Individual ID	Sample type	Osteological sex	Genetic sex	Age-at-death
28	T	-	F	8-12
29	PB	PF	M	36-55
30	PB	F	M	33-57
31	PB	-	M	>30
32	PB	F	M	>21
33	PB	M	M	35-50
34	T [#]	F	-	38-58

T - tooth; PB - petrous bone; [#] - contaminated; F - female; M - male; PF - probably female; PM - probably male

Determination of leprosy cases and controls

When assigning the case vs. control status to each individual, various lines of evidence were examined. Firstly, all samples were screened for the presence of *M. leprae* DNA, which was considered as indicative of leprosy. Secondly, information on osteological lesions was assessed for two sites where it was available: Melaten cemetery and Lübeck St. Jürgen. Although Melaten was associated with a leprosarium at some point in the past, the graveyard was in use after the institution changed into a general infirmary. The hospital in Lübeck is also thought to have housed leprosy-affected individuals. However, not exclusively. Thus, it is possible that the graveyards accommodated people who did not suffer from leprosy. Only individuals exhibiting pathognomonic for leprosy lesions and/or molecular evidence of infection were considered as cases.

Osteological analysis of the remains from Lübeck St. Jürgen did not exhibit lesions indicative of leprosy. Moreover, pathogen screening of the metagenomic samples showed no evidence of *M. leprae* infection. Thus, individuals from Lübeck were considered controls in this study.

Molecular traces of *M. leprae* were found in samples belonging to nine individuals from Melaten. For seven of these, pathognomonic for leprosy lesions were also noted. Additionally, osteological records suggested the case status for 23 individuals. The majority of the remaining individuals exhibited nonspecific lesions of infection and their case vs. control status was unclear. They were thus excluded from further analysis.

For the individuals buried at the cemetery associated with the St. Jørgen leprosarium in Odense, leprosy was confirmed in previous analyses (Boldsen, 2009; Krause-Kyora et al., 2018). Moreover, *M leprae* DNA was found in 30 samples belonging to 29 individuals analyzed in this study.

Pathogen screening did not reveal the presence of *M. leprae* at the remaining 14 archaeological sites. In the absence of osteological evaluation, all individuals were considered as controls.

Genetic sex determination and kinship analysis

Genetic sex was determined for 367/380 individuals (162 females and 205 males). For 137 out of these, information about the estimated osteological sex was available. The level of agreement between the two methods of sex assessment was 81.6%. For cases where genetic sex was not determined (N=13), osteological sex was available for 8 individuals (seven females and one male). Taken together, the sample comprised 380 individuals (169 females, 206 males and 5 indeterminate).

Second degree kinship was reliably detected for one pair of individuals from Tirup (G272 and G557) (Tab. S11). Individual G557 was excluded from the association analysis.

Population genetic analysis

Genotyping resulted in 182 individuals passing the threshold of 20000 SNPs set for the PCA. Individuals from all 17 sites clustered within the diversity of modern day populations of western and northern Europe (Fig. 2, S2). When cases and controls were considered separately, the groups overlapped on the PCA plot along both axes (Fig. 3 and S3). These results indicate that there is no

substantial genetic diversity between the groups and a comparison of allelic frequencies in a case-control association study would not lead to spurious results due to population stratification.

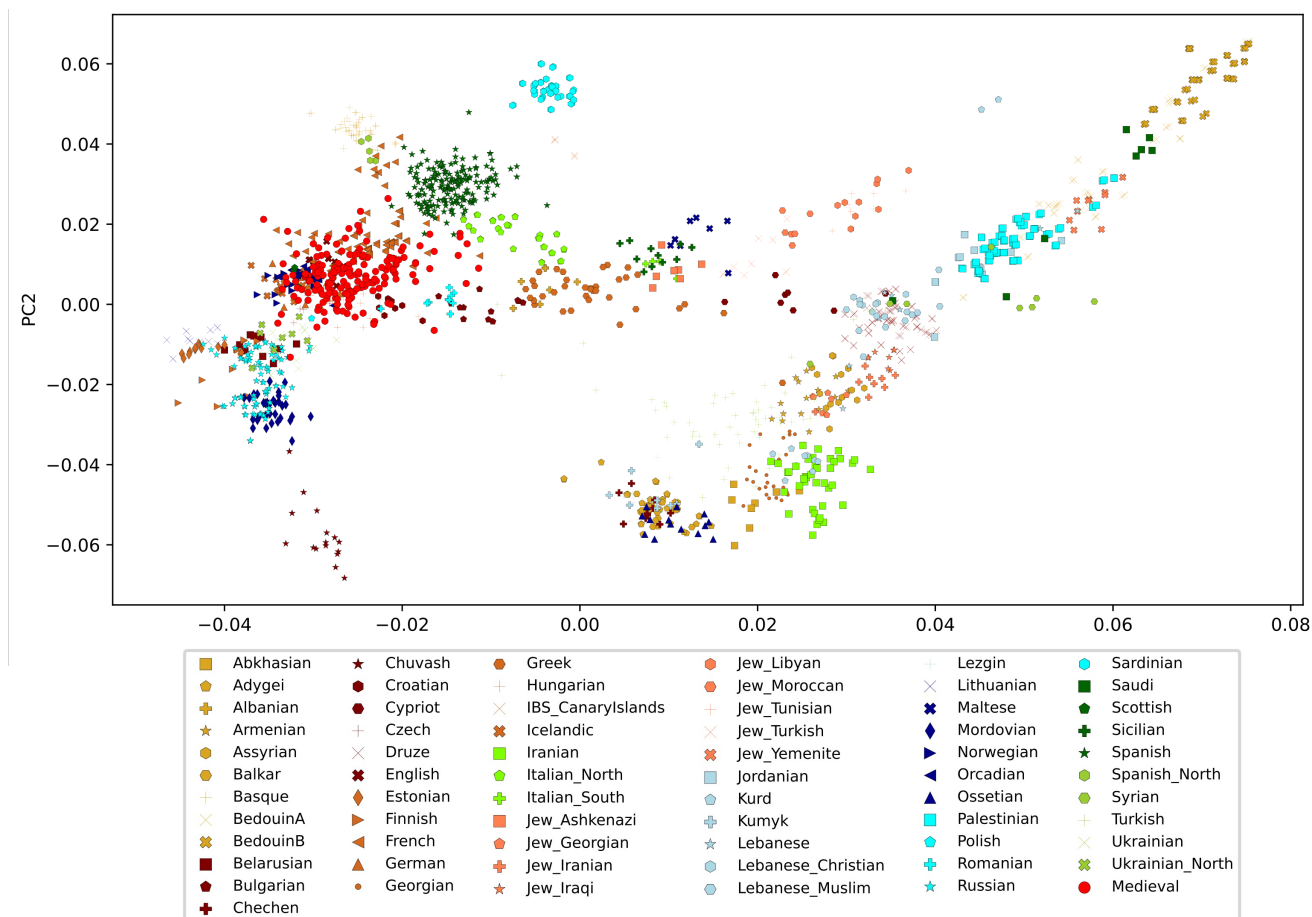


Figure 2. PCA of 182 medieval individuals (red circles) and 66 present-day West-Eurasian populations.

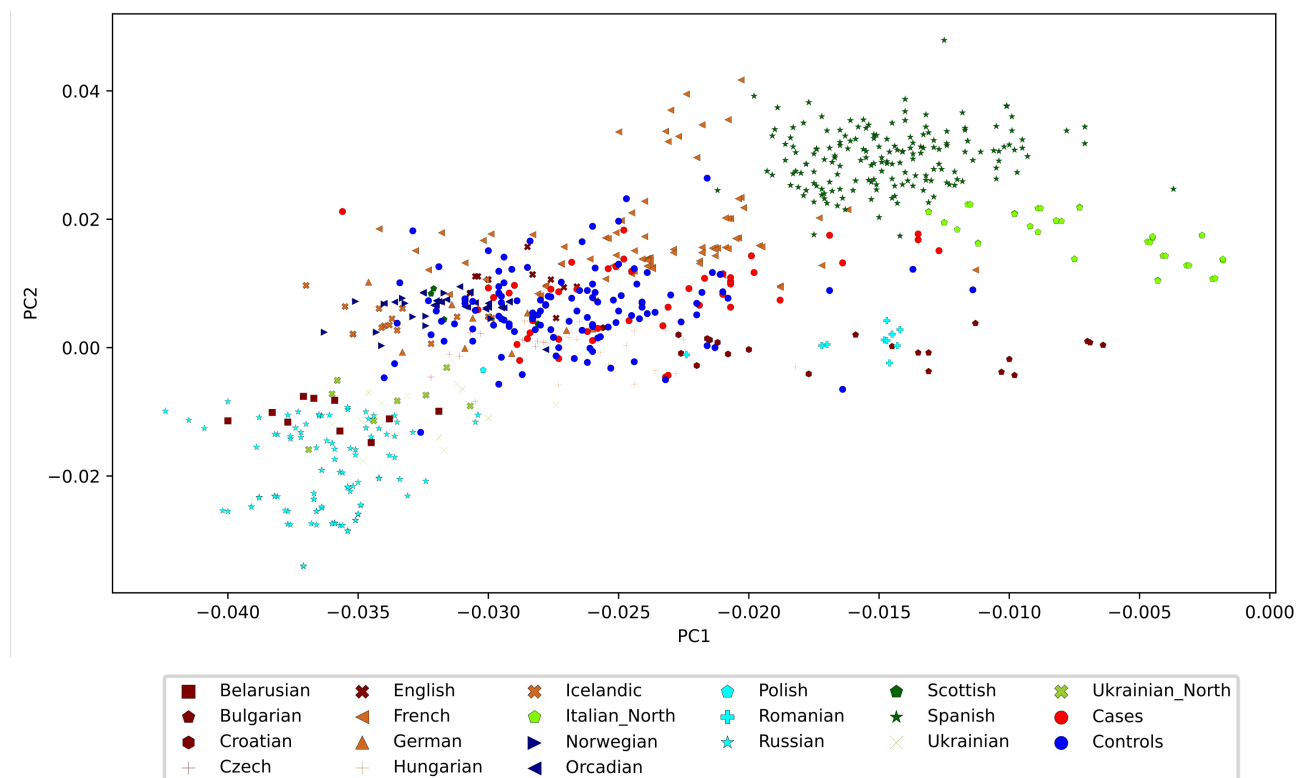


Figure 3. PCA of 76 cases (red circles) and 106 controls (blue circles) were projected onto the first two principal components calculated from 66 present-day West-Eurasian populations. For clarity, only the closely-clustering 19 populations were plotted.

HLA allele typing

HLA targeted enrichment was successfully performed for 323 samples. Using TARGT and OptiType, information on HLA profiles was produced for 271 individuals. Additionally, HLA data generated in a previous study (Pierini et al., 2020) was incorporated (Tab. S5, S6 and S7). Eighty-two alleles (32 individuals) were independently called in this and the previous study (Pierini et al., 2020). Discrepancy between the calls at the first-field resolution was noted for 10 alleles (7 - different allele, 3 - a homozygous call was made in this study and heterozygous in the previous study) (Tab. S4). The level of agreement at the first-field resolution was 87.8%.

Reproducibility of the manual typing using TARGT was tested by two observers of different experience levels on 86 randomly chosen alleles (Tab. S12). At the first-field resolution, the overall agreement was 82.6% (71/86). In 13 instances where resolution differences were noted between the observers, higher resolution was obtained by the more experienced observer for 12 alleles. Furthermore, the overall disagreement rate of 17.4% (15/86) arose from the inability to make a call by the less experienced observer. These results indicate the influence of the observer's experience on the success rate of the manual HLA typing (quantity of the obtained data).

Case-control HLA association analysis

In total, the frequency of HLA variants was analyzed in a sample of 300 individuals (99 cases and 201 controls, Tab. 3 and S13). Statistical analysis was performed for alleles present in both groups (Tab. 4-7).

Table 3. The number of medieval cases (N=99) and controls (N=201) included in the association analysis.

Site	Location	Status	N of individuals
St. Jørgen cemetery	Odense	cases	73 [#]
St. Knuds Plads		controls	2
Albani Torv		controls	6
Klosterbakken		controls	1
St. Trinitatis	Viborg	controls	13
St. Morten		controls	4
St. Mathias		controls	2
St. Mikkel		controls	10
Grabodre		controls	2
Faldborg Kirke		controls	3
Klosterkirken	Horsens	controls	32
Ole Worm Gade		controls	20
Odekirkegard	Sejet	controls	27
Tirup	Tirup	controls	28
Lindegarden	Ribe	controls	21
St. Jürgen	Lübeck	controls	30
Melaten	Aachen	cases	26
Total	-	-	300

[#] - including individuals from Pierini et al., 2020

Table 4. Differences in HLA-A allele frequencies between medieval leprosy cases and controls.

HLA-A					
Allele	Cases (N=155)	Frequency [%]	Controls (N=358)	Frequency [%]	p-value
01	20	12.9	65	18.2	0.145
02	49	31.6	117	32.7	0.78
03	26	16.8	50	14.0	0.404
11	11	7.1	22	6.1	0.681
23	7	4.5	1	0.3	0.001*
24	10	6.5	33	9.2	0.303
25	0	0	5	1.4	-
26	8	5.2	6	1.7	0.037*
29	2	1.3	5	1.4	1
30	2	1.3	9	2.5	0.518
31	3	1.9	9	2.5	1
32	4	2.6	12	3.4	0.786
33	0	0	2	0.6	-
36	1	0.6	1	0.3	0.513
66	0	0	3	0.8	-
68	10	6.5	17	4.7	0.423
69	1	0.6	0	0.0	-
74	1	0.6	1	0.3	0.513

N - number of alleles; * - nominally significant;

Table 5. Differences in HLA-B allele frequencies between medieval leprosy cases and controls.

HLA-B					
Allele	Cases (N=141)	Frequency [%]	Controls (N=344)	Frequency [%]	p-value
07	26	18.4	49	14.2	0.241
08	15	10.6	42	12.2	0.633
13	0	0	8	2.3	-
14	0	0	2	0.6	-
15	21	14.9	41	11.9	0.418
18	3	2.1	17	4.9	0.158
27	4	2.8	21	6.1	0.141
35	6	4.3	17	4.9	0.751
37	1	0.7	8	2.3	0.458
38	5	3.5	1	0.3	0.009*
39	2	1.4	12	3.5	0.369
40	14	9.9	32	9.3	0.823
41	1	0.7	1	0.3	0.496

Continuation of Table 5. Differences in HLA-B allele frequencies between medieval leprosy cases and controls.

HLA-B					
Allele	Cases (N=141)	Frequency [%]	Controls (N=344)	Frequency [%]	p-value
42	1	0.7	0	0	-
44	19	13.5	27	7.8	0.054
45	1	0.7	1	0.3	0.496
48	1	0.7	5	1.5	0.677
49	0	0	2	0.6	-
50	1	0.7	1	0.3	0.496
51	9	6.4	29	8.4	0.451
52	1	0.7	0	0	-
55	2	1.4	8	2.3	0.731
56	1	0.7	2	0.6	1
57	7	5	13	3.8	0.547
58	0	0	4	1.2	0.328
81	0	0	1	0.3	-

N - number of alleles; * - nominally significant;

Table 6. Differences in HLA-C allele frequencies between medieval leprosy cases and controls.

HLA-C					
Allele	Cases (N=147)	Frequency [%]	Controls (N=341)	Frequency [%]	p-value
01	5	3.4	4	1.2	0.136
02	6	4.1	17	5.0	0.675
03	37	25.2	86	25.2	0.996
04	11	7.5	21	6.2	0.582
05	10	6.8	18	5.3	0.502
06	6	4	33	9.7	0.044*
07	50	34	109	32.0	0.643
08	0	0	7	2.1	-
11	0	0	1	0.3	-
12	10	6.8	15	4.4	0.337
14	0	0	8	2.3	-
15	6	4	10	2.9	0.58
16	5	3.4	10	2.9	0.779
17	1	0.7	1	0.3	0.511
18	0	0	1	0.3	-

N - number of alleles; * - nominally significant;

Across all loci examined in this study (HLA-A, B, C, DRB1 and DQB1), frequencies of five HLA alleles differed significantly (based on nominal p-values) between cases and controls (Tab. 8). HLA-A*23 (p=0.001, OR=16.9), A*26 (p=0.037, OR=3.2) and B*38 (p=0.009, OR=12.6) were more frequent in leprosy-affected individuals. Conversely, the frequency of variants C*06 (p=0.044, OR=0.4) and DRB1*13 (p=0.022, OR=0.48) was higher in controls.

Table 7. Differences in HLA class II allele frequencies between medieval leprosy cases and controls.

HLA-DQB1					
Allele	Cases (N=150)	Frequency [%]	Controls (N=335)	Frequency [%]	p-value
02	23	15.3	68	20.3	0.200
03	53	35.3	102	30.4	0.307
04	4	2.7	19	5.7	0.152
05	19	12.7	39	11.6	0.739
06	51	34	107	31.9	0.735
HLA-DRB1					
Allele	Cases (N=147)	Frequency [%]	Controls (N=334)	Frequency [%]	p-value
01	15	10.2	30	9.0	0.672
03	14	9.5	52	15.6	0.076
04	34	23.1	53	15.9	0.057
07	16	10.9	30	9.0	0.513
08	4	2.7	22	6.6	0.084
09	2	1.4	2	0.6	0.589
10	1	0.7	1	0.3	0.518
11	9	6.1	16	4.8	0.544
12	0	0	4	1.2	-
13	13	8.8	56	16.8	0.022*
14	3	2	8	2.4	1
15	34	23.1	55	16.5	0.083
16	4	2.7	5	1.5	0.466

N - number of alleles; * - nominally significant;

Table 8. HLA alleles associated with leprosy in medieval Europeans.

Allele	Phenotype	p-value*	OR	95% C.I.
A*23	susceptibility	0.001	16.885	2.059-138.444
A*26	susceptibility	0.037	3.193	1.089-9.363
B*38	susceptibility	0.009	12.61	1.46-108.931
C*06	protection	0.044	0.397	0.163-0.969
DRB1*13	protection	0.022	0.482	0.255-0.911

* - nominally significant; C.I. – confidence interval;

Furthermore, although not statistically significant, trends were observed for frequencies of other DRB1 alleles, such as DRB1*03, 04, 08 and 15. HLA-DRB1*04 and 15 were more frequent in cases, whilst DRB1*03 and 08 were less frequent (Tab. 7). Statistical analysis with G*Power indicated that this study is underpowered to detect frequency shifts for these alleles at a significant level of 0.05 and the sample size would have to be doubled to increase the achieved statistical power of approx. 0.5 to the standard 0.8.

HLA allele frequencies in medieval vs. modern Europeans

Confirmation of genetic continuity between the medieval individuals and modern Germans (Fig. 2 and 3) allowed an investigation of potential shifts in HLA diversity over time. An exploratory assessment of the allele frequencies in the medieval controls and a large cohort of modern individuals (AFND: Germany, population 8 (HLA-A, B, C and DRB1) and population 3 (HLA-DQB1)) (Gonzalez-Galarza et al., 2020) was performed (Tab. S14). Frequency differences of above 5% were considered as a substantial shift. Interestingly, the number of B*35, C*04, DQB1*03 and DRB1*11 seems to have increased since the Middle Ages. On the contrary, A*02, C*03 and DQB1*06 are less common in present-day Germans relative to medieval individuals (Tab. 9, Fig. 4). The shift is particularly pronounced for the DQB1*03 and C*03, with a difference of over 10%.

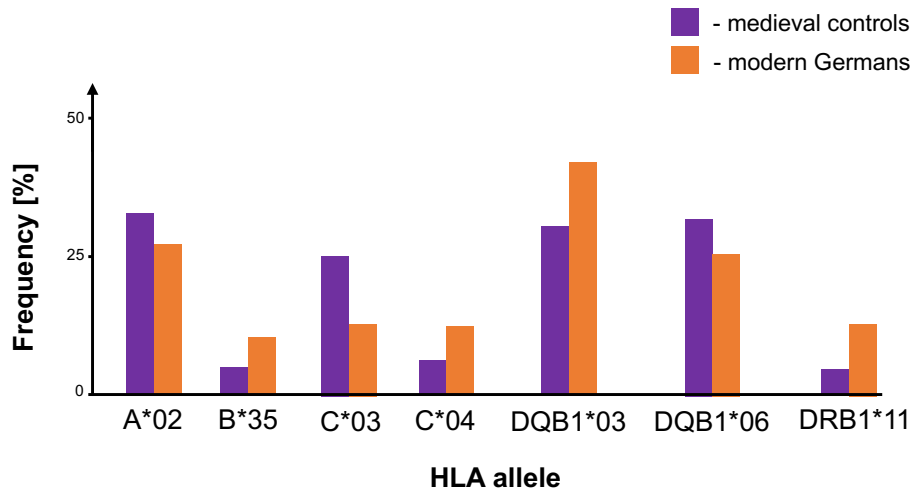


Figure 4. HLA alleles which exhibit a temporal shift (>5%) in frequency.

Table 9. HLA alleles which exhibit a temporal shift (>5%) in frequency.

Allele	Frequency [%]	
	Medieval controls	Modern Germans
A*02	32.7	27
B*35	4.9	10.4
C*03	25.2	13
C*04	6.2	12.5
DQB1*03	30.4	41.9
DQB1*06	31.9	25.6
DRB1*11	4.8	13

Discussion

Here, a large sample of medieval individuals was examined in the context of HLA-associated susceptibility to leprosy. Whilst the previous HLA case-control association analysis performed on medieval Europeans focused on a single allele (DRB1*15) (Krause-Kyora et al., 2018), this study represents the first association analysis between multiple HLA alleles and leprosy in Europeans. Five alleles were found to be associated with the disease (HLA-A*23, A*26, B*38, C*06 and DRB1*13) (Tab. 8). HLA-B*38 was previously identified as a risk factor for leprosy in southern

Brazilians (2.7% vs. 1.1%, $p=0.04$, $OR=2.44$) (Franceschi et al., 2011). Moreover, a nominally significant association between HLA-C*06:02 and leprosy protection was made for the population of India (12.5% vs. 37.3%, $p=0.01$, $OR=0.24$) (Shankankumar et al., 2003). Although, in this study, the association analysis was performed at a first-field resolution (C*06), C*06:02 was the most common C*06 allele called for the medieval individuals (Tab. S13). All but one calls were either C*06 (N=20) or C*06:02:01G (N=18). The remaining three alleles associated with leprosy in this study (A*23, A*26 and DRB1*13) have not been previously linked to disease susceptibility in modern non-European populations. It is possible that these are novel risk/protective variants. However, due to the limited sample size and a large number of HLA alleles, an exploratory approach was taken and the p-values reported here are nominal. Therefore, the findings presented here have to be confirmed in a larger independent sample. Replication of the results may be difficult to achieve due to the existing interpopulation variation in leprosy susceptibility (White and Franco-Paredes, 2015) and the virtual lack of leprosy cases in modern-day Europe.

The allele, which is the most consistently associated with leprosy across various populations, HLA-DRB1*15:01, (Gorodezky et al., 2004; Krause-Kyora et al., 2018; Joko et al., 2000; Mehra et al., 1995; Vanderborght et al., 2007; Zerva et al., 1996; Zhang et al., 2009) did not differ significantly in its frequency between cases and controls. Thus, the previous association of this allele with leprosy risk for medieval Europeans (Krause-Kyora et al., 2018) could not be confirmed in this study. This is, however, likely due to the insufficient sample size (achieved power = 0.51) as a strong trend was noted with a frequency of 23.1% in cases and 16.5% in controls ($p=0.083$). Substantially more frequent in leprosy-affected individuals was also DRB1*04 (23.1 vs. 15.9%, $p=0.057$). Interestingly, it was identified as a protective allele in Argentinians, Brazilians, Japanese, Taiwanese and Vietnamese (da Silva et al., 2009; Hsieh et al., 2010; Joko et al., 2000; Motta et al., 1007; Vanderborght et al., 2007). This discrepancy in the direction of the association (susceptibility vs. protection) could possibly be explained by interpopulation variation. Moreover, differences in predicted *M. leprae* antigen repertoires between HLA-DRB1*04:04 and 04:01 were reported (Krause-Kyora et al., 2018), with DRB1*04:01 exhibiting a highly limited repertoire. Such a reduced ability to recognize *M. leprae* peptides may lead to a less efficient immune response and thus to disease progression. Whilst frequency of DRB1*04:04 was 2.1% in both cases and controls, DRB1*04:01 was more common in leprosy-affected individuals (9.5% vs.

4.8%) (Tab. S13). It is thus possible that the trend observed in this study relates to DRB1*04:01. Nevertheless, it is important to consider that 8.8% and 6.6% of the DRB1:04 calls for cases and controls respectively were made at first-field resolution, obscuring the true values of DRB1*04:04 and 04:01 frequencies. HLA-DRB1*08, which was less frequent in cases (2.7% vs. 6.6%, $p=0.084$), was previously linked to decreased leprosy susceptibility in Mexican Mestizo (Escamilla-Tilch et al., 2013) and Argentineans (DRB1*08:08) (Borras et al., 2008). More common in medieval controls relative to cases was also DRB1*03 (9.5% vs. 15.6%, $p=0.076$). This allele has not been associated with protection against the *M. leprae* infection in any modern non-European populations until now. It was, however, more frequent in leprosy cases in a study conducted among Mexicans (Gorodezky et al., 1987). Interestingly, computational prediction of HLA-presented antigenic *M. leprae* peptides for common DRB1 alleles (Krause-Kyora et al., 2018) revealed that DRB1*03:01, 08:01 and 13:01 are among the variants exhibiting the largest *M. leprae* antigen repertoires, making these alleles relatively efficient in their recognition. This could possibly explain the trends and the association with leprosy protection observed in this study.

Frequency comparison of the associated in this study alleles (Tab. 8) between medieval controls (Tab. 4-7) and modern-day Germans (Tab. S14) revealed similar values (max. difference of 3.3% noted for DRB1*13). Associations between these variants and infectious as well as autoimmune conditions were previously made (Tab. 10). This could suggest shared mechanisms underlying these diseases and possible balancing selection acting on the alleles.

Table 10. Previous associations reported for the HLA alleles associated with leprosy in this study.

Allele	<i>In this study</i>	protection	risk	ref.
A*23	<i>risk</i>	-	-	-
A*26	<i>risk</i>	-	Behçet's disease (Saudi Arabia, Japan)	Al-Okaily et al., 2016; Kaburaki et al., 2010; Nakamura et al., 2019
B*38	<i>risk[#]</i>	-	leprosy (Brazil)	Franceschi et al., 2011
C*06	<i>protection[#]</i>	leprosy (India)	psoriasis (Finland)	Haapasalo et al., 2018; Shankankumar et al., 2003
DRB1*13	<i>protection</i>	chronic hepatitis B (Caucasians, Thailand) microscopical polyangiitis (Japan) systemic lupus erythematosus (Japan) rheumatoid arthritis (Japan) systemic sclerosis (Japan) dementia (Western Europe) cervical cancer (North American whites) tuberculosis (Poland)	chronic hepatitis C (Egypt) autoimmune hepatitis (Argentina) colorectal carcinoma (Caucasians) Parkinson's disease (India)	Aureli et al., 2015; Dubaniewicz et al., 2000; El Bendary et al., 2019; Furukawa et al., 2014; 2015; 2016; Garcia Borrás et al., 2011; Hohler et al., 1997; James & Georgopoulos 2019; Kawasaki et al., 2016; Kummee et al., 2007; Madeleine et al., 2002; Pandi et al., 2021

[#] - this study supports previous association of the allele with leprosy susceptibility/protection

Epidemics of leprosy greatly affected the European population in the Middle Ages, which is reflected in the number of institutions that housed individuals infected with *M. leprae* at the time. It was estimated that approx. 19,000 leprosaria were established across the continent by the 13th cent. AD (Raffcliffe, 2006). Nowadays, however, leprosy is virtually absent in Europe and only around 5-10% of infections lead to disease development. Furthermore, clinical manifestation of leprosy is highly variable and mediated by the immune response of the host (Lockwood and

Saunderson, 2012). In the light of the high stability of the *M. leprae* genome since the early Middle Ages (Monot et al., 2005; Schuenemann et al., 2013; 2018; Krause-Kyora et al., 2018; Pfrengle et al., 2021), it is likely that environmental and host genetic factors played a major role in leprosy disappearance from the European continent. In addition to the social isolation of the leprosy-affected individuals in leprosaria and their rights to marry being questioned, hormonal changes associated with the disease (Gunawan et al., 2020; Khanna et al., 2014; Saporta & Yuksel, 1994) likely had a detrimental effect on reproduction. Moreover, a wave of tuberculosis, which swept through Europe in the late Middle Ages (Daniel, 1997), could have contributed to leprosy's disappearance. Interestingly, the decrease of leprosy prevalence coincided with a peak incidence of tuberculosis (Manchester, 1984). As the response to infection with *M. leprae* and *M. tuberculosis* is similar (Aufderheide and Rodriguez-Martin, 1998; Lurie, 2008; Ohara et al., 2000), it is possible that exposure to a far more deadly disease, such as tuberculosis, might have led to selection of resistant genotypes (Manchester, 1984; Donoghue et al., 2005; Monot et al., 2009).

A temporal shift in frequencies of several HLA alleles was observed in this study (Tab. 7). Interestingly, these alleles were previously associated with susceptibility to infectious diseases that are highly prevalent in modern-day non-European populations, such as hepatitis, malaria, acquired immunodeficiency syndrome (AIDS) and tuberculosis, as well as Chagas disease and dengue fever (Tab. 11). The majority of the associations were made in non-European populations, due to the current epidemiology of the diseases. With that in mind, whilst the frequency of HLA alleles protective against hepatitis, malaria, dengue fever and chagas disease has increased, so has the frequency of risk factors related to HIV and SarsCov2. Moreover, a decrease in frequency was noted for the alleles protective against severe COVID19, Epstein-Barr Virus-positive (EBV⁺) Hodgkin lymphoma and hepatitis. Based on these associations, however, it is not possible to infer possible reasons as to why the shifts in frequency have occurred since the Middle Ages. Interestingly, the same alleles were also previously associated with various autoimmune and degenerative conditions, such as systemic sclerosis, lupus, alopecia, Grave's disease, osteoporosis and Parkinson's disease (Tab. 11). The majority of the associations were made for alleles whose frequency has increased since the Middle Ages, particularly class II HLA alleles DQB1*03 and DRB1*11. Although there is no simple directional link between the associations with infectious diseases and inflammatory conditions, these findings highlight the pleiotropy of the HLA system

and suggest a possible evolutionary trade-off between susceptibility to infections and autoimmunity. In all probability, ancient epidemics of infectious diseases helped to shape the present-day diversity of HLA alleles in the European population, possibly influencing the levels of genetic predisposition to autoimmune conditions. Although evidence of such epidemics may be obscured by the lack of osteological lesions and written sources, degradation of pathogens' DNA over time, inability to detect RNA viruses and pathogens outside the bloodstream, major outbreaks of infectious diseases are likely to have left their mark on the human genome. To decipher the associations and their underlying mechanisms will undoubtedly be a challenging task of future studies.

Table 11. Previous associations reported for the HLA alleles exhibiting a temporal shift in frequency.

Allele	frequency in Modern Germans[#]	risk	protection	ref.
A*02	decreased	childhood acute leukemia (Romania)	HCV (China) EBV ⁺ Hodgkin lymphoma (Netherlands) severe COVID19 (Russia)	Huang et al., 2016; Gai et al., 2010; Niens et al., 2007; Shkurnikov et al., 2021
B*35	increased	HIV progression (European, African American and Hispanic populations) systemic sclerosis (Brazil, Italy, Choctaw tribe) postmenopausal osteoporosis (China)	dengue fever (Mexico) malaria (Ghana)	Arnett et al., 1996; Carrington et al., 1999; Falcón-Lezama et al., 2009; Li and Liu, 2014; Gao et al., 2001; Pereyra et al., 2010; Santaniello et al., 2006; Toledo Del Rio et al., 2013; Yamazaki et al., 2011
C*03	decreased	cardiac involvement in Chagas disease (Venezuela) HIV escape mutant (Africans) Grave's disease (Poland: C*03:02)	Grave's disease (Poland: C*03:04, UK)	Honeyborne et al., 2010; Layrisse et al., 2000; Simmonds et al., 2007; Zawadzka-Starczewska et al., 2022

Continuation of Table 11. Previous associations reported for the HLA alleles exhibiting a temporal shift in frequency.

Allele	frequency	risk	protection	ref.
	in Modern Germans[#]			
C*03	decreased	Parkinson's disease (Europeans) multiple myeloma (China)		Wang et al., 2008; Wisseman et al., 2013
C*04	increased	HIV progression (Peru) SarsCov2 (Armenia, Germany, Spain, Switzerland) narcolepsy (Europeans)	-	Hovhannisyan et al., 2022; Olvera et al., 2015; Tafti et al., 2016; Weiner et al., 2020
DQB1*03	increased	HBV progression (African Americans) tuberculosis (South East Asia) recurrent respiratory papillomatosis (Korea) Grave's disease (Poland: DQB1*03:01) bullous pemphigoid (China, Iran, Japan, Brazil, Caucasians) systemic lupus erythematosus (India) pityriasis rosea (China) alopecia areata (Northwestern Europeans)	dengue fever (Mexico) HBV (Gambia) Chagas disease (Venezuela) HCV clearance (France, UK, Brazil, African Americans) Grave's disease (Poland: DQB1*03:03) celiac disease (Venezuela)	Ahmed et al., 1991; Alric et al., 1997; Cangussu et al., 2011; Chagury et al., 2018; Delgado et al., 1996; Esmali et al., 2013; Falcón-Lezama et al., 2009; Fernandez-Mestre et al., 1998; 2022 Fu et al., 2018; Harcourt et al., 2001; Song et al., 2013; Shankankumar et al., 2003; Sun et al., 2018; Thio et al., 2001; Thursz et al., 1995; Toyo-oka et al., 2017; Welsh et al., 1994; Zawadzka-Starczewska et al., 2022
DQB1*06	decreased	HCV (Korea) Parkinson's disease (India)	HCV (Egypt) pemphigus vulgaris (India) pityriasis rosea (China)	Al-Motwee et al., 2013; El-Chennawi et al., 2008; Fu et al., 2018; Pandi et al., 2021; Priyadarshini et al., 2018; Shankankumar et al., 2003;; Yoon et al., 2005

Continuation of Table 11. Previous associations reported for the HLA alleles exhibiting a temporal shift in frequency.

Allele	frequency in Modern Germans[#]	risk	protection	ref.
DQB1*06	decreased	multiple sclerosis (Norway, Spain, China) systemic lupus erythematosus (Egypt, Saudi Arabia, India)		Fernandez et al., 2004; Mokbel et al., 2015; Serjeantson et al., 1992; Spurkland et al., 2008
DRB1*11	increased	recurrent respiratory papillomatosis (Korea), thrombotic thrombocytopenic purpura (France) systemic sclerosis (North American whites) juvenile idiopathic arthritis (Western Europeans) asthma (Venezuela) cervical cancer (North American whites) Grave's disease (Poland) Scleroderma (China)	Chagas disease (Argentina) HCV clearance (France, UK) asymptomatic HCV (Italy) rheumatoid arthritis (Pakistan)	Alric et al., 1997; Asti et al., 1999; Berangere et al., 2020; Garcia-Borras et al., 2009; Harcourt et al., 2001; He et al., 2014; Kuwana et al., 1999; Lara-Marquez et al., 1999; Madeleine et al., 2002; Ombrello et al., 2015; Muazzam et al., 2013; Song et al., 2013; Zawadzka-Starczewska et al., 2022

[#] - Table shows the frequency change in relation to medieval controls.

Although five HLA alleles were found to be associated with leprosy in this analysis, the study is not without limitations, which have to be considered when interpreting the results. Firstly, this study was explorative and considered the complete HLA profiles without a correction for multiple testing. Although such analysis can detect potentially novel associations, confirmation of the findings in an independent sample is necessary. Secondly, as the sample examined here is of a relatively small size in comparison to genome wide association studies (GWAS) (Wang et al.,

2016; Zhang et al., 2009), the variants of small effect size were likely undetected. The statistical power could be potentially further reduced by the presence of leprosy-affected individuals in the control group. Although stringent selection criteria were applied for cases, osteological information about possible leprosy lesions was not available for the controls. It is thus probable that pathogen screening alone was not sufficient to reliably determine the “case vs. control” status. Thirdly, errors in genotyping could have theoretically influenced the HLA frequencies. Levels of agreement between manual HLA calls made by two independent observers for two sets of alleles were 87.8% (Tab. S4) and 82.6% (Tab. S12). The highly degraded nature of aDNA makes the HLA typing challenging. Although the semi-automated TARGT pipeline is currently the most suitable method for this kind of data, it involves some amount of subjectivity and experience, which seems to influence the success rate of HLA typing in general (Tab. S12). Although genotyping error should not have a substantial effect on the association of alleles of a relatively high frequency, the low-frequency alleles, such as A*23 and B*38 (Tab. 4 and 5) could possibly be affected. In this study, however, Optitype (Szolek et al., 2014) was used to support or reject a call in cases where the manual calls were ambiguous. Lastly, HLA genes are in strong genetic linkage and the entire complex is inherited as an HLA haplotype from each parent (Choo, 2007). Such a non-random association of alleles at multiple loci is known as linkage disequilibrium (LD). It is thus possible that identified alleles are not causal for leprosy, but rather are in LD with a different HLA allele or a SNP at a different locus linked to the disease (Ajioka et al., 1997; Cardoso et al., 2002; Ellis et al., 1997). This is, however, uncommon for HLA-disease associations, as existing data supports the role of the associated allele in the majority of cases (Holoshitz, 2013).

References

- Ahmed, A. R., Foster, S., Zaltas, M., Notani, G., Awdeh, Z., Alper, C. A., & Yunis, E. J. (1991). Association of DQw7 (DQB1*0301) with ocular cicatricial pemphigoid. *Proceedings of the National Academy of Sciences*, 88(24), 11579–11582. <https://doi.org/10.1073/pnas.88.24.11579>
- Ajioka, R. S., Jorde, L. B., Gruen, J. R., Yu, P., Dimitrova, D., Barrow, J., Radisky, E., Edwards, C. Q., Griffen, L. M., & Kushner, J. P. (1997). Haplotype Analysis of Hemochromatosis: Evaluation of Different Linkage-Disequilibrium Approaches and Evolution of Disease Chromosomes. *The American Journal of Human Genetics*, 60(6), 1439–1447. <https://doi.org/10.1086/515466>
- Ali, S., Srivastava, A. K., Chopra, R., Aggarwal, S., Garg, V. K., Bhattacharya, S. N., & Bamezai, R. N. K. (2013). *IL12B* SNPs and copy number variation in *IL23R* gene associated with susceptibility to leprosy. *Journal of Medical Genetics*, 50(1), 34–42. <https://doi.org/10.1136/jmedgenet-2012-101214>
- Al-Motwee, S., Jawdat, D., Jehani, G. S., Anazi, H., Shubaili, A., Sutton, P., Uyar, A. F., & Hajeer, A. H. (2013). Association of HLA-DRB1*15 and HLA-DQB1*06 with SLE in Saudis. *Annals of Saudi Medicine*, 33(3), 229–234. <https://doi.org/10.5144/0256-4947.2013.229>
- Al-Okaily, F., Al-Rashidi, S., Al-Balawi, M., Mustafa, MD., Arfin, M., & Al-Asmari, A. (2016). Genetic Association of *HLA-A*26*, *-A*31*, and *-B*51* with Behcet's Disease in Saudi Patients. *Clinical Medicine Insights: Arthritis and Musculoskeletal Disorders*, 9, CMAMD.S39879. <https://doi.org/10.4137/CMAMD.S39879>
- Alric, L., Izopet, J., Fort, M., Vinel, J.-P., Fontenelle, P., Orfila, C., Payen, J.-L., Sandres, K., Desmorat, H., Charlet, J.-P., Duffaut, M., & Abbal, M. (1999). Study of the association between major histocompatibility complex class II genes and the response to interferon alpha in patients with chronic hepatitis C infection. *Human Immunology*, 60(6), 516–523. [https://doi.org/10.1016/S0198-8859\(99\)00021-X](https://doi.org/10.1016/S0198-8859(99)00021-X)
- Alvarado-Navarro, A., Montoya-Buelna, M., Muñoz-Valle, J. F., López-Roa, R. I., Guillén-Vargas, C., & Fafutis-Morris, M. (2008). The 3'UTR 1188 A/C polymorphism in the interleukin-12p40 gene (IL-12B) is associated with lepromatous leprosy in the west of Mexico. *Immunology Letters*, 118(2), 148–151. <https://doi.org/10.1016/j.imlet.2008.03.015>
- Andersen, J. G., & Manchester, K. (1992). The rhinomaxillary syndrome in leprosy: A clinical, radiological and palaeopathological study. *International Journal of Osteoarchaeology*, 2(2), 121–129. <https://doi.org/10.1002/oa.1390020204>
- Anderson, D., Thompson, G., & Popovich, F. (1976). Age of attainment of mineralization stages of the permanent dentition. *Journal of Forensic Sciences*, 21(1), 191–200.

Arnett, F. C., Howard, R. F., Tan, F., Moulds, J. M., Bias, W. B., Durban, E., Cameron, H. D., Paxton, G., Hodge, T. J., Weathers, P. E., & Reveille, J. D. (1996). Increased prevalence of systemic sclerosis in a native american tribe in oklahoma. Association with an amerindian HLA haplotype. *Arthritis & Rheumatism*, 39(8), 1362–1370. <https://doi.org/10.1002/art.1780390814>

Ascadi, G., & Nemeskeri, J. (1970). *History of Human Life Span and Mortality*. Akadémiai Kiadó.

Asti, M., Martinetti, M., Zavaglia, C., Cuccia, M. C., Gusberti, L., Tinelli, C., Cividini, A., Bruno, S., Salvaneschi, L., Ideo, G., Mondelli, M. U., & Silini, E. M. (1999). Human leukocyte antigen class II and III alleles and severity of hepatitis C virus-related chronic liver disease. *Hepatology*, 29(4), 1272–1279. <https://doi.org/10.1002/hep.510290445>

Aufderheide, A., & Rodriguez-Martin, C. (1998). *The Cambridge encyclopedia of human palaeopathology*. Cambridge University Press.

Aureli, A., Canossi, A., del Beato, T., Franceschilli, L., Buonomo, O., Papola, F., de Sanctis, F., Lanzilli, G., Sileri, P., Coppola, A., Caratelli, S., Arriga, R., Orlandi, A., Lauro, D., Rossi, P., & Sconocchia, G. (2015). HLA-DRB1*13:01 allele in the genetic susceptibility to colorectal carcinoma. *International Journal of Cancer*, 136(10), 2464–2468. <https://doi.org/10.1002/ijc.29285>

Bérangère S. Joly, Pascale Loiseau, Michael Darmon, Thierry Leblanc, Hervé Chambost, Fanny Fouyssac, Vincent Guignonis, Jérôme Harambat, Alain Stepanian, Paul Coppo, & Agnès Veyradier. (2020). HLA-DRB1*11 is a strong risk factor for acquired thrombotic thrombocytopenic purpura in children. *Haematologica*, 105(10), e531. <https://doi.org/10.3324/haematol.2019.241968>

Blackwell, J. M., Jamieson, S. E., & Burgner, D. (2009). HLA and Infectious Diseases. *Clinical Microbiology Reviews*, 22(2), 370–385. <https://doi.org/10.1128/CMR.00048-08>

Boldsen, J. L. (2018). *Personal communication on archaeological dating of the Danish sites*.

Boldsen, J. L. (2001). Epidemiological approach to the paleopathological diagnosis of leprosy. *American Journal of Physical Anthropology*, 115(4), 380–387. <https://doi.org/10.1002/ajpa.1094>

Boldsen, J. L. (2008). Leprosy in the early medieval Lauchheim community. *American Journal of Physical Anthropology*, 135(3), 301–310. <https://doi.org/10.1002/ajpa.20744>

Boldsen, J. L. (2009). Leprosy in Medieval Denmark Osteological and epidemiological analyses. *Anthropologischer Anzeiger*, 67(4), 407–425. <https://doi.org/10.1127/0003-5548/2009/0031>

Boldsen, J. L., & Møllerup, L. (2006a). Outside St. Jørgen: Leprosy in the medieval Danish city of Odense. *American Journal of Physical Anthropology*, 130(3), 344–351. <https://doi.org/10.1002/ajpa.20363>

- Boldsen, J. L., & Møllerup, L. (2006b). Outside St. Jørgen: Leprosy in the medieval Danish city of Odense. *American Journal of Physical Anthropology*, 130(3), 344–351. <https://doi.org/10.1002/ajpa.20363>
- Boldsen, J. L., Rasmussen, K. L., Riis, T., Dittmar, M., & Weise, S. (2013). Schleswig: Medieval leprosy on the boundary between Germany and Denmark. *Anthropologischer Anzeiger*, 70(3), 273–287. <https://doi.org/10.1127/0003-5548/2013/0318>
- Borrás, S. G., Cotorruelo, C., Racca, L., Recarte, M., Garcías, C., Biondi, C., Bottasso, O., & Racca, A. (2008). Association of leprosy with *HLA-DRB1* in an Argentinean population. *Annals of Clinical Biochemistry: International Journal of Laboratory Medicine*, 45(1), 96–98. <https://doi.org/10.1258/acb.2007.007156>
- Borrás, S. G., Racca, L., Cotorruelo, C., Biondi, C., Beloscar, J., & Racca, A. (2009). Distribution of *HLA-DRB1* Alleles in Argentinean Patients with Chagas' Disease Cardiomyopathy. *Immunological Investigations*, 38(3–4), 268–275. <https://doi.org/10.1080/08820130902766589>
- Britton, W. (1998). The management of leprosy reversal reactions. *Leprosy Review*, 69(3), 225–234. <https://doi.org/10.5935/0305-7518.19980024>
- Brooks, S., & Suchey, J. M. (1990). Skeletal age determination based on the os pubis: A comparison of the Acsádi-Nemeskéri and Suchey-Brooks methods. *Human Evolution*, 5, 227–238. <https://doi.org/10.1007/BF02437238>
- Bryceson, A., & Pfaltzgraff, R. (1990). *Leprosy* (3rd edition). Churchill Livingstone.
- Buikstra, J. E., & Ubelaker, D. H. (1994). Standards for Data Collection from Human Skeletal Remains: Proceedings of a Seminar at the Field Museum of Natural History. In *Arkansas Archeological Survey research series*.
- Cangussu, L. O. F., Teixeira, R., Campos, E. F., Rampim, G. F., Mingoti, S. A., Martins-Filho, O. A., & Gerbase-DeLima, M. (2011). HLA Class II Alleles and Chronic Hepatitis C Virus Infection. *Scandinavian Journal of Immunology*, 74(3), 282–287. <https://doi.org/10.1111/j.1365-3083.2011.02568.x>
- Cardoso, C. S., Alves, H., Mascarenhas, M., Gonçalves, R., Oliveira, P., Rodrigues, P., Cruz, E., de Sousa, M., & Porto, G. (2002). Co-selection of the H63D mutation and the HLA-A29 allele: a new paradigm of linkage disequilibrium? *Immunogenetics*, 53(12), 1002–1008. <https://doi.org/10.1007/s00251-001-0414-8>
- Carrington, M., Nelson, G. W., Martin, M. P., Kissner, T., Vlahov, D., Goedert, J. J., Kaslow, R., Buchbinder, S., Hoots, K., & O'Brien, S. J. (1999). *HLA* and HIV-1: Heterozygote Advantage and *B*35 - Cw*04* Disadvantage. *Science*, 283(5408), 1748–1752. <https://doi.org/10.1126/science.283.5408.1748>

Chagury, A. A., Sennes, L. U., Gil, J. M., Kalil, J., Rodrigues, H., Rosales, C. B., & Miziara, I. D. (2018). HLA-C*17, DQB1*03:01, DQA1*01:03 and DQA1*05:05 Alleles Associated to Bullous Pemphigoid in Brazilian Population. *Annals of Dermatology*, 30(1), 8. <https://doi.org/10.5021/ad.2018.30.1.8>

Choo, S. Y. (2007). The HLA System: Genetics, Immunology, Clinical Testing, and Clinical Implications. *Yonsei Medical Journal*, 48(1), 11. <https://doi.org/10.3349/ymj.2007.48.1.11>

Cole, S. T., Eiglmeier, K., Parkhill, J., James, K. D., Thomson, N. R., Wheeler, P. R., Honoré, N., Garnier, T., Churcher, C., Harris, D., Mungall, K., Basham, D., Brown, D., Chillingworth, T., Connor, R., Davies, R. M., Devlin, K., Duthoy, S., Feltwell, T., ... Barrell, B. G. (2001). Massive gene decay in the leprosy bacillus. *Nature*, 409, 1007-1011. <https://doi.org/10.1038/35059006>

da Silva, S. A., Mazini, P. S., Reis, P. G., Sell, A. M., Tsuneto, L. T., Peixoto, P. R., & Visentainer, J. el. (2009). HLA-DR and HLA-DQ alleles in patients from the south of Brazil: markers for leprosy susceptibility and resistance. *BMC Infectious Diseases*, 9(1), 134. <https://doi.org/10.1186/1471-2334-9-134>

Daniel, T. (1997). *Captain of Death, The Story of Tuberculosis*. University of Rochester Press.
de Vries, J., & Perry, B. (1985). Leprosy case detection rates by age, sex, and polar type under leprosy control conditions. *American Journal of Epidemiology*, 121(3), 403-413. <https://doi.org/10.1093/oxfordjournals.aje.a114012>

del Rio, A. P. T., Sachetto, Z., Sampaio-Barros, P. D., Marques-Neto, J. F., Londe, A. C. S., & Bertolo, M. B. (2013). HLA Markers for Poor Prognosis in Systemic Sclerosis Brazilian Patients. *Disease Markers*, 35, 73-78. <https://doi.org/10.1155/2013/301415>

Delgado, J. C., Turbay, D., Yunis, E. J., Yunis, J. J., Morton, E. D., Bhol, K., Norman, R., Alper, C. A., Good, R. A., & Ahmed, R. (1996). A common major histocompatibility complex class II allele HLA-DQB1* 0301 is present in clinical variants of pemphigoid. *Proceedings of the National Academy of Sciences*, 93(16), 8569-8571. <https://doi.org/10.1073/pnas.93.16.8569>

Donoghue, H. D., Marcsik, A., Matheson, C., Vernon, K., Nuorala, E., Molto, J. E., Greenblatt, C. L., & Spigelman, M. (2005). Co-infection of Mycobacterium tuberculosis and Mycobacterium leprae in human archaeological samples: A possible explanation for the historical decline of leprosy. *Proceedings of the Royal Society B: Biological Sciences*, 272(1561), 389-394. <https://doi.org/10.1098/rspb.2004.2966>

Dubaniewicz, A., Lewko, B., Moszkowska, G., Zamorska, B., & Stepinski, J. (2000). Molecular subtypes of the HLA-DR antigens in pulmonary tuberculosis. *International Journal of Infectious Diseases*, 4(3), 129-133. [https://doi.org/10.1016/S1201-9712\(00\)90073-0](https://doi.org/10.1016/S1201-9712(00)90073-0)

El-Bendary, M., Neamatallah, M., Elalfy, H., Besheer, T., Kamel, E., Mousa, H., Eladl, A.-H., El-Setouhy, M., El-Gilany, A.-H., El-Waseef, A., & Esmat, G. (2019). HLA Class II-DRB1 Alleles with Hepatitis C Virus Infection Outcome in Egypt: A Multicentre Family-based Study. *Annals of Hepatology*, 18(1), 68-77. <https://doi.org/10.5604/01.3001.0012.7864>

El-Chennawi, F. A., Auf, F. A., Metwally, S. S., Mosaad, Y. M., El-Wahab, M. A., & Tawhid, Z. E. (2008). HLA-Class II Alleles in Egyptian Patients with Hepatocellular Carcinoma. *Immunological Investigations*, 37(7), 661–674. <https://doi.org/10.1080/08820130802111605>

Ellis, M. C., Hetisimer, A. H., Ruddy, D. A., Hansen, S. L., Kronmal, G. S., McClelland, E., Quintana, L., Drayna, D. T., Aldrich, M. S., & Mignot, E. (1997). HLA class II haplotype and sequence analysis support a role for DQ in narcolepsy. *Immunogenetics*, 46(5), 410–417. <https://doi.org/10.1007/s002510050295>

Escamila-Tilch, M., Torres-Carillo, N. M., Payan, R. R., Aguilar-Medina, M., Salazar, M. I., Fafutis-Morris, M., Arenas-Guzman, R., Estrada-Parra, S., Estrada-Garcia, I., & Granados, J. (2013). Association of genetic polymorphism of HLA-DRB1 antigens with the susceptibility to lepromatous leprosy. *Biomedical Reports*, 1(6), 945–949. <https://doi.org/10.3892/br.2013.167>

Esmali, N., Mortazavi, H., Chams-Davatchi, S., Daneshpazoooh, M., Damavandi, M., Aryanian, Z., & Amirzargar, A. (2013). Association between HLA-DQB1*03:01 and Bullous Pemphigoid in Iranian Patients. *Iranian Journal of Immunology*, 10(1), 1–9.

Falcón-Lezama, J. A., Ramos, C., Zuñiga, J., Juárez-Palma, L., Rangel-Flores, H., García-Trejo, A. R., Acunha-Alonzo, V., Granados, J., & Vargas-Alarcón, G. (2009). HLA class I and II polymorphisms in Mexican Mestizo patients with dengue fever. *Acta Tropica*, 112(2), 193–197. <https://doi.org/10.1016/j.actatropica.2009.07.025>

Faul, F., Erdfelder, E., Buchner, A., & Lang, A.-G. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods*, 41(4), 1149–1160. <https://doi.org/10.3758/BRM.41.4.1149>

Fernandez, O., Fernandez, V., Alonso, A., Caballero, A., Luque, G., Bravo, M., Leon, A., Mayorga, C., Leyva, L., & de Ramon, E. (2004). DQB1*0602 allele shows a strong association with multiple sclerosis in patients in Malaga, Spain. *Journal of Neurology*, 251(4), 440–444. <https://doi.org/10.1007/s00415-004-0350-2>

Fernandez-Mestre, M., Layrisse, Z., Montagnani, S., Acquatella, H., Catalioti, F., Matos, M., Balbas, O., Makhatadze, N., Dominguez, E., & Madrigal, A. (1998). Influence of the HLA class II polymorphism in chronic Chagas' disease. *Parasite Immunology*, 20(4), 197–203.

Fernández-Mestre, M., Padrón-Lowe, D., Salazar-Alcalá, E., & Blanco-Pérez, F. (2022). Role of HLA-DQB1 alleles in the risk, signs and symptoms, and severity of celiac disease in a Venezuelan population. *Revista de Gastroenterología de México (English Edition)*. <https://doi.org/10.1016/j.rgmxen.2022.03.008>

Fox, J. (2008). *Applied Regression Analysis and Generalized Linear Models* (2nd edition). Sage Publications.

Franceschi, D. S. A., Tsuneto, L. T., Mazini, P. S., Sacramento, W. S. do, Reis, P. G., Rudnick, C. C. C., Clementino, S. L., Sell, A. M., & Visentainer, J. E. L. (2011a). Class-I human leukocyte alleles in leprosy patients from Southern Brazil. *Revista Da Sociedade Brasileira de Medicina Tropical*, 44(5), 616–620. <https://doi.org/10.1590/S0037-86822011000500018>

Franceschi, D. S. A., Tsuneto, L. T., Mazini, P. S., Sacramento, W. S. do, Reis, P. G., Rudnick, C. C. C., Clementino, S. L., Sell, A. M., & Visentainer, J. E. L. (2011b). Class-I human leukocyte alleles in leprosy patients from Southern Brazil. *Revista Da Sociedade Brasileira de Medicina Tropical*, 44(5), 616–620. <https://doi.org/10.1590/S0037-86822011000500018>

Fu, L.-Y., Xu, C.-C., Zheng, X.-D., Chen, G., Zhu, J., & Wang, P.-G. (2018). Association of HLA-DQB1*03:03 with pityriasis rosea in Chinese patients. *Clinical and Experimental Dermatology*, 43(4), 389–392. <https://doi.org/10.1111/ced.13313>

Fu, Q., Posth, C., Hajdinjak, M., Petr, M., Mallick, S., Fernandes, D., Furtwängler, A., Haak, W., Meyer, M., Mittnik, A., Nickel, B., Peltzer, A., Rohland, N., Slon, V., Talamo, S., Lazaridis, I., Lipson, M., Mathieson, I., Schiffels, S., ... Reich, D. (2016). The genetic history of Ice Age Europe. In *Nature* (Vol. 534, pp. 200–205). <https://doi.org/10.1038/nature17993>

Furukawa, H., Kawasaki, A., Oka, S., Ito, I., Shimada, K., Sugii, S., Hashimoto, A., Komiya, A., Fukui, N., Kondo, Y., Ito, S., Hayashi, T., Matsumoto, I., Kusaoi, M., Amano, H., Nagai, T., Hirohata, S., Setoguchi, K., Kono, H., ... Tohma, S. (2014). Human Leukocyte Antigens and Systemic Lupus Erythematosus: A Protective Role for the HLA-DR6 Alleles DRB1*13:02 and *14:03. *PLoS ONE*, 9(2), e87792. <https://doi.org/10.1371/journal.pone.0087792>

Furukawa, H., Oka, S., Kawasaki, A., Shimada, K., Sugii, S., Matsushita, T., Hashimoto, A., Komiya, A., Fukui, N., Kobayashi, K., Osada, A., Ihata, A., Kondo, Y., Nagai, T., Setoguchi, K., Okamoto, A., Okamoto, A., Chiba, N., Suematsu, E., ... Tsuchiya, N. (2016). Human Leukocyte Antigen and Systemic Sclerosis in Japanese: The Sign of the Four Independent Protective Alleles, DRB1*13:02, DRB1*14:06, DQB1*03:01, and DPB1*02:01. *PLOS ONE*, 11(4), e0154255. <https://doi.org/10.1371/journal.pone.0154255>

Furukawa, H., Oka, S., Shimada, K., Hashimoto, A., & Tohma, S. (2015). Human leukocyte antigen polymorphisms and personalized medicine for rheumatoid arthritis. *Journal of Human Genetics*, 60(11), 691–696. <https://doi.org/10.1038/jhg.2015.36>

Gai, E., Gruia, A., Boldeanu, F., Cristea, M., Serban, M., Arghirescu, S., Jinca, C., Gib, L., Ordodi, V., Tatu, C., & Paunescu, V. (2010). Relationship between HLA A*02 A*03 association and the risk of childhood acute leukemia. *Analele Științifice Ale Universității Alexandru Ioan Cuza Din Iași, Secțiunea II A : Genetica Si Biologie Moleculara*, 11.

Gao, X., Nelson, G. W., Karacki, P., Martin, M. P., Phair, J., Kaslow, R., Goedert, J. J., Buchbinder, S., Hoots, K., Vlahov, D., O'Brien, S. J., & Carrington, M. (2001). Effect of a Single Amino Acid Change in MHC Class I Molecules on the Rate of Progression to AIDS. *New England Journal of Medicine*, 344(22), 1668–1675. <https://doi.org/10.1056/NEJM200105313442203>

García Borrás, S., Moreno, J., Abraham, N., Tanno, H., García Laplaca, M., Rossi, M. C., Racca, L., & Racca, A. (2011). HLA-DRB1 as a risk or resistance factor in Autoimmune Hepatitis. *Inmunología*, 30(4), 115–118. <https://doi.org/10.1016/j.inmuno.2011.07.003>

Gindhart, P. S. (1973). Growth standards for the tibia and radius in children aged one month through eighteen years. *American Journal of Physical Anthropology*, 39(1), 41–48. <https://doi.org/10.1002/ajpa.1330390107>

Gorodezky, C., Alaez, C., Munguía, A., Cruz, R., Vazquez, A., Camacho, A., Flores, O., Rodriguez, M., & Rodriguez, O. (2004). Molecular mechanisms of MHC linked susceptibility in leprosy: towards the development of synthetic vaccines. *Tuberculosis*, 84(1–2), 82–92. <https://doi.org/10.1016/j.tube.2003.08.009>

Gorodezky, C., Flores, J., Arevalo, N., Castro, L., Silva, A., & Rodrigues, O. (1987). Tuberculoid leprosy in Mexicans is associated with HLA-DR3. *Leprosy Review*, 58(4), 401–406. <https://doi.org/10.5935/0305-7518.19870042>

Gunawan, H., Achdiat, P. A., Rahardjo, R. M., Hindritiani, R., & Suwarsa, O. (2020). Frequent testicular involvement in multibacillary leprosy. *International Journal of Infectious Diseases*, 90, 60–64. <https://doi.org/10.1016/j.ijid.2019.10.013>

Haak, W., Lazaridis, I., Patterson, N., Rohland, N., Mallick, S., Llamas, B., Brandt, G., Nordenfelt, S., Harney, E., Stewardson, K., Fu, Q., Mittnik, A., Bánffy, E., Economou, C., Francken, M., Friederich, S., Pena, R. G., Hallgren, F., Khartanovich, V., ... Reich, D. (2015). Massive migration from the steppe was a source for Indo-European languages in Europe. *Nature*, 522, 207–211. <https://doi.org/10.1038/nature14317>

Haapasalo, K., Koskinen, L. L. E., Suvilehto, J., Jousilahti, P., Wolin, A., Suomela, S., Trembath, R., Barker, J., Vuopio, J., Kere, J., Jokiranta, T. S., & Saavalainen, P. (2018). The Psoriasis Risk Allele *HLA-C*06:02* Shows Evidence of Association with Chronic or Recurrent Streptococcal Tonsillitis. *Infection and Immunity*, 86(10), e00304–18. <https://doi.org/10.1128/IAI.00304-18>

Harcourt, G., Hellier, S., Bunce, M., Satsangi, J., Collier, J., Chapman, R., Phillips, R., & Klenerman, P. (2001). Effect of HLA class II genotype on T helper lymphocyte responses and viral control in hepatitis C virus infection. *Journal of Viral Hepatitis*, 8(3), 174–179. <https://doi.org/10.1046/j.1365-2893.2001.00289.x>

He, D., Wang, J., Yi, L., Guo, X., Guo, S., Guo, G., Tu, W., Wu, W., Yang, L., Xiao, R., Li, Y., Chu, H., Lai, S., Jin, L., Zou, H., Reveille, J. D., Assassi, S., Mayes, M. D., & Zhou, X. (2014). Association of the HLA-DRB1 with Scleroderma in Chinese Population. *PLoS ONE*, 9(9), e106939. <https://doi.org/10.1371/journal.pone.0106939>

Herbig, A., Maixner, F., Bos, K., Zink, A., Krause, J., & Huson, D. (2016). MALT: Fast alignment and analysis of metagenomic DNA sequence data applied to the Tyrolean Iceman. *BioRxiv*. <https://doi.org/10.1101/050559>

Höhler, T., Gerken, G., Notghi, A., Lubjuhn, R., Taheri, H., Protzer, U., Löhr, H. F., Schneider, P. M., Meyer zum Büschenfelde, K.-H., & Rittner, C. (1997). HLA-DRB1*1301 AND *1302 protect against chronic hepatitis B. *Journal of Hepatology*, 26(3), 503–507. [https://doi.org/10.1016/S0168-8278\(97\)80414-X](https://doi.org/10.1016/S0168-8278(97)80414-X)

Holoshitz, J. (2013). The quest for better understanding of HLA-disease association: scenes from a road less travelled by. *Discovery Medicine*, 16(87), 93–101.

Honeyborne, I., Codoñer, F. M., Leslie, A., Tudor-Williams, G., Luzzi, G., Ndung'u, T., Walker, B. D., Goulder, P. J., & Prado, J. G. (2010). HLA-Cw*03-Restricted CD8 + T-Cell Responses Targeting the HIV-1 Gag Major Homology Region Drive Virus Immune Escape and Fitness Constraints Compensated for by Intracodon Variation. *Journal of Virology*, 84(21), 11279–11288. <https://doi.org/10.1128/JVI.01144-10>

Hovhannisyan, A., Madelian, V., Avagyan, S., Nazaretyan, M., Hyussyan, A., Sirunyan, A., Arakelyan, R., Manukyan, Z., Yepiskoposyan, L., Mayilyan, K. R., & Jordan, F. (2022). HLA-C*04:01 Affects HLA Class I Heterozygosity and Predicted Affinity to SARS-CoV-2 Peptides, and in Combination With Age and Sex of Armenian Patients Contributes to COVID-19 Severity. *Frontiers in Immunology*, 13, 769900. <https://doi.org/10.3389/fimmu.2022.769900>

Hsieh, N.-K., Chu, C.-C., Lee, N.-S., Lee, H.-L., & Lin, M. (2010). Association of HLA-DRB1*0405 with resistance to multibacillary leprosy in Taiwanese. *Human Immunology*, 71(7), 712–716. <https://doi.org/10.1016/j.humimm.2010.03.007>

Huang, J., Huang, K., Xu, R., Wang, M., Liao, Q., Xiong, H., Li, C., Tang, X., Shan, Z., Zhang, M., Rong, X., Nelson, K., & Fu, Y. (2016). The Associations of HLA-A*02:01 and DRB1*11:01 with Hepatitis C Virus Spontaneous Clearance Are Independent of IL28B in the Chinese Population. *Scientific Reports*, 6(1), 31485. <https://doi.org/10.1038/srep31485>

Huson, D. H., Auch, A. F., Qi, J., & Schuster, S. C. (2007). MEGAN analysis of metagenomic data. *Genome Research*, 17, 377–386. <https://doi.org/10.1101/gr.5969107>

Immel, A., Pierini, F., Rinne, C., Meadows, J., Barquera, R., Szolek, A., Susat, J., Böhme, L., Dose, J., Bonczarowska, J., Drummer, C., Fuchs, K., Ellinghaus, D., Kässens, J. C., Furholt, M., Kohlbacher, O., Schade-Lindig, S., Franke, A., Schreiber, S., ... Krause-Kyora, B. (2021). Genome-wide study of a Neolithic Wartberg grave community reveals distinct HLA variation and hunter-gatherer ancestry. *Communications Biology*. <https://doi.org/10.1038/s42003-020-01627-4>

James, L., & Georgopoulos, A. (2019). The Human Leukocyte Antigen (HLA) DRB1*13:02 Allele Protects against Dementia in Continental Western Europe. *Journal of Neurology & Neuromedicine*, 4(5), 1–6. <https://doi.org/10.29245/2572.942X/2019/5.1253>

Joko, S., Numaga, J., Kawashima, H., Namisato, M., & Maeda, H. (2000). Human leukocyte antigens in forms of leprosy among Japanese patients. *International Journal of Leprosy and Other Mycobacterial Diseases*, 68(1), 49–56.

Kaburaki, T., Takamoto, M., Numaga, J., Kawashima, H., Makoto, A., Ohnogi, Y., Harihara, S., Kuwata, S., & Fujito, T. (2010). Genetic association of HLA-A*2601 with ocular Behçet's disease in Japanese patients. *Clinical and Experimental Rheumatology*, 28, 39–44.

Kawasaki, A., Hasebe, N., Hidaka, M., Hirano, F., Sada, K., Kobayashi, S., Yamada, H., Furukawa, H., Yamagata, K., Sumida, T., Miyasaka, N., Tohma, S., Ozaki, S., Matsuo, S., Hashimoto, H., Makino, H., Arimura, Y., Harigai, M., & Tsuchiya, N. (2016). Protective Role of HLA-DRB1*13:02 against Microscopic Polyangiitis and MPO-ANCA-Positive Vasculitides in a Japanese Population: A Case-Control Study. *PLOS ONE*, 11(5), e0154393. <https://doi.org/10.1371/journal.pone.0154393>

Khanna, N., Singh, M., Rasool, S., Ammini, A., Bhatla, N., Garg, V., Rao, S., & Bhattacharya, S. N. (2014). Menstrual irregularities, fertility status, and ovarian function in female patients with leprosy in India. *International Journal of Dermatology*, 53(9), 1114–1118. <https://doi.org/10.1111/j.1365-4632.2012.05771.x>

Klepinger, L. (2006). *Fundamentals of forensic anthropology*. John Wiley & Sons.

Klioze, A. M., & Ramos-Caro, F. A. (2000). Visceral leprosy. In *International Journal of Dermatology*, 39(9), 641-658. <https://doi.org/10.1046/j.1365-4362.2000.00860.x>

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*, 194(1), 3-6. <https://doi.org/10.1016/j.aanat.2011.03.008>

Knussmann, R., & Martin, R. (1988). *Anthropologie : Handbuch der vergleichenden Biologie des Menschen*. Fischer.

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, 1569. <https://doi.org/10.1038/s41467-018-03857-x>

Kummee, P., Tangkijvanich, P., Poovorawan, Y., & Hirankarn, N. (2007). Association of HLA-DRB1*13 and TNF-alpha gene polymorphisms with clearance of chronic hepatitis B infection and risk of hepatocellular carcinoma in Thai population. *Journal of Viral Hepatitis*, 14(12), 841-848. <https://doi.org/10.1111/j.1365-2893.2007.00880.x>

Kuwana, M., Kaburaki, J., Arnett, F. C., Howard, R. F., Medsger, T. A., & Wright, T. M. (1999). Influence of ethnic background on clinical and serologic features in patients with systemic sclerosis and anti-DNA topoisomerase I antibody. *Arthritis & Rheumatism*, 42(3), 465–474. [https://doi.org/10.1002/1529-0131\(199904\)42:3<465::AID-ANR11>3.0.CO;2-Y](https://doi.org/10.1002/1529-0131(199904)42:3<465::AID-ANR11>3.0.CO;2-Y)

Lara-Marquez, M., Yunis, J., Layrisse, Z., Ortega, F., Carvallo-Gil, E., Montagani, N., Makhatadze, N., Pocino, M., Granja, C., & Yunis, E. (1999). Immunogenetics of atopic asthma:

association of DRB1*1101 DQA1*0501 DQB1*0301 haplotype with *Dermatophagoides* spp.-sensitive asthma in a sample of the Venezuelan population. *Clinical & Experimental Allergy*, 29(1), 60–71. <https://doi.org/10.1046/j.1365-2222.1999.00461.x>

Layrisse, Z., Fernandez, M. T., Montagnani, S., Matos, M., Balbas, O., Herrera, F., Colorado, I. A., Catalioti, F., & Acquatella, H. (2000). HLA-C*03 is a risk factor for cardiomyopathy in chagas disease. *Human Immunology*, 61(9), 925–929. [https://doi.org/10.1016/S0198-8859\(00\)00161-0](https://doi.org/10.1016/S0198-8859(00)00161-0)

Lazaridis, I., Patterson, N., Mittnik, A., Renaud, G., Mallick, S., Kirsanow, K., Sudmant, P. H., Schraiber, J. G., Castellano, S., Lipson, M., Berger, B., Economou, C., Bollongino, R., Fu, Q., Bos, K. I., Nordenfelt, S., Li, H., de Filippo, C., Prüfer, K., ... Krause, J. (2014). Ancient human genomes suggest three ancestral populations for present-day Europeans. *Nature*, 513, 409–413. <https://doi.org/10.1038/nature13673>

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>

Li, S., Zhou, D., & Liu, M.-Y. (2014). Associations between polymorphisms of HLA-B gene and postmenopausal osteoporosis in Chinese Han population. *International Journal of Immunogenetics*, 41(4), 324–329. <https://doi.org/10.1111/iji.12130>

Lipatov, M., Sanjeev, K., Patro, R., & Veeramah, K. (2015). Maximum Likelihood Estimation of Biological Relatedness from Low Coverage Sequencing Data. *BioRxiv*. <https://doi.org/10.1101/023374>

Lockwood, D. N., & Saunderson, P. R. (2012). Nerve damage in leprosy: a continuing challenge to scientists, clinicians and service providers. *International Health*, 4(2), 77–85. <https://doi.org/10.1016/j.inhe.2011.09.006>

Lovejoy, C. O., Meindl, R. S., Pryzbeck, T. R., & Mensforth, R. P. (1985). Chronological metamorphosis of the auricular surface of the ilium: A new method for the determination of adult skeletal age at death. *American Journal of Physical Anthropology*, 68(1), 15–28. <https://doi.org/10.1002/ajpa.1330680103>

Lurie, M. B. (2008). *A Pathogenetic Relationship Between Tuberculosis and Leprosy: The Common Denominators in the Tissue Response to Mycobacteria* (pp. 340–343). <https://doi.org/10.1002/9780470718933.ch24>

Mace, A. (1962). *Sample Size Determination*. Chapman & Hall.

Madeleine, M., Brumback, B., Cushing-Haugen, K., Schwartz, S., Daling, J., Smith, A., Nelson, J., Porter, P., Shera, K., McDougall, J., & Galloway, D. (2002). Human Leukocyte Antigen Class II and Cervical Cancer Risk: A Population-Based Study. *The Journal of Infectious Diseases*, 186(11), 1565–1574.

Malhotra, D., Darvishi, K., Sood, S., Sharma, S., Grover, C., Relhan, V., Reddy, B. S. N., & Bamezai, R. N. K. (2005). IL-10 promoter single nucleotide polymorphisms are significantly associated with resistance to leprosy. *Human Genetics*, 118(2), 295–300. <https://doi.org/10.1007/s00439-005-0042-8>

Manchester, K. (1984). Tuberculosis and leprosy in antiquity: An interpretation. *Medical History*, 28, 162–173. <https://doi.org/10.1017/S0025727300035705>

Mathieson, I., Lazaridis, I., Rohland, N., Mallick, S., Patterson, N., Roodenberg, S. A., Harney, E., Stewardson, K., Fernandes, D., Novak, M., Sirak, K., Gamba, C., Jones, E. R., Llamas, B., Dryomov, S., Pickrell, J., Arsuaga, J. L., de Castro, J. M. B., Carbonell, E., ... Reich, D. (2015). Genome-wide patterns of selection in 230 ancient Eurasians. *Nature*, 528, 499–503. <https://doi.org/10.1038/nature16152>

Mazzi, S. (1992). *Salute e societa` durante il Medioevo*. Mailand.

Meindl, R. S., & Lovejoy, C. O. (1985). Ectocranial suture closure: A revised method for the determination of skeletal age at death based on the lateral-anterior sutures. *American Journal of Physical Anthropology*, 68(1), 57–66. <https://doi.org/10.1002/ajpa.1330680106>

Miles, A. (1963). The Dentition in the Assessment of Individual Age of Skeletal Material. In D. Brothwell (Ed.), *Dental Anthropology. Symposia of the Society for the Study of Human Biology* (pp. 191–209). Brothwell (Hrsg.).

Mokbel, A. N., Al-Zifzaf, D. S., ElSawy, W. S., & ElGabarty, S. (2015). Association of HLA-DQB1*06 with susceptibility to systemic lupus erythematosus in Egyptians. *The Egyptian Rheumatologist*, 37(1), 17–22. <https://doi.org/10.1016/j.ejr.2014.06.005>

Møller-Christensen, V. (1952). Case of leprosy from the Middle Ages of Denmark. *Acta Medica Scandinavica. Supplementum*. <https://doi.org/10.1111/j.0954-6820.1952.tb13354.x>

Møller-Christensen, V. (1974). Changes in the anterior nasal spine and the alveolar process of the macillae in leprosy a clinical examination. *International Journal of Leprosy and Other Mycobacterial Diseases*, 42(4), 431–435.

Møller-Christensen, V. (1978). *Leprosy Changes of the Skull*. Odense University Press.

Monot, M., Honoré, N., Garnier, T., Araoz, R., Coppée, J.-Y., Lacroix, C., Sow, S., Spencer, J. S., Truman, R. W., Williams, D. L., Gelber, R., Virmond, M., Flageul, B., Cho, S.-N., Ji, B., Paniz-Mondolfi, A., Convit, J., Young, S., Fine, P. E., ... Cole, S. T. (2005). On the Origin of Leprosy. *Science*, 308(5724), 1040–1042. <https://doi.org/10.1126/science/1109759>

Monot, M., Honoré, N., Garnier, T., Zidane, N., Sherafi, D., Paniz-Mondolfi, A., Matsuoka, M., Taylor, G. M., Donoghue, H. D., Bouwman, A., Mays, S., Watson, C., Lockwood, D., Khamesipour, A., Dowlati, Y., Jianping, S., Rea, T. H., Vera-Cabrera, L., Stefani, M. M., ... Cole,

- S. T. (2009). Comparative genomic and phylogeographic analysis of *Mycobacterium leprae*. *Nature Genetics*, 41, 1282–1289. <https://doi.org/10.1038/ng.477>
- Monroy Kuhn, J. M., Jakobsson, M., & Günther, T. (2018). Estimating genetic kin relationships in prehistoric populations. *PLOS ONE*, 13(4), e0195491. <https://doi.org/10.1371/journal.pone.0195491>
- Moorrees, C. F. A., Fanning, E. A., & Hunt, E. E. (1963). Age Variation of Formation Stages for Ten Permanent Teeth. *Journal of Dental Research*, 42(6), 1490–1502. <https://doi.org/10.1177/00220345630420062701>
- Motta, P. M., Cech, N., Fontan, C., Fernando Giménez, M., Lodeiro, N., Marinic, K., Laura Molinari, M., Graciela Sotelob, M., & de Sorrentino, A. H. (2007). Papel de las moléculas HLA-DR y HLA-DQ en la lepra multibacilar y paucibacilar en la provincia del Chaco; Argentina. *Enfermedades Infecciosas y Microbiología Clínica*, 25(10), 627–631. <https://doi.org/10.1157/13112938>
- Muazzam, A., Mansoor, A., Ali, L., Siddiqi, S., Hameed, A., Ajmal, M., & Mazhar, K. (2013). Association of HLA-DRB1 and -DQB1 alleles and haplotypes with rheumatoid arthritis in a Pakistani population. *Arthritis Research & Therapy*, 15(4), R95. <https://doi.org/10.1186/ar4275>
- Nakamura, J., Meguro, A., Ishii, G., Mihara, T., Takeuchi, M., Mizuki, Y., Yuda, K., Yamane, T., Kawagoe, T., Ota, M., & Mizuki, N. (2019). The association analysis between HLA-A*26 and Behçet's disease. *Scientific Reports*, 9(1), 4426. <https://doi.org/10.1038/s41598-019-40824-y>
- 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>
- Niens, M., Jarrett, R. F., Hepkema, B., Nolte, I. M., Diepstra, A., Platteel, M., Kouprie, N., Delury, C. P., Gallagher, A., Visser, L., Poppema, S., te Meerman, G. J., & van den Berg, A. (2007). HLA-A*02 is associated with a reduced risk and HLA-A*01 with an increased risk of developing EBV+ Hodgkin lymphoma. *Blood*, 110(9), 3310–3315. <https://doi.org/10.1182/blood-2007-05-086934>
- Noordeen, S. (1994). Epidemiology of Leprosy. In *Mycobacteria. hapman & Hall Medical Microbiology Series*. Springer.
- Novotny, V. (1975). Diskriminanzanalyse der Geschlechtsmerkmale auf dem Os coxae beim Menschen. *XIII Czechoslovakian Anthropological Congress*.
- Ohara, N., Matsuoka, M., Nomaguchi, H., Naito, M., & Yamada, T. (2000). Inhibition of multiplication of *Mycobacterium leprae* in mouse foot pads by recombinant Bacillus Calmette-Guérin (BCG). *Vaccine*, 18(14), 1294–1297. [https://doi.org/10.1016/S0264-410X\(99\)00420-X](https://doi.org/10.1016/S0264-410X(99)00420-X)

Olvera, A., Pérez-Álvarez, S., Ibarrondo, J., Ganoza, C., Lama, J. R., Lucchetti, A., Cate, S., Hildebrand, W., Bernard, N., Gomez, L., Sanchez, J., & Brander, C. (2015). The HLA-C*04. *AIDS*, 29(5), 507–517. <https://doi.org/10.1097/QAD.0000000000000574>

Ombrello, M. J., Remmers, E. F., Tachmazidou, I., Grom, A., Foell, D., Haas, J.-P., Martini, A., Gattorno, M., Özen, S., Prahalad, S., Zeff, A. S., Bohnsack, J. F., Mellins, E. D., Ilowite, N. T., Russo, R., Len, C., Hilario, M. O. E., Oliveira, S., Yeung, R. S. M., ... Yeung, R. S. (2015). *HLA-DRB1*11* and variants of the MHC class II locus are strong risk factors for systemic juvenile idiopathic arthritis. *Proceedings of the National Academy of Sciences*, 112(52), 15970–15975. <https://doi.org/10.1073/pnas.1520779112>

Ortner, D. J. (2003). Identification of Pathological Conditions in Human Skeletal Remains. In *Identification of Pathological Conditions in Human Skeletal Remains*. <https://doi.org/10.1016/B978-0-12-528628-2.X5037-6>

Pandi, S., Chinniah, R., Sevak, V., Ravi, P. M., Raju, M., Vellaippan, N. A., & Karuppiyah, B. (2021). Association of HLA–DRB1, DQA1 and DQB1 alleles and haplotype in Parkinson’s disease from South India. *Neuroscience Letters*, 765, 136296. <https://doi.org/10.1016/j.neulet.2021.136296>

Patterson, N., Moorjani, P., Luo, Y., Mallick, S., Rohland, N., Zhan, Y., Genschoreck, T., Webster, T., & Reich, D. (2012). Ancient admixture in human history. *Genetics*, 192(3), 1065–1093. <https://doi.org/10.1534/genetics.112.145037>

Patterson, N., Price, A. L., & Reich, D. (2006). Population structure and eigenanalysis. *PLoS Genetics* 2(12), e190. <https://doi.org/10.1371/journal.pgen.0020190>

Peltzer, A., Jäger, G., Herbig, A., Seitz, A., Knip, C., Krause, J., & Nieselt, K. (2016). EAGER: efficient ancient genome reconstruction. *Genome Biology*. <https://doi.org/10.1186/s13059-016-0918-z>

Pereyra, F., Jia, X., McLaren, P., Telenti, A., de Bakker, P., Walker, B., Ripke, S., Brumme, C., Pulit, S., Carrington, M., ... Zhao, M. (2010). The Major Genetic Determinants of HIV-1 Control Affect HLA Class I Peptide Presentation. *Science*, 330(6010), 1551–1557. <https://doi.org/10.1126/science.1195271>

Pfrenge, S., Neukamm, J., Guellil, M., Keller, M., Molak, M., Avanzi, C., Kushniarevich, A., Montes, N., Neumann, G. U., Reiter, E., Tukhbatova, R. I., Berezina, N. Y., Buzhilova, A. P., Korobov, D. S., Suppersberger Hamre, S., Matos, V. M. J., Ferreira, M. T., González-Garrido, L., Wasterlain, S. N., ... Schuenemann, V. J. (2021). Mycobacterium leprae diversity and population dynamics in medieval Europe from novel ancient genomes. *BMC Biology*, 19(1), 220. <https://doi.org/10.1186/s12915-021-01120-2>

Phenice, T. W. (1969). A newly developed visual method of sexing the os pubis. *American Journal of Physical Anthropology*, 30(2), 297–301. <https://doi.org/10.1002/ajpa.1330300214>

Pierini, F., Nutsua, M., Böhme, L., Özer, O., Bonczarowska, J., Susat, J., Franke, A., Nebel, A., Krause-Kyora, B., & Lenz, T. L. (2020). Targeted analysis of polymorphic loci from low-coverage shotgun sequence data allows accurate genotyping of HLA genes in historical human populations. *Scientific Reports* 10, 7339. <https://doi.org/10.1038/s41598-020-64312-w>

Pilli, E., Modi, A., Serpico, C., Achilli, A., Lancioni, H., Lippi, B., Bertoldi, F., Gelichi, S., Lari, M., & Caramelli, D. (2013). Monitoring DNA Contamination in Handled vs. Directly Excavated Ancient Human Skeletal Remains. *PLoS ONE*, 8(1), e52524. <https://doi.org/10.1371/journal.pone.0052524>

Piontek, J. (1996). *Biologia populacji pradziejowych, Skrypt dla studentów biologii i archeologii*. Wydawnictwo Naukowe Uniwersytetu Im. Adama Mickiewicza w Poznaniu.

Prescher, A., & Wagner, P. (2016). *Aachen, Melaten. Der Friedhof des mittelalterlichen Leprosoriums an der Via Regia*. Verlag Philipp von Zabern.

Priyadarshini, A., George, R., Daniel, D., Varughese, S., & Jayaseelan, V. (2018). Association between human leukocyte antigen-DRB1 and human leukocyte antigen-DQB1 alleles and pemphigus vulgaris in Indian patients: A case-control study. *Indian Journal of Dermatology, Venereology and Leprology*, 84(3), 280. https://doi.org/10.4103/ijdv.IJDVL_1014_16

Raffcliffe, C. (2006). *Leprosy in Medieval England*. Boydell Press.

Rambukkana, A. (2004). Mycobacterium leprae-induced demyelination: a model for early nerve degeneration. *Current Opinion in Immunology*, 16(4), 511–518. <https://doi.org/10.1016/j.coi.2004.05.021>

Santaniello, A., Salazar, G., Lenna, S., Antonioli, R., Colombo, G., Beretta, L., & Scorza, R. (2006). HLA-B35 upregulates the production of endothelin-1 in HLA-transfected cells: a possible pathogenetic role in pulmonary hypertension. *Tissue Antigens*, 68(3), 239–244. <https://doi.org/10.1111/j.1399-0039.2006.00657.x>

Saporta, L., & Yuksel, A. (1994). Androgenic status in patients with lepromatous leprosy. *British Journal of Urology*, 74(2), 221–224. <https://doi.org/10.1111/j.1464-410X.1994.tb16590.x>

Scheuer, L., & Black, S. (2004). The Juvenile Skeleton. In *The Juvenile Skeleton*.

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), e1006997. <https://doi.org/10.1371/journal.ppat.1006997>

Schuenemann, V. J., Singh, P., Mendum, T. A., Krause-Kyora, B., Jäger, G., Bos, K. I., Herbig, A., Economou, C., Benjak, A., Busso, P., Nebel, A., Boldsen, J. L., Kjellström, A., Wu, H., Stewart, G. R., Taylor, G. M., Bauer, P., Lee, O. Y. C., Wu, H. H. T., ... Krause, J. (2013).

Genome-wide comparison of medieval and modern *Mycobacterium leprae*. *Science*, 341(6142), 179-183. <https://doi.org/10.1126/science.1238286>

Schulze, W. (1993). *St. Jürgen. Lepra in Schleswig-Holsten und Hamburg*. Dreves.

Serjeantson, S. W., Gao, X., Hawkins, B. R., Higgins, D. A., & Yu, Y. L. (1992). Novel HLA-DR2-related haplotypes in Hong Kong Chinese implicate the DQB1*0602 allele in susceptibility to multiple sclerosis. *European Journal of Immunogenetics*, 19(1-2), 11-19. <https://doi.org/10.1111/j.1744-313X.1992.tb00043.x>

Shankarkumar, U. (2003). HLA-DRB1*03 and DQB1*0302 associations in a subset of patients severely affected with systemic lupus erythematosus from western India. *Annals of the Rheumatic Diseases*, 62(1), 92-93. <https://doi.org/10.1136/ard.62.1.92>

Shankarkumar, U., Ghosh, K., Badakere, S., & Mohanty, D. (2003). Novel HLA Class I Alleles Associated with Indian Leprosy Patients. *Journal of Biomedicine and Biotechnology*, 2003(3), 208-211. <https://doi.org/10.1155/S1110724303210019>

Shkurnikov, M., Nersisyan, S., Jankevic, T., Galatenko, A., Gordeev, I., Vechorko, V., & Tonevitsky, A. (2021). Association of HLA Class I Genotypes With Severity of Coronavirus Disease-19. *Frontiers in Immunology*, 12, 641900. <https://doi.org/10.3389/fimmu.2021.641900>

Siegfried, K. (2018). *Grabungsbericht. Fundstelle 5.02.73. Ratzeburger Allee zwischen Kronsfordter Allee und Am Brink (EBHL)*.

Simmonds, M. J., Howson, J. M. M., Heward, J. M., Carr-Smith, J., Franklyn, J. A., Todd, J. A., & Gough, S. C. L. (2007). A novel and major association of HLA-C in Graves' disease that eclipses the classical HLA-DRB1 effect. *Human Molecular Genetics*, 16(18), 2149-2153. <https://doi.org/10.1093/hmg/ddm165>

Song, E. Y., Shin, S., Park, K. U., Park, M. H., Sung, M.-W., Kim, K. H., & Kwon, T.-K. (2013). Associations of HLA-DRB1 and -DQB1 alleles with severe recurrent respiratory papillomatosis in Korean patients. *Human Immunology*, 74(8), 961-964. <https://doi.org/10.1016/j.humimm.2013.04.020>

Sousa, A. L. M., Fava, V. M., Sampaio, L. H., Martelli, C. M. T., Costa, M. B., Mira, M. T., & Stefani, M. M. A. (2012). Genetic and Immunological Evidence Implicates Interleukin 6 as a Susceptibility Gene for Leprosy Type 2 Reaction. *The Journal of Infectious Diseases*, 205(9), 1417-1424. <https://doi.org/10.1093/infdis/jis208>

Spurkland, A., Celius, E. G., Knutsen, I., Belske, A., Thorsby, E., & Vartdal, F. (1997). The HLA-DQ(α 1*0102, β 1*0602) heterodimer may confer susceptibility to multiple sclerosis in the absence of the HLA-DR(α 1*01, β 1*1501) heterodimer. *Tissue Antigens*, 50(1), 15-22. <https://doi.org/10.1111/j.1399-0039.1997.tb02828.x>

Stloukal, M., & Hanáková, H. (1978). Die Länge der Längsknochen Altslawischer Bevölkerungen - Unter besonderer Berücksichtigung von Wachstumsfragen. *HOMO- Journal of Comparative Human Biology*, 29, 53–69.

Stone, A. C., Wilbur, A. K., Buikstra, J. E., & Roberts, C. A. (2009). Tuberculosis and leprosy in perspective. *American Journal of Physical Anthropology*, 140(S49), 66–94. <https://doi.org/10.1002/ajpa.21185>

Sun, Y., Liu, H., Wang, Z., Fu, X., Wang, C., Mi, Z., Sun, L., Bao, F., Yu, G., Zhou, G., & Zhang, F. (2018). The HLA-DQB1*03:01 Is Associated with Bullous Pemphigoid in the Han Chinese Population. *Journal of Investigative Dermatology*, 138(8), 1874–1877. <https://doi.org/10.1016/j.jid.2018.02.021>

Szolek, A., Schubert, B., Mohr, C., Sturm, M., Feldhahn, M., & Kohlbacher, O. (2014). OptiType: Precision HLA typing from next-generation sequencing data. *Bioinformatics*, 30(23), 3310–3316. <https://doi.org/10.1093/bioinformatics/btu548>

Tafti, M., Lammers, G. J., Dauvilliers, Y., Overeem, S., Mayer, G., Nowak, J., Pfister, C., Dubois, V., Eliaou, J.-F., Eberhard, H.-P., Liblau, R., Wierzbicka, A., Geisler, P., Bassetti, C. L., Mathis, J., Lecendreux, M., Khatami, R., Heinzer, R., Haba-Rubio, J., ... Tiercy, J.-M. (2016). Narcolepsy-Associated HLA Class I Alleles Implicate Cell-Mediated Cytotoxicity. *Sleep*, 39(3), 581–587. <https://doi.org/10.5665/sleep.5532>

Taylor, G. M., Tucker, K., Butler, R., Pike, A. W. G., Lewis, J., Roffey, S., Marter, P., Lee, O. Y. C., Wu, H. H. T., Minnikin, D. E., Besra, G. S., Singh, P., Cole, S. T., & Stewart, G. R. (2013). Detection and Strain Typing of Ancient Mycobacterium leprae from a Medieval Leprosy Hospital. *PLoS ONE*, 8(4), e62406. <https://doi.org/10.1371/journal.pone.0062406>

Thio, C. L., Thomas, D. L., Goedert, J. J., Vlahov, D., Nelson, K. E., Hilgartner, M. W., O'Brien, S. J., Karacki, P., Marti, D., Astemborski, J., & Carrington, M. (2001). Racial Differences in HLA Class II Associations with Hepatitis C Virus Outcomes. *The Journal of Infectious Diseases*, 184(1), 16–21. <https://doi.org/10.1086/321005>

Thursz, M., Yallop, R., Goldin, R., Trepo, C., & Thomas, H. C. (1999). Influence of MHC class II genotype on outcome of infection with hepatitis C virus. *The Lancet*, 354(9196), 2119–2124. [https://doi.org/10.1016/S0140-6736\(99\)91443-5](https://doi.org/10.1016/S0140-6736(99)91443-5)

Todd, T. W. (1921). Age changes in the pubic bone. *American Journal of Physical Anthropology*, 4(1), 1–70. <https://doi.org/10.1002/ajpa.1330040102>

Toyo-oka, L., Mahasirimongkol, S., Yanai, H., Mushiroda, T., Wattanapokayakit, S., Wichukchinda, N., Yamada, N., Smittipat, N., Juthayothin, T., Palittapongarnpim, P., Nedsuwan, S., Kantipong, P., Takahashi, A., Kubo, M., Sawanpanyalert, P., & Tokunaga, K. (2017). Strain-based HLA association analysis identified *HLA-DRB1*09:01* associated with modern strain tuberculosis. *HLA*, 90(3), 149–156. <https://doi.org/10.1111/tan.13070>

Vanderborght, P. R., Pacheco, A. G., Moraes, M. E., Antoni, G., Romero, M., Verville, A., Thai, V. H., Huong, N. T., Ba, N. N., Schurr, E., Sarno, E. N., & Moraes, M. O. (2007). HLA-DRB1*04 and DRB1*10 are associated with resistance and susceptibility, respectively, in Brazilian and Vietnamese leprosy patients. *Genes & Immunity*, 8(4), 320–324. <https://doi.org/10.1038/sj.gene.6364390>

Vogelsang, Th. M. (1965). Leprosy in Norway. *Medical History*, 9(1), 29–35. <https://doi.org/10.1017/S0025727300030118>

Wang, X., An, G., Wang, J., Zhang, Y., Li, Q., wei, H., Qiu, L., & Ru, K. (2018). The association of HLA-C alleles with multiple myeloma in Chinese patients. *Experimental Hematology & Oncology*, 7(1), 19. <https://doi.org/10.1186/s40164-018-0112-y>

Wang, Z., Sun, Y., Fu, X., Yu, G., Wang, C., Bao, F., Yue, Z., Li, J., Sun, L., Irwanto, A., Yu, Y., Chen, M., Mi, Z., Wang, H., Huai, P., Li, Y., & et al. (2016). A large-scale genome-wide association and meta-analysis identified four novel susceptibility loci for leprosy. *Nature Communications*, 7, 13760. <https://doi.org/https://doi.org/10.1038/ncomms13760>

Weiner, J., Suwalski, P., Holtgrewe, M., Rakitko, A., Thibeault, C., Müller, M., Patriki, D., Quedenau, C., Krüger, U., Ilinsky, V., Popov, I., Balnis, J., Jaitovich, A., Helbig, E. T., Lippert, L. J., Stubbemann, P., Real, L. M., Macías, J., Pineda, J. A., ... Heidecker, B. (2021). Increased risk of severe clinical course of COVID-19 in carriers of HLA-C*04:01. *EClinicalMedicine*, 40, 101099. <https://doi.org/10.1016/j.eclinm.2021.101099>

Welsh, E. A., Clark, H. H., Epstein, S. Z., Reveille, J. D., & Duvic, M. (1994). Human Leukocyte Antigen-DQB1*03 Alleles Are Associated with Alopecia Areata. *Journal of Investigative Dermatology*, 103(6), 758–763. <https://doi.org/10.1111/1523-1747.ep12412584>

White, C., & Franco-Paredes, C. (2015). Leprosy in the 21st Century. *Clinical Microbiology Reviews*, 28(1), 80–94. <https://doi.org/10.1128/CMR.00079-13>

White, T., & Folkens, P. (2005). The Human Bone Manual. In *The Human Bone Manual*. <https://doi.org/10.1016/C2009-0-00102-0>

Wissemann, W. T., Hill-Burns, E. M., Zabetian, C. P., Factor, S. A., Patsopoulos, N., Hoglund, B., Holcomb, C., Donahue, R. J., Thomson, G., Erlich, H., & Payami, H. (2013). Association of Parkinson Disease with Structural and Regulatory Variants in the HLA Region. *The American Journal of Human Genetics*, 93(5), 984–993. <https://doi.org/10.1016/j.ajhg.2013.10.009>

Yamazaki, A., Yasunami, M., Ofori, M., Horie, H., Kikuchi, M., Helegbe, G., Takaki, A., Ishii, K., Omar, A. H., Akanmori, B. D., & Hirayama, K. (2011). Human leukocyte antigen class I polymorphisms influence the mild clinical manifestation of Plasmodium falciparum infection in Ghanaian children. *Human Immunology*, 72(10), 881–888. <https://doi.org/10.1016/j.humimm.2011.06.007>

Yang, D. Y., & Watt, K. (2005). Contamination controls when preparing archaeological remains for ancient DNA analysis. *Journal of Archaeological Science*, 32(30), 331-336. <https://doi.org/10.1016/j.jas.2004.09.008>

Yoon, S. K., Han, J. Y., Pyo, C.-W., Yang, J. M., Jang, J. W., Kim, C. W., Chang, U. I., Bae, S. H., Choi, J. Y., Chung, K. W., Sun, H. S., Choi, H. B., & Kim, T.-G. (2005). Association between human leukocytes antigen alleles and chronic hepatitis C virus infection in the Korean population. *Liver International*, 25(6), 1122–1127. <https://doi.org/10.1111/j.1478-3231.2005.01105.x>

Zawadzka-Starczewska, K., Tymoniuk, B., Stasiak, B., Lewiński, A., & Stasiak, M. (2022). Actual Associations between HLA Haplotype and Graves' Disease Development. *Journal of Clinical Medicine*, 11(9), 2492. <https://doi.org/10.3390/jcm11092492>

Zerva, L., Cizman, B., Mehra, N., Alahari, S., Murali, R., Zmijewski, C., Kamoun, M., & Monos, D. (1996). Arginine at positions 13 or 70-71 in pocket 4 of HLA-DRB1 alleles is associated with leprosy susceptibility to tuberculoid leprosy. *Journal of Experimental Medicine*, 183(3), 829–836.

Zhang, F., Liu, H., Chen, S., Wang, C., Zhu, C., Zhang, L., Chu, T., Liu, D., Yan, X., & Liu, J. (2009). Evidence for an association of HLA-DRB1*15 and DRB1*09 with leprosy and the impact of DRB1*09 on disease onset in a Chinese Han population. *BMC Medical Genetics*, 10(1), 133. <https://doi.org/10.1186/1471-2350-10-133>

PART II: PATHOGEN EVOLUTION

Chapter 3: Pathogen genomics study of an early medieval community in Germany reveals extensive co-infections

Genome Biology 2022

doi.org/10.1186/s13059-022-02806-8


Supplementary material for this chapter can be found here: doi.org/10.1186/s13059-022-02806-8

RESEARCH

Open Access



Pathogen genomics study of an early medieval community in Germany reveals extensive co-infections

Joanna H. Bonczarowska^{1†}, Julian Susat^{1†}, Barbara Mühlemann^{2,3†}, Isabelle Jasch-Boley⁴, Sebastian Brather⁵, Benjamin Höke⁶, Susanne Brather-Walter⁵, Valerie Schoenenberg⁷, Jonathan Scheschekwitz⁶, Gabriele Graenert⁶, Dirk Krausse⁶, Michael Francken⁸, Terry C. Jones^{2,9}, Joachim Wahl⁴, Almut Nebel¹ and Ben Krause-Kyora^{1*} 

[†]Joanna H. Bonczarowska, Julian Susat, and Barbara Mühlemann contributed equally to this work and share the first authorship.

*Correspondence:
b.krause-kyora@ikmb.uni-kiel.de

¹Institute of Clinical Molecular Biology, Kiel University, 24105 Kiel, Germany
Full list of author information is available at the end of the article

Abstract

Background: The pathogen landscape in the Early European Middle Ages remains largely unexplored. Here, we perform a systematic pathogen screening of the rural community Lauchheim “Mittelhofen,” in present-day Germany, dated to the Merovingian period, between fifth and eighth century CE. Skeletal remains of individuals were subjected to an ancient DNA metagenomic analysis. Genomes of the detected pathogens were reconstructed and analyzed phylogenetically.

Results: Over 30% of the individuals exhibit molecular signs of infection with hepatitis B virus (HBV), parvovirus B19, variola virus (VARV), and *Mycobacterium leprae*. Seven double and one triple infection were detected. We reconstructed four HBV genomes and one genome each of B19, VARV, and *M. leprae*. All HBV genomes are of genotype D4 which is rare in Europe today. The VARV strain exhibits a unique pattern of gene loss indicating that viruses with different gene compositions were circulating in the Early Middle Ages. The *M. leprae* strain clustered in branch 3 together with the oldest to-date genome from the UK.

Conclusions: The high burden of infectious disease, together with osteological markers of physiological stress, reflect a poor health status of the community. This could have been an indirect result of the climate decline in Europe at the time, caused by the Late Antique Little Ice Age (LALIA). Our findings suggest that LALIA may have created an ecological context in which persistent outbreaks set the stage for major epidemics of severe diseases such as leprosy and smallpox hundreds of years later.

Keywords: Ancient DNA, Pathogen genomics, Ancient genomics, Variola virus, Smallpox, Leprosy, Hepatitis B virus, Parvovirus B19, Pathogen evolution, Infectious diseases



© The Author(s) 2022. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

Background

The pathogen landscape of the Early Middle Ages in Europe (late fifth–tenth century CE) remains largely unexplored. Bacterial and viral genomes from this period have so far only been detected in a few cases across the continent, such as a hepatitis B virus (HBV) [1, 2], variola virus (VARV) [3], and *Mycobacterium leprae* (*M. leprae*) [4]. These findings, although important in the context of pathogen evolution, are widely dispersed in time and space and do not significantly contribute to the understanding of the prevalence of infectious diseases in a specific locale during that period.

Prevalence rates of infectious diseases in pre-modern populations are uncertain, as they are usually based only on historical records [5] and paleopathological analysis of human skeletal and mummified remains [6]. Ancient pathogen genomics is an additional approach that can confirm the infection, even when there are no lesions present on skeletal remains or when disease manifestation is atypical. Combined analysis of the written record, skeletal remains, and genomic information can provide valuable knowledge of historical infectious disease presence.

Here, we perform a systematic metagenomic pathogen screening of 70 individuals representing 89% of all people buried within the Lauchheim “Mittelhofen” village (present-day Germany) in the Alemannic duchy of the Merovingian kingdom. The Merovingian period (late fifth century to 751 CE) was an important transitional phase, paving the way for what became known as the culture of Medieval Europe [7]. Here, we present molecular evidence of infection with four different pathogens (human parvovirus B19 (B19), HBV, variola virus, and *M. leprae*) in 22 individuals. This study provides insights into the diversity and prevalence of infectious diseases in a rural community during the late seventh and eighth century CE.

Results

Seventy-nine burials, in six burial groups (Fig. 1), were excavated at the Lauchheim “Mittelhofen” settlement. Based on archeological and radiocarbon evidence, the burials date to the late seventh and eighth century CE. It is thought that inhabitants buried the deceased at the edge of their property [8]. In total, metagenomic samples from 70 individuals were sequenced in this study (Additional file 1: Tab. S1).

Pathogen screening

Bacterial and viral screening of metagenomic samples revealed that DNA of four pathogens was present in 22 individuals: B19 ($n=21$), HBV ($n=8$), VARV ($n=1$), and *M. leprae* ($n=1$) (Table 1 and Additional file 1: Tab. S2 and S3). In six cases, a double infection with B19 and HBV was observed. A double infection with B19 and variola virus was found in one individual (G31) and three different pathogens (B19, HBV, and *M. leprae*) were detected in a sample of an adolescent male (G83). Data enrichment allowed the reconstruction and analysis of four HBV (G78, G79, G27, and G83), one B19 (G83), the VARV (G31) and the *M. leprae* genomes (G83) (Table 1). Based on these data, we report on the crude period prevalence (CPP) of infectious diseases in the Lauchheim settlement [6]. CPP is defined as the percentage of individuals in a population (%) that exhibit a molecular sign of infection over the period

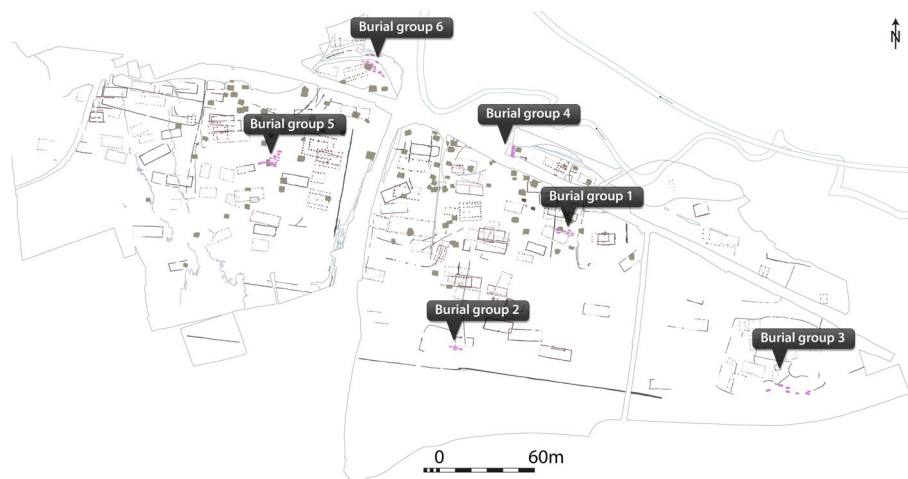


Fig. 1 Map of the archeological features of the Lauchheim “Mittelhofen” settlement. Burial groups are indicated. Burial groups are thought to be associated with distinguishable farmsteads. Purple rectangles represent the graves, while gray squares indicate pit houses and dotted-line rectangles post pits of houses

of time in which the settlement was used as a burial ground (approx. a century). Thus, the CPP indicates the prevalence of infections over three to four human generations, during which the pathogen landscape can change. The CPP calculated here for the Lauchheim population was 31.4%. When adult and subadult individuals are considered separately, the CPP was 29.1% for adults and 40% for subadults. Authenticity of the ancient origin of the pathogens was confirmed by the presence of typical ancient DNA (aDNA) damage patterns (Additional file 1: Tab. S4) and the ancestral positions of the pathogen genomes in phylogenetic trees (see below).

Genome-wide analysis of the identified pathogens

Hepatitis B virus

From the eight HBV-positive samples (CPP = 11.4%), we generated four almost complete viral genomes (G78, G79, G27, and G83) ranging in coverage from 96 to 99.9% (Additional file 1: Tab. S5). In the phylogeny, all four strains were basal to genotype D4 (Fig. 2 and Additional file 1: Fig. S1), clustering in two pairs of identical strains (G78 and G79, G27 and G83). Each genome carried the 33nt deletion in *preS1* typical of genotype D and a complete sequence of the HBe-antigen.

Human parvovirus B19

From the 21 parvovirus-positive samples (CPP = 30%), one complete B19 genome (G83) was reconstructed, with highest nucleotide identity to the modern genotype 2 (Additional file 1: Tab. S5). In the phylogenetic tree, G83 was basal to the cluster of several ancient European and modern genotype-2 strains (Fig. 3 and Additional file 1: Fig. S2).

Table 1 Individuals from the Lauchheim “Mittelhofen” settlement with detected infection

No.	Grave number/ individual ID	Sex	Age [years]	Pathogen	Burial group	¹⁴ C (2 σ) [years cal CE]	Physiological stress indicators [9]
1	5	M	40–50	B19	1	-	<i>Cribra orbitalia</i>
2	7	M	Around 25	B19	1	-	<i>Cribra orbitalia</i> , <i>cribra femoris</i> , enamel hypoplasia
3	22†	M	Around 15	B19	3	-	<i>Cribra orbitalia</i>
4	23	M	21–23	B19	3	-	<i>Cribra cranii</i> , <i>cribra femoris</i>
5	24	F	Senile	B19	3	-	<i>Cribra orbitalia</i> , <i>cribra cranii</i> , <i>cribra femoris</i> ,
6	37	M	Around 25	B19	4	674–820	<i>Cribra orbitalia</i> , <i>cribra cranii</i> and <i>cribra femoris</i> , enamel hypoplasia
7	56	M	Early mature	B19	5	-	<i>Cribra orbitalia</i> , <i>cribra cranii</i>
8	63	M	25–30	B19	6	-	<i>Cribra cranii</i> , enamel hypoplasia
9	67†	F	16–20	B19	6	-	Enamel hypoplasia
10	69	F	40–50	B19	6	-	-
11	70	F	30–40	B19	6	-	Enamel hypoplasia
12	82	F	Mature	B19	6	-	<i>Cribra orbitalia</i> , enamel hypoplasia
13	86	M	Around 60	B19	6	644–770	Enamel hypoplasia
14	78	M	40–50	HBV ⁹	6	647–771	<i>Cribra femoris</i>
15	31	F	25–30	B19, VARV ⁹	4	685–879	<i>Cribra cranii</i>
16	16	F	20–25	B19, HBV	2	-	<i>Cribra orbitalia</i> , <i>cribra femoris</i>
17	27	M	30–35	B19, HBV ⁹	3	703/704 [#]	<i>Cribra orbitalia</i> , <i>cribra cranii</i> , enamel hypoplasia
18	29	F	20–25	B19, HBV	4	665–773	<i>Cribra cranii</i> , <i>cribra femoris</i> , enamel hypoplasia
19	46†	PF ⁵	12–14	B19, HBV	5	-	<i>Cribra orbitalia</i> , <i>cribra cranii</i> , <i>cribra femoris</i>
20	66†	M	6–8	B19, HBV	6	-	Enamel hypoplasia
21	79†	M	13–14	B19, HBV ⁹	6	650–771	<i>Cribra orbitalia</i> , <i>cribra femoris</i> , enamel hypoplasia
22	83†	M	13–15	B19 ⁹ , HBV ⁹ , <i>M. leprae</i> ⁹	6	656–772	<i>Cribra orbitalia</i> , <i>cribra cranii</i>

⁹ Genome reconstructed, *M* male, *F* female, *PF* probable female, ⁵based on osteological analysis [9], [†]subadult, [#]based on dendrochronology

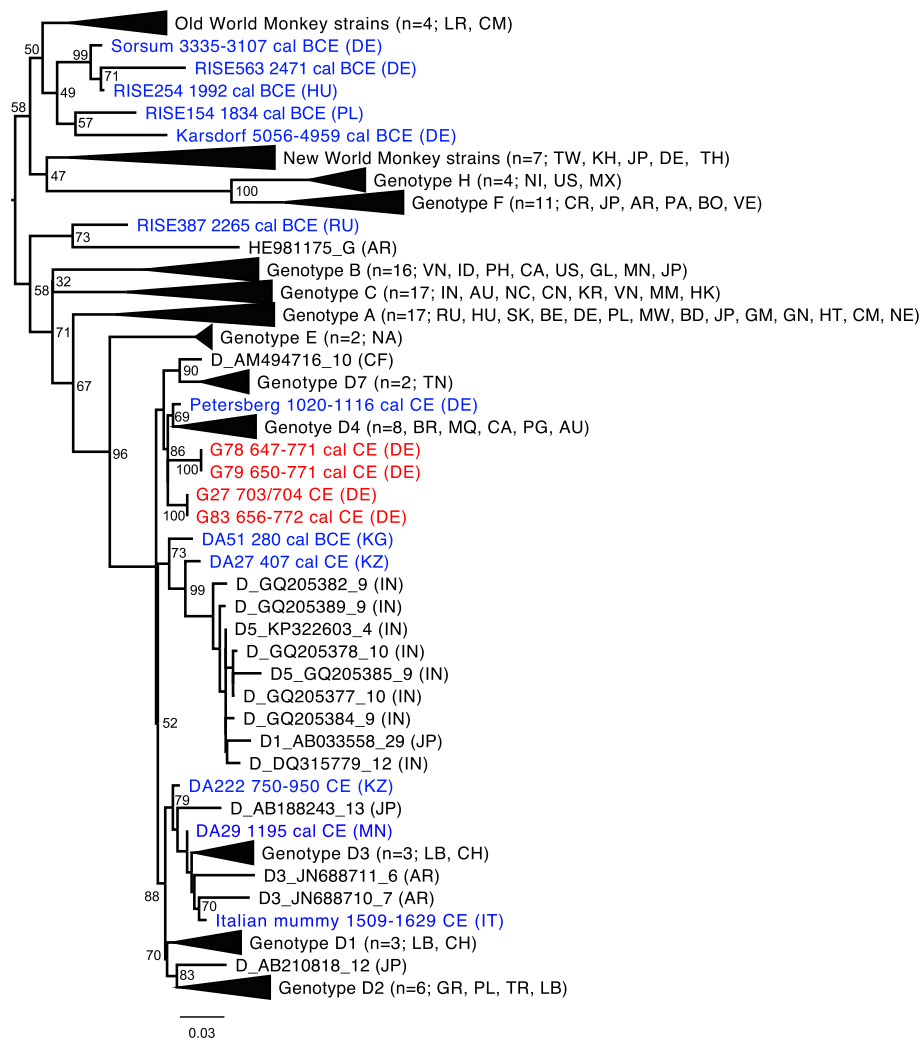


Fig. 2 A maximum-likelihood tree illustrating the phylogenetic position of the HBV strains from Lauchheim (red). Other ancient strains are marked in blue and modern HBV are shown in black. Numbers at each node are bootstrap support over 500 replications. The tree includes 130 strains (109 modern and 21 ancient). Country codes can be found in the Additional file 1: Table S6. Dates are provided for ancient strains

Variola virus

Sample G31 (685–879 cal CE) had 81 reads mapping to VARV in the initial screening (NC_001611.1). Subsequent capture for viral sequences resulted in the recovery of a near complete genome (98%, 60.8 x coverage) (Additional file 1: Tab. S7). Phylogenetically, the G31 genome clusters within the ancient VARV (aVARV) clade, basal to the strains VK281 and VK480 found in Denmark and Russia, respectively (Fig. 4A) [3]. The dates of the most recent common ancestor of all aVARVs (~1500 years ago), as well as all VARVs (~1700 years ago), are consistent with previous publications [3]. G31 exhibits a pattern of gene inactivation with features from both the VK388-VK382 and VK281-VK470 clades (Fig. 4B and Additional file 1: Fig. S3), congruent with its position in the phylogeny (Fig. 4A). Concerning the genes showing differences in activation patterns across the two branches, G31 shares gene-inactivating mutations with the VK382-VK388 branch in

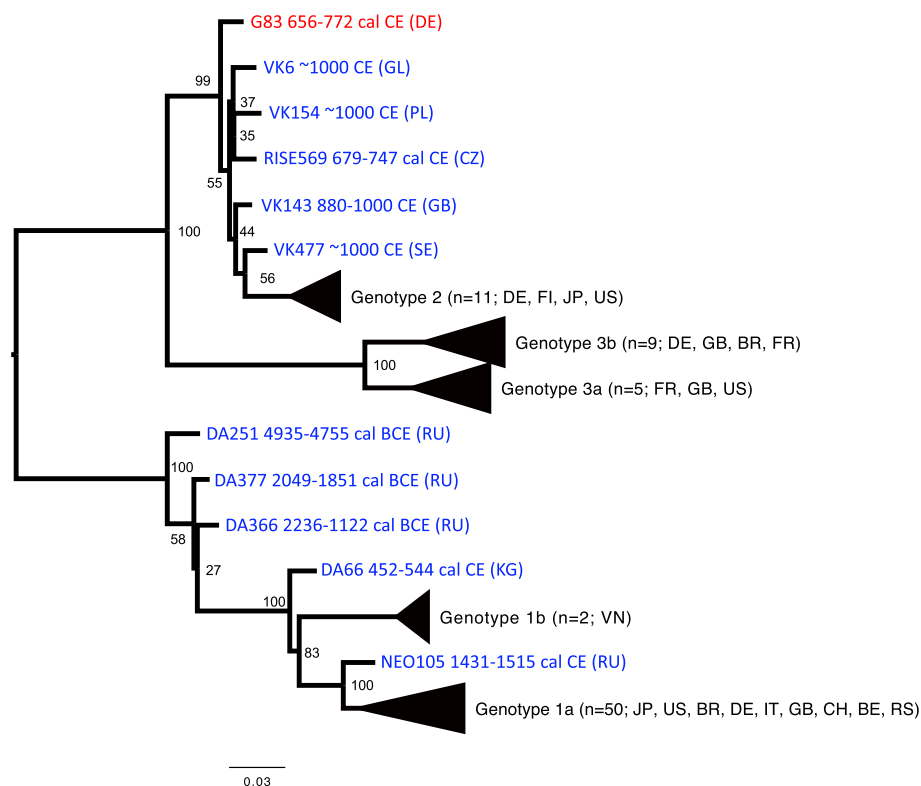


Fig. 3 A maximum-likelihood tree illustrating the phylogenetic position of the B19 strain from Lauchheim (red). Other ancient strains are marked in blue and modern B19 are shown in black. Numbers at each node are bootstrap support over 500 replications. The tree includes 88 strains (77 modern and 11 ancient). Country codes can be found in the Additional file 1: Table S6. Dates are provided for ancient strains

three genes (A39R, C2L, F3L) and with the VK281-VK470 branch in four genes (22682/CPXV020, K1L, 168145/CPXV181, A57R).

Mycobacterium leprae

A genome of *M. leprae* (4 x) was assembled (Additional file 1: Tab. S8) from the skeletal remains of an adolescent male (G83) (Table 1) which showed several pathological lesions (Additional file 1: Tab. S9, Fig. S4 and Fig. S5). In the phylogenetic analysis, G83 was positioned in branch 3, together with modern strains from the Americas (human and nine-banded armadillo) and the UK (red squirrel) as well as human medieval strains from Denmark and the UK (Fig. 5 and Additional file 1: Fig. S6). G83 clustered with the British GC96 sequence dating to the fifth–sixth century CE, forming a separate subbranch.

Discussion

Pathogen screening of the community buried at the early medieval settlement Lauchheim “Mittelhofen” revealed that more than 30% of the population was positive for at least one pathogen at the time of death. We identified seven individuals with

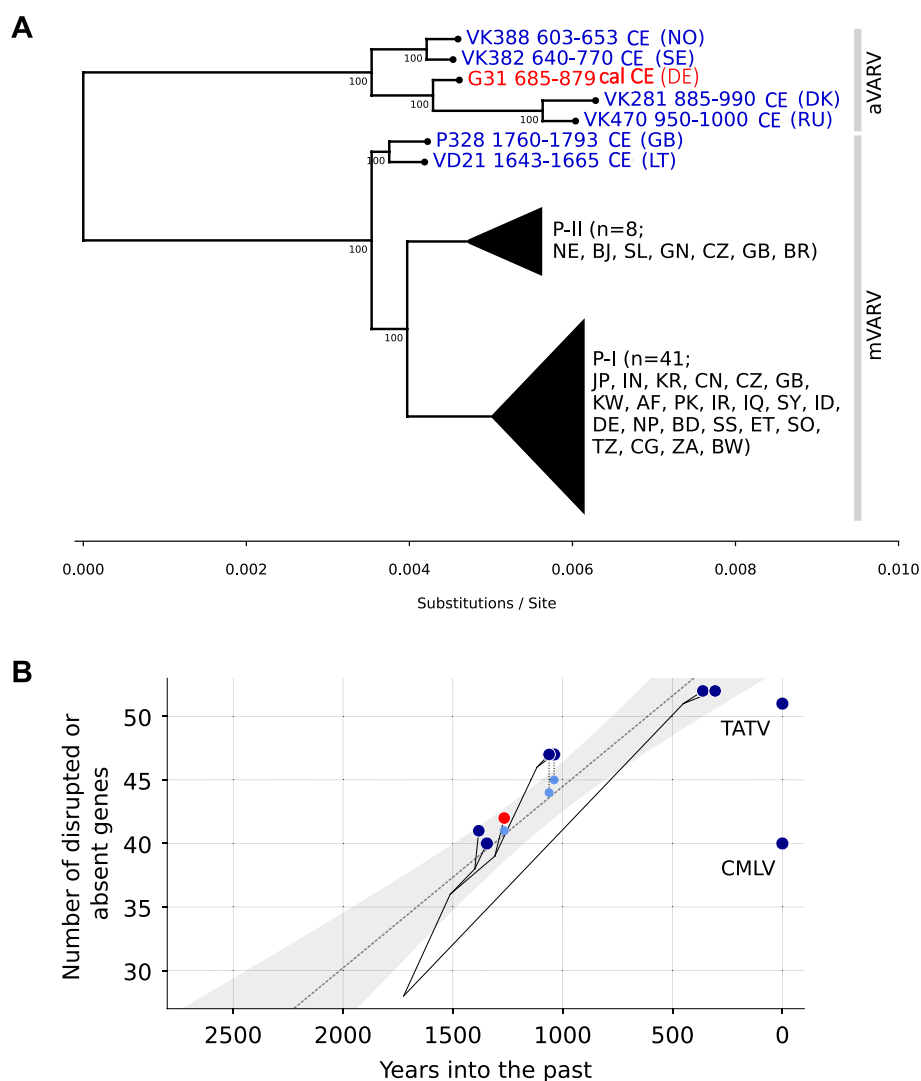


Fig. 4 G31 variola virus from Lauchheim exhibits intermediate genetic characteristics in relation to other ancient strains. **A** A maximum-likelihood tree illustrating the phylogenetic position of the VARV strain from Lauchheim (red) among other ancient (blue) and modern (black) strains. Numbers at each node are bootstrap support. The tree was inferred using IQTree, with 10,000 ultrafast bootstrap replicates. The tree includes 56 sequences (49 modern and 7 ancient). P-I and P-II refer to VARV primary clade I (corresponding to VARV major and East, Central and South African VARV sequences) and primary clade II (comprising of VARV alastrim minor and West African sequences), respectively. Country codes can be found in the Additional file 1: Table S6. Dates are provided for ancient strains. **B** Temporal trend of gene inactivation, with the number of inactivated genes shown on the y-axis and the x-axis indicating time into the past. Phylogenetic relationships between viruses are indicated by gray lines. Light blue circles indicate the number of inactivated genes when uncertain inactivations are not taken into account. The gray dotted line and shaded area indicate a linear regression of number of genes against time into the past, indicating that a hypothetical virus with a full set of genes existed ~4116 years into the past ($R: 0.92$, $R^2: 0.85$, slope: 0.014, y-intercept: 58.76, stderr: 0.002, $P: 0.00002$)

concurrent infections with two viruses (G16, G27, G29, G31, G46, G66, and G79) and one individual with a triple infection of HBV, B19, and *M. leprae* (G83) (Table 1).

Seven percent of adults and 26.6% of subadults had an HBV infection (Table 1), a prevalence which today is considered endemic [10]. Furthermore, the proportion of

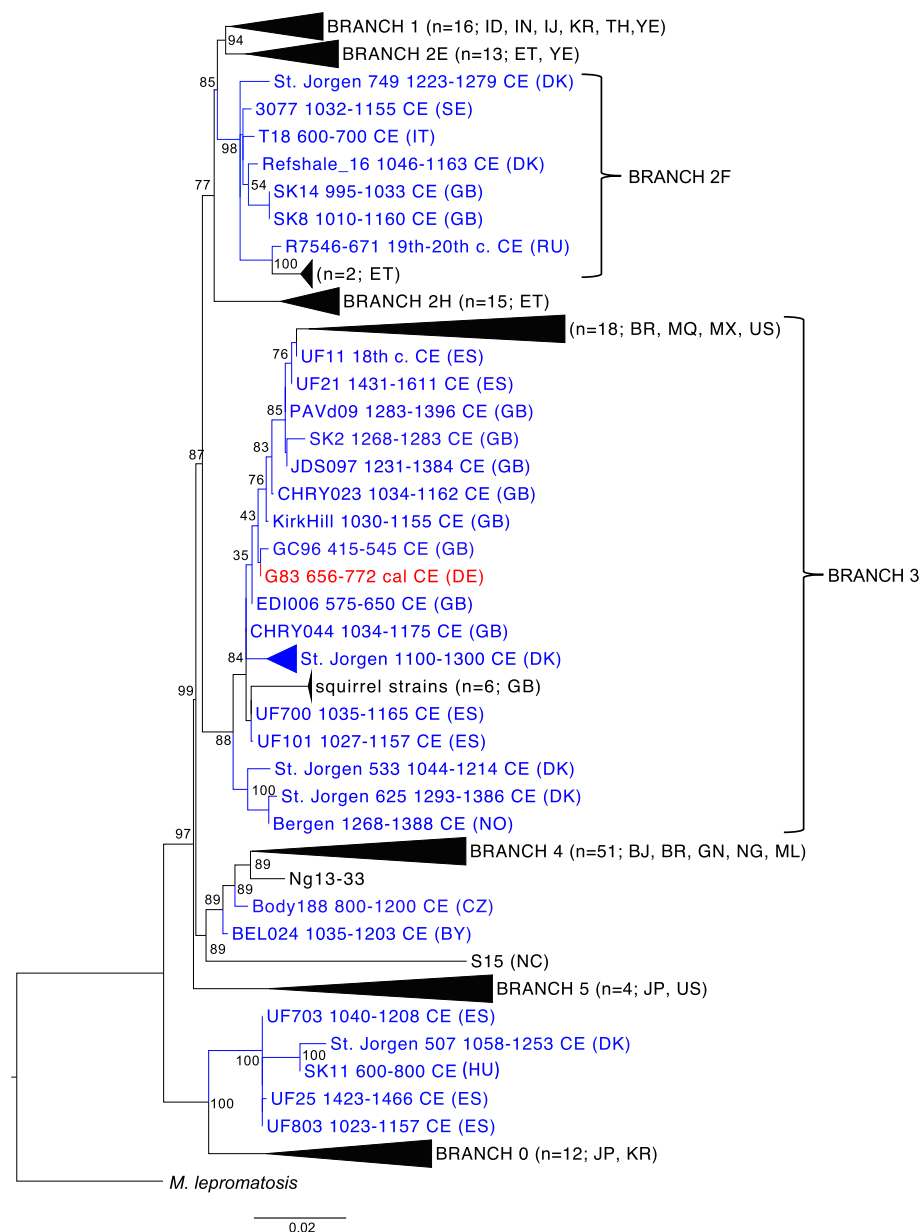


Fig. 5 A maximum-likelihood tree illustrating the phylogenetic position of the *M. leprae* strain from Lauchheim (red). Other ancient strains are marked in blue and modern *M. leprae* are shown in black. Numbers at each node are bootstrap support over 500 replications. The tree includes 177 strains (139 modern and 38 ancient). Country codes can be found in the Additional file 1: Table S6. Dates are provided for ancient strains

HBV-infected subadults in Lauchheim was over three times as high as that of adults, which is in accordance with the age-related infection pattern in endemic areas [10]. Although today most HBV infections are asymptomatic or cause mild symptoms, individuals with a compromised immune system are at risk of developing severe complications such as liver failure, cirrhosis, and liver cancer. All HBV genomes were of sub-genotype D4—as previously found in eleventh–twelfth century Petersberg (Germany) [11]. To date, sub-genotype D4 is mainly found in Australasia and Northern America. Thus, the available data point to a temporal change in the geographical

distribution of HBV genotype D sub-genotypes in Europe—with early medieval D4 subtypes being replaced with those that are prevalent today, such as D1, D2, and D3 [12]. Interestingly, in sixteenth century Italy, an HBV strain of genotype D3 was already present [13] (Fig. 2). Two pairs (G78 and G79; G27 and G83) of identical genomes were reconstructed (based on sequence comparison). Although the diversity of HBV at the time is unknown, this finding could possibly suggest a direct transmission between these individuals or a common source of infection.

In comparison to HBV, B19 was more abundant, with 30% of the Lauchheim community identified as positive. The CPP estimated for subadults (40%) falls within the prevalence seen today (15–60%). The adult CPP of 27.3% is slightly below the rate expected in the present day (30–60%), suggesting that it might likely be an underestimate. B19 infections can lead to a mild disease in children (fifth disease) and serious fetal conditions in pregnant women. In immunocompromised individuals, the infection may lead to anemia and life-threatening aplastic crisis [14]. Thirteen out of the 21 B19-positive individuals in Lauchheim died before reaching an estimated 30 years of age (62%, in comparison to 41% in the overall population). Although not particularly likely, premature death could have possibly occurred due to complications developed after a B19 infection. In the phylogenetic analysis, the G83 strain (genotype 2) is basal to all modern and ancient genotype 2 genomes recovered to date. Genotype 2 is today mainly found in elderly northern Europeans [15].

Variola virus (VARV) was detected in one individual (G31). VARV is the causative agent of the now-eradicated smallpox. The disease is characterized by high mortality (reaching up to 30%), with symptoms including fever, vomiting, and skin rash [16]. In the phylogenetic analysis, the G31 strain clustered within the known diversity of early medieval VARVs from northern and eastern Europe [3]. These VARVs are distinguished from modern VARVs by a larger number of intact genes. G31 exhibits a pattern of gene loss that is a combination of unique features found in both the VK382-VK388 and VK281-VK470 branch of the medieval VARV clade, showing that viruses with different gene compositions were circulating in the Early Middle Ages. The reductive evolution in modern VARVs is thought to be a result of changed selection pressures due to adaptation to humans as exclusive hosts [17]. Given the transmissibility of modern VARV, it is interesting that we only detected a VARV infection in one Lauchheim inhabitant. The absence of additional VARV-positive samples in Lauchheim may be due to a lower fatality rate [18], or the fact that viremia is not thought to last for the duration of infection [19], and thus the viral DNA was not preserved. Furthermore, genomic differences between the modern and medieval VARVs (Fig. 4 and Additional File 1: Fig. S3) could possibly reflect a reduced efficiency of human-human transmission in ancient VARVs. The epidemiology of VARV in the Early Middle Ages, however, remains unknown.

In addition to viral infections, the pathogen screening revealed the presence of *M. leprae* (causative agent of leprosy) in an adolescent male (G83). Leprosy mainly affects the skin, mucous membranes, and peripheral nerves. Subsequent sensory loss in the hands and feet makes these areas susceptible to injuries and secondary infections, usually leading to deformation and in some cases even loss of extremities. Leprosy per se is not deadly and only approximately 5% of infected individuals exhibit clinical symptoms. As leprosy can cause severe disability as well as visible disfigurement of the limbs and face

[20], social stigma of the disease likely surpassed its clinical outcome in terms of suffered consequences. The skeleton of individual G83 exhibits signs suggestive of lepromatous leprosy that take years to develop. Although the adolescent likely had visible facial disfigurement caused by leprosy, he was buried together with the rest of the community, indicating that he might not have been ostracized. Although leprosy was already present in Europe in the Early Middle Ages [21, 22], its prevalence peaked between the twelfth and fourteenth centuries CE. The G83 strain is the oldest found in Germany to date. Two British strains dating to a similar time period were included in the phylogenetic analysis (GC96, fifth–sixth century and EDI006, sixth–seventh century CE). The Lauchheim strain groups in branch 3 together with both British strains, suggesting low diversity of *M. leprae* in the Early Medieval Period [23]. Based on the observed variation (Fig. 5) as well as the temporal and geographic distribution of the medieval strains (Additional file 1: Fig. S7), we can hypothesize an early appearance of branch 3 strains in Europe. The spread may have been facilitated by the expanding Roman Empire [22, 23].

Although the dating of the Lauchheim settlement abuts the Justinian plague (541–543 CE), no traces of the causative *Yersinia pestis* (*Y. pestis*) were found in any of the analyzed samples. Moreover, in the case of the Lauchheim settlement, most likely not all inhabitants were buried within its borders and part of the population might have been buried in the traditional graveyard situated nearby or at another unknown location. Overall, the burials may not be representative of the general Lauchheim population. Excavation and preservation bias also influence the amount of data available for analysis. Inherent limitations of the collected molecular evidence must also be considered. Lack of infection cannot be inferred from pathogen-negative DNA samples as degradation of the genetic material could have erased traces of the microorganism. Thus, the observed infectious disease prevalence in Lauchheim is probably an underestimate.

In the early medieval community of Lauchheim, we observed a high prevalence of infectious diseases caused by four different pathogens. Pathogens were detected in all six burial groups, indicating a broad distribution of infections across the population. The extent of the disease burden becomes even more apparent when considering the fact that pathogens which do not enter the bloodstream are most likely undetectable, due to lack of DNA preservation. Moreover, other types of diseases, such as hereditary, metabolic, and nutritional diseases or cancer, were not accounted for in this investigation. Interestingly, 95.5% of the pathogen-positive individuals (Table 1) and 84.4% of the individuals with no molecular sign of infection exhibited skeletal lesions suggestive of physiological stress. This extremely high burden of lesions indicates poor health in Lauchheim. In one scenario, malnutrition could have increased the level of physiological stress, which in turn resulted in a weakened immune system and a higher probability of successful pathogen transmission and zoonoses. In another scenario, a high burden of infectious diseases in the community compromised the immune system leading to increased physiological stress and metabolic disturbances. The cause-and-effect relationship is unclear and it is likely that multiple factors simultaneously contributed to the elevated stress in the Lauchheim community. Interestingly, Europe experienced a major climate decline between the fifth and seventh century CE, i.e., the Late Antique Little Ice Age (LALIA), which was likely an environmental driver of crop failure, famine, and disease [24]. Although the degree to which LALIA affected the Lauchheim community

is unknown, it presents one possible explanation for the high disease burden. The far-reaching contacts and networks during that time, also documented for Lauchheim [25], might have contributed to the spread of infections.

It is notable that LALIA coincides with several changes in the pathogen landscape. A case in point is the Justinian plague (541–543 CE) that affected millions of people across Europe. Another example is that the dating of the G31 VARV genome together with VK388 and VK382 (Fig. 4A) falls within LALIA. Moreover, the early medieval *M. leprae* infections were the starting point of the leprosy epidemic during the eleventh–fourteenth centuries CE. This scenario is supported by the ancestral positions of the early Lauchheim (G83) and the British (GC96) genomes in clade 3 of the phylogeny (Fig. 5), which were to dominate the later epidemic (Additional file 1: Fig. S7) [23]. Taken together, we would like to hypothesize that LALIA may have created an ecological context in which persistent outbreaks set the stage for severe epidemics hundreds of years later. This speculation, however, needs corroboration.

The pathogens detected in Lauchheim are responsible for some of the most feared diseases of the last millennium and, with the exception of variola virus, still represent major health burdens today.

Conclusions

Systematic pathogen screening of skeletal remains belonging to 70 individuals from the Lauchheim “Mittelhofen” provided molecular evidence for the high burden of infectious diseases in this early medieval community (> 30% of the individuals were positive for at least one pathogen at the time of death). Such a high frequency of infection together with physiological stress markers on the bones indicate an overall poor health status of the community that lived in a time period characterized by rapid cultural transition and major climate changes (e.g., LALIA). As seen today, climate change is an eminent driver of emergence, re-emergence and transmission of infectious diseases.

Material

The material comprised bone and tooth samples of 70 individuals (Additional file 1: Tab. S1) excavated at the Lauchheim “Mittelhofen” settlement located in today’s Baden-Württemberg (southern Germany). The settlement was inhabited from the sixth to twelfth century CE (dating derived primarily from ceramic findings) [8, 26]. However, the settlement was used as a burial ground only during the late seventh and eighth century CE. Seventy-nine burials containing the remains of 77 individuals were found within the settlement borders [26].

Methods

Processing of samples

Processing of samples and DNA extraction were carried out in a dedicated ancient DNA facility at the University of Kiel. DNA extraction, library preparation, sequencing, and post-processing were performed according to a previously published protocol for non-UGD treated samples [27]. Subsequently, all generated metagenomic datasets were

screened for the presence of pathogens with Megan Alignment Tool 0.3.0 (MALT) [28] (SemiGlobal alignment mode, identity threshold = 85%), using a custom database containing all bacterial and viral genomes from the NCBI website (download: 24.01.2019). Output alignments were inspected visually in MEGAN 6 [29]. Samples showing presence of pathogen DNA were further sequenced to obtain more data (Additional file 1: Tab. S3). Additionally, virus (111 human pathogens available on the NCBI website in 2018) targeted capture was performed on the VARV-positive sample (G31). Enriched data were then mapped against reference genomes of the respective pathogens to confirm the findings.

***M. leprae* phylogeny**

The pre-processed reads were mapped against the *M. leprae* TN reference genome (NC_002677) with the Burrows-Wheeler Aligner (BWA) v0.7.12 [30]. Output SAM files were converted to BAM format and merged, sorted and indexed with SAMtools v1.3 with default parameters [31]. Duplicated reads were removed using DeDup v0.12.2 [32] and the BAM file was filtered for quality 30. To generate VCF files, UnifiedGenotyper module from the Genome Analysis Toolkit (GATK) v3.6 was used. All the reads in the input were assigned to a single new read group with the Picard AddOrReplaceReadGroups tool with default parameters [33].

Based on a previously published dataset [21], 3124 SNPs with a minimum coverage of 3 x, a quality score of 30 and a 90% majority call were extracted. The phylogenetic *M. leprae* tree was produced with RaxML [34] with a bootstrap of 500 replicates (GTR-GAMMA model). The tree consisted of 176 previously published *M. leprae* strains (37 ancient and 139 modern) and the G83 individual. *M. lepromatosis* was used as an out-group (Additional file 1: Tab. S10).

HBV analysis

Using BWA [30], the HBV positive samples were mapped against a reference dataset representing the modern and known ancient diversity of HBV (Additional file 1: Tab. S11) as previously described [11]. BAM files belonging to one individual were merged using SAMtools [31], and duplicates were removed using the software DeDup with the default options [32]. BamUtil was used to clip two nucleotides from the 5' and 3' end. After a second round of duplicate removal, FASTQ was extracted from the BAM files using SAMtools. Subsequently a de novo assembly for HBV-positive samples was performed and resulting contigs were mapped (using BWA) against the respective sequence dataset containing all the references (Additional file 1: Tab. S11). For genomic reconstruction, the results of the remapping were visually inspected and, if possible, the consensus sequence was directly exported. Subsequent curation of the genomes was done manually. Using the Needleman-Wunsch algorithm, the ancient HBV genomes from Lauchheim were aligned to strains listed in the Additional file 1: Table S12.

A dataset containing 109 representative modern HBV genomes, 17 ancient HBV genomes and the four genomes from Lauchheim was used for the generation of a phylogenetic tree. A multiple sequence alignment was made using MAFFT [35] and phylogenetic informative blocks were extracted using GBLOCKS with default parameters [36]. A maximum likelihood phylogeny was generated using PhyML with the GTR

substitution model, estimated proportion of invariable sites, NNI tree topology search, and 500 bootstrap replicates [37].

B19 analyses

Parvovirus B19-positive samples were aligned against a dataset representing the modern and known ancient diversity of B19 diversity (Additional file 1: Tab. S13). Using samtools idxstats, AJ717293 was identified as the best-fitting reference, and subsequent mapping against AJ717293 was performed (Additional file 1: Tab. S14). Further processing of BAM files was done as described in the HBV section. For genomic reconstruction, the results of the mapping against AJ717293 were visually inspected and the consensus sequence was directly exported. Using the Needleman-Wunsch algorithm, the ancient B19 genome from Lauchheim was aligned to strains in the Additional file 1: Tab. S15. A MAFFT alignment followed by GBLOCKS was performed. A maximum likelihood phylogeny was generated using PhyML with the TN93 substitution model, estimated proportion of invariable sites, NNI tree topology search and 500 bootstrap replicates.

Generation of VARV consensus sequence

Reads from individual G31 were mapped against the sequence of VK382 (<https://doi.org/10.6084/m9.figshare.12185466.v1>) in Geneious Prime (<https://www.geneious.com>). The alignment was manually checked for insertions and deletions. The consensus sequence was called at 1 x coverage, if identical to the reference. SNPs were called if present in at least two reads.

VARV gene loss and phylogenetic analyses

Gene loss analysis was performed as described [3]. For phylogenetic analyses, a multiple sequence alignment was constructed using available modern VARV and previously published ancient sequences [3, 38]. Only the central region of the genome (VACV-Cop *F4L* to VACV-Cop *A24R*) was used. A maximum likelihood (ML) tree was constructed and a regression of sampling dates and root-to-tip distances (extracted from the ML tree) was performed. Dating analyses were conducted with the same alignment that was used for the ML tree. A TPM1 substitution model with unequal base frequencies and invariant sites was selected, and the clock rate was constrained using a uniform distribution with bounds (1×10^{-9} – 10×10^{-3} substitutions/site/year). After path sampling, likelihood values between different combinations of priors and clock models were compared with a Bayes factor test. Median node ages and substitution rates were inferred from trees run for 10M (for strict clock models) or 100M (for log-normal relaxed clock models) generations.

Genetic sex determination

The genetic sex of the analyzed individuals was determined based on the ratio of sequences aligning to the X and Y chromosomes compared to the autosomes [39].

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s13059-022-02806-8>.

Additional file 1: Supplementary figures S1-S5, and supplementary tables S1-S15.

Additional file 2: Review history.

Peer review information

Andrew Cosgrove was the primary editor of this article and managed its editorial process and peer review in collaboration with the rest of the editorial team.

Review history

The review history is available as Additional file 2

Authors' contributions

B.K.-K., J.W., and I.J.-B. developed the idea for this study. I.J.-B., J.W., and M.F. analyzed the human skeletal remains. S.B., B.H. S.B.-W., V.S., J.Sch., G.G., and D.K. performed the archeological and historical contextualization and dating. B.K.-K. and J.H.B. generated ancient DNA data. J.H.B., J.S., B.M., T.C.J., and B.K.-K. analyzed the ancient DNA data. J.H.B., J.S., B.M., I.J.-B., S.B., B.H. S.B.-W., V.S., J.Sch., G.G., D.K., M.F., T.C.J., J.W., A.N., and B.K.-K. interpreted the findings. J.H.B., J.S., B.M., and B.K.-K. wrote the manuscript with input from all other authors. The authors read and approved the final manuscript.

Funding

Open Access funding enabled and organized by Projekt DEAL. This study was funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) Project-ID 290391021—SFB1266 and under Germany's Excellence Strategy—EXC 2167—390884018 and EXC 2150—390870439. T.C.J. is in part funded through NIAID-NIH CEIRS contract HHSN272201400008C. J.H.B. was funded by the International Max Planck Research School for Evolutionary Biology.

Availability of data and materials

Analyzed sequences are available through the European Nucleotide Archive under Accession Number PRJEB49149: <http://www.ebi.ac.uk/ena/browser/view/<accession>> [40].

Declarations

Ethics approval and consent to participate

Due to the nature of the samples analyzed, ethical approval was not needed.

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests.

Author details

¹Institute of Clinical Molecular Biology, Kiel University, 24105 Kiel, Germany. ²Institute of Virology, Charité – Universitätsmedizin Berlin, corporate member of Freie Universität Berlin, Humboldt-Universität zu Berlin, and Berlin Institute of Health, 10117 Berlin, Germany. ³German Centre for Infection Research (DZIF), partner site Charité, 10117 Berlin, Germany. ⁴Institute for Archaeological Sciences, Palaeoanthropology Working Group, University of Tübingen, Rümelinstrasse 23, 72070 Tübingen, Germany. ⁵Institute of Archaeology, Freiburg University, Belfortstraße 22, 79085 Freiburg, Germany. ⁶Landesamt für Denkmalpflege im Regierungspräsidium Stuttgart, Berliner Straße 12, 73728 Esslingen, Germany. ⁷Museum im Ritterhaus, Ritterstraße 10, 77652 Offenburg, Germany. ⁸Landesamt für Denkmalpflege im Regierungspräsidium Stuttgart, Konstanz, Germany. ⁹Center for Pathogen Evolution, Department of Zoology, University of Cambridge, Cambridge CB2 3EJ, UK.

Received: 27 January 2022 Accepted: 27 October 2022

Published online: 13 December 2022

References

- Mühlemann B, Jones TC, de Barros DP, Allentoft ME, Shevnina I, Logvin A, et al. Ancient hepatitis B viruses from the Bronze Age to the Medieval period. *Nature*. 2018;557:418–23. <https://doi.org/10.1038/s41586-018-0097-z>.
- Kocher A, Papac L, Barquera R, Key FM, Spyrou MA, Hübner R, et al. Ten millennia of hepatitis B virus evolution. *Science*. 2021;374:182–8. <https://doi.org/10.1126/science.abi5658>.
- Mühlemann B, Vinner L, Margaryan A, Wilhelmson H, de la Fuente CC, Allentoft ME, et al. Diverse variola virus (smallpox) strains were widespread in northern Europe in the Viking Age. *Science*. 2020;369:eaaw8977. <https://doi.org/10.1126/science.aaw8977>.
- Inskip SA, Taylor GM, Zakrzewski SR, Mays SA, Pike AWG, Llewellyn G, et al. Osteological, biomolecular and geochemical examination of an early Anglo-Saxon case of lepromatous leprosy. *PLoS One*. 2015;10:e0124282. <https://doi.org/10.1371/journal.pone.0124282>.

5. McCullagh CB. Bias in historical description, interpretation, and explanation. *Hist Theory*. 2000;39:39–66.
6. Waldron T. Rates for the job. Measures of disease frequency in palaeopathology. *Int J Osteoarchaeol*. 1991;1:17–25. <https://doi.org/10.1002/oa.1390010104>.
7. Geary PJ. Before France and Germany: the creation and transformation of the Merovingian world. Oxford: Oxford University Press; 1988.
8. Schoenenberg V. Die frühmittelalterliche Siedlung in Lauchheim, Gewinn "Mittelhofen". Ostalbkreis [Doctoral dissertation]. University of Freiburg; 2014. <https://doi.org/10.6094/UNIFR/175620>.
9. Jasch-Boley I. Lepra, Krebs und Mangelerscheinung – Die Leiden der frühmittelalterlichen Bevölkerung von Lauchheim "Mittelhofen" anhand ausgewählter Fallbeispiele [Master thesis]. University of Tübingen; 2020.
10. MacLachlan JH, Cowie BC. Hepatitis B virus epidemiology. *Cold Spring Harb Perspec Med*. 2015;5:a021410. <https://doi.org/10.1101/cshperspect.a021410>.
11. Krause-Kyora B, Susat J, Key FM, Kühnert D, Bosse E, Immel A, et al. Neolithic and medieval virus genomes reveal complex evolution of hepatitis B. *eLife*. 2018;7:e36666. <https://doi.org/10.7554/eLife.36666.001>.
12. Yousif M, Kramvis A. Genotype D of hepatitis B virus and its subgenotypes: an update. *Hepatol Res*. 2013;4:355–64. <https://doi.org/10.1111/j.1872-034X.2012.01090.x>.
13. Patterson Ross Z, Klunk J, Giuffra V, Duchene S, Duggan AT, Poinar D, et al. The paradox of HBV evolution as revealed from a 16th century mummy. *PLoS Pathog*. 2018;14(2):e1006887. <https://doi.org/10.1371/journal.ppat.1006887>.
14. Heegaard ED, Brown KE. Human parvovirus B19. *Clin Microbiol Rev*. 2002;15:485–505. <https://doi.org/10.1128/CMR.15.3.485-505.2002>.
15. Norja P, Hokynar K, Aaltonen LM, Chen R, Ranki A, Partio EK, et al. Bioportfolio: lifelong persistence of variant and prototypic erythrovirus DNA genomes in human tissue. *Proc Natl Acad Sci U S A*. 2006;103:7450–3. <https://doi.org/10.1073/pnas.0602259103>.
16. Thèves C, Biagini P, Crubézy E. The rediscovery of smallpox. *Clin Microbiol Infect*. 2014;20:210–8. <https://doi.org/10.1111/1469-0691.12536>.
17. Hendrickson RC, Wang C, Hatcher EL, Lefkowitz EJ. Orthopoxvirus genome evolution: the role of gene loss. *Viruses*. 2010;2:1933–67. <https://doi.org/10.3390/v2091933>.
18. Carmichael AG, Silverstein AM. Smallpox in Europe before the seventeenth century: virulent killer or benign disease? *J Hist Med Allied Sci*. 1987;42:147–68. <https://doi.org/10.1093/jhmas/42.2.147>.
19. Elwood JM. Smallpox and its eradication. *J Epidemiol Community Health*. 1989;42:92.
20. Britton WJ, Lockwood DNJ. Leprosy. *Lancet*. 2004;363:1209–19. [https://doi.org/10.1016/S0140-6736\(04\)15952-7](https://doi.org/10.1016/S0140-6736(04)15952-7).
21. Schuenemann VJ, Avanzi C, Krause-Kyora B, Seitz A, Herbig A, Inskip S, et al. Ancient genomes reveal a high diversity of *Mycobacterium leprae* in medieval Europe. *PLoS Pathog*. 2018;14:e1006997. <https://doi.org/10.1371/journal.ppat.1006997>.
22. Boldsen JL. Leprosy in the early medieval Lauchheim community. *Am J Phys Anthropol*. 2008;135:301–10. <https://doi.org/10.1002/ajpa.20744>.
23. Pfrengle S, Neukamm J, Guellil M, Keller M, Molak M, Avanzi C, et al. *Mycobacterium leprae* diversity and population dynamics in medieval Europe from novel ancient genomes. *BMC Biol*. 2021;19:220. <https://doi.org/10.1186/s12915-021-01120-2>.
24. Büntgen U, Myglan VS, Ljungqvist FC, McCormick M, Di Cosmo N, Sigl M, et al. Cooling and societal change during the Late Antique Little Ice Age from 536 to around 660 AD. *Nat Geosci*. 2016;9:231–6. <https://doi.org/10.1038/ngeo2652>.
25. Stork I. Lauchheim. In: Beck H, et al., editors. *Hoops Reallexikon der Germanischen Altertumskunde A 18*. Berlin: De Gruyter; 2001.
26. Wahl J, Stork I. Außergewöhnliche Gräber beim Herrenhof. Merowingerzeitliche Siedlungsbestattungen aus Lauchheim "Mittelhofen". In: Biel J, et al., editors. *Landesarchäologie. Festschrift für Dieter Planck zum 65. Geburtstag. Forschungen und Berichte zur Vor- und Frühgeschichte in Baden-Württemberg 100*. Stuttgart: wbg Theiss in Wissenschaftliche Buchgesellschaft; 2009.
27. Krause-Kyora B, Nutsua M, Boehme L, Pierini F, Pedersen DD, Kornell SC, et al. Ancient DNA study reveals HLA susceptibility locus for leprosy in medieval Europeans. *Nat Commun*. 2018;9:1569. <https://doi.org/10.1038/s41467-018-03857-x>.
28. Herbig A, Maixner F, Bos K, Zink A, Krause J, Huson D. MALT: Fast alignment and analysis of metagenomic DNA sequence data applied to the Tyrolean Iceman. *BioRxiv*. 2016:050559. <https://doi.org/10.1101/050559>.
29. Huson DH, Auch AF, Qi J, Schuster SC. MEGAN analysis of metagenomic data. *Genome Res*. 2007;17(3):377–86. <https://doi.org/10.1101/gr.5969107>.
30. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics*. 2009;25(14):1754–60. <https://doi.org/10.1093/bioinformatics/btp324>.
31. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. The Sequence Alignment/Map format and SAMtools. *Bioinformatics*. 2009;25:2078–9. <https://doi.org/10.1093/bioinformatics/btp352>.
32. Peltzer A, Jäger G, Herbig A, Seitz A, Knip C, Krause J, et al. EAGER: efficient ancient genome reconstruction. *Genome Biol*. 2016;17:60. <https://doi.org/10.1186/s13059-016-0918-z>.
33. McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernysky A, et al. The genome analysis toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res*. 2010;20(9):1297–303. <https://doi.org/10.1101/gr.107524.110>.
34. Stamatakis A. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics*. 2014;30(9):1312–3. <https://doi.org/10.1093/bioinformatics/btu033>.
35. Nakamura Y, Katoh T. Parallelization of MAFFT for large-scale multiple sequence alignments. *Bioinformatics*. 2018;34:2490–2.
36. Castresana J. Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. *Mol Biol Evol*. 2000;17(4):540–52. <https://doi.org/10.1093/oxfordjournals.molbev.a026334>.
37. Guindon S, Gascuel O. A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Syst Biol*. 2003;52(5):696–704. <https://doi.org/10.1080/10635150390235520>.

38. Duggan AT, Perdomo MF, Piombino-Mascali D, Marciniak S, Poinar D, Emery MV, et al. 17th century variola virus reveals the recent history of smallpox. *Curr Biol*. 2016;26:3407–12. <https://doi.org/10.1016/j.cub.2016.10.061>.
39. Fu Q, Posth C, Hajdinjak M, Petr M, Mallick S, Fernandes D, et al. The genetic history of Ice Age Europe. *Nature*. 2016;534:200–5. <https://doi.org/10.1038/nature17993>.
40. Bonczarowska JH, Susat J, Mühlemann B, et al. Pathogen genomics study of an early medieval community in Germany reveals extensive co-infections - sequencing data. PRJEB49149. European Nucleotide Archive. 2022. <https://www.ebi.ac.uk/ena/browser/view/PRJEB49149>.

Publisher's Note

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

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

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

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions



**Chapter 4: *Yersinia pestis* strains from Latvia
show depletion of the *pla* virulence gene
at the end of the second plague pandemic**

Scientific Reports 2020

doi: 10.1038/s41598-020-71530-9

Supplementary material can be found here: <https://doi.org/10.1038/s41598-020-71530-9>



OPEN

Yersinia pestis strains from Latvia show depletion of the *pla* virulence gene at the end of the second plague pandemic

Julian Susat¹, Joanna H. Bonczarowska¹, Elīna Pētersone-Gordina², Alexander Immel¹, Almut Nebel¹, Guntis Gerhards² & Ben Krause-Kyora¹✉

Ancient genomic studies have identified *Yersinia pestis* (*Y. pestis*) as the causative agent of the second plague pandemic (fourteenth–eighteenth century) that started with the Black Death (1,347–1,353). Most of the *Y. pestis* strains investigated from this pandemic have been isolated from western Europe, and not much is known about the diversity and microevolution of this bacterium in eastern European countries. In this study, we investigated human remains excavated from two cemeteries in Riga (Latvia). Historical evidence suggests that the burials were a consequence of plague outbreaks during the seventeenth century. DNA was extracted from teeth of 16 individuals and subjected to shotgun sequencing. Analysis of the metagenomic data revealed the presence of *Y. pestis* sequences in four remains, confirming that the buried individuals were victims of plague. In two samples, *Y. pestis* DNA coverage was sufficient for genome reconstruction. Subsequent phylogenetic analysis showed that the Riga strains fell within the diversity of the already known post-Black Death genomes. Interestingly, the two Latvian isolates did not cluster together. Moreover, we detected a drop in coverage of the pPCP1 plasmid region containing the *pla* gene. Further analysis indicated the presence of two pPCP1 plasmids, one with and one without the *pla* gene region, and only one bacterial chromosome, indicating that the same bacterium carried two distinct pPCP1 plasmids. In addition, we found the same pattern in the majority of previously published post-Black Death strains, but not in the Black Death strains. The *pla* gene is an important virulence factor for the infection of and transmission in humans. Thus, the spread of *pla*-depleted strains may, among other causes, have contributed to the disappearance of the second plague pandemic in eighteenth century Europe.

Yersinia pestis (*Y. pestis*) is the causative agent of plague that evolved from a relatively benign pathogen—*Yersinia pseudotuberculosis*—thousands of years ago^{1–4}. Whilst plague is a zoonotic disease with rodents being the primary hosts of the pathogen^{5–7}, it has also afflicted humans for at least 5,000 years¹. In the last two millennia, *Y. pestis* was responsible for three major epidemics, of which the second was the most infamous. This pandemic began with the Black Death in the fourteenth century (1,346–1,353). It is thought to have originated in Asia from where it quickly spread to and across the European continent killing approximately 30–50% of the population within a few years^{8,9}. The second pandemic persisted for over 400 years¹⁰ causing, among others, the Great Plague of London (1,665–1,666) and Marseille (1,720–1,722)¹¹. In contrast to western Europe, the first major outbreaks of plague in the north-eastern part of the continent took place mainly in the post-Black Death period (after 1,353), for instance in Riga (1,601–3, 1,621–23 and 1,657) or during the Russian Plague (1,770–1,772)^{12–14}.

Ancient DNA (aDNA) studies have shown that all second-plague pandemic strains cluster together in branch 1 of the *Y. pestis* phylogeny, close to present-day representatives^{15–17}. Interestingly, pathogens involved in the post-Black Death outbreaks derived from two Black Death strains from London and Barcelona^{15,18}. This post-Black Death subset includes strains from Germany (Starnberg 1,433–1,523, Ellwangen 1,486–1,627),

¹Institute of Clinical Molecular Biology, Kiel University, Rosalind-Franklin-Straße 12, 24105 Kiel, Germany. ²Institute of Latvian History, University of Latvia, Kalpaka bulvāris 4, Riga 1050, Latvia. ✉email: b.krause-kyora@ikmb.uni-kiel.de

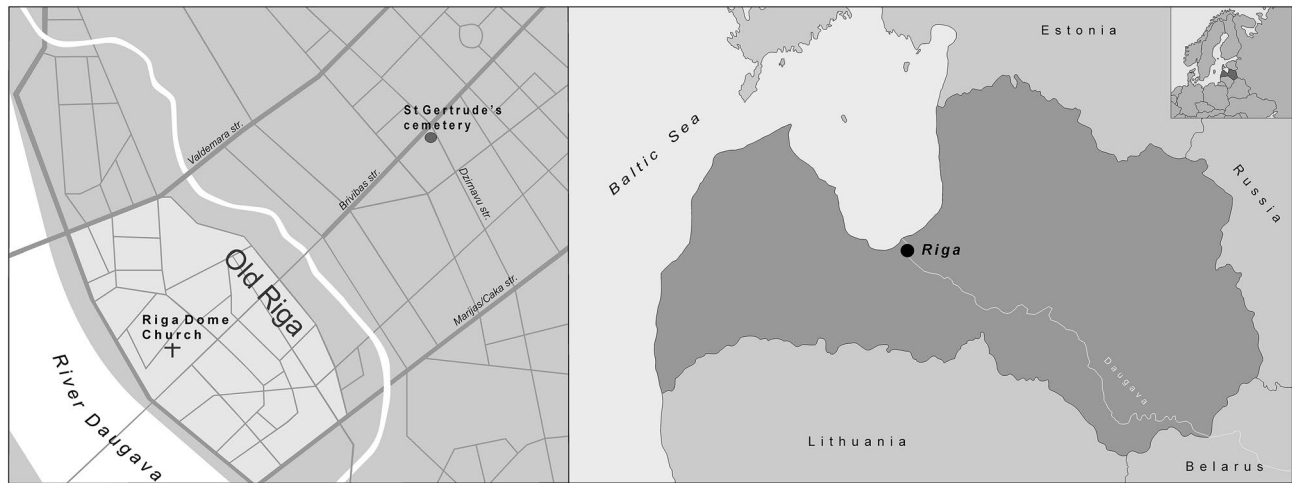


Figure 1. Cemeteries in Riga from which the samples in this study were obtained (left picture). Map of Latvia and Europe (right picture, upper right corner) with the location of Riga. (Figure was created using CorelDRAW Home & Student X8, URL: <https://www.coreldraw.com/en/>).

Cemetery	Context	Burial no.	Osteological sex	Genetic sex	Age (years)	<i>Yersinia pestis</i>
St. Gertrude's Church cemetery	Mass grave 1	G83	ND	No DNA	11–13	No DNA
		G157	M	No DNA	30–35	No DNA
		G488	M	M	30–35	+
		G645	F	F	20–25	+
	Mass grave 2	G103	ND	M	10–11	+
		G177	ND	M	10–12	–
		G687	M	M	45–50	–
	Burial pit	G691	ND	F	9–11	–
		G701	F	F	35–40	+
	General cemetery	G41	ND	M	14–15	–
G92		M	M	55–60	–	
G100		M	M	20–25	–	
Riga Dome Church cemetery	Mass burial pit	G143b	F	F	25–30	–
		G161c	F	M	50–60	–
		G171c	F	ND	40–50	–
	Collective burial pit	G184b	M	M	40–50	–

Table 1. List of samples from St. Gertrude's Church cemetery and Riga Dome Church cemetery. *M* male, *F* female, *ND* not determinable; + /– *Y. pestis*-positive or negative.

England (Cambridge 1,433–1,523, London 1,560–1,635), Sweden (Stans 1,485–1,635) and France (Marseille 1,720–1,722)¹⁹. Since no modern descendants have yet been identified, this lineage probably went extinct¹⁰.

There are ongoing debates about the reasons as to why the disease disappeared from Europe in the eighteenth century. Improvements in living conditions, hygiene, and quarantine facilities as well as a replacement of the rodent host from black to brown rat have been suggested²⁰. Other explanations revolved around the idea that evolutionary processes may have favored less virulent pathogen strains²¹ or may have resulted in increased immunity of the human population²⁰.

Most of the second-pandemic strains analyzed to date have been isolated from human remains excavated in western Europe, with exception of two strains found in Russia (Bolgar 1,362–1,400¹⁸, Laishovo 1,300–1,400¹⁹). Therefore, not much is known about the diversity and microevolution of the strains responsible for disease outbreaks in eastern Europe, which could have served as a gateway for the spread of the pathogen towards the west.

At least three plague outbreaks in Riga (Latvia) dating to the seventeenth century have been previously reported^{12,13}. To confirm the presence of *Y. pestis* in Riga, we analysed samples from 16 human skeletons excavated from two cemeteries (fifteenth–seventeenth century¹³) (Fig. 1 and Table 1). The individuals had been interred in multiple or mass burials and single graves (Table 1). Four of the samples yielded positive screening results for the presence of *Y. pestis*, leading to the successful reconstruction of two complete bacterial genomes. This new data allowed us to expand the geographical range of known *Y. pestis* strains to north-eastern Europe.

Sample	Aligned reads	Coverage ≥ 1x (%)	Coverage ≥ 2x (%)	Coverage ≥ 3x (%)	Mean coverage whole genome
Chromosome					
G701	237,557	92.74	75.66	53.77	3.09x
G488	597,757	98.85	95.83	89.76	6.79x
pMT1					
G701	6,540	97.74	89.33	75.03	4.16x
G488	17,450	99.65	99.04	97.52	9.68x
pPCP1					
G701	4,544	98.75	94.42	88.24	30.07x
G488	11,893	99.80	99.14	98.72	68.28x
pCD1					
G701	6,659	99.35	96.30	89.55	5.9x
G488	26,728	99.99	99.14	98.72	20.99x

Table 2. Basic mapping statistics for the strongly positive *Y. pestis* samples after duplicate removal and quality filtering. Mapping against the CO92 chromosome (NC_003143.1), pMT1 (NC_003134.1), pPCP1 (NC_003132.1) and pCD1 (NC_003131.1).

Results

Sixteen individuals excavated from two cemeteries located in Riga, Latvia were submitted for aDNA analysis. For 14 samples, authentic ancient shotgun sequences were generated. Of these, we identified four *Y. pestis*-positive individuals using MALT in BlastN mode (Table 1 and S1). Three of them (G103, G645, G488) were buried in two mass graves and one (G701) in the burial pit of the St. Gertrude’s Church cemetery. The presence of *Y. pestis*-specific reads was confirmed by mapping against a multi fasta reference containing several representatives of the genus *Yersinia* (Table S2). Furthermore, we calculated an endogenous based score (i.e. the number of *Y. pestis*-specific reads with respect to the number of total reads in the sample)² (Table S3). Two of the samples (G488, G701) were identified as strong positives (endogenous based score > 0.005, Table S3). For these candidates, we subsequently generated 330 and 383 million reads, respectively, to reconstruct the complete bacterial genomes. The obtained reads were mapped against the *Y. pestis* reference (CO92, NC_003143.1) containing the chromosome and all three plasmids (pCD1, NC_003131.1; pMT1, NC_003134.1; pPCP1, NC_003132.1). Both samples exhibited a genome coverage above 92% for the chromosome as well as for the three plasmids (Table 2 and Fig. S1). Terminal deamination patterns displayed damage profiles typical of authentic aDNA (Fig. S2).

We computed maximum likelihood and Bayesian trees (Figs. 2, S3 and S4) with the two genomes from Riga (G488, G701) and previously published *Y. pestis* strains (228 modern (Table S4), 36 ancient (Table S5)) and a *Y. pseudotuberculosis* genome (NZ_CP008943.1). The strains from Riga clustered with other second-plague pandemic strains from Starnberg (STA001) and Cambridge (NMS002)¹⁹ with a strong bootstrap support of > 97% and a posterior probability of 1 (Figs. 2, S3 and S4).

Both Riga isolates had SNP alleles (p1 and p2¹⁴) diagnostic for branch 1 (Table S6). Furthermore, they showed differences in the SNPs p3 and p4 relative to the CO92 reference sequence. These substitutions also characterize other previously reported second-plague pandemic strains from London (England), Barcelona (Spain), Ellwangen (Germany), and l’Observance (France)^{10,15,18}.

The two reconstructed genomes from Riga exhibited a drastic drop in coverage of sequences mapping to the *pla* gene region (NC_003132.1:6,428–8,530), which consists of the *pla* gene, a putative transcriptional regulator and a hypothetical protein on the pPCP1 plasmid (Fig. 3). We confirmed the presence of two variants of the plasmid, i.e. one with (*pla*+) and one without (*pla*–) the *pla* region, by identifying gap-spanning reads in each of the strains (Figs. 3C,D, S5). In contrast, we noted only one bacterial chromosome per strain, indicating that the two distinct pPCP1 plasmids existed in the same bacterium. To rule out that the *pla* gene originated from contamination, a competitive mapping against the *pla* gene from environmental bacteria (e.g. *E. coli* and *Citrobacter koseri*) and *Y. pestis* showed that the reads in this region were indeed specific for *Y. pestis* and exhibited the expected damage patterns (Table S7). This observation was verified by a competitive mapping against a *pla*+ and a *pla*– reference (Fig. 3 and Table S8). The observed ratio between the coverage of the region in *pla*+ and *pla*– plasmids was about 1:10 (Table S9). To further explore this finding, we re-analyzed 28 previously published *Y. pestis* genomes from the second pandemic^{10,15,18,19}. Interestingly, in 14 strains (marked in blue, Fig. 2) we detected gap-spanning reads and a depletion of the *pla* region, also indicating the presence of both *pla*+ and *pla*– plasmids (Fig. S6 and Table S10). Four more strains showed no signs of *pla*– plasmids (marked in red, Figs. 2 and S6), whereas ten others, for which the overall pPCP1 plasmid coverage was poor, were excluded as their *pla* region could not be examined reliably (marked in black, Fig. 2). The possibility that the different plasmids could be the result of a co-infection with two different strains is highly unlikely as this pattern was found in several *Y. pestis* genomes from various sites and times, and there was no evidence for such an infection on the chromosomal level.

With regard to (potentially) functional SNPs, we identified two variants of interest. First, both Riga genomes carried the T259 allele in the *pla* gene that is important for massive bacterial dissemination in the bubonic form of plague²². Second, in sample G488, SnpEff²³ detected a variant in the *ssrA-smpB* region.

For five of the samples (G92, G143b, G161c, G488 and G701), the human endogenous DNA content in the metagenomic data was sufficient to investigate the population genetic ancestry using PCA, f3 and admixture

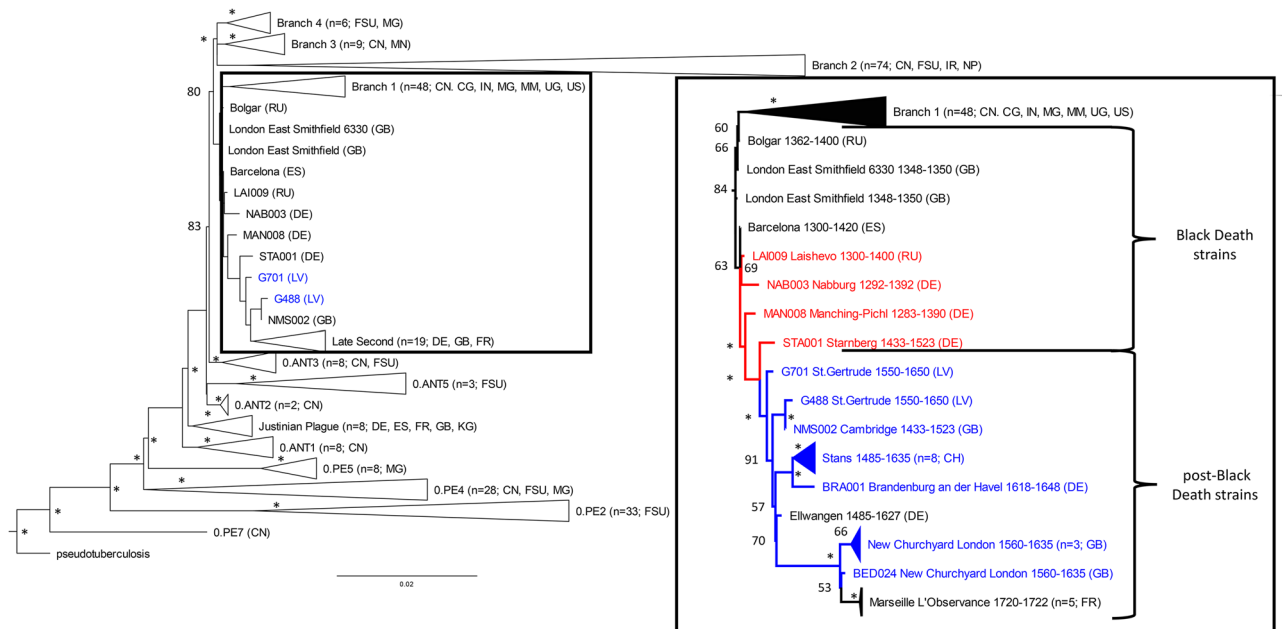


Figure 2. Maximum-likelihood tree. The tree is based on the SNP alignment (15,976 positions) of 228 modern *Y. pestis* genomes, 36 published ancient *Y. pestis* strains, one *Y. pseudotuberculosis* genome, and the two genomes G488 and G701 from St. Gertrude's Church cemetery (left panel, blue). Country abbreviation is given in brackets (DE = Germany, ES = Spain, FR = France, GB = Great Britain, RU = Russia, LV = Latvia, CN = China, CG = Congo, FSU = Former Soviet Union, IN = India, IR = Iran, MG = Madagascar, MM = Myanmar, MN = Mongolia, NP = Nepal, UG = Uganda, KG = Kyrgyzstan, CH = Switzerland). Bootstrap values are shown on the nodes for 500 replicates. An asterisk (*) indicates a bootstrap support above 94. Strains in red exhibit no signs of the *pla*-negative plasmid. Strains in blue exhibit clear signs of the *pla*-negative plasmid. For the strains in black (within the Black Death and post-Black Death subclades), the *pla* status was inconclusive or could not be determined due to a general lack of coverage.

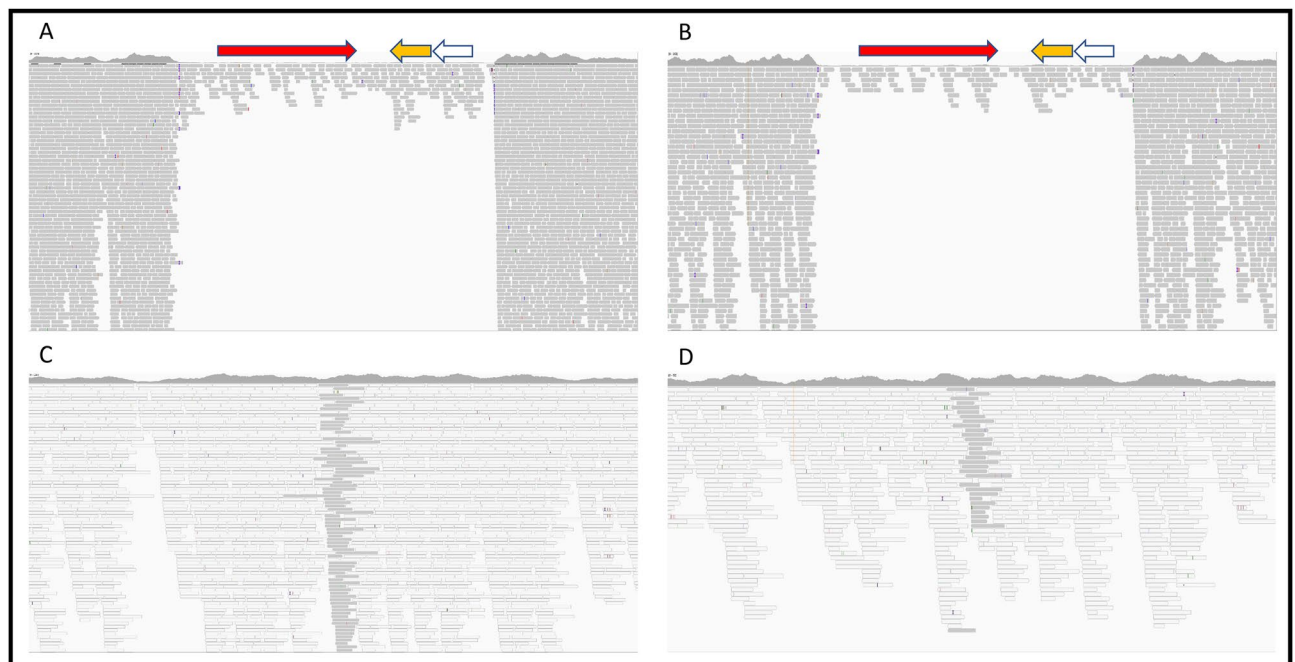


Figure 3. Coverage plots of pPCP1. (A, B) Coverage plot of the *pla* region in G701 (A) and G488 (B) depicting the *pla* gene in red, the putative transcriptional regulator in yellow and the hypothetical protein in white. G488 average coverage (ac) before the *pla* region 107x, ac of the *pla* region 7x; G701 ac before the *pla* region 47x, ac of the *pla* region 3x. Each grey line represents an obtained aDNA sequence. (C) Coverage plot showing reads of G701 (sequences in dark grey) specifically mapping to the gap-spanning region. A white color indicates sequences that could be mapped to both the *pla*+ and *pla*- references as they are not specific to the newly created region. (D) Coverage plot showing reads of G488 (sequences in dark grey) specifically mapping to the gap-spanning region.

analyses. The results indicate a northern European origin of the individuals who were genetically fairly homogeneous (Figs. S7, S8, S9 and S10).

Discussion

Here we analyzed the remains of 16 individuals who had been interred in two cemeteries (Dome Church cemetery, St. Gertrude's Church cemetery) in Riga, Latvia (fifteenth–seventeenth century) (Fig. 1 and Table 1). Four of the individuals buried in two mass graves and a burial pit of St. Gertrude's Church cemetery were infected with the *Y. pestis* bacterium. Previous analyses led to the hypothesis that people inhumed in the mass graves were rural immigrants. Our human population genetic data confirmed their northern European descent. Most likely they came to Riga during the famine of 1601–2 and stayed in shelters in the vicinity of St. Gertrude's Church¹³. Inadequate nutrition and crowded, unhygienic conditions may have increased their susceptibility to infectious diseases, including plague^{24,25}. Our finding showed that the individuals buried in the mass graves as well as the small burial pit were indeed victims of plague^{26,27}. Thus, we corroborate historical reports of local plague outbreaks in Riga at the beginning of the seventeenth century (1,601–3, 1,621–23 and 1,657)¹².

From two (G488 and G701) of the four infected individuals, high-quality whole genomes of *Y. pestis* were generated. They formed a subclade in branch 1 with other strains dating to the post-Black Death pandemic (Fig. 2). They all originated from Black Death genomes from London¹⁵ and Barcelona¹⁸. Although the two Riga strains were found in the same cemetery (in one mass grave and the burial pit), they showed a surprisingly large genetic distance. Based on this observation and the close proximity of G488 to NMS002 (Cambridge¹⁹) (Fig. 2), we assume that G488 had not evolved from G701 during a single outbreak. This may indicate that the two different strains were responsible for two subsequent plague events^{12,13}.

As no modern strains fall into the post-Black Death subclade, which includes the two genomes from Riga, it could represent a *Y. pestis* lineage that went extinct¹⁰. Although the strains differ widely in terms of time and geography, their positions within the subclade—like pearls on a string—suggest continuous in situ evolution. This observation agrees with the hypothesis that the second plague pandemic reached Europe from Asia rather than in one than in multiple waves. If the pandemic had multiple sources, one would expect a higher diversity. Thus, the post-Black Death outbreaks were likely due to established plague reservoirs within Europe. Little is known about natural plague foci and hosts during the second pandemic. Several rodent species, e.g. marmots and squirrels, from the Alps and the steppe have been mentioned as possible reservoirs²⁰.

The majority (ratio 10:1) of the pPCP1 plasmid copies carried by the two Riga strains G488 and G701 did not possess the *pla* region. Based on the chromosome data we did not see any signs of coinfection with different strains that could explain the presence of two different pPCP1 plasmids. In a re-evaluation of already published second-plague pandemic genomes, we identified 14 strains carrying both plasmid types (*pla*– and *pla*+) (Table S10, Fig. S6). All 14 strains were dated to the post-Black Death period whereas four strains which only had copies of the *pla*– plasmid were present during the Black Death (Fig. 2). The *pla* region consists of the *pla* gene, a putative transcriptional regulator and a hypothetical protein. Up to now, the role of the putative transcriptional regulator and the hypothetical protein with respect to *pla* expression and function has remained unclear. Several studies have shown that *pla* is an important factor for the infection of and transmission in humans. The *pla* protease activates host fibrinolysis and allows the spread of the bacteria to the lymphatic nodes²⁸ as well as their extensive growth in the lower airways^{29,30}. *Pla*– mutants of *Y. pestis* have been associated with reduced dissemination and virulence^{22,30,31}. In mice, *pla*– strains transmitted by fleas were reported to cause primary septicemic plague while no bubonic forms were detected³¹. Therefore, *pla* seems to influence the infection routes leading to bubonic and pneumonic plague, while the septicemic route remains unaffected³². Therefore, the extremely high mortality of the Black Death could possibly be explained by the functional features of *pla*. The pandemic was mainly associated with a rapid spread of the severe bubonic form via the flea vector. In contrast, subsequent outbreaks could have been influenced by a slower transmission rate and a different disease course (primarily septicemic rather than bubonic) due to the *pla*-depleted strains (Fig. 2).

Horizontal gene transfer, selective gene loss, and changes in regulatory pathways have been described in the evolution and adaptation of *Y. pestis* to new host environments^{33,34}. These mechanisms together with the genomic structure (e.g. IS-like elements or transposases) of the plasmids could have led to a loss of *pla* or a *pla*-depletion. *Pla*– *Y. pestis* strains have been isolated worldwide^{35,36} and *pla*– plasmids are known from *Pestoides* strains³⁷—enzoonotic pathogens that infect rodents³⁸, but rarely humans³⁹. The *pla*-depleted post-Black Death strains could therefore be an adaptation to a specific host environment, perhaps in rodents.

The strain from individual G488 possessed a mutation in the *ssrA* gene, which has not yet been described in other strains. Together with the *smgB* protein, *ssrA* creates a complex responsible for maintaining cellular homeostasis^{40–42}. The important role of this system in bacterial pathogenicity was shown in previous studies on several bacteria such as *Salmonella*, *Y. pseudotuberculosis* and *Y. pestis*^{43–45}. Interestingly, in a mouse model, *Y. pestis* strains with a deletion of *ssrA* and *smgB* had limited ability to disseminate and colonize target organs. Moreover, mice which were intranasally immunized with this mutated strain were protected against pulmonary infection with a fully virulent strain⁴⁶. In view of these findings, a SNP in the *ssrA* gene is intriguing. However, the functional implications of this single point mutation remain unclear.

Y. pestis strains responsible for the Post-Black Death plague showed a *pla*-depletion in the pPCP1 plasmid. This change in the pathogen genome can probably be attributed to an adaptation of the bacterium to a new host environment (e.g. rodents). In humans, the depletion may have favorably influenced the disease course, the probability of disease manifestation after a flea bite and/or the transmission rate. These positive effects fit with the evolutionary hypothesis that less virulent strains may, at least in part, explain the easing and ultimate disappearance of plague from Europe²¹. Nevertheless, it is imperative to obtain more *Y. pestis* genomes and conduct

additional functional studies to create a more comprehensive picture of the bacteria responsible for the second plague pandemic.

Material and methods

Historical background and sample selection. Individuals excavated from the cemeteries Riga Dome Church and St. Gertrude's Church cemetery, Riga, were selected for a bacterial pathogen screening (Fig. 1). At Riga Dome Church cemetery, three mass burial pits were found at the northern wall of the church. A total of 87 individuals were identified and archaeological finds suggest that they date to the seventeenth century⁴⁷. Unlike people buried elsewhere in the cemetery, the individuals in the mass burial pits had been buried in coffins, and none of them showed evidence of peri-mortem trauma⁴⁸. Therefore, it is likely that they had succumbed to a mass mortality event that was not related to violence. Indeed, they could be victims of one of the seventeenth century plague epidemics. Historical sources suggest that during the plague epidemic of 1,657, the churches of Riga ran out of intramural burial space, which was normally reserved for people of a higher social status⁴⁸. It is possible that during the epidemic, the dead were buried in larger pits outside the church, albeit in a dedicated area and in coffins to signify their higher social status. In addition, smaller collective burials close to the aforementioned mass burial pits were found. In these, four to nine people were buried in several layers, without coffins. For DNA analysis, three samples were selected from the mass burial pits with coffins, and one from the smaller collective burial pits (Table 1).

The other site selected for aDNA analysis was St. Gertrude's Church cemetery. The graveyard was in use between the fifteenth and seventeenth century, and back then located outside the city wall (Fig. 1). The cemetery was divided into four distinct burial areas: a general cemetery, two mass graves, and a small mass burial pit. In total, 709 individuals were excavated from this cemetery, and 283 of these were buried in the two mass graves (south-eastern mass grave, or mass grave 1, and north-western mass grave, or mass grave 2)¹³. The bodies in the mass graves were placed on top of each other, reaching seven to eight layers in total. Both archaeological (based on artefacts) and radiocarbon analysis dated the two mass graves to the end of the sixteenth or the beginning of the seventeenth century, although they are not necessarily contemporaneous⁴⁹. The chronology of the small mass burial pit with 15 buried individuals is not entirely clear. The demographic profile of the mass graves with comparatively few children < 7 years of age and, again, the lack of evidence of peri-mortem trauma, suggested that these people had died in a non-violent mass mortality event^{13,48}. Moreover, the date of both mass graves coincides with two historically documented plague epidemics (1,601–3 and 1,621–23) and a famine (1,601–2)¹². For DNA analysis, four samples were selected from mass grave 1, three from mass grave 2, two from the grave pit and three from the general cemetery (Table 1).

DNA isolation, sequencing and postprocessing. All lab work was carried out in a dedicated ancient DNA facility. DNA extraction of 16 tooth samples from Riga, sequencing (Illumina HiSeq4000 (2 × 75 bp)) and subsequent clip and merge was performed as described in Krause-Kyora et al.⁵⁰. Reads that were shorter than 25 bp were discarded after merging. Two samples (G83 and G157) did not yield sufficient DNA and were not processed further (Table 1).

Pathogen screening. Screening of the 14 samples was carried out with the software MALT using a custom database consisting of complete bacterial genomes downloaded from NCBI (24.01.2019)⁵¹. A sequence identity threshold of 90% was set and the alignment mode was changed to SemiGlobal. The analysis was done using the following command:

```
malt-run -mode BlastN -e 0.001 -id 85 -alignmentType SemiGlobal -index $index -inFile $FASTQCM -output $OUT
```

where \$index is the index file, \$FASTQCM is the clipped and merged file and \$OUT is the output file.

The resulting alignments were visually inspected using MEGAN 6⁵².

***Y. pestis* alignment.** All four positive samples from the screening (G103, G488, G645, G701) were mapped against the *Y. pestis* reference genome including the three plasmids pCD1, pMT1 and pPCP1 (NC_003143.1; NC_003131.1; NC_003134.1; NC_003132.1). The mapping was conducted using BWA with the following command line⁵³:

```
bwa aln -n 0.01 -l 300 $INDEX $FASTQCM $OUT
```

where \$INDEX is the reference, \$FASTQCM is the input file and \$OUT is the output file. Minimum mapping quality was set to 0. Subsequent duplicate removal was performed using DeDup version 0.11.3, part of the EAGER pipeline⁵⁴, with default parameters. Subsequently, the Binary Alignment Map (BAM) files were filtered for a mapping quality of 30 using the following command in samtools⁵⁵:

```
samtools view -bq $INT $IN > $OUT
```

where \$INT is the mapping quality, \$IN is the input BAM file and \$OUT is the output BAM file.

Sample authentication. To further authenticate the positive samples, an endogenous score was used to calculate the number of *Y. pestis*-specific reads with respect to the number of total reads in the sample². Positive samples were mapped against a multi fasta reference containing several representatives of the genus *Yersinia* (Table S2) and the three *Y. pestis*-specific plasmids (pCD1, pMT1, pPCP1). The mapping was done as described

above followed by a duplicate removal step. Mapped reads were filtered for a mapping quality greater than 30 using samtools as described above.

Applying samtools idxstats, an endogenous based score was used to evaluate the potential of the sample being positive for *Y. pestis*. The score was calculated with the following formula:

$$\frac{(YPS - (YS))}{M} \times 1000$$

where YPS is the number of reads specifically mapping to *Y. pestis*; YS is the maximum number of reads mapping specifically to any *Yersinia* species with the exception of *Y. pestis* and M is the total number of merged reads in the sample. By using the maximum number of reads mapping to another species of the genus *Yersinia*, the score takes into account different sources of contamination other than *Y. pseudotuberculosis*. Two of the four positive samples exhibited a score higher than 0.005, a value which indicates a strong positive candidate (Table S3).

VCF-generation and phylogenetic analysis. VCF files were generated for the two positive samples using GATK and the *Y. pestis* reference with the following command line⁵⁶:

GenomeAnalysisTK.jar -T UnifiedGenotyper -R \$REF -I \$IN -out_mode EMIT_ALL_SITES -o \$OUT

where \$REF is the reference, \$IN is the input BAM file and \$OUT is the VCF output. These two VCFs together with a VCF dataset of 264 strains (Table S4 and S5, mapped (see *Y. pestis* alignment) and generated as described above) and one *Y. pseudotuberculosis* (NZ_CP008943.1) genome were used as an input for the MultiVCFAnalyzer to generate a SNP-based multiple alignment⁵⁷. The analysis was carried out with the following command:

Java -jar MultiVCFAnalyzer_0-85-1.jar NA \$REF NA \$OUT F 30 3 0.9 0.9 NA

where \$REF is the *Y. pestis* reference and \$OUT is the output folder. *Y. pseudotuberculosis*-specific variants were removed from the SNP table to improve visual resolution of the tree. The resulting SNP table was also filtered for SNPs in repetitive regions and homoplasies¹⁶. The SNP table was reformatted to multi fasta and nexus format that were used as input for RAxML and MrBayes^{58,59}. RAxML was executed with the following command line:

RaxmlHPC -f a -x 12345 -p 12345 -#500 -m GTRGAMMA -s \$IN -o \$OG -n \$NAME

where \$IN is the input file, \$OG is the name of the outgroup and \$NAME is the outfile name.

MrBayes was executed applying the GTR model with the default number of generations.

Resulting trees were visualized using FigTree⁶⁰.

SNP effect analysis. VCFs from the *Y. pestis*-positive samples were filtered for SNPs with a coverage of at least 3x, a calling quality of 30 and a 90% majority call. This set of filtered SNPs was provided to SNPEff to analyze the effects of SNPs in the ancient genomes²³. SNPEff was executed with the following command line:

java -jar snpEff.jar \$REF \$IN > \$OUT

where \$REF is the SNPEff reference, \$IN is the input VCF file and \$OUT is the annotated SNPEff output file. The annotated VCF file was filtered for high-impact SNPs and evaluated.

pPCP1 and *pla* analysis. To ascertain that the coverage in the low-coverage region was not introduced artificially by environmental contamination, we carried out a competitive mapping as described above. The fasta reference that was used contained the *Y. pestis* reference sequence of the pPCP1 plasmid and also the *pla* genes of *Citrobacter* and *E. coli* (Table S7). Gap-bridging reads were analyzed by removing the low-coverage region from the pPCP1 plasmid. Seqkit was used with the following command lines⁶¹:

seqkit subseq -r 1:6428 \$IN > \$OUT1

seqkit subseq -r 8530:9612 \$IN > \$OUT2

where \$IN is the pPCP1 reference (NC_003132.1) and \$OUT1/\$OUT2 are parts of the new plasmid without the *pla* region. Both output files were concatenated to form a complete plasmid. An additional mapping with the pPCP1 plasmid from which the *pla* region was removed (*pla*-) and the regular pPCP1 plasmid (*pla*+) as references was performed (Table S8). These references were used to identify reads which were truly unique for the new region in the *pla*- plasmid. All the non-specific reads received a mapping quality of zero as the exact same region was present in both references. The reads mapping exclusively to the new region were extracted and compared with the reads from the initial mapping against *Y. pestis* and gap-bridging reads were identified. Based on these exclusive mapping reads, the coverage for the *pla*- plasmid was calculated. The coverage of the *pla* region (NC_003132.1, 6,428–8,529) in the regular mapping was used to estimate the number of *pla*+ plasmids. Based on these two values a ratio was calculated to estimate the proportions.

Human genetic analyses. *Mapping and aDNA damage patterns.* Shotgun sequence data was mapped to the human genome build hg19 using BWA with a reduced mapping stringency parameter “-n 0.01” to account for mismatches in aDNA^{54,62}. Duplicates were removed. Authenticity of the aDNA fragments was assessed using mapDamage 2.0^{63,64}. After validation of the terminal damage patterns, the first two positions from the 5' end of the fastq-reads were trimmed.

Genotyping. Alleles were drawn at random from a pool of 1,233,013 SNPs (commonly used for population genetic analyses) in a pseudo-haploid manner using a custom script as described previously^{65–68}. Datasets had to have at least 10,000 SNPs to be considered for further analysis⁶⁸. Five individuals from Riga (G92, G143b, G161c, G488 and G701) passed this filtering step.

Population genetic analyses. The five genotype datasets from Riga were merged with a dataset of previously published 59 West Eurasian populations genotyped on the aforementioned 1,233,013 SNPs using the program mergeit from the EIGENSOFT package⁶⁹. Principal component analysis (PCA) was performed using the software smartpca and the 'lsqproject' option⁶⁹. Admixture modelling was done with the software ADMIXTURE on the same populations as used in the PCA. The number of ancestral components k ranged from 4 to 12⁷⁰. Cross-validation was performed for every admixture model and the model with the highest accuracy was determined by the lowest cross-validation error.

f_3 outgroup statistics. f_3 outgroup statistics were run as a part of the Admixtools package in the form of f_3 (Riga; test; Mbuti) using for test the same populations as used in the PCA and ADMIXTURE analysis⁷⁰.

Material transfer agreement. The authors confirm that the sampling and the ancient DNA analysis of human bones from the fifteenth–seventeenth century archaeological features unearthed in Riga, Latvia were performed with permission from the institution which headed the excavations and the anthropological analysis (Institute of Latvian History, University of Latvia).

Data availability

The aligned sequences are available through the European Nucleotide Archive under Accession Number PRJEB36413.

Received: 20 March 2020; Accepted: 4 August 2020

Published online: 03 September 2020

References

- Rasmussen, S. *et al.* Early divergent strains of *Yersinia pestis* in Eurasia 5,000 years ago. *Cell* **163**, 571–582 (2015).
- Valtuna, A. A. *et al.* The Stone Age plague and its persistence in Europe. *Curr. Biol.* **27**, 3683–3691 (2017).
- Spyrou, M. A. *et al.* Analysis of 3800-year old *Yersinia pestis* genomes suggests Bronze Age origin for bubonic plague. *Nat. Commun.* **9**, 2234. <https://doi.org/10.1038/s41467-018-04550-9> (2018).
- Rascovan, N. *et al.* Emergence and spread of basal lineages of *Yersinia pestis* during the Neolithic Decline. *Cell* **176**, 295–305 (2019).
- Anisimov, A. P., Lindler, L. E. & Pier, G. B. Intraspecific diversity of *Yersinia pestis*. *Clin. Microbiol. Rev.* **17**, 434–464 (2004).
- Gage, K. L. & Kosoy, M. Y. Natural history of plague: perspectives from more than a century of research. *Annu. Rev. Entomol.* **50**, 505–528 (2005).
- Roelle, J. E., Miller, B. J., Godbey, J. L. & Biggins, D. E. Recovery of the black-footed ferret progress and continuing challenges. *US Geol. Surv. Sci. Invest. Rep.* **5293**, 213–231 (2006).
- Morelli, G. *et al.* *Yersinia pestis* genome sequencing identifies patterns of global phylogenetic diversity. *Nat. Genet.* **42**, 1140–1143 (2010).
- Ziegler, P. *The Black Death* (Harper Collins, New York, 1969).
- Bos, K. I. *et al.* Eighteenth century *Yersinia pestis* genomes reveal the long-term persistence of an historical plague focus. *Elife* **5**, e12994. <https://doi.org/10.7554/eLife.12994> (2016).
- Byrne, J. P. *Encyclopedia of the Black Death* (ABC-CLIO, Santa Barbara, 2012).
- Napiersky, J. G. L. *Bodeckers Chronik livländischer und rigascher Ereignisse [Bodecker's chronicle of Livonian and Riga events]* (in German) (Kymmel's Buchhandlung, Riga, 1890).
- Petersone-Gordina, E., Roberts, C., Millard, A. R., Montgomery, J. & Gerhards, G. Dental disease and dietary isotopes of individuals from St Gertrude Church cemetery, Riga, Latvia. *PLoS ONE* **13**, e0191757. <https://doi.org/10.1371/journal.pone.0191757> (2018).
- John, T. A. *Bubonic Plague in Early Modern Russia: Public Health and Urban Disaster* (Oxford University Press, Oxford, 2003).
- Bos, K. I. *et al.* A draft genome of *Yersinia pestis* from victims of the Black Death. *Nature* **478**, 506–510 (2011).
- Cui, Y. *et al.* Historical variations in mutation rate in an epidemic pathogen, *Yersinia pestis*. *Proc. Natl. Acad. Sci. USA* **110**, 577–582 (2013).
- Namouchi, A. *et al.* Integrative approach using *Yersinia pestis* genomes to revisit the historical landscape of plague during the Medieval Period. *Proc. Natl. Acad. Sci. USA* **115**, E11790–E11797. <https://doi.org/10.1073/pnas.1812865115> (2018).
- Spyrou, M. A. *et al.* Historical *Y. pestis* genomes reveal the European Black Death as the source of ancient and modern plague pandemics. *Cell Host Microbe* **19**, 874–881 (2016).
- Spyrou, M. A. *et al.* Phylogeography of the second plague pandemic revealed through analysis of historical *Yersinia pestis* genomes. *Nat. Commun.* **10**, 4470. <https://doi.org/10.1038/s41467-019-12154-0> (2019).
- Appleby, A. B. The disappearance of plague: a continuing puzzle. *Econ. Hist. Rev.* **33**, 161–173 (1980).
- Lenski, R. E. Evolution of plague virulence. *Nature* **334**, 473–474 (1988).
- Zimble, D. L., Schroeder, J. A., Eddy, J. L. & Lathem, W. W. Early emergence of *Yersinia pestis* as a severe respiratory pathogen. *Nat. Commun.* **6**, 7487. <https://doi.org/10.1038/ncomms8487> (2015).
- Cingolani, P. *et al.* A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff. *Fly* **6**, 80–92 (2012).
- Dirks, R. *et al.* Social responses during severe food shortages and famine. *Cur. Anthropol.* **21**, 21–44 (1980).
- Wisner, B., Adams, J. Control of communicable diseases and prevention of epidemics. Rigas pilsētas valde WHO. www.who.int/entity/water_sanitation_health/hygiene/emergencies/em2002chap11.pdf (2019).
- Pirangs, H. *Rigas kapsētas [Graveyards of Riga]* (in Latvian) (Rigas pilsētas valde, Riga, 1932).
- Rusovs, B. *Livonijas kronika [Chronicle of Livonia]* (in Latvian) (Valters un Rapa, Riga, 1926).
- Degen, J. L., Bugge, T. H. & Goguen, J. D. Fibrin and fibrinolysis in infection and host defense. *J. Thromb. Haemost.* **5**, 24–31 (2007).
- Lathem, W. W. *et al.* Post-transcriptional regulation of the *Yersinia pestis* cAMP receptor protein Crp and impact on virulence. *MBio* **5**, e01038-13. <https://doi.org/10.1128/mBio.01038-13> (2014).
- Perry, R. D. & Fetherston, J. D. *Yersinia pestis*—etiologic agent of plague. *Clin. Microbiol. Rev.* **10**, 35–66 (1997).

31. Sebbane, F., Jarrett, C. O., Gardner, D., Long, D. & Hinnebusch, B. J. Role of the *Yersinia pestis* plasminogen activator in the incidence of distinct septicemic and bubonic forms of flea-borne plague. *Proc. Natl. Acad. Sci. USA* **103**, 5526–5530 (2006).
32. Sebbane, F. *et al.* Adaptive response of *Yersinia pestis* to extracellular effectors of innate immunity during bubonic plague. *Proc. Natl. Acad. Sci. USA* **103**, 11766–11771 (2006).
33. Zhou, D. & Yang, R. Molecular Darwinian evolution of virulence in *Yersinia pestis*. *Infect. Immun.* **77**, 2242–2250 (2009).
34. Sun, Y. C., Jarrett, C. O., Bosio, C. F. & Hinnebusch, B. J. Retracing the evolutionary path that led to flea-borne transmission of *Yersinia pestis*. *Cell Host Microbe* **15**, 578–584 (2014).
35. Cavalcanti, Y. V. N., Leal, N. C. & De Almeida, A. M. P. Typing of *Yersinia pestis* isolates from the state of Ceará, Brazil. *Lett. Appl. Microbiol.* **35**, 543–547 (2002).
36. Filippov, A. A., Solodovnikov, N. S., Kookleva, L. M. & Protsenko, O. A. Plasmid content in *Yersinia pestis* strains of different origin. *FEMS. Microbiol. Lett.* **67**, 45–48 (1990).
37. Garcia, E. *et al.* Pestoides F, an atypical *Yersinia pestis* strain from the former Soviet Union. *Adv. Exp. Med. Biol.* **603**, 17–22 (2007).
38. Achtman, M. *et al.* Microevolution and history of the plague bacillus, *Yersinia pestis*. *Proc. Natl. Acad. Sci. USA* **101**, 17837–17842 (2004).
39. Rajanna, C. *et al.* Characterization of pPCP1 plasmids in *Yersinia pestis* strains isolated from the former Soviet Union. *Int. J. Microbiol.* <https://doi.org/10.1155/2010/760819> (2010).
40. Gueneau de Novoa, P. & Williams, K. P. The tmRNA website: reductive evolution of tmRNA in plastids and other endosymbionts. *Nucl. Acids. Res.* **32**, D104–D108. <https://doi.org/10.1093/nar/gkh102> (2004).
41. Keiler, K. C. Biology of trans-translation. *Annu. Rev. Microbiol.* **62**, 133–151 (2008).
42. Dulebohn, D., Choy, J., Sundermeier, T., Okan, N. & Karzai, A. W. Trans-translation: the tmRNA-mediated surveillance mechanism for ribosome rescue, directed protein degradation, and nonstop mRNA decay. *Biochemistry* **46**, 4681–4693 (2007).
43. Okan, N. A., Bliska, J. B. & Karzai, A. W. A role for the SmpB-SsrA system in *Yersinia pseudotuberculosis* pathogenesis. *PLoS Pathog.* **2**, e6. <https://doi.org/10.1371/journal.ppat.0020006> (2006).
44. Baumler, A. J., Kusters, J. G., Stojiljkovic, I. & Heffron, F. Salmonella typhimurium loci involved in survival within macrophages. *Infect. Immun.* **62**, 1623–1630 (1994).
45. Julio, S. M., Heithoff, D. M. & Mahan, M. J. ssrA (tmRNA) plays a role in Salmonella enterica serovar Typhimurium pathogenesis. *J. Bacteriol.* **182**, 1558–1563 (2000).
46. Okan, N. A., Mena, P., Benach, J. L., Bliska, J. B. & Karzai, A. W. The smpB-ssrA mutant of *Yersinia pestis* functions as a live attenuated vaccine to protect mice against pulmonary plague infection. *Infect. Immun.* **78**, 1284–1293 (2010).
47. Tilko, S. Apbedišanas tradīcijas Doma viduslaiku kapsētā [Mortuary Practices in the Dom Medieval Cemetery] (in Latvian). Senā Rīga: Pētījumi pilsētas arheoloģijā un vēsturē [Ancient Riga: Studies in the archaeology and history of the city]. *Rīga: Latvijas vēstures institūta apgāds*. 273–289 (1998).
48. Gerhards, G. Traumas un ievainojumi Rīgas 13.–18. gadsimta iedzīvotājiem [Trauma and Injuries of Riga 13th–18th Century Population] (in Latvian). Senā Rīga: pētījumi pilsētas arheoloģijā un vēsturē [Ancient Riga: Studies in the archaeology and history of the city] (in Latvian). *Rīga: Latvijas vēstures institūta apgāds*. 128–148 (2012).
49. Petersone-Gordina, E., *et al.* Investigating dietary life histories and mobility in children buried in St Gertrude Church cemetery, Riga, Latvia (15th–17th centuries AD). *Archaeometry* **62**, 3–18 (2020).
50. Krause-Kyora, B. *et al.* Ancient DNA study reveals HLA susceptibility locus for leprosy in medieval Europeans. *Nat. Commun.* **9**, 1569. <https://doi.org/10.1038/s41467-018-03857-x> (2018).
51. Vāgene, A. J. *et al.* Salmonella enterica genomes from victims of a major sixteenth-century epidemic in Mexico. *Nat. Ecol. Evol.* **2**, 520–528 (2018).
52. Huson, D. H., Auch, A. F., Qi, J. & Schuster, S. C. MEGAN analysis of metagenomic data. *Genome Res.* **17**, 377–386 (2007).
53. Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **25**, 1754–1760 (2009).
54. Peltzer, A. *et al.* EAGER: efficient ancient genome reconstruction. *Genome Biol.* **17**, 60. <https://doi.org/10.1186/s13059-016-0918-z> (2016).
55. Li, H. *et al.* The sequence alignment/map format and SAMtools. *Bioinformatics* **25**, 2078–2079 (2009).
56. McKenna, A. *et al.* The genome analysis toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* **20**, 1297–1303 (2010).
57. Bos, K. I. *et al.* Pre-Columbian mycobacterial genomes reveal seals as a source of New World human tuberculosis. *Nature* **514**, 494–497 (2014).
58. Stamatakis, A. RAXML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* **30**, 1312–1313 (2014).
59. Ronquist, F. *et al.* MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. *Syst. Biol.* **61**, 539–542 (2012).
60. Rambaut, A. FigTree version 1.4.3. <https://tree.bio.ed.ac.uk/software/figtree/>.
61. Shen, W., Le, S., Li, Y. & Hu, F. SeqKit: A cross-platform and ultrafast toolkit for FASTA/Q file manipulation. *PLoS ONE* **11**, e0163962. <https://doi.org/10.1371/journal.pone.0163962> (2016).
62. Lander, E. S. *et al.* Initial sequencing and analysis of the human genome. *Nature* **409**, 860–921 (2001).
63. Jónsson, H., Ginolhac, A., Schubert, M., Johnson, P. L. F. & Orlando, L. MapDamage2.0: fast approximate Bayesian estimates of ancient DNA damage parameters. *Bioinformatics* **29**, 1682–1684 (2013).
64. Briggs, A. W. *et al.* Patterns of damage in genomic DNA sequences from a Neandertal. *Proc. Natl. Acad. Sci. USA* **104**, 14616–14621 (2007).
65. Lazaridis, I. *et al.* Ancient human genomes suggest three ancestral populations for present-day Europeans. *Nature* **513**, 409–413 (2014).
66. Haak, W. *et al.* Massive migration from the steppe was a source for Indo-European languages in Europe. *Nature* **522**, 207–211 (2015).
67. Mathieson, I. *et al.* Genome-wide patterns of selection in 230 ancient Eurasians. *Nature* **528**, 499–503 (2015).
68. Lamnidis, T. C. *et al.* Ancient Fennoscandian genomes reveal origin and spread of Siberian ancestry in Europe. *Nat. Commun.* **9**, 5018. <https://doi.org/10.1038/s41467-018-07483-5> (2018).
69. Patterson, N., Price, A. L. & Reich, D. Population structure and eigenanalysis. *PLoS Genet.* **2**, e190. <https://doi.org/10.1371/journal.pgen.0020190> (2006).
70. Alexander, D. H., Novembre, J. & Lange, K. Fast model-based estimation of ancestry in unrelated individuals. *Genome Res.* **19**, 1655–1664 (2009).

Acknowledgements

J.B. was funded by the International Max Planck Research School for Evolutionary Biology. This study was supported by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) through the following projects: numbers 2901391021 (CRC 1266) and 390870439 (EXC 2150 - ROOTS) to A.N. and B.K.-K.

Author contributions

B.K.-K., G.G. and E.P. conceived and designed the research. J.S., A.I. and B.K.-K. generated and analyzed the ancient DNA data. J.S., J.B., E.P., G.G., A.N. and B.K.-K. interpreted the findings. J.S., J.B., E.P., G.G., A.N. and B.K.-K. wrote the manuscript.

Funding

Open Access funding provided by Projekt DEAL.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at <https://doi.org/10.1038/s41598-020-71530-9>.

Correspondence and requests for materials should be addressed to B.K.-K.

Reprints and permissions information is available at www.nature.com/reprints.

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



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

© The Author(s) 2020

Chapter 5: *Yersinia pestis* responsible for ancient plague outbreaks lacked zonula occludens toxin (Zot) present in strains of the modern pandemic

Currently submitted to **Proceedings of the Royal Society B** Status: major revision (09.05.2022)

The chapter is formatted according to the guidelines of the journal.

Supplementary material can be found here: <https://doi.org/10.57892/100-4>

***Yersinia pestis* responsible for ancient plague outbreaks lacked zonula occludens toxin
(Zot) present in strains of the modern pandemic**

Joanna H. Bonczarowska¹, Julian Susat¹, Ben Krause-Kyora¹, Almut Nebel^{1,†}, Daniel Unterweger^{2,3,†,*}

¹ Institute of Clinical Molecular Biology, Kiel University, Rosalind-Franklin-Straße 12, 24105 Kiel, Germany

² Institute for Experimental Medicine, Kiel University, Michaelisstraße 5, 24105 Kiel, Germany

³ Max Planck Institute for Evolutionary Biology, August-Thienemann-Straße 2, 24306 Plön, Germany

* corresponding author:

Daniel Unterweger

Institute for Experimental Medicine, Kiel University, Michaelisstraße 5, 24105 Kiel, Germany

email: d.unterweger@iem.uni-kiel.de, phone: +49 (0431) 500 30312

† These authors contributed equally to this article.

Abstract

Yersinia pestis is the causative agent of at least three major plague pandemics (Justinianic, Medieval and Modern). Previous studies on ancient *Y. pestis* genomes revealed that the pathogen achieved its remarkable virulence via several genomic alterations, which occurred ~5000 to 3000 years ago. How a subset of strains arose to cause the Modern pandemic is little understood. Here, we examined the prophage (YpfΦ), which is exclusively present in strains associated with the Modern pandemic. One of the proteins encoded by YpfΦ exhibits homology to a known virulence factor – zonula occludens toxin (Zot). YpfΦ was not found in the ancient *Y. pestis* genomes analysed here, indicating its chromosomal insertion approx. 500 years ago. Our findings present an updated view on the distribution of YpfΦ and Zot, which might help to explain the altered host spectrum, wide geographic spread and seemingly increased virulence of *Y. pestis* responsible for the Modern pandemic.

Keywords: *Yersinia pestis*, plague, pandemic, prophage, virulence factor, zonula occludens toxin,

Introduction

Yersinia pestis is the pathogenic agent of plague – a zoonotic disease that can be transmitted from its main host (rodents) to humans via a bite of an infected flea. This route of infection leads to bubonic plague (infection of the lymphatic system). *Y. pestis* can also spread between humans via respiratory droplets, resulting in the infection of the lower airways (pneumonic plague). When the disease is untreated, the bacterium enters the bloodstream causing sepsis (septicaemic plague). Although mortality rates of plague vary depending on the clinical form (30% -100%), *Y. pestis* infection is almost always fatal when no antibiotics are administered in a timely manner. The bacterium is responsible for at least three major pandemics in human history: the Justinianic plague (6th-8th cent. AD), the Medieval plague (14th-18th cent. AD) and the Modern plague (end of 19th – mid 20th cent. AD) [1-3]. Nowadays, *Y. pestis* still persists in environmental reservoirs around the globe [4] remaining a major threat to public health, as evidenced by the recent outbreaks of plague in Africa [5-7] and Asia [8].

Y. pestis evolved from *Yersinia pseudotuberculosis* approximately 7000 years ago [9] and achieved its remarkably high virulence by a number of genomic alterations. These changes include the inactivation of genes and the acquisition of additional virulence genes [10-12]. The impact of those genetic alterations on disease is most drastically demonstrated by their effect on the transmission routes of *Y. pestis* and the affected organs in human infection. Acquiring *Yersinia* murine toxin (*ymt*), together with inactivating several other genes (*pde2*, *pde3*, *ureD*, *rcaA*, *flhD*), enabled efficient transmission of the pathogen via fleas from rodents to humans and resulted in infection of the lymphatic system [12,13]. The acquisition of the plasminogen activator *pla* enabled the bacterium to be transmitted by droplets from humans to humans resulting in airway infections and pneumonic plague [14,15]. Taken together, these findings describe the early stages of *Y. pestis* evolution – dating all the way back to the Neolithic period (approx. 3300 – 1200 BC) [9, 12, 13, 16]. As much as genetic alterations can aid the rise of pandemics, they might contribute to their fall and the extinction of strains. We recently reported depletion of *pla* in medieval post-Black Death strains that have no modern descendants and seem to have gone extinct [17].

Unlike the genetic changes that contributed to the early evolution of the pathogen [18], recent genome alterations affecting the virulence of *Y. pestis* are little understood. Among multiple diverse biovars of modern *Y. pestis*, only strains of the *orientalis* biovar are thought to have caused the Modern plague pandemic [4]. Here, we applied a comparative genomics approach to identify putative virulence factors of modern *Y. pestis* that might have contributed to the rise of the Modern pandemic. Over 300 modern and ancient strains were systematically analysed, representing a timespan of over 5000 years. We report that a virulence-associated filamentous phage (YpfΦ), previously found in modern *orientalis* strains [19-21], was absent in all ancient strains. The phage was likely stably inserted into the chromosome after the Medieval pandemic. Homology of a phage-encoded protein with the widespread virulence factor zonula occludens toxin could suggest differences in pathogenicity of *Y. pestis* and clinical manifestation of plague between the pandemics.

Results

Ancient and medieval *Y. pestis* strains lack a prophage carrying a toxin-encoding gene

To explore genomic differences between modern and medieval *Y. pestis* isolates, sequencing reads of one randomly chosen medieval strain (G488) [17] (approx. 1550-1650 AD) were mapped to the sequence of the modern reference strain CO92 (end of the 20th cent. AD). An 8734bp-long genomic region was found in the modern strain that was not covered by any reads of the medieval strain (Fig. 1A). This region encompasses 13 genes (Tab. S1) and was previously identified as *Y. pestis* phage YpfΦ [21]. Interestingly, one of the YpfΦ genes (*zot*) encodes zonula occludens toxin (Zot), which had recently been detected in *Y. pestis* CO92 strain independently [22]. Zot is a known virulence factor in *Vibrio cholerae* [23].

To test if the lack of YpfΦ is specific to strain G488 or common among medieval and ancient strains, available sequencing reads of 45 medieval (6th – 17th cent. AD, Tab. S2) and nine ancient (Neolithic and Bronze Age, Tab. S3) strains were additionally analysed. None of the ancient and medieval strains carry YpfΦ in their genomes (Fig. 1B). This finding indicates that the absence of the phage seems to be a generalizable pattern among isolates spanning over five thousand years, including the first two plague pandemics.

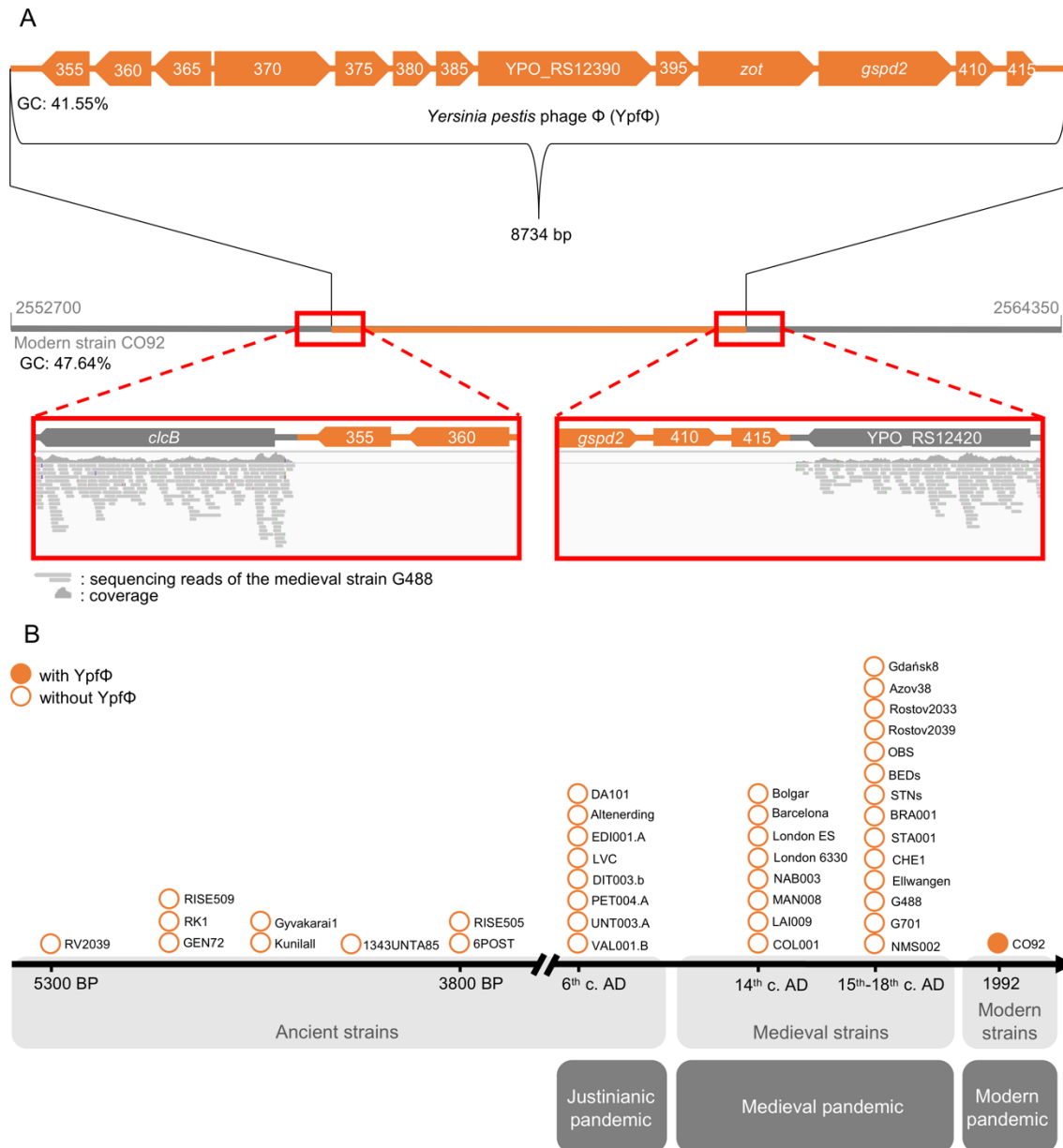


Figure 1. Ypf Φ of the modern strain CO92 is absent from ancient and medieval *Y. pestis* strains.

A - Ypf Φ of modern *Y. pestis* strain CO92. Sequencing reads of the medieval strain G488 (as a representative strain) map to adjacent genes (coloured in grey) but not to the phage (coloured in orange) of the modern reference strain. Individual genes are labelled with their names or locus tags. The positions of the visualized genomic region are indicated. Shown G+C ratio was calculated using the online GC Content Calculator [24].

B – Ancient and medieval strains lack Ypf Φ (empty circles). Individual strains are indicated along a timeline based on their estimated dating. BP before present, c century, AD anno domini.

A subset of modern *Y. pestis* strains stably integrated YpfΦ in their genome

To examine the prevalence of YpfΦ among modern *Y. pestis* strains, 255 published genomes were screened for the phage. The screening revealed YpfΦ in the chromosomes of a subset of strains (N = 46, Tab. S4), which clustered together in one subbranch of branch 1 in the phylogenetic tree (Fig. 2 and S1). All of the isolates with YpfΦ belonged to the *orientalis* biovar that is thought to be responsible for the Modern plague pandemic. All remaining branch 1 strains belong to the *antiqua* biovar and did not carry the phage. Although two previous studies suggested an extrachromosomal presence of YpfΦ in several modern non-*orientalis* strains with a PCR approach [21, 25], this finding could not be confirmed for our larger dataset using whole genome sequencing (WGS) that provides higher resolution. This discrepancy could be explained by differences in the applied methods and the strains examined, as no YpfΦ-positive non-*orientalis* strains from previous studies were included in this analysis.

To further characterise the phage among the YpfΦ-positive strains, the intraspecific diversity of YpfΦ was analysed. Genome sequences of the phage were aligned and compared. Our results revealed the existence of multiple YpfΦ variants (Fig. S2, Tab. S5). Some strains encoded two copies of the phage in a row whereas other strains encoded one copy of the entire phage followed by an incomplete second copy. Moreover, pseudogenes in the YpfΦ of several strains were observed. In strains with two copies of YpfΦ, one copy often contained a pseudogene whereas the other copy was marked as functional. This observation might indicate selection for one fully functional YpfΦ.

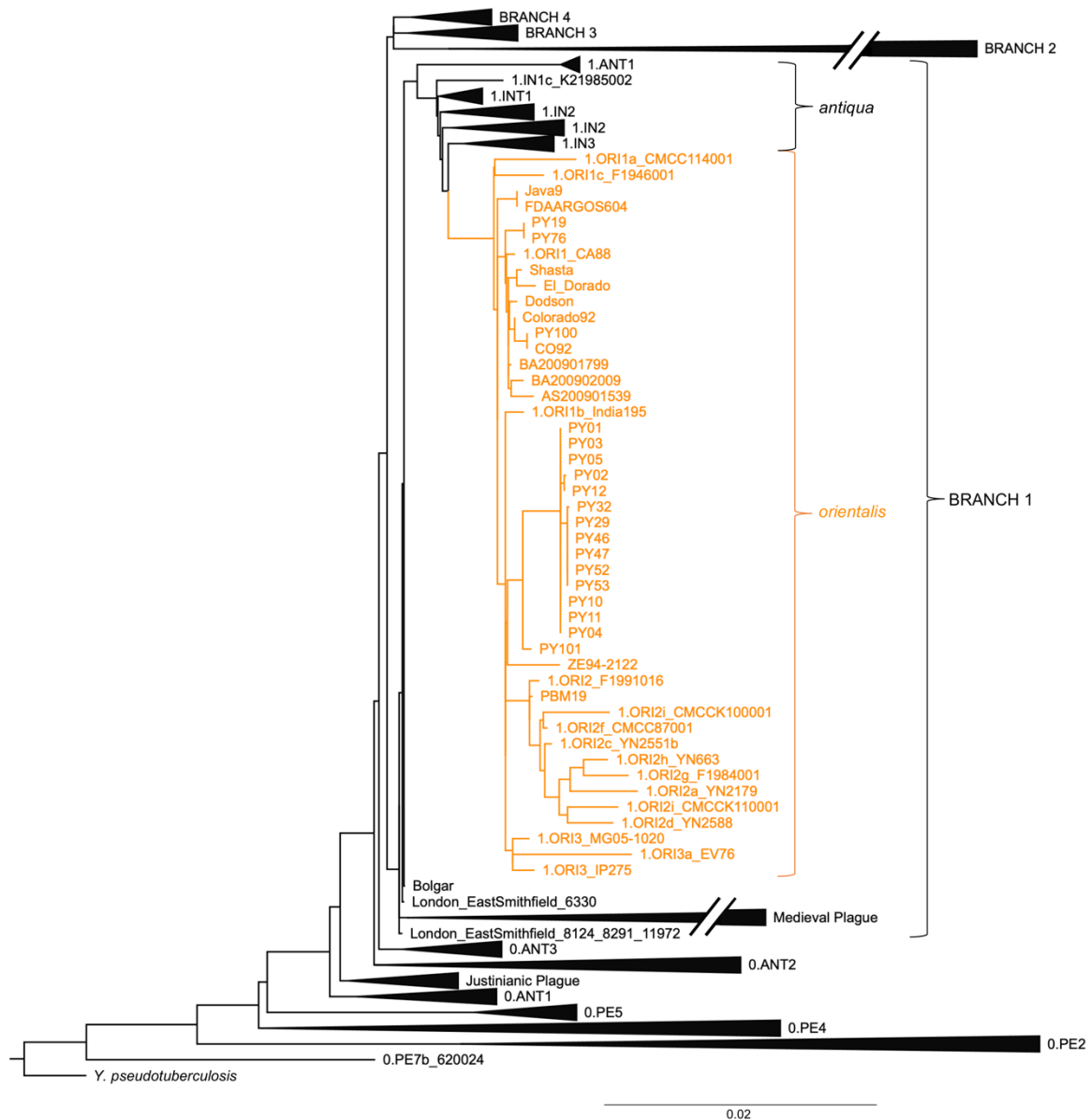


Figure 2. Maximum likelihood phylogeny of 255 modern and 46 medieval *Y. pestis* strains. Strains carrying YpfΦ in their genome are marked in orange and all belong to the *orientalis* biovar. Remaining modern branch 1 strains represent the biovar *antiqua*. The tree is based on the SNP alignment of 13638 positions with a bootstrap of 500 replicates and includes *Y. pseudotuberculosis* as an outgroup. Full tree with uncollapsed branches and bootstrap values can be found in the supplementary material (Fig. S1).

Common ancestry of YpfΦ in *Y. pestis* and other *Enterobacteria*

To determine a possible origin of YpfΦ in *Y. pestis*, the NCBI database was screened for bacterial strains carrying analogous sequences. We found a highly similar genetic region in the *E. coli* strain MOD1-EC6770 (Fig. 3A). Unlike in previous observations [20], sequencing reads and contigs of the here identified *E. coli* isolate aligned to full-length YpfΦ of *Y. pestis* CO92 at high nucleotide identities (Fig. 3A, S3). This finding suggests a common ancestry of the phages in *E. coli* and *Y. pestis*.

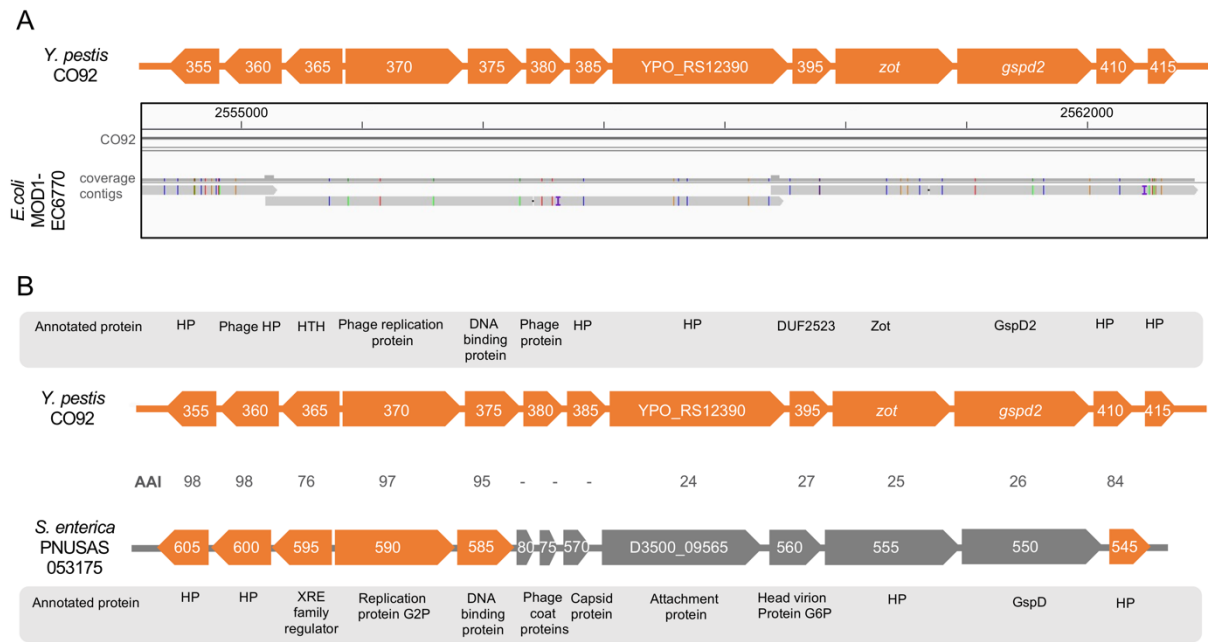


Figure 3. YpfΦ is encoded by other *Enterobacteria*.

A – Alignment of YpfΦ from *E. coli* MOD1-EC6770 (three contigs) and *Y. pestis* CO92 (NC_003143.1). SNPs are marked in colours along the contigs. A graphical depiction of the phage is shown above the alignment in orange.

B – *S. enterica* PNUSAS053175 carries a putative phage with a similar genetic architecture to YpfΦ. A graphical depiction of the respective genes and annotation is shown. Genes in the phage of *S. enterica* that encode proteins exhibiting at least 30% amino acid identity in pairwise alignments with the respective proteins of *Y. pestis* are marked in orange. Amino acid sequence identities (AAI [%]) are indicated. HP stands for hypothetical protein.

Slightly truncated forms of YpfΦ are also found in other *Enterobacteria* such as *Enterobacter ludwigii* and *hormarcheri*, *Morganella morganii*, *Klebsiella aerogenes*, *Salmonella enterica*, *Citrobacter amalonaticus*, *koseri*, and *portucalensis* as well as *Escherichia coli* and *Cedecea davisiae*. Those truncated forms encode 11 genes and lack equivalents to the two open reading frames at the 5' end of *Y. pestis* YpfΦ, namely YPO_RS12355 and YPO_RS12360. In *Citrobacter freundii*, *Enterobacter cloacae* and *Shigella sonnei* strains, we also found multiple copies of the phage. An alignment of the YpfΦs of various species reveals over 98.7% nucleotide identity over the 11 conserved genes (Tab. S6) and suggests high conservation across a wide range of enterobacterial isolates of mostly human origin.

The YpfΦ of *Y. pestis* might belong to a large family of phages with a modular genomic architecture. The YpfΦ-like phage in *S. enterica* PNUSAS053175 encodes common and distinct features to YpfΦ of *Y. pestis* (Fig. 3B). Whereas the first five genes (supposedly involved in phage regulation and replication) share high homology between the two phages, the proteins encoded in the remaining eight predicted open reading frames (mediating phage morphogenesis and secretion) differ strongly and share less than 30% amino acid identity. Interestingly, the *S. enterica* strain does not encode a homologue to *zot* and future work will need to resolve whether the proteins encoded by the YpfΦ-like phages also act as toxins. The potential role of YpfΦ in *Y. pestis* could be better understood by examination of its genomic architecture and other proteins encoded by the phage.

YpfΦ encodes a protein with structural homology to the T2SS secretin

In addition to *Zot*, YpfΦ encodes a gene for a protein that is annotated as a type II secretion system (T2SS) secretin and that we call GspD2 (Fig. 1A, 4A). The chromosome of *Y. pestis* already encodes a protein called GspD in a gene outside of YpfΦ (Fig. 4A). To assess the difference between the two proteins, their size was compared and an alignment of the amino acid sequences of GspD and GspD2 was performed. Whilst GspD is built of 640 amino acids (aa), GspD2 is smaller (414 aa). Only 26% amino acid identity was detected between the two proteins in the alignment, suggesting that they are unrelated.

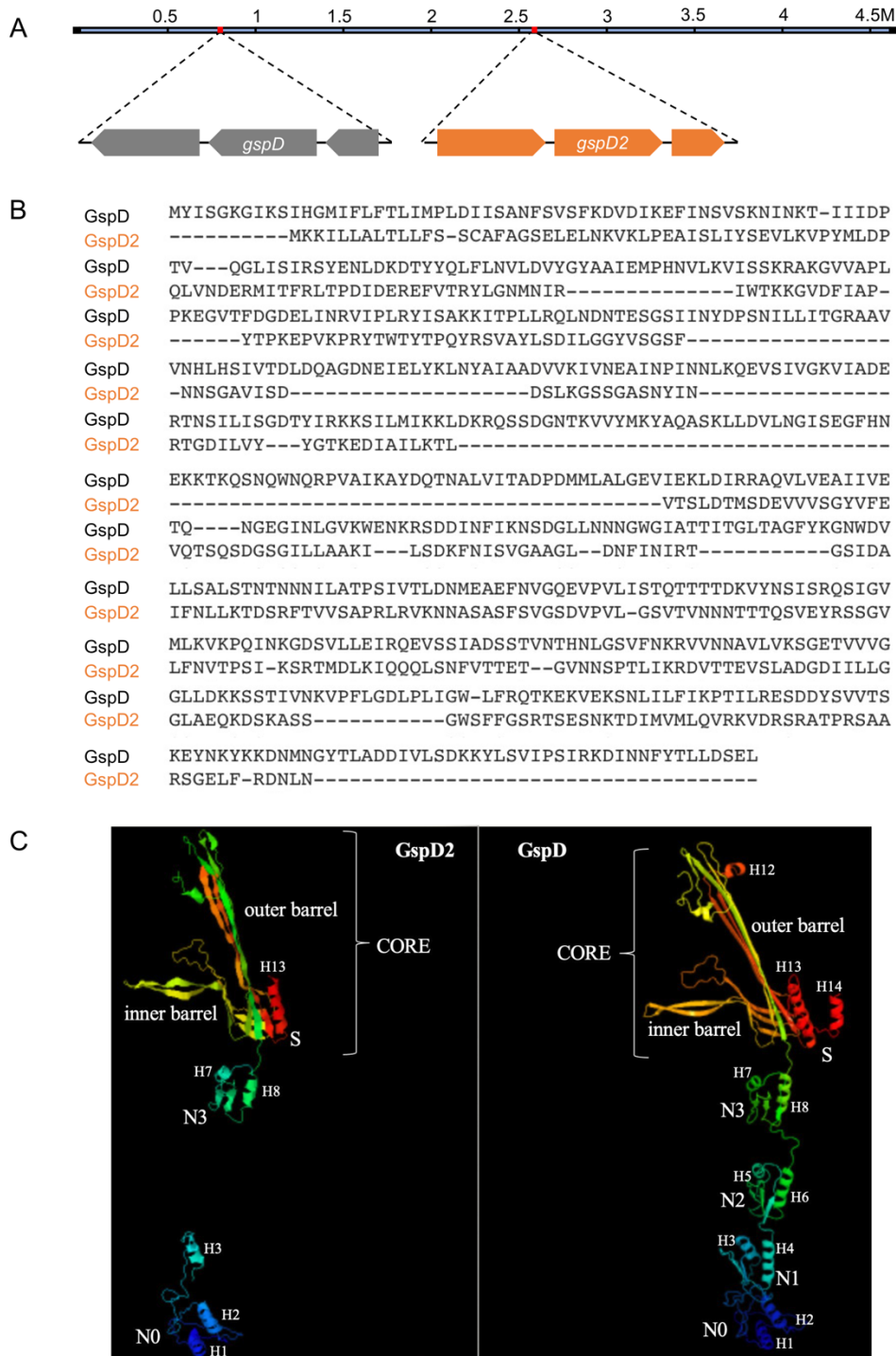


Figure 4. YpfΦ-encoded GspD2 shows structural similarity to GspD

A – *Y. pestis* CO92 carries two *gspD* genes (*gspD* and *gspD2*). *GspD2* is part of YpfΦ (coloured in orange) whilst *gspD* (coloured in grey) is not. The bacterial chromosome is indicated as a horizontal line. B – Amino-acid alignment of GspD and GspD2; C – Predicted 3D protein structures reveal differences in the protein core as well as in the periplasmic domain, which is relevant for substrate specificity. Helices (Hx) and domains (Nx and S) are labelled.

To further test if they share similar 3D structures despite sequence dissimilarities, the structure of the two *Y. pestis* proteins was modelled with 100% confidence (90% and 84% coverage for GspD and GspD2, respectively) onto the existing crystal structure of the T2SS secretin of *Klebsiella pneumoniae* (template c6hcgF) [26]. The 3D models of GspD and GspD2 look more similar than expected based on the amino acid sequences (Fig. 4). Both proteins form a core with an inner and outer barrel as well as the N-terminal extensions of alpha helices. Several minor differences were noted. The N1 and N2 domains, elements of the core domain (S16, H9, H11 and H12) and elements of the S domain (H14 and H15) were absent in GspD2 but present in GspD. Moreover, an additional β -strand between the bends S19 and S20 was noted in GspD2 (Fig. S4). In sum, the high similarity in structure suggests that GspD and GspD2 both function as secretins, although with different substrate specificities as determined by the N-terminal domains in the periplasm. Based on a possible role of Zot (in reference 27 - YPO2279) in phage assembly and release [27], GspD2 might form a complex with Zot to facilitate the secretion of Ypf Φ .

Y. pestis strains with and without Ypf Φ differ in their host spectrum

To better understand the potential impact of Ypf Φ -encoded proteins on bacteria and their behaviour, we determined features that are associated with Ypf Φ -positive modern strains (of the biovar *orientalis*) and are absent from Ypf Φ -negative modern branch 1 strains (of the biovar *antiqua*). Of particular interest is the source of the analysed isolates, as *Y. pestis* is found in multiple animal hosts besides humans. These animal hosts serve as natural reservoirs of this pathogenic bacterium and play a key role during transmission. When comparing the host spectrum between strains with and without the phage, we found Ypf Φ -positive strains (*orientalis*) predominantly among animals associated with human habitats, like rodents, whereas Ypf Φ -negative isolates are found among more diverse animals in the wild (Fig. 5). Furthermore, the proportion of the human host was higher for the Ypf Φ -positive strains relative to the Ypf Φ -negative strains (Fig. 5). Here, it is imperative to mention that the proportion of human isolates among Ypf Φ -positive bacteria that is seen here is likely an underestimate, considering that the analysed isolates were chosen to represent maximal host diversity. Although a possible influence of the sampling bias cannot be excluded, based on available data, we observe an altered host range among isolates with and without the phage. Another feature

of the YpfΦ-positive strains is that they all belong to the *orientalis* biovar which has been linked with the Modern plague pandemic and subsequent disease outbreaks.

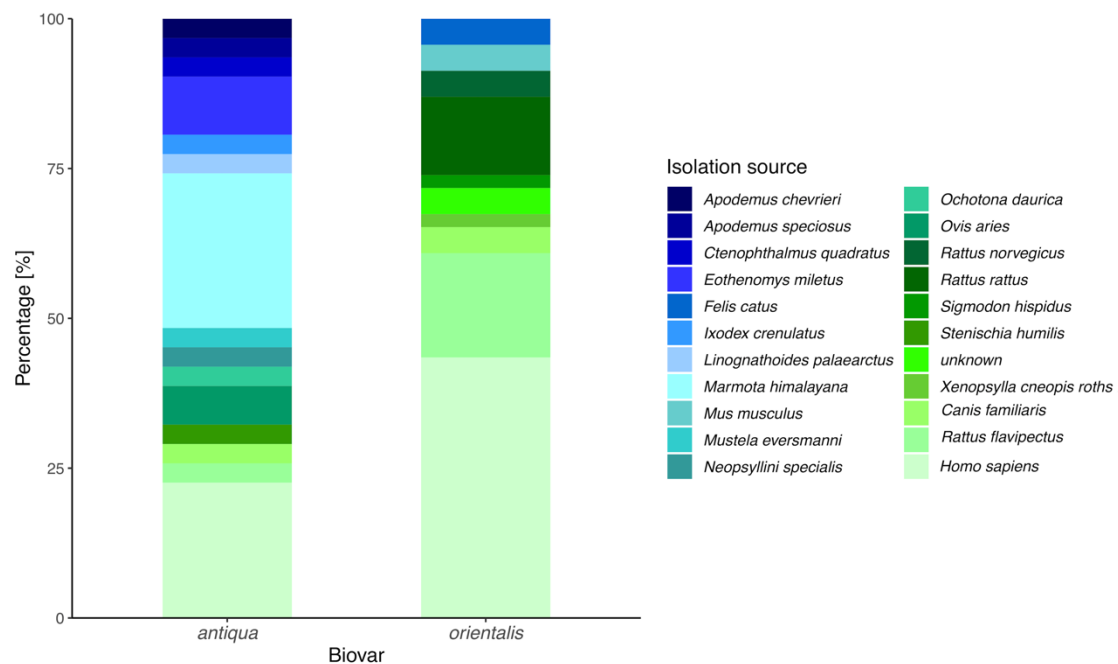


Figure 5. Different host spectrum between YpfΦ-negative strains of the biovar *antiqua* (n=31) and YpfΦ-positive strains of the biovar *orientalis* (n=46). Isolation sources are indicated with respective colours.

Discussion

In this study, we confirm previous indications that YpfΦ was associated with *Y. pestis* responsible for the Modern plague pandemic (*orientalis* biovar) [21] and was absent in strains which caused the Medieval and Justinianic pandemics (Fig. 6A). Our results suggest that the phage-encoded Zot has a dual function: for YpfΦ assembly and as a putative virulence factor that permeabilizes epithelial barriers (Fig. 6B). Together with experimental studies that show an enhanced virulence of YpfΦ-positive strains in mammals [21], our bioinformatic analysis suggests differences in pathogenicity of modern and ancient *Y. pestis*, pointing to possible changes in clinical manifestation and/or course of plague between the pandemics. The presence of YpfΦ and *zot* in modern *Y. pestis* strains of the *orientalis* biovar correlates with their seemingly altered host spectrum, global spread and pandemic potential relative to other branch 1 modern *Y. pestis* (*antiqua*, YpfΦ-negative) (Fig. 6C).

Zot, as we find it here in Ypf Φ , is a homologue of the endogenous human zonulin [28]. Elevated levels of zonulin are linked with functional loss of the intestinal barrier resulting in uncontrolled antigen trafficking that triggers the innate and eventually the adaptive immune responses, leading to chronic inflammation [29]. The role of zonulin in various inflammatory diseases, such as celiac disease (e.g., [30]), diabetes (e.g., [31]), inflammatory bowel disease (IBD) (e.g., [32]) and multiple sclerosis (e.g., [30]) has been widely documented.

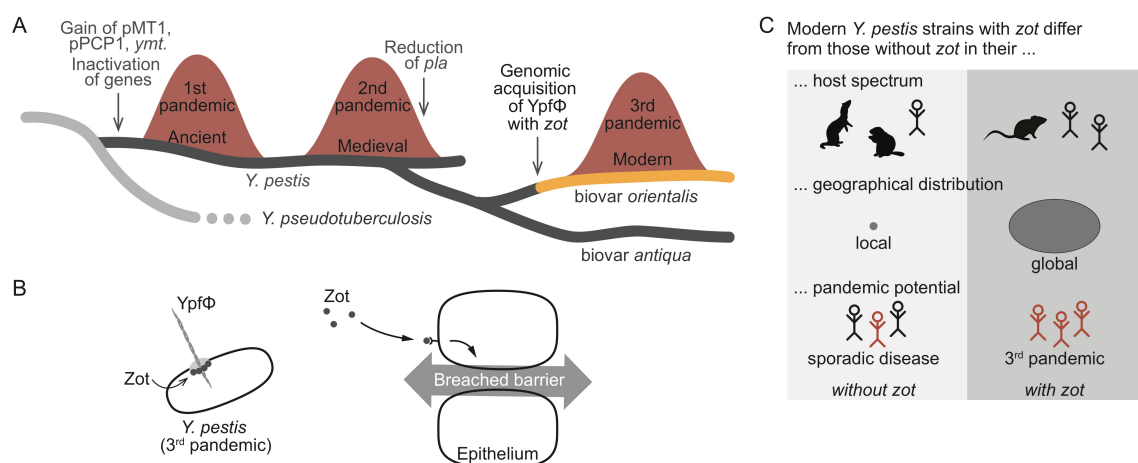


Figure 6. Genomic acquisition of Ypf Φ by modern *Y. pestis* might enable pandemic strains to secrete Zot and permeabilize host membranes.

A – Chromosomal insertion of Ypf Φ in the context of virulence-associated genetic changes in *Y. pestis*' evolution and known plague pandemics. Ypf Φ is present as a prophage among strains of the modern *orientalis* biovar, which is thought to be responsible for the Modern plague pandemic.

B – Dual role of Zot in assembly of the filamentous phage and as virulence factor that permeabilizes epithelial barriers

C – Differences between the analyzed here Ypf Φ -positive and Ypf Φ -negative *Y. pestis* strains with regards to the host spectrum, geographic distribution and pathogenicity. Animal depictions originate from <http://phylopic.org>.

Apart from *Y. pestis*, a wide variety of bacterial pathogens harbour Zot as a virulence factor. Among them are *Neisseria meningitidis*, *Acinetobacter baumannii*, *Salmonella enterica*, *Burkholderia cenocepacia*, and *Vibrio cholerae*, in which it had first been identified [22,23]. Zot was shown to increase permeability of the gut epithelia [23] as well as the mucosa [33] and

the endothelial cells in the brain [34]. In addition to typical bacterial pathogens, Zot has been detected in various Enterobacterial species of the microbiome, some of which were linked to chronic inflammation. For instance, Zot-harboured *Campylobacter concisus* was isolated from saliva samples of inflammatory bowel disease patients [35].

For *Y. pestis*, chromosomal acquisition of YpfΦ and subsequent release of Zot might have provided the pathogen with a new virulence mechanism inducing damage of the host barriers. A previous functional study done on mice showed that YpfΦ allows better colonisation of the host, conferring increased fitness during infection of mammals [21]. YpfΦ could thus affect the pathogenicity of *Y. pestis* and the clinical manifestation of plague, suggesting possible differences between the ancient (Justinianic and Medieval) and Modern pandemics. This highlights the introduction of potential bias when interpreting various historical sources on plague epidemiology in the past based on the inference from knowledge about modern plague. Differences in pathogenicity do not necessarily have to be reflected in the estimated mortality rates of the pandemics, as various factors, such as proximity to zoonotic reservoirs, climate, settlement type/crowding, health status and behaviour-related exposure, can influence the outbreak outcomes [e.g. 36-40]. Based on our current knowledge on *Y. pestis*' pathogenesis and YpfΦ, the phage may be advantageous for the pathogen during the initial phase of infection when the bacterium enters the macrophages and acquires resistance to phagocytosis [18,41,42]. It was shown that not only the free-floating (armed with resistance to phagocytosis) but also cell-borne (still hidden within the host cells) *Y. pestis* are involved in pathogen's migration through the lymphatic system [43]. As examination of the Zot function in *Campylobacter concisus* revealed Zot-enhanced response of macrophages to *E. coli* [44], it is possible that YpfΦ is involved the initial entry of *Y. pestis* into macrophages. This could facilitate the protection against phagocytosis as well as the "silent" spread of the pathogen, resulting in more efficient colonisation of various organs. This hypothesis, however, would have to be further explored by functional studies. The YpfΦ-induced damage to the epithelial layer may also be advantageous for the oral route of infection. Oral administration of an attenuated *Y. pestis* strain EV76 (YpfΦ-positive) to a group of vervet monkeys resulted in gastrointestinal disturbance in several of the animals [45], as it would be expected from Zot activity. Instances of gastroenteritis and pharyngeal plague in humans, following consumption of *Y. pestis*-contaminated meat, were also reported [46]. Unfortunately, it is not known if the causative strains were YpfΦ-positive. Nevertheless, YpfΦ may affect *Y. pestis* pathogenicity [21] and

has the potential to reflect a recent evolutionary change in the species that could be comparable, however likely to a lesser extent, to the previously known High Pathogenicity Island (HPI) – characteristic for highly pathogenic *Yersinia* [47].

Our analysis demonstrated that YpfΦ was absent in all strains responsible for previous pandemics and other ancient strains. A lack of coverage of YpfΦ genes in medieval strains can also be seen in the data published by Seguin-Orlando and colleagues (Figure 4A in Seguin-Orlando et al., 2021 [48]). Published molecular dating of the splits in the phylogeny of *Y. pestis* suggests that stabilisation of the phage in the bacterium's chromosome occurred relatively recently, approx. 500 years ago [9]. Presence of the complete phage in *E. coli* may suggest this species as a possible source of the YpfΦ in *Y. pestis*. Based on the spread of YpfΦ among various *Enterobacteria* (Tab. S6) [20,22], it is possible that the human host represents an environment, that is shared by various bacteria, in which the acquisition occurred.

A better knowledge on the recent evolution of *Y. pestis*' virulence is key to the understanding of the bacterium's pathogenicity, as plague begins to occur in unusual and thus challenging to diagnose clinical forms. For instance, during the plague-related outbreak of gastroenteritis in Afghanistan, 20.5 % of the infected individuals (17/83) died before the appropriate treatment was applied [46]. Although neural involvement in plague is very rare, cases of *Y. pestis*-induced meningitis were previously described [49]. Identifying YpfΦ-encoded Zot as a potential virulence factor that breaches host epithelial barriers should make us rethink these clinical cases and work on developing new therapeutic targets for the future.

Material

The data analysed here comprised of 255 previously published modern and 55 ancient (46 medieval, 8 Bronze Age and 1 Neolithic) *Y. pestis* strains downloaded from online repositories such as the European Nucleotide Archive (ENA) and the National Center for Biotechnology Information (NCBI) (Tab. S1 and S2).

Methods

Alignment and detection of the YpfΦ

Preprocessed sequencing data from all *Y. pestis* strains (Tab. S1 and S2) were mapped against the CO92 reference genome (NC_003143.1, NC003131.1, NC_003132.1, NC_003134.1,) with Burrows-Wheeler Aligner (BWA) v0.7.12 (n = 0.01, l = 300) [50]. Regions with zero coverage were identified and the samples with no reads mapping to the NC_003143.1:2554178-2562912 region were classified as strains without the YpfΦ. Moreover, sequencing data of a randomly chosen example strain CMCC10012 (not carrying the phage) was mapped in a competitive mapping against the CO92 reference genome and the CO92 chromosome without the 2554178-2562912 region. This way, unique gap-bridging reads were identified for the CMCC10012 (Fig. S5), confirming the lack of YpfΦ.

Analysis of YpfΦ

Reads mapping to the CO92 YpfΦ region were extracted from a randomly chosen example strain (EV76). The phage reads were then mapped against previously published genome assemblies and contigs of the YpfΦ-positive *Y. pestis* strains (Tab. S3) with Burrows-Wheeler Aligner (BWA) v0.7.12 (n = 0.01, l = 300) [45] to locate the position of YpfΦ within the chromosome of each strain. Knowing the phage locus for the strains, genome annotation was inspected in the graphic view panel of the sequences in NCBI to identify the YpfΦ variants. Association with a particular biovar was made based on available data [51] or the position in the phylogeny (Fig. 2). To identify similar sequences in other bacterial species, the extracted phage reads were blasted with the BLASTn online tool (default parameters) against the Nucleotide collection (nt).

Comparative analysis of *Y. pestis* secretins

Three-dimensional (3D) structure of CO92 *Y. pestis* secretins (GspD and GspD2) was predicted with the Protein Homology/analogY Recognition Engine V 2.0 (Phyre2) [52]. BLASTn and BLASTp online tools (default parameters) were used to compare nucleotide and amino-acid sequences of the two secretins in *Y. pestis*.

Phylogenetic analysis

Phylogenetic analysis was performed with RAxML [53] using the GTRGAMMA model with 500 bootstrap replicates. MultiVCFAnalyzer [54] was used to generate a SNP-based multiple alignment of 255 previously published modern and 46 medieval *Y. pestis* strains with *Y. pseudotuberculosis* (NZ_CP008943.1) as an outgroup (Tab. S1). The input VCF files were generated with the UnifiedGenotyper module from the Genome Analysis Toolkit (GATK) v3.6 [55]. A SNP was called if the position was covered by at least 3 reads, the genotype quality was at least 30 and the fraction of mapped reads containing the SNP was at least 90 %.

Acknowledgements

We thank the Kiel Evolution Center for facilitating this collaborative project. J.H.B. was funded by the International Max Planck Research School for Evolutionary Biology and the Precision Medicine in Chronic Inflammation (PMI) Cluster (excellence strategy – EXC 2167 – 390884018). Work in the Unterweger Lab is supported by the German Federal Ministry for Education and Research (grant 01KI2020).

References

1. Haensch S, Bianucci R, Signoli M, Rajerison M, Schultz M, Kacki S, et al. Distinct clones of *Yersinia pestis* caused the black death. *PLoS Pathog.* 2010;6(10): e1001134. 10.1371/journal.ppat.1001134
2. Harbeck M, Seifert L, Hensch S, Wagner DM, Birdsell D, Parise KL, et al. *Yersinia pestis* DNA from Skeletal Remains from the 6th Century AD Reveals Insights into Justinianic Plague. *PLoS Pathog.* 2013;9(5): e1003349. 10.1371/journal.ppat.1003349
3. Pryor E. The great plague of Hong Kong. *J Hong Kong Branch R Asiat Soc.* 1975;15: 61–70.
4. Vogler AJ, Chan F, Nottingham R, Andersen G, Drees K, Beckstrom-Sternberg SM, et al. A Decade of Plague in Mahajanga, Madagascar: Insights into the Global Maritime Spread of Pandemic Plague. *mBio.* 2013;4(1): e00623-12. 10.1128/mBio.00623-12
5. Abedi AA, Shako J-C, Gaudart J, Sudre B, Ilunga BK, Shamamba SKB, et al. Ecologic Features of Plague Outbreak Areas, Democratic Republic of the Congo, 2004–2014. *Emerg Infect Dis.* 2018;24(2): 210-220. 10.3201/eid2402.160122
6. Randremanana R, Andrianaiwoarimanana V, Nikolay B, Ramasindrazana B, Paireau J, ten Bosch QA, et al. Epidemiological characteristics of an urban plague epidemic in Madagascar, August–November, 2017: an outbreak report. *Lancet Infect Dis.* 2019;19(5): 537-545. 10.1016/S1473-3099(18)30730-8
7. Respicio-Kingry LB, Yockey BM, Acayo S, Kaggwa J, Apangu T, Kugeler KJ, et al. Two Distinct *Yersinia pestis* Populations Causing Plague among Humans in the West Nile Region of Uganda. *PLoS Negl Trop Dis.* 2016;10(2): e0004360. 10.1371/journal.pntd.0004360
8. Shi L, Yang G, Zhang Z, Xia L, Liang Y, Tan H, et al. Reemergence of human plague in Yunnan, China in 2016. *PLoS One.* 2018;13(6): e0198067. 10.1371/journal.pone.0198067
9. Susat J, Lübke H, Immel A, Brinker U, Macāne A, Meadows J, et al. A 5,000-year-old hunter-gatherer already plagued by *Yersinia pestis*. *Cell Rep.* 2021;35(13): 109278. 10.1016/j.celrep.2021.109278
10. Rasmussen S, Allentoft M, Nielsen K, Orlando L, Sikora M, Sjögren K, et al. Early divergent strains of *Yersinia pestis* in Eurasia 5,000 years ago. *Cell.* 2015;163(3): 571–82. 10.1016/j.cell.2015.10.009
11. Spyrou MA, Keller M, Tikhbatova RI, Scheib CL, Nelson EA, Andrades Valtueña A, et al. Phylogeography of the second plague pandemic revealed through analysis of historical *Yersinia pestis* genomes. *Nat Commun.* 2019;10: 4470. 10.1038/s41467-019-12154-0
12. Sun Y-C, Jarrett CO, Bosio CF, Hinnebusch BJ. Retracing the Evolutionary Path that Led to Flea-Borne Transmission of *Yersinia pestis*. *Cell Host Microbe.* 2014;15(5): 578-586. 10.1016/j.chom.2014.04.003
13. Chouikha I, Hinnebusch BJ. Silencing urease: A key evolutionary step that facilitated the adaptation of *Yersinia pestis* to the flea-borne transmission route. *Proc Natl Acad Sci USA.* 2014;111(52): 18709-18714. 10.1073/pnas.1413209111
14. Lathem WW, Schroeder JA, Bellows LE, Ritzert JT, Koo JT, Price PA, et al. Posttranscriptional Regulation of the *Yersinia pestis* Cyclic AMP Receptor Protein Crp and Impact on Virulence. *mBio.* 2014;5(1): e01038-13. 10.1128/mBio.01038-13
15. Perry RD, Fetherston JD. *Yersinia pestis*—etiologic agent of plague. *Clin Microbiol Rev.* 1997;10(1): 35-66. 10.1128/CMR.10.1.35

16. Zimble DL, Schroeder JA, Eddy JL, Lathem WW. Early emergence of *Yersinia pestis* as a severe respiratory pathogen. *Nat Commun.* 2015 Nov;6(1): 7487. 10.1038/ncomms8487
17. Susat J, Bonczarowska JH, Pētersone-Gordina E, Immel A, Nebel A, Gerhards G, et al. *Yersinia pestis* strains from Latvia show depletion of the *pla* virulence gene at the end of the second plague pandemic. *Sci Rep.* 2020;10: 14628. 10.1038/s41598-020-71530-9
18. Demeure CE, Dussurget O, Mas Fiol G, le Guern A-S, Savin C, Pizarro-Cerdá J. *Yersinia pestis* and plague: an updated view on evolution, virulence determinants, immune subversion, vaccination, and diagnostics. *Genes Immun.* 2019;20(5): 357-370. 10.1038/s41435-019-0065-0
19. Gonzalez MD, Lichtensteiger CA, Vimr ER. Adaptation of signature-tagged mutagenesis to *Escherichia coli* K1 and the infant-rat model of invasive disease. *FEMS Microbiol Lett.* 2001;198(2): 125-128. 10.1111/j.1574-6968.2001.tb10630.x
20. Gonzalez MD, Lichtensteiger CA, Caughlan R, Vimr ER. Conserved Filamentous Prophage in *Escherichia coli* O18:K1:H7 and *Yersinia pestis* Biovar orientalis. *J Bacteriol.* 2002;184(21): 6050-6055. 10.1128/JB.184.21.6050-6055.2002
21. Derbise A, Chenal-Francisque V, Pouillot F, Fayolle C, Prévost M, Médigue C, et al. A horizontally acquired filamentous phage contributes to the pathogenicity of the plague bacillus. *Mol Microbiol.* 2007;63(4): 1145-1157. 10.1111/j.1365-2958.2006.05570.x
22. Yap KP, Gan HM, Teh CSJ, Chai LC, Thong KL. Comparative genomics of closely related serovar Typhi strains reveals genome dynamics and the acquisition of novel pathogenic elements. *BMC Genomics.* 2014;15: 1007. 10.1186/1471-2164-15-1007
23. Fasano A, Baudry B, Pumplun DW, Wasserman SS, Tall BD, Ketley JM, et al. *Vibrio cholerae* produces a second enterotoxin, which affects intestinal tight junctions. *Proc Natl Acad Sci USA.* 1991;88(12): 5242-5246. 10.1073/pnas.88.12.5242
24. Guerra ÁP, Calvo EP, Wasserman M, Chaparro-Olaya J. Production of recombinant proteins from *Plasmodium falciparum* in *Escherichia coli*. *Biomédica.* 2016;36:97-108. 10.7705/biomedica.v36i3.3011
25. Li Y, Dai E, Cui Y, Li M, Zhang Y, Wu M, et al. Different Region Analysis for Genotyping *Yersinia pestis* Isolates from China. *PLoS One.* 2008;3(5): e2166. 10.1371/journal.pone.0002166
26. Chernyatina AA, Low HH. Core architecture of a bacterial type II secretion system. *Nat Commun.* 2019;10(1): 5437. 10.1038/s41467-019-13301-3
27. Chouikha I, Charrier L, Filali S, Derbise A, Carniel E. Insights into the infective properties of YpfΦ, the *Yersinia pestis* filamentous phage. *Virology.* 2010;407(1): 43–52. 10.1016/j.virol.2010.07.048
28. Fasano A, Not T, Wang W, Uzzau S, Berti I, Tommasini A, et al. Zonulin, a newly discovered modulator of intestinal permeability, and its expression in coeliac disease. *Lancet.* 2000;355(9214): 1518-1519. 10.1016/S0140-6736(00)02169-3
29. Sturgeon C, Fasano A. Zonulin, a regulator of epithelial and endothelial barrier functions, and its involvement in chronic inflammatory diseases. *Tissue Barriers.* 2016;4(4): e1251384. 10.1080/21688370.2016.1251384
30. Fasano A. Zonulin and Its Regulation of Intestinal Barrier Function: The Biological Door to Inflammation, Autoimmunity, and Cancer. *Physiol Rev.* 2011;91(1): 151-175. 10.1152/physrev.00003.2008
31. Leffler DA, Kelly CP, Green PHR, Fedorak RN, DiMarino A, Perrow W, et al. Larazotide Acetate for Persistent Symptoms of Celiac Disease Despite a Gluten-Free

- Diet: A Randomized Controlled Trial. *Gastroenterology*. 2015;148(7): 1311-9.e6. 10.1053/j.gastro.2015.02.008
32. Arrieta MC, Madsen K, Doyle J, Meddings J. Reducing small intestinal permeability attenuates colitis in the IL10 gene-deficient mouse. *Gut*. 2009;58(1): 41-48. 10.1136/gut.2008.150888
 33. Gopalakrishnan S, Pandey N, Tamiz A, Vere J, Carrasco R, Somerville R, et al. Mechanism of action of ZOT-derived peptide AT-1002, a tight junction regulator and absorption enhancer. *Int J Pharm*. 2009;365(1-2): 121-130. 10.1016/j.ijpharm.2008.08.047
 34. Karyekar CS, Fasano A, Raje S, Lu R, Dowling TC, Eddington ND. Zonula Occludens Toxin Increases the Permeability of Molecular Weight Markers and Chemotherapeutic Agents Across the Bovine Brain Microvessel Endothelial Cells. *J Pharm Sci*. 2003;92(2): 414-423. 10.1002/jps.10310
 35. Mahendran V, Tan YS, Riordan SM, Grimm MC, Day AS, Lemberg DA, et al. The Prevalence and Polymorphisms of Zonula Occluden Toxin Gene in Multiple *Campylobacter concisus* Strains Isolated from Saliva of Patients with Inflammatory Bowel Disease and Controls. *PLoS One*. 2013;8(9): e75525. 10.1371/journal.pone.0075525
 36. Slavin P. Out of the West: formation of a permanent plague reservoir in South-Central Germany (1349–1356) and its implications. *Past Present*. 2021;252: 3–51. 10.1093/pastj/gtaa028
 37. Stenseth NC, Samia NI, Viljugrein H, Kausrud KL, Begon M, Davis S, et al. Plague dynamics are driven by climate variation. *Proc Natl Acad Sci USA*. 2006;103(35): 13110-13115. 10.1073/pnas.0602447103
 38. Izdebski A, Guzowski P, Poniat R, Masci L, Palli J, Vignola C, et al. Palaeoecological data indicates land-use changes across Europe linked to spatial heterogeneity in mortality during the Black Death pandemic. *Nat Ecol Evol*. 2022;6: 297–306. 0.1038/s41559-021-01652-4
 39. DeWitte SN. Sex differentials in frailty in medieval England. *Am J Phys Anthropol*. 2010;143(2): 285–297. 10.1002/ajpa.21316
 40. Cleri DJ, Vernaleo JR, Lombardi LJ, Rabbat MS, Mathew A, Marton R, et al. Plague pneumonia disease caused by *Yersinia pestis*. *Semin Respir Infect*. 1997;12(1): 12–23.
 41. Ke Y, Chen Z, Yang R. *Yersinia pestis*: mechanisms of entry into and resistance to the host cell. *Front Cell Infect Microbiol*. 2013;3: 106. 10.3389/fcimb.2013.00106
 42. Pujol C, Klein KA, Romanov GA, Palmer LE, Ciota C, Zhao Z, et al. *Yersinia pestis* Can Reside in Autophagosomes and Avoid Xenophagy in Murine Macrophages by Preventing Vacuole Acidification. *Infect Immun*. 2009;77(6): 2251-2261. 10.1128/IAI.00068-09
 43. St. John AL, Ang WYG, Huang M-N, Kunder CA, Chan EW, Gunn MD, et al. S1P-Dependent Trafficking of Intracellular *Yersinia pestis* through Lymph Nodes Establishes Buboes and Systemic Infection. *Immunity*. 2014;41(3): 440-450. 10.1016/j.immuni.2014.07.013
 44. Mahendran V, Liu F, Riordan SM, Grimm MC, Tanaka MM, Zhang L. Examination of the effects of *Campylobacter concisus* zonula occludens toxin on intestinal epithelial cells and macrophages. *Gut Pathog*. 2016;8(1): 18. 10.1186/s13099-016-0101-9
 45. Chen TH, Elberg SS, Eisler DM. Immunity in plague: protection of the vervet (*Cercopithecus aethiops*) against pneumonic plague by the oral administration of live attenuated *Yersinia pestis*. *J Infect Dis*. 1977;135(2): 289-293. 10.1093/infdis/135.2.289

46. Leslie T, Whitehouse CA, Yingst S, Baldwin C, Kakar F, Mofleh J, et al. Outbreak of gastroenteritis caused by *Yersinia pestis* in Afghanistan. *Epidemiol Infect.* 2011;139(5): 728-735. 10.1017/S0950268810001792
47. Carniel E. Plasmids and Pathogenicity Islands of *Yersinia*. In: Hacker J, Kaper JB, editors. *Pathogenicity Islands and the Evolution of Pathogenic Microbes. Current Topics in Microbiology and Immunology*, vol 264/2. Berlin: Springer; 2002. p. 89-108. 10.1007/978-3-642-56031-6_6
48. Seguin-Orlando A, Costedoat C, der Sarkissian C, Tzortzis S, Kamel C, Telmon N, et al. No particular genomic features underpin the dramatic economic consequences of 17th century plague epidemics in Italy. *iScience.* 2021;24(4): 102383. 10.1016/j.isci.2021.102383
49. Becker TM, Poland JD, Quan TJ, White ME, Mann JM, Barnes AM. Becker, T M et al. Plague meningitis--a retrospective analysis of cases reported in the United States, 1970-1979. *West J Med.* 1987;147(5): 554-7.
50. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics.* 2009;25(14): 1754-1760. 10.1093/bioinformatics/btp324
51. Cui Y, Yu C, Yan Y, Li D, Li Y, Jombart T, et al. Historical variations in mutation rate in an epidemic pathogen, *Yersinia pestis*. *Proc Natl Acad Sci USA.* 2013;110(2): 577-582. 10.1073/pnas.1205750110
52. Kelley LA, Mezulis S, Yates CM, Wass MN, Sternberg MJE. The Phyre2 web portal for protein modeling, prediction and analysis. *Nat Protoc.* 2015;10(6): 845-58. 10.1038/nprot.2015.053
53. Stamatakis A. RAxML version 8: A tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics.* 2014;30(9): 1312-1313. 10.1093/bioinformatics/btu033
54. Bos KI, Harkins KM, Herbig A, Coscolla M, Weber N, Comas I, et al. Pre-Columbian mycobacterial genomes reveal seals as a source of New World human tuberculosis. *Nature.* 2014;514(7523): 494-497. 10.1038/nature13591
55. McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytsky A, et al. The genome analysis toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* 2010;20(9): 1297-1303. 10.1101/gr.107524.110

Chapter 6: Genomic and phylogenetic analysis of medieval *Yersinia pestis*

The chapter was prepared as a hybrid of a manuscript and a dissertation chapter.

Supplementary material can be found here: <https://doi.org/10.57892/100-4>

Genomic and phylogenetic analysis of medieval *Yersinia pestis*

Abstract

Yersinia pestis (*Y. pestis*) is a highly virulent pathogen whose beginnings can be traced back to approximately 7000 years ago. Since then, the bacterium has evolved its remarkable ability to utilize various transmission modes and complex infectious mechanisms. Several previous studies on ancient and modern *Y. pestis* genomes reported on the important genetic changes that shaped the pathogen's virulence and highlighted the value of ancient genomes in the study of pathogenicity and evolution. Pathogen screening of 526 samples from 392 medieval individuals revealed presence of *Y. pestis* DNA in 17 individuals. This finding presented an opportunity to explore the bacterium's diversity and identify recent genomic alterations associated with virulence. Two *Y. pestis* genomes were successfully reconstructed and analyzed, confirming previous findings presented in chapters 4 (pla depletion) and 5 (YpfΦ phage). Furthermore, for one strain, deletions of genes potentially involved in virulence was observed. This may possibly reflect local microevolutionary changes that could have contributed to the disappearance of plague from Europe in the late Middle Ages. The frequency of *Y. pestis* infection was estimated to be 4.3-7.2%, which suggests a high prevalence of plague across medieval parish cemeteries. In the future, not only plague-associated mass burials but also individual graves can aid the study of *Y. pestis* diversity, evolution and plague epidemiology.

Introduction

Yersinia pestis (*Y. pestis*) is the causative agent of plague – a deadly disease that is known to have decimated human populations in three pandemics (Justinianic plague (6th-8th cent. AD), the Medieval plague (14th-18th cent. AD) and the Modern plague (19th-20th cent. AD). The remarkably high virulence of *Y. pestis* stems from several genomic alterations that occurred at the early stages of the pathogen's evolution – the earliest of them dating all the way back to the Neolithic/Bronze Age period (approx. 3300 – 1200 BC) (Demeure et al., 2019; Rascovan et al., 2019; Susat et al., 2021). These changes enabled the pathogen to utilize various transmission modes (vector borne, respiratory and oral infection routes) and pathogenesis

(bubonic, pneumonic and septicemic forms). Although *Y. pestis* continues to pose a major threat to public health, it is unclear whether more recent genome alterations also significantly affected the virulence of *Y. pestis*. The findings presented in previous chapters (4 and 5) of this dissertation highlight the importance of medieval genomes in the research of *Y. pestis* pathogenicity and evolution.

In the study described in this chapter, metagenomic samples of 392 individuals buried at 17 medieval cemeteries (~6th to 18th cent. AD) across Denmark and Germany were screened for the presence of *Y. pestis* DNA. Molecular evidence of plague was detected in samples from 17 individuals. Genomic and phylogenetic analysis of the reconstructed genomes were performed to 1) explore the genetic diversity of medieval *Y. pestis* and to 2) possibly identify recent genomic changes related to virulence. This analysis will also serve as a validation of the findings from chapters 4 and 5.

Material

The material comprised 526 samples (215 teeth, 289 petrous bones and 22 postcranial elements) belonging to 392 individuals. The skeletal remains were excavated from 17 cemeteries (Denmark and Germany) dating to approx. 6th – 18th cent. AD (Baldsen, 2018; Prescher & Wagner, 2016; Siegfried, 2018). A detailed description of the archaeological sites and the samples used for pathogen screening can be found in the material section and the supplement of chapter 2.

For the analysis of the *Y. pestis* isolates, additional data from a total of 228 previously published modern (Tab. S1) and 48 ancient *Y. pestis* strains was used (Tab. S2).

Methods

The methods used for the DNA extraction, library preparation and shotgun sequencing as well as the genomic and phylogenetic analyses of the detected *Y. pestis* strains were described in the methods section of the previous chapters (3 - wet lab; 4 and 5 - bioinformatics).

Y. pestis-specific reads were obtained in competitive mapping against *Y. pestis* (NC_003143.1, NC_003131.1, NC_003134.1, NC_003132.1) and *Yersinia pseudotuberculosis* (NC_006155.1) reference genomes, using Burrows-Wheeler Aligner (BWA) v0.7.12 (n = 0.01, l = 300) (Li & Durbin, 2009). Output BAM files were then filtered for quality 30 with SAMtools (Li et al., 2009) (*samtools view -bq 30 input.bam > output.bam*). Subsequently, the number of *Y. pestis*-specific reads was noted with *samtools idxstats*.

To determine the genetic sex of the individuals, the metagenomic reads were mapped to the human genome (build hg19) and the ratio of sequences aligning to the X and Y chromosomes compared to the autosomes was assessed (Fu et al., 2016).

Results

Evidence of plague was detected in four parish cemeteries from medieval Denmark

Pathogen screening revealed the presence of *Y. pestis* DNA in 17 metagenomes from four sites (St. Trinitatis in Viborg, Ole Worm Gade in Horsens, Sejet and Ribe). The “positive” status of each sample was confirmed based on the presence of unique *Y. pestis*-specific sequences, including reads aligning to the pPCP1 plasmid (plasmid that is exclusively present in *Y. pestis*) (Tab. 1). For three samples (A5465/X3003, G52 and G27), the number of *Y. pestis* reads was sufficient for further analysis (> 50% of the genome covered by at least one read) (Tab. 2).

For three out of 17 *Y. pestis*-positive skeletal remains (G89, A5465/X3003 and A16/X3013), only a foot bone or a tooth sample were screened for pathogen DNA. In the case of the remaining 14 skeletons, metagenomes isolated from both teeth and petrous bones were analyzed. For these individuals, evidence of *Y. pestis* infection was detected only in DNA material extracted from teeth and no trace of the pathogen was found in the petrous bone samples.

Table 1. List of *Y. pestis*-positive samples.

Site	Grave N	Sample type	Genetic sex	N of <i>Y. pestis</i> -specific reads			
				chr	pPCP1	pMT1	pCD1
St. Trinitatis (Viborg)	G52	tooth	F	33516	9451	14924	17006
	G89	foot bone	M	53	4	9	0
Odekirkedgard (Sejet)	A5465/X3003	tooth	M	607783	36459	71360	131045
	A569/X569	tooth	M	140	20	36	48
	A146/X146	tooth	M	95	8	31	30
	A106/X106	tooth	M	159	26	91	114
	A146/X3011	tooth	M	500	100	272	148
	A196/X196	tooth	M	77	7	34	18
	A16/X3013	tooth	M	60	10	30	18
	A407/X407	tooth	F	19	6	12	2
Ole Worm Gade (Horsens)	A1063/X1063.1	tooth	F	52	6	6	0
	A1513/X1513.1	tooth	M	101	22	28	26
	A1643/X1643.1	tooth	F	327	29	209	148
	A1513/X1513.2	tooth	M	174	42	88	69
	A1248/X1248.1	tooth	M	38	2	14	6
	A1209/X1209.1	tooth	M	155	19	75	74
Lindegard (Ribe)	G27	tooth	M	4931	3016	984	1733

chr – chromosome; F – female; M – male; N – number;

Genomic and phylogenetic analyses of *Y. pestis* isolates

To examine the diversity of the three strains (A5465/X3003, G52 and G27) with sufficient coverage, genomic analysis was performed. This included evaluation of the previously reported *pla* depletion in the pPCP1 plasmid (chapter 4 – Susat et al., 2020) and presence of the *Y. pestis* prophage (YpfΦ) (chapter 5; Derbise et al., 2007). Additionally, coverage gaps in the alignment between the medieval strains and the modern CO92 reference genome were explored.

Table 2. Mapping statistics for three *Y. pestis* isolates from medieval Denmark (Viborg and Sejet) with a quality filter set to 30.

Strain / Site	Genomic region	N of aligned reads	Mean coverage	Coverage $\geq 1x$ [%]	Coverage $\geq 2x$ [%]	Coverage $\geq 3x$ [%]
St. Trinitatis (Viborg) G52	chromosome	516891	7.0	94.8	93.1	89.6
	pCD1	20202	18.6	94.0	93.3	92.9
	pPCP1	10707	72.7	80.9	80.8	80.8
	pMT1	17892	11.6	94.0	94.0	93.3
Odekirkedgard (Sejet) A5465/X3003	chromosome	1215553	17.42	95.2	94.9	94.7
	pCD1	53758	52.04	95.3	94.9	94.4
	pPCP1	14843	103.82	81.3	80.9	80.9
	pMT1	24813	16.79	94.2	94.1	94.1
Lindegarden (Ribe) G27	chromosome	68304	0.7	49.1	17.2	4.8
	pCD1	3062	2.2	80.2	59.6	39.0
	pPCP1	977	5.2	79.5	76.9	71.5
	pMT1	1726	0.9	56.0	22.5	6.9

Analysis of the pPCP1 plasmid revealed that, similarly to the strains responsible for the post-Black Death outbreaks, two plasmid variants were present in the Viborg strain. Via competitive mapping, the gap-bridging reads were identified in the *pla*-negative plasmid, which together with reads mapping specifically to the *pla* region are shown in Figure S1. The *pla*-negative to *pla*-positive plasmid ratio was 8.6:1. On the contrary, the *Y. pestis* isolates from Sejet and Ribe did not exhibit the depletion of the *pla* region in the pPCP1 plasmid.

After mapping the *Y. pestis* reads of the samples from Viborg, Sejet and Ribe against the CO92 genome, a careful examination of the alignments revealed the presence of several coverage gaps in the medieval strains in relation to the modern reference. All three genomes lacked the 8734 bp-long region in the chromosome (CO92, NC_003143.1:2554178-2562912) corresponding to Ypf Φ . Moreover, three additional gaps in coverage were observed in the near proximity to the phage locus for the Sejet isolate, suggesting that this strain lacked several genes in comparison to the CO92 reference strain (Tab. 3, Fig. 1). The missing genes are involved in the bacterium's virulence, response to stress and biosynthesis (Ackermann et al., 2008; Boin et al., 2004; Chain et al., 2004). These gaps in coverage were interpreted as deletions as these regions were covered in the alignment of all other medieval strains.

Table 3. Deletions observed in the chromosome of the Black Death strain from Sejet, Denmark.

Gap start position	Gap end position	Gene	Annotated protein
1917182	1925454	<i>cheZ</i>	protein phosphatase CheZ
		<i>YPO_RS09385</i>	N-acetylmuramoyl-L-alanine amidase
		<i>YPO_RS09390</i>	cell surface protein / coiled stalk of trimeric autotransporter adhesin family protein
		<i>YPO_RS09395</i>	YadA family autotransporter adhesin
		<i>YPO_RS09400</i>	putative exported protein
		<i>YPO_RS09405</i>	alanine racemase
		<i>YPO_RS09410</i>	hypothetical protein
2568948	2572633	<i>YPO_RS12445</i>	AMP nucleosidase
		<i>YPO_RS12450</i>	hypothetical protein / virulence factor (transmembrane protein)
		<i>YPO_RS12455</i>	putative virulence factor SrfB
2579140	2582342	<i>ilvB</i>	acetolactate synthase large subunit
		<i>ilvN</i>	acetolactate synthase small subunit
		<i>yjgG</i>	pyrimidine 5'-nucleotidase
		<i>YPO_RS12490</i>	erB-C domain protein

(A) NC_003143.1: 1917182-1925454



(B) NC_003143.1: 2568948-2572633



(C) NC_003143.1: 2579140-2582342



Figure 1. Gene content of the three chromosomal regions (A-C) missing in the Sejet strain.

Based on the mean genome coverage, only two strains (Viborg (G52) and Sejet (A5465/X3003)) (Tab. 2) were included in the phylogenetic analysis. The two medieval Danish strains were included in a phylogenetic analysis with 276 previously published *Y. pestis* genomes (228 modern and 48 medieval) (Fig. 2). The isolate from Viborg clustered with the Latvian strain G701 within the known diversity of the post-Black Death *Y. pestis* strains. The

genome from Sejet positioned itself in close proximity to the Black Death strains from Germany (MAN008) and Italy (COL001).

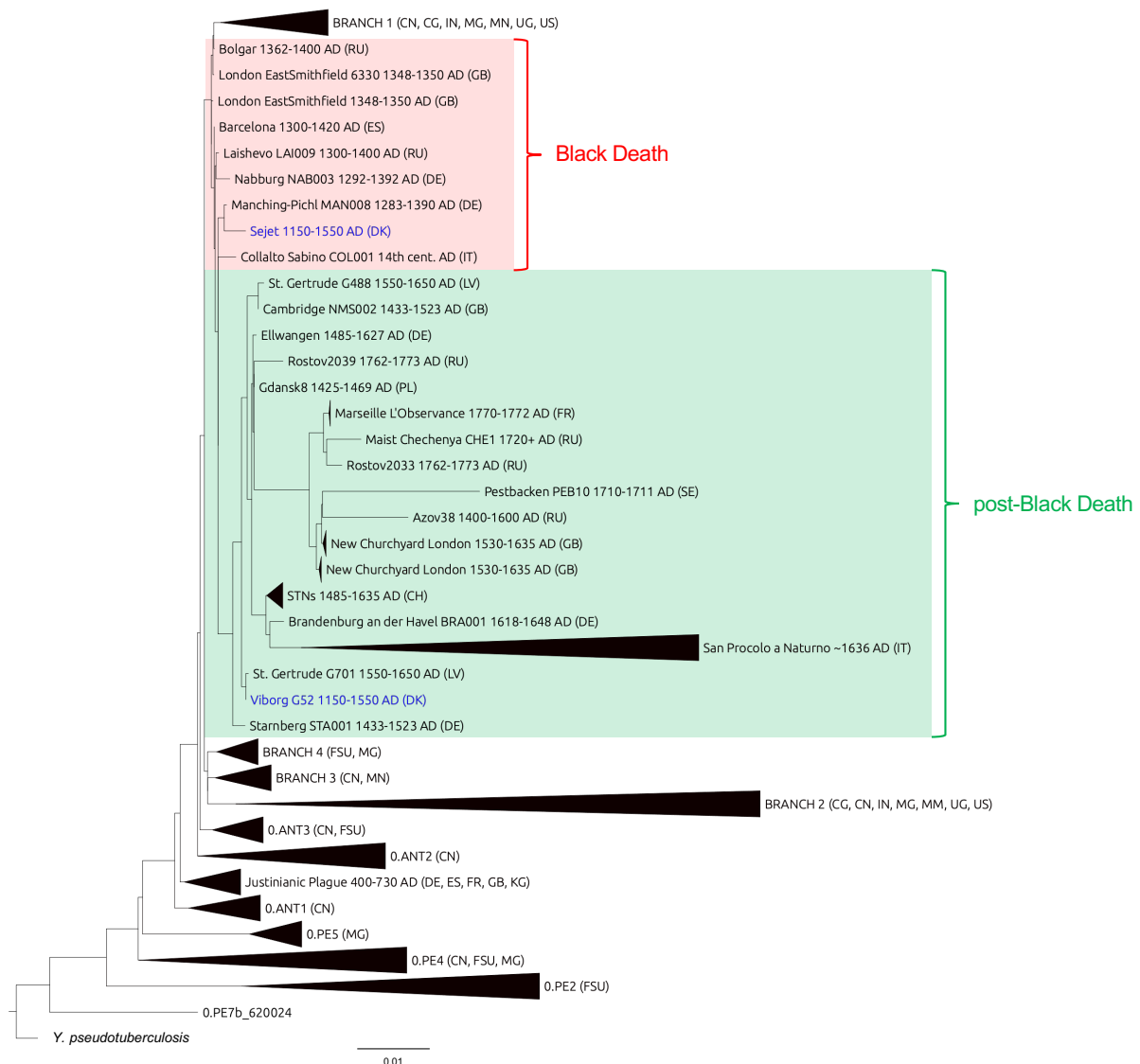


Figure 2. Maximum likelihood *Y. pestis* phylogeny based on the SNP alignment of 14900 positions with a bootstrap of 500 replicates. The tree includes 278 strains (228 modern and 50 ancient) with *Y. pseudotuberculosis* as an outgroup. The genomes from Viborg and Sejet are shown in blue. Dates are provided for the ancient strains. Country codes can be found in Table S3.

Discussion

The pathogen screening of 526 metagenomic samples (392 individuals) revealed that 17 individuals (4.3%) were infected with *Y. pestis*, making plague the likely cause of their death. Evidence of plague was detected in the metagenomes extracted from 16 teeth and one postcranial element. No traces of the pathogen were found in the petrous bone samples belonging to the same individuals. These results are in accordance with previous findings of Margaryan and colleagues (Margaryan et al., 2018) who detected *Y. pestis* DNA in teeth of five individuals from Bronze Age and Iron Age (Rasmussen et al., 2015) with no evidence of plague found in the corresponding petrous bones. Although the petrous portion exhibits high levels of endogenous DNA content (Gamba et al., 2014; Pinhasi et al., 2015), it does not prove to be particularly useful in ancient microbial genetics, as it usually carries insufficient amounts of exogenous DNA (Margaryan et al., 2018). Interestingly, all previously published ancient *Y. pestis* genomes were reconstructed from teeth samples (Bos et al., 2016; Bos et al., 2011; Morozova et al., 2020; Rascovan et al., 2019; Seguin-Orlando et al., 2021; Spyrou et al., 2016, 2018; Susat et al., 2020; Valtueña et al., 2017; Wagner et al., 2014) with a few studies screening various bones (Morozova et al., 2020; Seguin-Orlando et al., 2021; Valtueña et al., 2017). Thus, it would be sensible to assume that metagenomes extracted from petrous bones could present as false negatives in pathogen screening. In this study, only the petrous bone was analyzed for 157 individuals. For the remaining 235 individuals, 215 teeth and 22 postcranial elements were screened for the presence of *Y. pestis* (17/235, 7.2%). Moreover, preservation of the pathogen's DNA in the human skeleton is relatively poor, as only a small fraction of samples from plague-associated mass burials showed evidence of *Y. pestis*' DNA in previous studies (e.g., Seguin-Orlando et al., 2021; Spyrou et al., 2016; Susat et al., 2020). With all that in mind, it is possible that the true prevalence rate of plague in the examined sample was higher than reported.

All *Y. pestis*-positive individuals were buried in singular graves across four parish cemeteries, with the majority of cases (eight (47.1%) and six (35.3%)) detected at two sites (Sejet and Horsens, respectively). It is possible that these two cemeteries were dedicated, at least partially, for burials of plague victims. However, no historical evidence corroborating this scenario was discovered. A large proportion of previous *Y. pestis* findings dating to the Middle Ages were made based on pathogen screening of skeletal remains buried in mass graves, for which plague was already suspected *a priori*. In medieval Europe, outbreaks of deadly diseases often resulted

in mass inhumations (e.g., Haller et al., 2021; Willmott et al., 2020). Thus, in an instance when a medieval mass grave is excavated, plague is often suspected as the likely culprit – which is not always the case (Haller et al., 2021). Here, the aDNA-based crude period prevalence (CPP) of plague across common cemeteries is reported. The time frame of the *Y. pestis*-positive burials can be narrowed down to 1346-18th cent. AD, assuming that the infection occurred during the Medieval plague pandemic. Thus, the CPP (4.3-7.2%) does not relate to a specific point in time, but rather indicates the prevalence over approx. 400 years. A similar analysis was previously done for 191 individuals buried across five medieval graveyards in England (Cessford et al., 2021). In their study, Cessford and colleagues reported a *Y. pestis*' prevalence of 4.2-5.2%, which is comparable to the CPP noted in this study. When burials dating to the time period of known plague epidemics in the area were exclusively considered, the prevalence increased to 12-16% (Cessford et al., 2021). Our findings and those of Cessford and colleagues highlight the importance of pathogen screening of common cemeteries and individual graves. It has long been thought that many plague victims were buried at parish cemeteries (Grainger et al., 2008). With the use of aDNA methods, at least part of these otherwise indistinguishable plague burials, can be detected. Furthermore, such analyses could shed light on the genetic diversity of *Y. pestis* strains responsible for sporadic plague outbreaks of a relatively limited scale – possibly zoonotic events (chapter 4; Bos et al., 2016; Spyrou et al., 2016). These outbreaks, which did not result in mass inhumations, could have possibly been caused by strains less virulent for humans. The isolate from Viborg exhibited the *pla* depletion in the pPCP1 plasmid that is thought to be associated with reduced transmission of the pathogen (Susat et al., 2020). Although the Sejet genome carried only *pla*-positive pPCP1 plasmids, it lacked several seemingly useful genes (*YPO_RS09395*, *YPO_RS12445*, *YPO_RS12455*) involved in virulence (Ackermann et al., 2008; Chain et al., 2004) and chemotaxis (Boin et al., 2004) (Tab. 3, Fig. 1). It cannot, however, be excluded that the preventive measures used in Denmark at the time (Christensen, 2003) were effective in containing the outbreaks. Together with historical data, detection of *Y. pestis*-positive cases in singular graves followed by analysis of the reconstructed genomes can provide us with a better understanding of plague epidemiology and the bacterium's micro-evolution during the medieval pandemic.

The *Y. pestis*-positive samples belonged to four females and 13 males. Various literature and bioarcheological studies explored the potential sex selectivity of plague in the Middle Ages. Whilst the Black Death was known as the “universal killer” (e.g., DeWitte, 2009; Castex &

Kacki, 2016), studies of the post Black Death outbreaks suggested that the mortality rates were mostly skewed against females (e.g., Ell, 1989; Jimenez-Brobeil, 2020; Signoli et al., 2002). A consensus on this matter, however, has not been reached, as more studies reported conflicting results (e.g., Alexander, 1980; Bramanti et al., 2018; Curtis & Roosen, 2017; Hollingsworth & Hollingsworth, 1971; Whittles & Didelot, 2016). This inconsistent pattern of sex distribution among plague victims suggests that the underlying reason for these observations is likely of an extrinsic nature (unrelated to physiological sexual dimorphism), resulting from increased exposure or health status. For instance, the landmark study of 299 individuals from the Black Death cemetery (East Smithfield, London, UK) revealed an increased mortality associated with osteological stress markers in males relative to females. This suggests that previous physiological stress elevated the risk of death for men. However, higher mortality of generally healthy females in comparison to healthy males cannot be excluded (DeWitte, 2010). In a series of plague outbreaks in the Italian city of Milan (1452-1523 AD), mortality rates were skewed towards females. The authors proposed poor and overcrowded living conditions of female immigrants as a possible explanation for this observation (Cohn & Alfani, 2007). Data collected from modern plague outbreaks also seem to support the considerable influence of the environment on plague infection rates in males and females. In the United States, certain activities that are more common for men today, such as hunting and ranching, lead to increased exposure to infection (Butler, 1989; Cleri et al., 1997). Conversely, differences in sleeping conditions, and thus exposure to fleas, in Tanzania were shown to result in higher infection rates in women – who often sleep on the floor (Davis et al., 2006; Kamugisha et al., 2007). It is possible that the larger number of infected males reported here in comparison to females represents gender differences in behavior, for instance – occupational exposure. In the Middle Ages, the female role was often limited to household chores, while males provided sustenance for the family leading to an increased contact with contaminated sources. Historical records suggest that although preventive measures were applied during plague outbreaks in medieval Denmark, a large proportion of the population could not afford to abandon work and quarantine, which led to continuous disease spread (Christensen, 2003).

Archeological dating of the *Y. pestis*-positive burials included in this study (Boldsen, 2018) provided us with wide time frames of at least 1000 years (6th – 18th cent. AD). Detection of *Y. pestis*’ DNA, followed by the genomic and phylogenetic analyses of the three genomes (Viborg, Sejet and Ribe) allowed us to narrow down the time frame for each strain, assigning

them to two separate periods: the Black Death (Sejet and Ribe: 1346-1353 AD) and the post-Black Death (Viborg: 1354 – 1550 AD) (Fig. 2). As the *Y. pestis* phylogeny reflects a temporal distribution of the medieval strains, together with the pPCP1 *pla* depletion (associated with the post-Black Death medieval strains), phylogenetic analysis of medieval *Y. pestis* genomes may serve as an additional source of information on burial dating.

References

- Ackermann, N., Tiller, M., Anding, G., Roggenkamp, A., & Heesemann, J. (2008). Contribution of Trimeric Autotransporter C-Terminal Domains of Oligomeric Coiled-Coil Adhesin (Oca) Family Members YadA, UspA1, EibA, and Hia to Translocation of the YadA Passenger Domain and Virulence of *Yersinia enterocolitica*. *Journal of Bacteriology*, 190(14), 5031–5043. <https://doi.org/10.1128/JB.00161-08>
- Alexander, J. (1980). Bubonic plague in early modern Russia: Public health and urban disaster. Baltimore: John Hopkins University Press.
- Andrades Valtueña, A., Mitnik, A., Key, F. M., Haak, W., Allmäe, R., Belinskij, A., Daubaras, M., Feldman, M., Jankauskas, R., Janković, I., Massy, K., Novak, M., Pfrengle, S., Reinhold, S., Šlaus, M., Spyrou, M. A., Szécsényi-Nagy, A., Törv, M., Hansen, S., ... Krause, J. (2017). The Stone Age Plague and Its Persistence in Eurasia. *Current Biology*, 27(23), 3683–3691. <https://doi.org/10.1016/j.cub.2017.10.025>
- Boin, M. A., Austin, M. J., & Håase, C. C. (2004). Chemotaxis in *Vibrio cholerae*. *FEMS Microbiology Letters*, 239(1), 1–8. <https://doi.org/10.1016/j.femsle.2004.08.039>
- Boldsen, J., (2018). Personal communication on archaeological dating of the Danish sites.
- Bos, K., Herbig, A., Sahl, J., Waglechner, N., Fourment, M., Forrest, S. A., Klunk, J., Schuenemann, V. J., Poinar, D., Kuch, M., Golding, 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, e12994. <https://doi.org/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, 506–510. <https://doi.org/10.1038/nature10549>
- Bramanti, B., Zedda, N., Rinaldo, N., & Gualdi-Russo, E. (2018). A critical review of anthropological studies on skeletons from European plague pits of different epochs. *Scientific Reports*, 8(1), 17655. <https://doi.org/10.1038/s41598-018-36201-w>
- Butler, T. (1989). The black death past and present. 1. Plague in the 1980s. *Transactions of the Royal Society of Tropical Medicine and Hygiene*, 83(4), 458–460. [https://doi.org/10.1016/0035-9203\(89\)90246-0](https://doi.org/10.1016/0035-9203(89)90246-0)
- Chain, P. S. G., Carniel, E., Larimer, F. W., Lamerdin, J., Stoutland, P. O., Regala, W. M., Georgescu, A. M., Vergez, L. M., Land, M. L., Motin, V. L., Brubaker, R. R., Fowler, J., Hinnebusch, J., Marceau, M., Medigue, C., Simonet, M., Chenal-Francisque, V., Souza, B., Dacheux, D., ... Garcia, E. (2004). Insights into the evolution of *Yersinia pestis* through whole-genome comparison with *Yersinia pseudotuberculosis*. *Proceedings of the National Academy of Sciences*, 101(38), 13826–13831. <https://doi.org/10.1073/pnas.0404012101>

Castex, D., & Kacki, S. (2016). Demographic Patterns Distinctive of Epidemic Cemeteries in Archaeological Samples. *Microbiology Spectrum*, 4(4). <https://doi.org/10.1128/microbiolspec.PoH-0015-2015>

Cessford, C., Scheib, C. L., Guellil, M., Keller, M., Alexander, C., Inskip, S. A., & Robb, J. E. (2021). Beyond Plague Pits: Using Genetics to Identify Responses to Plague in Medieval Cambridgeshire. *European Journal of Archaeology*, 24(4), 496–518. <https://doi.org/10.1017/ear.2021.19>

Christensen, P. (2003). “In these perilous times”: plague and plague policies in early modern Denmark. *Medical History*, 47, 413–450.

Cleri, D. J., Vernaleo, J. R., Lombardi, L. J., Rabbat, M. S., Mathew, A., Marton, R., & Reyelt, M. C. (1997). Plague pneumonia disease caused by *Yersinia pestis*. *Seminars of Respiratory Infections*, 12(1), 12–23.

Cohn, S. K., & Alfani, G. (2007). Nonantola 1630: Anatomia di una pestilenza e meccanismi del contagio. *Popolazione e Storia*, 2, 99–138.

Curtis, D. R., & Roosen, J. (2017). The sex-selective impact of the Black Death and recurring plagues in the Southern Netherlands, 1349–1450. *American Journal of Physical Anthropology*, 164(2), 246–259. <https://doi.org/10.1002/ajpa.23266>

Davis, S., Makundi, R. H., Machang'u, R. S., & Leirs, H. (2006). Demographic and spatio-temporal variation in human plague at a persistent focus in Tanzania. *Acta Tropica*, 100(1–2), 133–141. <https://doi.org/10.1016/j.actatropica.2006.10.006>

Demeure, C. E., Dussurget, O., Mas Fiol, G., le Guern, A.-S., Savin, C., & Pizarro-Cerdá, J. (2019). *Yersinia pestis* and plague: an updated view on evolution, virulence determinants, immune subversion, vaccination, and diagnostics. *Genes & Immunity*, 20(5), 357–370. <https://doi.org/10.1038/s41435-019-0065-0>

Derbise, A., Chenal-Francisque, V., Pouillot, F., Fayolle, C., Prévost, M., Médigue, C., Hinnebusch, B. J., & Carniel, E. (2007). A horizontally acquired filamentous phage contributes to the pathogenicity of the plague bacillus. *Molecular Microbiology*, 63(4), 1145–1157. <https://doi.org/10.1111/j.1365-2958.2006.05570.x>

DeWitte, S. N. (2010). Sex differentials in frailty in medieval England. *American Journal of Physical Anthropology*, 143(2), 285–297. <https://doi.org/10.1002/ajpa.21316>

DeWitte, S. N. (2009). The effect of sex on risk of mortality during the Black Death in London, A.D. 1349–1350. *American Journal of Physical Anthropology*, 139(2), 222–234. <https://doi.org/10.1002/ajpa.20974>

Ell, S. R. (1989). Three Days in October of 1630: Detailed Examination of Mortality During an Early Modern Plague Epidemic in Venice. *Clinical Infectious Diseases*, 11(1), 128–139. <https://doi.org/10.1093/clinids/11.1.128>

Fu, Q., Posth, C., Hajdinjak, M., Petr, M., Mallick, S., Fernandes, D., Furtwängler, A., Haak, W., Meyer, M., Mittnik, A., Nickel, B., Peltzer, A., Rohland, N., Slon, V., Talamo, S.,

- Lazaridis, I., Lipson, M., Mathieson, I., Schiffels, S., ... Reich, D. (2016). The genetic history of Ice Age Europe. In *Nature* (Vol. 534, pp. 200–205). <https://doi.org/10.1038/nature17993>
- 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>
- Grainger, I., Hawkins, D., Cowal, L., & Mikulski, R. (2008). *The Black Death Cemetery, East Smithfield*. Museum of London Archaeology Service.
- Haller, M., Callan, K., Susat, J., Flux, A. L., Immel, A., Franke, A., Herbig, A., Krause, J., Kupczok, A., Fouquet, G., Hummel, S., Rieger, D., Nebel, A., & Krause-Kyora, B. (2021). Mass burial genomics reveals outbreak of enteric paratyphoid fever in the Late Medieval trade city Lübeck. *IScience*, 24(5), 102419. <https://doi.org/10.1016/j.isci.2021.102419>
- Hollingsworth, M. F., & Hollingsworth, T. H. (1971). Plague mortality rates by age and sex in the parish of St. Botolph's without Bishopsgate, London, 1603. *Population Studies*, 25(1), 131–146. <https://doi.org/10.1080/00324728.1971.10405789>
- Jimenez-Brobeil, S. A. (2020). Sex-differential mortality during the 1679 plague epidemic in Granada (Spain). *Revista española de antropología física*, 42, 13–23.
- Kamugisha, M. L., Gesase, S., Minja, D., Mgema, S., Mlwilo, T. D., & Mayala, B. K. (2007). Pattern and spatial distribution of plague in Lushoto, north-eastern Tanzania. *Tanzania Journal of Health Research*, 9(1), 12–18. <https://doi.org/10.4314/thrb.v9i1.14286>
- 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>
- Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., & Durbin, R. (2009). The Sequence Alignment/Map format and SAMtools. *Bioinformatics*, 25, 2078–2079. <https://doi.org/10.1093/bioinformatics/btp352>
- Margaryan, A., Hansen, H. B., Rasmussen, S., Sikora, M., Moiseyev, V., Khoklov, A., Epimakhov, A., Yepiskoposyan, L., Kriiska, A., Varul, L., Saag, L., Lynnerup, N., Willerslev, E., & Allentoft, M. E. (2018). Ancient pathogen DNA in human teeth and petrous bones. *Ecology and Evolution*, 8(6), 3534–3542. <https://doi.org/10.1002/ece3.3924>
- Morozova, I., Kasianov, A., Bruskin, S., Neukamm, J., Molak, M., Batieva, E., Pudło, A., Rühli, F. J., & Schuenemann, V. J. (2020). New ancient Eastern European *Yersinia pestis* genomes illuminate the dispersal of plague in Europe. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 375(1812), 20190569. <https://doi.org/10.1098/rstb.2019.0569>
- 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>

Rascovan, N., Sjögren, K.-G., Kristiansen, K., Nielsen, R., Willerslev, E., Desnues, C., & Rasmussen, S. (2019). Emergence and Spread of Basal Lineages of *Yersinia pestis* during the Neolithic Decline. *Cell*, 176(1–2), 295–305. <https://doi.org/10.1016/j.cell.2018.11.005>

Rasmussen, S., Allentoft, M., Nielsen, K., Orlando, L., Sikora, M., Sjögren, K., Pedersne, A., Schubert, M., Vam Dam, A., Kapel, C., Nielsen, H., Btunak, S., Avetisyan, P., Epimakhov, A., Khalyapin, M., Guni, A., Kriiska, A., Lasak, I., Metspalu, M., ... Willerslev, E. (2015). Early divergent strains of *Yersinia pestis* in Eurasia 5,000 years ago. *Cell*, 163(3), 571–582.

Prescher, A., & Wagner, P. (2016). *Aachen, Melaten. Der Friedhof des mittelalterlichen Leprosoriums an der Via Regia*. Verlag Philipp von Zabern.

Seguin-Orlando, A., Costedoat, C., der Sarkissian, C., Tzortzis, S., Kamel, C., Telmon, N., Dalén, L., Thèves, C., Signoli, M., & Orlando, L. (2021). No particular genomic features underpin the dramatic economic consequences of 17th century plague epidemics in Italy. *IScience*, 24(4), 102383. <https://doi.org/10.1016/j.isci.2021.102383>

Siegfried, K. (2018). *Grabungsbericht. Fundstelle 5.02.73. Ratzeburger Allee zwischen Kronsfordter Allee und Am Brink (EBHL)*.

Signoli, M., Seguy, I., Biraben, J.-N., Dutour, O., & Belle, P. (2002). Paleodemography and Historical Demography in the Context of an Epidemic: Plague in Provence in the Eighteenth Century. *Population* (English Edition, 2002-), 57(6), 829. <https://doi.org/10.2307/3246618>

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>

Spyrou, M. A., Tukhbatova, R. I., Wang, C.-C., Valtueña, A. A., Lankapalli, A. K., Kondrashin, V. v., Tsybin, V. A., Khokhlov, A., Kühnert, D., Herbig, A., Bos, K. I., & Krause, J. (2018). Analysis of 3800-year-old *Yersinia pestis* genomes suggests Bronze Age origin for bubonic plague. *Nature Communications*, 9, 2234. <https://doi.org/10.1038/s41467-018-04550-9>

Susat, J., Bonczarowska, J. H., Pētersone-Gordina, E., Immel, A., Nebel, A., Gerhards, G., & Krause-Kyora, B. (2020). *Yersinia pestis* strains from Latvia show depletion of the pla virulence gene at the end of the second plague pandemic. *Scientific Reports*, 10, 14628. <https://doi.org/10.1038/s41598-020-71530-9>

Susat, J., Lübke, H., Immel, A., Brinker, U., Macāne, A., Meadows, J., Steer, B., Tholey, A., Zagorska, I., Gerhards, G., Schmöcke, U., Kalniņš, M., Franke, A., Pētersone-Gordina, E., Teßman, B., Törv, M., Schreiber, S., Andree, C., Bērziņš, V., ... Krause-Kyora, B. (2021). A 5,000-year-old hunter-gatherer already plagued by *Yersinia pestis*. *Cell Reports*, 35(13), 109278. <https://doi.org/10.1016/j.celrep.2021.109278>

Valtueña, A., Mitnik, A., Key, F. M., Haak, W., Allmäe, R., Belinskij, A., Daubaras, M., Feldman, M., Jankauskas, R., Janković, I., Massy, K., Novak, M., Pfrenge, S., Reinhold, S.,

Šlaus, M., Spyrou, M. A., Szécsényi-Nagy, A., Törv, M., Hansen, S., ... Krause, J. (2017). The Stone Age Plague and Its Persistence in Eurasia. *Current Biology*, 27(23), 3683-3691. <https://doi.org/10.1016/j.cub.2017.10.025>

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](https://doi.org/10.1016/S1473-3099(13)70323-2)

Whittles, L. K., & Didelot, X. (2016). Epidemiological analysis of the Eyam plague outbreak of 1665–1666. *Proceedings of the Royal Society B: Biological Sciences*, 283(1830), 20160618. <https://doi.org/10.1098/rspb.2016.0618>

Willmott, H., Townend, P., Swales, D. M., Poinar, H., Eaton, K., & Klunk, J. (2020). A Black Death mass grave at Thornton Abbey: the discovery and examination of a fourteenth-century rural catastrophe. *Antiquity*, 94 (373), 179-196. <https://doi.org/10.15184/aqy.2019.213>

Discussion

Infectious diseases were an every-day threat in medieval Europe

The Justinianic plague (541-549 AD) is the only infectious disease outbreak of an epidemic scale known from the Early Middle Ages (6th-10th cent. AD) (Wagner et al., 2014) and little attention has been given to pathogens other than *Y. pestis* in the existing literature. Although historical sources and paleopathological studies indicate presence of various infectious diseases at the time, such as measles, malaria and leprosy (e.g., Belcastro et al., 2005; Newfield, 2017; Retief & Cilliers, 2010), only a few cases of single infections from this period have been detected with molecular analyses until now (*Plasmodium falciparum*, *Mycobacterium leprae* (*M. leprae*) and variola virus (VARV)) (Marciniak et al., 2016; Mühlemann et al., 2020; Pfrengle et al., 2021; Schuenemann et al., 2018). Thus, the diversity and prevalence of infectious diseases in early medieval Europe remains unexplored. With the aim of addressing this issue, a systematic screening of skeletal remains buried within the borders of the Lauchheim settlement (7th-8th cent. AD, Germany) was performed in chapter 3. Estimation of disease prevalence in past populations is, however, a challenging task, mainly due to missing data. Firstly, pathogen screening of metagenomic samples allows for discovery of blood-borne DNA pathogens. Thus, localized infections and RNA viruses will not be detected. Secondly, the degree of DNA preservation may also affect the results of pathogen screening, leading to underestimated rates of infection. Moreover, the examined burials cannot be ascribed to the same moment in time. Although it was not possible to overcome the existing technological limitations, the concept of crude period prevalence (CPP) was introduced in an effort to account for the type of data available. In CPP, the period prevalence refers to the total number of infected individuals over a specified period of time and the crude prevalence reflects the rate of infection in the total population (without accounting for the age and sex structure) (Waldron, 1991). The systematic pathogen screening of Lauchheim revealed that over 30% of inhabitants (CPP = 31.4%, over a period of about a century) were infected with minimum of one pathogen and at least four different infectious agents circulated within the settlement (hepatitis B virus (HBV), parvovirus B19, VARV and *M. leprae*). These findings represent the first evidence of a high burden of infectious diseases in the Early Middle Ages. As there is no comparable data in the existing literature, interpretation of these results in a wider context is difficult. It is not known whether the reported number is high, average or low for that particular time period. Nevertheless, considering that the CPP is likely an underestimate, the difference between the infectious disease prevalence in the early medieval Lauchheim and present-day Europe is stark.

Not only both leprosy and smallpox are absent in Europe today, the current prevalence of HBV (<2%) (Schmit et al., 2021) is nowhere near the endemic level noted for Lauchheim (HBV CPP=11.4%). Additional information provided by osteological analysis indicated that the inhabitants of the Lauchheim settlement suffered from a substantial amount of physiological stress. It is possible that the overall poor health was a consequence of the high infection prevalence. However, it is also possible that other factors, e.g., malnutrition, contributed to the health decline that was further exacerbated by the relatively easy transmission of pathogens.

Interestingly, Europe experienced a major climate decline between the 5th and 7th cent. AD, known as the Late Antique Little Ice Age (LALIA). LALIA was likely an environmental driver of crop failure, famine, and disease (Büntgen et al., 2016), which could have possibly contributed to the high disease burden at the time. As seen today, climate change is an eminent driver of emergence, re-emergence and transmission of infectious diseases (e.g., Caminade et al., 2018; Mora et al., 2022). The ecological context of that time together with the demographic and economic changes might have allowed for persistent outbreaks with the potential to set the stage for severe epidemics in the Late Middle Ages and early Modern Europe. For instance, although leprosy was already present in early medieval Europe (Belcastro et al., 2005; Pfrengle et al., 2021; Schuenemann et al., 2018), including Lauchheim, it reached epidemic proportions only hundreds of years later (between the 11th and 14th cent. AD) (Schuenemann et al., 2018). The medieval leprosy epidemic was dominated by strains belonging to the phylogenetic branch 3 – the same branch that contains a low-diversity cluster of the oldest European *M. leprae* genomes dating to the LALIA period (5th-8th cent. AD): Lauchheim G83 (chapter 3), GC96 (Schuenemann et al., 2018) and EDI006 (Pfrengle et al., 2021). Moreover, the phylogenetic position and gene inactivation pattern of the VARV G31 strain detected in Lauchheim were intermediate in relation to ancient VARV strains from the same period (7th-8th cent. AD) and strains dating to between the 9th and 11th cent. AD (Mühlemann et al., 2020). VARV went through reductive evolution over time, which is thought to be a consequence of the pathogen's adaptation to the human host (Hendrickson et al., 2010). The gene inactivation pattern noted for the G31 strain from Lauchheim could represent a step in VARV's adaptation to humans. It is, however, necessary to gather more evidence in order to evaluate the hypothesis on LALIA's influence on the infectious disease landscape and pathogen evolution.

The concept of CPP was also applied in chapter 6, in which pathogen screening of 17 graveyards (392 individuals) was performed with the aim of assessing the prevalence of plague in medieval Danish (N=15) and German (N=2) cemeteries. *Y. pestis*' DNA was detected at four different sites across Denmark (Viborg, Horsens, Sejet and Ribe) and the CPP was estimated to be between 4.3-7.2% over a period of 400 years. These results are comparable to the plague frequency of 4.2-5.2% that was reported for parish cemeteries in medieval England (Cessford et al., 2021). Although it had been previously hypothesized that plague victims were not exclusively buried in mass graves (Grainger et al., 2008), there was no clear evidence of their presence at parish cemeteries due to the lack of osteological lesions. Together with the study of Cessford and colleagues (Cessford et al., 2021), the findings from chapter 6 provided molecular evidence of *Y. pestis* in individual burials. It is possible that the plague victims buried in such graves died in plague outbreaks of a limited scale that did not result in emergency mass inhumations. Such sporadic and possibly zoonotic events (chapter 4; Bos et al., 2016; Spyrou et al., 2016) suggest an ever-present threat of disease in the Middle Ages. Interestingly, all three genomes from medieval Denmark that were successfully reconstructed in chapter 6 exhibited either a depletion of the virulence factor *pla* (chapter 4; see the following discussion section) or a deletion of genes seemingly useful for the pathogen's virulence (chapter 6; Ackermann et al., 2008; Chain et al., 2004). These findings highlight the potential of parish cemeteries as a source of data on genetic diversity of *Y. pestis*, the bacterium's micro-evolution during the Medieval pandemic and genomic alterations which might have contributed to the disappearance of plague from Europe in the 18th cent. AD.

Genomic changes in *Y. pestis*' genome might have contributed to the end of the Black Death and the outbreak of the Modern pandemic

Deadly outbreaks of plague in the Middle Ages often resulted in emergency mass burials (Grainger et al., 2008; Hartle, 2017; Pfizenmaier, 2016; Stoakley, 2019; Willmott et al., 2020). Thus, in an instance when a medieval mass grave is excavated, plague is often presumed as the likely culprit. As *Y. pestis* infection is acute, it does not leave skeletal lesions and detection of plague victims is only possible with reliable historical records or molecular methods. A large proportion of previous *Y. pestis* findings from that time period were made based on pathogen screening of skeletal remains buried in mass graves, for which plague was suspected *a priori*

(Bos et al., 2011; Spyrou et al., 2016). In some instances, however, *Y. pestis* could not be confirmed as responsible for the mass burials or a completely different infectious agent was detected (Haller et al., 2021). In chapter 4, pathogen screening of mass burials excavated on the grounds of the Riga Dome Church and the St. Gertrude's Church in Riga, Latvia was performed with the aim to determine whether these burials were a result of plague epidemics. Molecular analysis allowed to confirm that the graves associated with the St. Gertrude's Church were a resting place of plague victims (chapter 4). Two *Y. pestis* genomes from Riga were reconstructed and genomic analysis of the isolates together with previously published medieval strains revealed a substantial difference in the pPCP1 plasmids of the strains responsible for the Black Death (BD) and the post-Black Death (pBD) outbreaks. Contrary to the BD strains, the pBD strains exhibited a depletion of the *pla* gene, which is an important factor involved in the dissemination of the bacteria in the lymphatic system and the lower airways (Lathem et al., 2007). Interestingly, determination of the presence/absence of the *pla*-negative pPCP1 plasmids in medieval strains allows to assign them either to the BD or the pBD period. Thus, genomic and phylogenetic analysis of *Y. pestis* can be useful when estimating the archaeological date of burials (chapter 6). *pla* encodes the plasminogen activator protease (Pla) that is necessary for the full virulence of the pathogen in both the bubonic and pneumonic plague (Lathem et al., 2007; Smiley et al., 2019; Sodeinde et al., 1992). It is, however not required for bloodstream infection (Sodeinde et al., 1992). The reasons which might have driven the *pla* loss are speculative. It is possible that establishment of the pathogen in local rodent reservoirs led to the *pla* depletion over time. *Y. pestis* strains which do not carry the pPCP1 plasmid (*Pestoides*) were previously isolated from environmental reservoirs located in the former Soviet Union (Garcia et al., 2007). *Pestoides* can infect rodents via the flea vector (Rajanna et al., 2010). They are, however, unable to proliferate in the lungs (Zimmler et al., 2015) and rarely infect humans (Rajanna et al., 2010). Although little is known about natural plague reservoirs in medieval Europe, even a minor sylvatic focus with a small exposure is sufficient to cause an outbreak (Keeling & Giligan, 2000). It could be hypothesized that after the introduction of *Y. pestis* in the 14th cent. AD, the pathogen established in local rodent reservoirs and the sporadic pBD outbreaks were a result of environmental spillovers (chapter 4; Bos et al., 2016; Spyrou et al., 2016). Long circulation of the bacterium among rodent populations, where *pla* was not necessary for effective transmission, could then lead to the *pla* depletion. Virulence of the *pla*-depleted pBD strains was possibly limited relative to the BD strains, which in turn might have contributed to plague disappearance in the 18th cent. AD.

Whilst the pBD lineage went extinct, the BD strains are thought to have travelled back to Asia and become the source of the Modern pandemic (Spyrou et al., 2016). The Modern pandemic outbreaks are associated with one group of *Y. pestis* strains, i.e., the *orientalis* biovar (Vogler et al., 2013) that, unlike any ancient, medieval and other modern isolates, carries a chromosomal phage (YpfΦ) (chapter 5; Derbise et al., 2007). Genomic analysis of the YpfΦ performed in chapter 5 revealed that the phage encodes zonula occludens toxin (Zot), which is a known virulence factor in *Vibrio cholerae* (Fasano et al., 1991). It was shown to increase permeability of the gut epithelia (Fasano et al., 1991), mucosa (Gopalakrishnan et al., 2009) and endothelial cells in the brain (Karyekar et al., 2003). The acquisition of YpfΦ and subsequent release of Zot might have provided *Y. pestis* with a new virulence mechanism inducing damage to the host barriers. Moreover, a previous functional study done on mice showed that YpfΦ allows better colonisation of the host, conferring increased fitness during the infection of mammals (Debrise et al., 2007). YpfΦ could thus affect the pathogenicity of *Y. pestis* and the clinical manifestation of plague. Instances of rare plague forms and symptoms were previously reported, including the pharyngeal plague as well as *Y. pestis*-induced gastroenteritis and meningitis (Arbaji et al., 2005; Becker et al., 1987; Leslie et al., 2011). Such uncommon disease manifestations often delay a proper diagnosis, which can lead to fatal outcomes. For instance, during the plague-related outbreak of gastroenteritis in Afghanistan, 20.5 % of the infected individuals (17/83) died before the appropriate treatment was applied (Leslie et al., 2011). Although it is not known whether in these instances the causative strains were YpfΦ-positive, identification of the YpfΦ-encoded Zot as a potential virulence factor that breaches host epithelial barriers should make us rethink these clinical cases. Functional studies are necessary to determine the role of Zot in *Y. pestis* pathogenesis. As evidenced by recent plague outbreaks (Dutt et al., 1994; Leslie et al., 2011; Nguyen et al., 2018; Vogler et al., 2013), the disease remains a major public health concern. The relatively recent acquisitions of *zot* (chapter 5) and additional plasmids which confer antibiotic resistance (Lei & Kumar, 2022) reflect the evolution of the pathogen towards increased virulence. These genomic alterations of *Y. pestis* together with the ongoing climate change may present the bacterium with conditions that are conducive for large-scale plague outbreaks. It was estimated that a temperature rise of 1°C would increase the prevalence of *Y. pestis* infection by 50% (Stenseth et al., 2006). It is thus imperative to study the recent evolution of *Y. pestis*, especially in times of emerging and re-emerging infectious diseases.

Ancient DNA technology allows to study genetic predisposition to infectious diseases which disappeared from Europe hundreds of years ago

Nowadays, Europe is free of many infectious diseases which used to be common in the Middle Ages, such as leprosy, plague and enteric fever. Therefore, the genetic predisposition to these diseases in present-day Europeans cannot be investigated via case-control association studies, due to the virtual lack of cases. Development of the aDNA technology, however, allowed for such analyses to be performed on past populations. The first such study was conducted on the medieval sample of Danish and German individuals in which the association between the HLA-DRB1*15:01 and leprosy was tested (Krause-Kyora et al., 2018). DRB1*15:01 was the focus of the analysis as it is the most consistently leprosy-associated allele across various non-European populations, (Gorodezky et al., 2004; Joko et al., 2000; Vanderborght et al., 2007; Zerva et al., 1996; Zhang et al., 2009). In their work, Krause-Kyora and colleagues confirmed the association between DRB1*15:01 and leprosy in Europeans (Krause-Kyora et al., 2018). A similar approach was implemented in chapter 1, which describes a case-control association analysis between HLA-DRB1*03 and paratyphoid fever (caused by *S. enterica* Paratyphi C) in medieval Germans. This allele was previously linked to an increased risk of *S. enterica* Typhi infection (typhoid fever) in present-day Vietnamese (Dunstan et al., 2001). Therefore, the analysis performed in chapter 1 was a replication of the study linking DRB1*03 with *Salmonella* infection. It is the first analysis which investigated genetic predisposition to paratyphoid fever in Europeans and the only such study on Paratyphi C infection up to now. Such a focused approach was implemented due to the small sample size (29 medieval cases and 24 contemporaneous controls) and thus, the limited statistical power of the study which did not allow for testing of multiple alleles. On the other hand, a relatively large sample size in the case-control association study described in chapter 2 allowed for an exploratory analysis. In this chapter, the complete HLA profiles of 300 medieval Europeans (Danish and Germans) were analyzed in the context of leprosy susceptibility for the first time. This approach would potentially enable detection of novel associations. Indeed, a nominally significant association was noted for three alleles (A*23, A*26 and DRB1*13) which had not been linked to leprosy before. Moreover, the results obtained in chapter 2 support previous reports on B*38 and C*06 and leprosy susceptibility in Brazil and India, respectively (Franceschi et al., 2011; Shankankumar et al., 2003). Although the association between DRB1*15 and leprosy risk reported previously for medieval Europeans (Krause-Kyora et al., 2018) could not be

confirmed because of the limited statistical power, a strong trend was noted. Due to the exploratory nature of this study, the p-values were not corrected for multiple testing. Thus, these findings must be confirmed in a large independent sample. A similar exploratory approach was implemented in a case-control association study performed on medieval plague victims and contemporaneous controls (Immel et al., 2021). In their study, Immel and colleagues reported an association between the HLA-B*51:01 allele and plague risk. Although, multiple testing correction was applied, the sample size was modest (36 cases and 50 controls) (Immel et al., 2021) and no replication of this finding was performed. These four aDNA studies (chapters 1 and 2, Immel et al., 2021; Krause-Kyora et al., 2018) demonstrate the tremendous potential of the aDNA methodology in identification of genetic variants involved in predisposition to infectious diseases that are no longer present in Europe, but remain endemic in other parts of the world. Identification of such variants can improve our understanding concerning the contribution of specific HLA alleles to the disease under study and potentially allow to harness this knowledge to improve patient care.

Modern-day scourges: are chronic inflammation and autoimmunity the legacy of past epidemics?

Based on historical documentation, archaeological sources and molecular data, it is evident that medieval Europeans struggled with the high disease burden and the ever-present threat of infections (e.g., chapter 3, 4 and 6; Grainger et al., 2008; Haller et al., 2021; Hartle, 2017; Snellgrove, 1954). At the time, increasing population density, limited medical knowledge and poor hygiene fueled the spread of pathogens (O’Neil, 1993). Whilst some infections might have been relatively mild, others were deadly. It is known that the clinical outcome of infectious diseases may depend on various factors, including the nutritional status, co-occurrence of infections, stress level, previous exposure to the pathogen, strain type of the infectious agent as well as host genetics (Fava et al., 2012; Jia et al., 2017; Krawinkel, 2012; Verbrugghe et al., 2012). Leprosy is an epitome of the substantial influence of the host’s immune response on disease manifestation. Depending on the degree of the immunological reaction, *M. leprae* infection can lead to severe disability, mild disfigurement, painful chronic inflammation or a complete lack of symptoms (Joyce, 2012; Sales et al., 2011; Walker & Lockwood, 2007). The HLA system bears strong signs of pathogen-driven selection (Penn et al., 2002; Prugnolle et

al., 2005) and represents a genetic complex that is most often associated with various infectious diseases, including leprosy, typhoid fever, tuberculosis, malaria, acquired immunodeficiency syndrome (AIDS), hepatitis B and C, leishmaniasis and schistosomiasis (Blackwell et al., 2009; Sanchez-Mazas, 2020). aDNA case-control studies of the association between HLA alleles and leprosy, paratyphoid fever and plague confirmed the existence of a genetic component related to these infections in medieval Europeans (chapter 1 and 2; Immel et al., 2021; Krause-Kyora et al., 2018). Both plague and paratyphoid fever represent acute and deadly infections that do not discriminate based on demographic characteristics, such as age or sex (Castex & Kacki, 2016; Prechel, 1996). Whilst leprosy is a chronic disease that is not deadly *per se* (Bryceson and Pfaltzgraff, 1990), it presumably hindered the reproductive success of the affected individuals due to their stigmatization and isolation (Raffcliffe, 2006), leprosy-related hormonal disturbances (Gunawan et al., 2020; Khanna et al., 2014; Saporta & Yuksel, 1994) and occasional fatal complications (Klioze & Ramos-Caro, 2000). It is thus likely that large-scale epidemics of infectious diseases in the Middle Ages influenced the diversity of HLA alleles in the European population.

Many HLA alleles associated with genetic predisposition to infectious diseases were also shown to be linked to chronic inflammation and autoimmunity, which highlights the pleiotropy of the HLA complex. Existing data suggests an evolutionary trade-off between the decreased susceptibility to infectious diseases and increased risk of autoimmunity and/or inflammation (Brinkworth & Barreiro, 2014; Corona et al., 2010; Sironi & Clerici, 2010). Whilst a large repertoire of the recognized antigens provides a more efficient protection against infectious diseases, it also raises the risk of autoimmunity (Shevyrev et al., 2021). It was previously proposed that past epidemics of infectious diseases affected the frequency of variants associated with autoimmune and inflammatory conditions today (Karlsson et al., 2014). It could be hypothesized that alleles protective against various infections were favored in medieval Europe, where the burden of infectious disease was substantial, and *vice versa*. These same alleles that protected medieval Europeans from infections may predispose the present-day population to chronic inflammation and autoimmune conditions. For instance, the case-control association study described in chapter 2 identified the HLA-C*06 allele as a protective factor for leprosy, confirming the findings reported for Indians (Shankankumar et al., 2003). In addition, C*06 was previously linked to an increased risk of psoriasis in Finland (Haapasalo et al., 2018). This pattern, however, was not clear for the remaining alleles which were

associated with leprosy in chapter 2 (A*23, A*26, B*38 and DRB1*13) (Tab. 10 in chapter 2). This could be explained by the lack of known associations for these alleles as well as the existence of genetic differences in the populations examined across studies. Moreover, the complexity of the pathophysiological mechanisms underlying various diseases makes it challenging to interpret the HLA association patterns across different pathological conditions. For instance, DRB1*15:01 was found to be a susceptibility factor for leprosy (Krause-Kyora et al., 2018) and a protective factor for type I diabetes (DRB1*15:01-DQB1*06:02 haplotype) (Hu et al., 2015). Although DRB1*15:01 has decreased in the European population since the Middle Ages (from 18.4% to 13.8%) (Krause-Kyora et al., 2018), it remains a common allele that is linked to an increased risk of other autoimmune conditions, such as multiple sclerosis (Kular et al., 2018) and sarcoidosis (Rossman et al., 2003). Genetic variants that exhibit opposite risk profiles were also found in genes located outside of the HLA region and were suggested to be under balancing selection due to the antagonistic pleiotropic effects (Wang et al., 2010). A similar observation concerning the association pattern was made in chapter 1, where the link between DRB1*03 and an increased susceptibility to paratyphoid fever in medieval Germans (14th cent. AD) was reported. DRB1*03 was previously shown to be associated with protection against tuberculosis (Malkova et al., 2020), which was a highly prevalent disease in the 18-19th cent. AD (Barberis et al., 2017) and severe acute respiratory syndrome (SARS) in present-day Chinese (Ng et al., 2004; Wang et al., 2011). Moreover, the allele is a risk factor for various autoimmune conditions in modern populations across the globe (Chen et al., 1999; Cruz-Tapias et al., 2012; Fernando et al., 2007; Kwon et al., 1999; Malkova et al., 2020; Morris et al., 2012; Nakken et al., 2001; Shankarkumar et al., 2003; van Gerven et al., 2015). Like DRB1*15:01, the frequency of DRB1*03 has slightly decreased since the Middle Ages (by 2.8%), whilst remaining a prevalent allele (chapter 1). This finding indicates that DRB1*03 has been under balancing selection. Due to the temporal changes in the pathogen landscape, opposing selective pressures might have influenced the allele's frequency. It is also possible that DRB1*03 has been maintained due to its pleiotropy. Although seven HLA alleles (Tab. 11 in chapter 2) were shown to substantially shift in frequency (>5%) in the European population since the Middle Ages (chapter 2), the exact reasons for the changes are not clear based on the known disease associations. The prevalence of these alleles could have been influenced by changes in the pathogen diversity. However, the HLA system was also shown to be involved in other aspects of human physiology, such as pregnancy maintenance (Hviid & Christiansen, 2005; Kiseleva et al., 2019), development of the central nervous system

(Morimoto & Nakajima, 2019) and possibly mate choice (Havlicek & Roberts, 2009). Whilst identification of the associations between HLA variants and disease susceptibility can improve our understanding of how specific alleles contribute to various diseases, the action mode between the HLA molecules, relevant antigens and the receptors that bind them (T-cell receptors, TCRs) remains unknown. This leaves the exact mechanisms that are associated with the disease, and mediated by the HLA, unexplored, making the interpretation of the association patterns challenging. Thus, further studies characterizing the interactions between the HLA molecules, presented peptides and TCRs are necessary to harness the knowledge on disease associations for more in depth analyses of the mechanisms underlying the associations, interpretation of the frequency shifts and consequences of past epidemics.

Although there is a genetic component in autoimmune and chronic inflammatory diseases, it is believed that in many cases an environmental trigger is necessary for the onset of symptoms to occur (Khan & Wang, 2020). For instance, Epstein-Barr virus was identified as a leading cause of multiple sclerosis (Bjornevik et al., 2022). It has been also hypothesized that Crohn's disease (CD) is possibly caused by *Mycobacterium avium* spp. *paratuberculosis* (MAP) (Behr & Kapur, 2008; Schurr & Gros, 2009), as a chronic disease of the intestines in cattle and non-human primates that resembles CD was shown to be caused by MAP (McClure et al., 1987; Pérez et al., 1996). Another example of the connection between infectious disease and autoimmunity is the high clinical similarity between sarcoidosis and tuberculosis. Interestingly, certain HLA genotypes were shown to exhibit opposite effects on the predisposition to both diseases (Malkova et al., 2020). In chapter 5, a YpfΦ-encoded *zot* was discovered in various pathogenic and non-pathogenic Enterobacteria, such as *Y. pestis*, *Salmonella enterica*, *Escherichia coli* and *Citrobacter koseri* (chapter 5). Zot is a homologue of the endogenous human zonulin (Fasano et al., 2000). Elevated levels of zonulin play an important role in a wide variety of inflammatory diseases, such as celiac disease, diabetes, inflammatory bowel disease (IBD) and multiple sclerosis (e.g., Arrieta et al., 2009; Fasano, 2011; Sapone et al., 2006). Moreover, importance of the bacterial Zot for gastrointestinal pathogenesis was previously suggested as Zot-harboring *Campylobacter concisus* was isolated from saliva samples of IBD patients (Mahendran et al., 2013). Detection of YpfΦ, and thus *zot*, in strains from various Enterobacterial species provides further insight into the role of bacteria in chronic inflammation.

As evidenced above, autoimmunity and chronic inflammation seem to be intertwined with infectious diseases at various levels, including genetic susceptibility, etiology and pathophysiological mechanisms. It is possible that the HLA profiles that were advantageous a few centuries ago, when the frequency of infections was high, today contribute to the increasing prevalence of chronic inflammation in the regions with a relatively low pathogen burden. Further studies are necessary to investigate the complex interactions between autoimmune as well as chronic inflammation and infections.

References

- Ackermann, N., Tiller, M., Anding, G., Roggenkamp, A., & Heesemann, J. (2008). Contribution of Trimeric Autotransporter C-Terminal Domains of Oligomeric Coiled-Coil Adhesin (Oca) Family Members YadA, UspA1, EibA, and Hia to Translocation of the YadA Passenger Domain and Virulence of *Yersinia enterocolitica*. *Journal of Bacteriology*, 190(14), 5031–5043. <https://doi.org/10.1128/JB.00161-08>
- Arbaji, A., Kharabsheh, S., Al-Azab, S., Al-Kayed, M., Amr, Z. S., Abu Baker, M., & Chu, M. C. (2005). A 12-case outbreak of pharyngeal plague following the consumption of camel meat, in north-eastern Jordan. *Annals of Tropical Medicine & Parasitology*, 99(8), 789–793. <https://doi.org/10.1179/136485905X65161>
- Arrieta, M. C., Madsen, K., Doyle, J., & Meddings, J. (2009). Reducing small intestinal permeability attenuates colitis in the IL10 gene-deficient mouse. *Gut*, 58(1), 41–48. <https://doi.org/10.1136/gut.2008.150888>
- Barberis, I., Bragazzi, N., Galluzzo, L., & Martini, M. (2017). The history of tuberculosis: from the first historical records to the isolation of Koch's bacillus. *Journal of Preventive Medicine and Hygiene*, 58(1), E9–E12.
- Becker, T. M., Poland, J. D., Quan, T. J., White, M. E., Mann, J. M., & Barnes, A. M. (1987). Becker, T M et al. "Plague meningitis--a retrospective analysis of cases reported in the United States, 1970-1979." *The Western journal of medicine* vol. 147,5 (1987): 554-7. *The Western Journal of Medicine*, 147(5), 554–557.
- Behr, M. A., & Kapur, V. (2008). The evidence for *Mycobacterium paratuberculosis* in Crohn's disease. *Current Opinion in Gastroenterology*, 24(1), 17–21. <https://doi.org/10.1097/MOG.0b013e3282f1dcc4>
- Belcastro, M. G., Mariotti, V., Facchini, F., & Dutour, O. (2005). Leprosy in a skeleton from the 7th century necropolis of Vicenne-Campochiaro (Molise, Italy). *International Journal of Osteoarchaeology*, 15(6), 431–448. <https://doi.org/10.1002/oa.799>
- Bjornevik, K., Cortese, M., Healy, B. C., Kuhle, J., Mina, M. J., Leng, Y., Elledge, S. J., Niebuhr, D. W., Scher, A. I., Munger, K. L., & Ascherio, A. (2022). Longitudinal analysis reveals high prevalence of Epstein-Barr virus associated with multiple sclerosis. *Science*, 375(6578), 296–301. <https://doi.org/10.1126/science.abj8222>
- Blackwell, J. M., Jamieson, S. E., & Burgner, D. (2009). HLA and Infectious Diseases. *Clinical Microbiology Reviews*, 22(2), 370–385. <https://doi.org/10.1128/CMR.00048-08>
- 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, 506–510. <https://doi.org/10.1038/nature10549>

Brinkworth, J. F., & Barreiro, L. B. (2014). The contribution of natural selection to present-day susceptibility to chronic inflammatory and autoimmune disease. *Current Opinion in Immunology*, 31, 66–78. <https://doi.org/10.1016/j.coi.2014.09.008>

Bryceson, A., & Pfaltzgraff, R. (1990). *Leprosy* (3rd edition). Churchill Livingstone.

Büntgen, U., Myglan, V. S., Ljungqvist, F. C., McCormick, M., di Cosmo, N., Sigl, M., Jungclauss, J., Wagner, S., Krusic, P. J., Esper, J., Kaplan, J. O., de Vaan, M. A. C., Luterbacher, J., Wacker, L., Tegel, W., & Kirdyanov, A. v. (2016). Cooling and societal change during the Late Antique Little Ice Age from 536 to around 660 AD. *Nature Geoscience*, 9(3), 231–236. <https://doi.org/10.1038/ngeo2652>

Caminade, C., McIntyre, K. M., & Jones, A. E. (2019). Impact of recent and future climate change on vector-borne diseases. *Annals of the New York Academy of Sciences*, 1436(1), 157–173. <https://doi.org/10.1111/nyas.13950>

Castex, D., & Kacki, S. (2016). Demographic Patterns Distinctive of Epidemic Cemeteries in Archaeological Samples. *Microbiology Spectrum*, 4(4). <https://doi.org/10.1128/microbiolspec.PoH-0015-2015>

Cessford, C., Scheib, C. L., Guellil, M., Keller, M., Alexander, C., Inskip, S. A., & Robb, J. E. (2021). Beyond Plague Pits: Using Genetics to Identify Responses to Plague in Medieval Cambridgeshire. *European Journal of Archaeology*, 24(4), 496–518. <https://doi.org/10.1017/eea.2021.19>

Chain, P. S. G., Carniel, E., Larimer, F. W., Lamerdin, J., Stoutland, P. O., Regala, W. M., Georgescu, A. M., Vergez, L. M., Land, M. L., Motin, V. L., Brubaker, R. R., Fowler, J., Hinnebusch, J., Marceau, M., Medigue, C., Simonet, M., Chenal-Francisque, V., Souza, B., Dacheux, D., ... Garcia, E. (2004). Insights into the evolution of *Yersinia pestis* through whole-genome comparison with *Yersinia pseudotuberculosis*. *Proceedings of the National Academy of Sciences*, 101(38), 13826–13831. <https://doi.org/10.1073/pnas.0404012101>

Chen, Q.-Y., Huang, W., She, J.-X., Baxter, F., Volpe, R., & Maclaren, N. K. (1999). HLA-DRB1108, DRB1103/DRB310101, and DRB310202 Are Susceptibility Genes for Graves' Disease in North American Caucasians, Whereas DRB1107 Is Protective¹. *The Journal of Clinical Endocrinology & Metabolism*, 84(9), 3182–3186. <https://doi.org/10.1210/jcem.84.9.5991>

Corona, E., Dudley, J. T., & Butte, A. J. (2010). Extreme Evolutionary Disparities Seen in Positive Selection across Seven Complex Diseases. *PLoS ONE*, 5(8), e12236. <https://doi.org/10.1371/journal.pone.0012236>

Cruz-Tapias, P., Pérez-Fernández, O. M., Rojas-Villarraga, A., Rodríguez-Rodríguez, A., Arango, M. T., & Anaya, J. M. (2012). Shared HLA class II in six autoimmune diseases in Latin America: A meta-analysis. In *Autoimmune Diseases*. <https://doi.org/10.1155/2012/569728>

Derbise, A., Chenal-Francisque, V., Pouillot, F., Fayolle, C., Prévost, M., Médigue, C., Hinnebusch, B. J., & Carniel, E. (2007). A horizontally acquired filamentous phage contributes

to the pathogenicity of the plague bacillus. *Molecular Microbiology*, 63(4), 1145–1157. <https://doi.org/10.1111/j.1365-2958.2006.05570.x>

Dunstan, S. J., Stephens, H. A., Blackwell, J. M., Duc, C. M., Lanh, M. N., Dudbridge, F., Phuong, C. X. T., Luxemburger, C., Wain, J., Ho, V. A., Hien, T. T., Farrar, J., & Dougan, G. (2001). Genes of the class II and class III major histocompatibility complex are associated with typhoid fever in Vietnam. *Journal of Infectious Diseases*, 183(2), 261–268. <https://doi.org/10.1086/317940>

Dutt, A., Akhtar, R., & McVeigh, M. (1994). Surat plague of 1994 re-examined. *Southeast Asian Journal of Tropical Medicine and Public Health*, 37(4), 755–760.

Fasano, A. (2011). Zonulin and Its Regulation of Intestinal Barrier Function: The Biological Door to Inflammation, Autoimmunity, and Cancer. *Physiological Reviews*, 91(1), 151–175. <https://doi.org/10.1152/physrev.00003.2008>

Fasano, A., Baudry, B., Pumpllin, D. W., Wasserman, S. S., Tall, B. D., Ketley, J. M., & Kaper, J. B. (1991). *Vibrio cholerae* produces a second enterotoxin, which affects intestinal tight junctions. *Proceedings of the National Academy of Sciences of the United States of America* 88(12), 5242–5246. <https://doi.org/10.1073/pnas.88.12.5242>

Fasano, A., Not, T., Wang, W., Uzzau, S., Berti, I., Tommasini, A., & Goldblum, S. E. (2000). Zonulin, a newly discovered modulator of intestinal permeability, and its expression in coeliac disease. *The Lancet*, 355(9214). [https://doi.org/10.1016/S0140-6736\(00\)02169-3](https://doi.org/10.1016/S0140-6736(00)02169-3)

Fava, V., Orlova, M., Cobat, A., Alcaïs, A., Mira, M., & Schurr, E. (2012). Genetics of leprosy reactions: an overview. *Memórias Do Instituto Oswaldo Cruz*, 107(suppl 1), 132–142. <https://doi.org/10.1590/S0074-02762012000900020>

Fernando, M. M. A., Stevens, C. R., Sabeti, P. C., Walsh, E. C., McWhinnie, A. J. M., Shah, A., Green, T., Rioux, J. D., & Vyse, T. J. (2007). Identification of two independent risk factors for lupus within the MHC in United Kingdom families. *PLoS Genetics* 3(11), e192. <https://doi.org/10.1371/journal.pgen.0030192>

Franceschi, D. S. A., Tsuneto, L. T., Mazini, P. S., Sacramento, W. S. do, Reis, P. G., Rudnick, C. C. C., Clementino, S. L., Sell, A. M., & Visentainer, J. E. L. (2011). Class-I human leukocyte alleles in leprosy patients from Southern Brazil. *Revista Da Sociedade Brasileira de Medicina Tropical*, 44(5), 616–620. <https://doi.org/10.1590/S0037-86822011000500018>

Garcia, E., Chain, P., Worsham, P., Bearden, S. W., Malfatti, S., Lang, D., Larimer, F., & Lindler, L. (2007). *Pestoides F*, an Atypical *Yersinia pestis* Strain from the Former Soviet Union (pp. 17–22). https://doi.org/10.1007/978-0-387-72124-8_2

Gopalakrishnan, S., Pandey, N., Tamiz, A., Vere, J., Carrasco, R., Somerville, R., Tripathi, A., Ginski, M., Paterson, B., & Alkan, S. (2009). Mechanism of action of ZOT-derived peptide AT-1002, a tight junction regulator and absorption enhancer. *International Journal of Pharmaceutics*, 365(1–2). <https://doi.org/10.1016/j.ijpharm.2008.08.047>

Gorodezky, C., Alaez, C., Munguía, A., Cruz, R., Vazquez, A., Camacho, A., Flores, O., Rodriguez, M., & Rodriguez, O. (2004). Molecular mechanisms of MHC linked susceptibility

in leprosy: towards the development of synthetic vaccines. *Tuberculosis*, 84(1–2), 82–92. <https://doi.org/10.1016/j.tube.2003.08.009>

Grainger, I., Hawkins, D., Cowal, L., & Mikulski, R. (2008). *The Black Death Cemetery, East Smithfield*. Museum of London Archaeology Service.

Gunawan, H., Achdiat, P. A., Rahardjo, R. M., Hindritiani, R., & Suwarsa, O. (2020). Frequent testicular involvement in multibacillary leprosy. *International Journal of Infectious Diseases*, 90, 60–64. <https://doi.org/10.1016/j.ijid.2019.10.013>

Haapasalo, K., Koskinen, L. L. E., Suvilehto, J., Jousilahti, P., Wolin, A., Suomela, S., Trembath, R., Barker, J., Vuopio, J., Kere, J., Jokiranta, T. S., & Saavalainen, P. (2018). The Psoriasis Risk Allele *HLA-C*06:02* Shows Evidence of Association with Chronic or Recurrent Streptococcal Tonsillitis. *Infection and Immunity*, 86(10), e00304–18. <https://doi.org/10.1128/IAI.00304-18>

Haller, M., Callan, K., Susat, J., Flux, A. L., Immel, A., Franke, A., Herbig, A., Krause, J., Kupczok, A., Fouquet, G., Hummel, S., Rieger, D., Nebel, A., & Krause-Kyora, B. (2021). Mass burial genomics reveals outbreak of enteric paratyphoid fever in the Late Medieval trade city Lübeck. *IScience*, 24(5), 102419. <https://doi.org/10.1016/j.isci.2021.102419>

Hartle, R. (2017). *The New Churchyard: From Moorfields Marsh to Bethlem Burial Ground, Brokers Row and Liverpool Street (MOLA/Crossrail Archaeology 10)*. Museum of London Archaeology.

Havlicek, J., & Roberts, S. C. (2009). MHC-correlated mate choice in humans: A review. *Psychoneuroendocrinology*, 34(4), 497–512. <https://doi.org/10.1016/j.psyneuen.2008.10.007>

Hendrickson, R. C., Wang, C., Hatcher, E. L., & Lefkowitz, E. J. (2010). Orthopoxvirus genome evolution: The role of gene loss. *Viruses* 2(9), 1933–1967. <https://doi.org/10.3390/v2091933>

Hu, X., Deutsch, A. J., Lenz, T. L., Onengut-Gumuscu, S., Han, B., Chen, W.-M., Howson, J. M. M., Todd, J. A., de Bakker, P. I. W., Rich, S. S., & Raychaudhuri, S. (2015). Additive and interaction effects at three amino acid positions in HLA-DQ and HLA-DR molecules drive type 1 diabetes risk. *Nature Genetics*, 47(8), 898–905. <https://doi.org/10.1038/ng.3353>

Hviid, T. V. F., & Christiansen, O. B. (2005). Linkage Disequilibrium Between Human Leukocyte Antigen (HLA) Class II and HLA-G—Possible Implications for Human Reproduction and Autoimmune Disease. *Human Immunology*, 66(6), 688–699. <https://doi.org/10.1016/j.humimm.2005.03.003>

Immel, A., Key, F. M., Szolek, A., Barquera, R., Robinson, M. K., Harrison, G. F., Palmer, W. H., Spyrou, M. A., Susat, J., Krause-Kyora, B., Bos, K. I., Forrest, S., Hernández-Zaragoza, D. I., Sauter, J., Solloch, U., Schmidt, A. H., Schuenemann, V. J., Reiter, E., Kairies, M. S., ... Krause, J. (2021). Analysis of Genomic DNA from Medieval Plague Victims Suggests Long-Term Effect of *Yersinia pestis* on Human Immunity Genes. *Molecular Biology and Evolution*, 38(10), 4059–4076. <https://doi.org/10.1093/molbev/msab147>

Jia, L., Xie, J., Zhao, J., Cao, D., Liang, Y., Hou, X., Wang, L., & Li, Z. (2017). Mechanisms of Severe Mortality-Associated Bacterial Co-infections Following Influenza Virus Infection. *Frontiers in Cellular and Infection Microbiology*, 7, 388. <https://doi.org/10.3389/fcimb.2017.00338>

Joko, S., Numaga, J., Kawashima, H., Namisato, M., & Maeda, H. (2000). Human leukocyte antigens in forms of leprosy among Japanese patients. *International Journal of Leprosy and Other Mycobacterial Diseases*, 68(1), 49–56.

Joyce, M. P. (2012). Historic aspects of human susceptibility to leprosy and the risk of conjugal transmission. *Memórias Do Instituto Oswaldo Cruz*, 107(suppl 1), 17–21. <https://doi.org/10.1590/S0074-02762012000900004>

Karlsson, E. K., Kwiatkowski, D. P., & Sabeti, P. C. (2014). Natural selection and infectious disease in human populations. *Nature Reviews Genetics*, 15(6), 379–393. <https://doi.org/10.1038/nrg3734>

Karyekar, C. S., Fasano, A., Raje, S., Lu, R., Dowling, T. C., & Eddington, N. D. (2003). Zonula Occludens Toxin Increases the Permeability of Molecular Weight Markers and Chemotherapeutic Agents Across the Bovine Brain Microvessel Endothelial Cells. *Journal of Pharmaceutical Sciences*, 92(2), 414–423. <https://doi.org/10.1002/jps.10310>

Keeling, M. J., & Gilligan, C. A. (2000). Bubonic plague: a metapopulation model of a zoonosis. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 267(1458), 2219–2230. <https://doi.org/10.1098/rspb.2000.1272>

Khan, M. F., & Wang, H. (2020). Environmental Exposures and Autoimmune Diseases: Contribution of Gut Microbiome. *Frontiers in Immunology*, 10, 3094. <https://doi.org/10.3389/fimmu.2019.03094>

Khanna, N., Singh, M., Rasool, S., Ammini, A., Bhatla, N., Garg, V., Rao, S., & Bhattacharya, S. N. (2014). Menstrual irregularities, fertility status, and ovarian function in female patients with leprosy in India. *International Journal of Dermatology*, 53(9), 1114–1118. <https://doi.org/10.1111/j.1365-4632.2012.05771.x>

Kiseleva, A. N., Butina, E. v., Isaeva, N. v., Zaitseva, G. A., Pozdeev, N. M., & Ovchinnikov, V. v. (2019). Distribution of antigens of the HLA-system in married couples with reproductive disorders. *Obstetrics, Gynecology and Reproduction*, 13(2), 111–118. <https://doi.org/10.17749/2313-7347.2019.13.2.111-118>

Klioze, A. M., & Ramos-Caro, F. A. (2000). Visceral leprosy. In *International Journal of Dermatology*, 39(9), 641–658. <https://doi.org/10.1046/j.1365-4362.2000.00860.x>

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, 1569. <https://doi.org/10.1038/s41467-018-03857-x>

- Krawinkel, M. B. (2012). Interaction of Nutrition and Infections Globally: An Overview. *Annals of Nutrition and Metabolism*, 61(Suppl. 1), 39–45. <https://doi.org/10.1159/000345162>
- Kular, L., Liu, Y., Ruhrmann, S., Zheleznyakova, G., Marabita, F., Gomez-Cabrero, D., James, T., Ewing, E., Lindén, M., Górnikiewicz, B., Aeinehband, S., Stridh, P., Link, J., Andlauer, T. F. M., Gasperi, C., Wiendl, H., Zipp, F., Gold, R., Tackenberg, B., ... Jagodic, M. (2018). DNA methylation as a mediator of HLA-DRB1*15:01 and a protective variant in multiple sclerosis. *Nature Communications*, 9(1), 2397. <https://doi.org/10.1038/s41467-018-04732-5>
- Kwon, O. J., Karni, A., Israel, S., Brautbar, C., Amar, A., Meiner, Z., Abramsky, O., & Karussis, D. (1999). HLA class II susceptibility to multiple sclerosis among Ashkenazi and non-Ashkenazi Jews. *Archives of Neurology* 56(5), 555-560. <https://doi.org/10.1001/archneur.56.5.555>
- Lathem, W. W., Schroeder, J. A., Bellows, L. E., Ritzert, J. T., Koo, J. T., Price, P. A., Caulfield, A. J., & Goldman, W. E. (2014). Posttranscriptional Regulation of the *Yersinia pestis* Cyclic AMP Receptor Protein Crp and Impact on Virulence. *MBio*, 5(1), e01038-13. <https://doi.org/10.1128/mBio.01038-13>
- Lei, C., & Kumar, S. (2022). *Yersinia pestis* antibiotic resistance: a systematic review. *Osong Public Health and Research Perspectives*, 13(1), 24–36. <https://doi.org/10.24171/j.phrp.2021.0288>
- Leslie, T., Whitehouse, C. A., Yingst, S., Baldwin, C., Kakar, F., Mofleh, J., Hami, A. S., Mustafa, L., Omar, F., Ayazi, E., Rossi, C., Noormal, B., Ziar, N., & Kakar, R. (2011). Outbreak of gastroenteritis caused by *Yersinia pestis* in Afghanistan. *Epidemiology and Infection*, 139(5), 728-735. <https://doi.org/10.1017/S0950268810001792>
- Mahendran, V., Tan, Y. S., Riordan, S. M., Grimm, M. C., Day, A. S., Lemberg, D. A., Octavia, S., Lan, R., & Zhang, L. (2013). The Prevalence and Polymorphisms of Zonula Occluden Toxin Gene in Multiple *Campylobacter concisus* Strains Isolated from Saliva of Patients with Inflammatory Bowel Disease and Controls. *PLoS ONE*, 8(9), e75525. <https://doi.org/10.1371/journal.pone.0075525>
- Malkova, A., Starshinova, A., Zinchenko, Y., Basantsova, N., Mayevskaya, V., Yablonskiy, P., & Shoenfeld, Y. (2020). The opposite effect of human leukocyte antigen genotypes in sarcoidosis and tuberculosis: a narrative review of the literature. *ERJ Open Research*, 6, 00155. <https://doi.org/10.1183/23120541.00155-2020>
- 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 1 st –2 nd century CE southern Italy. *Current Biology*, 26(23), R1220–R1222. <https://doi.org/10.1016/j.cub.2016.10.016>
- McClure, H. M., Chiodini, R. J., Anderson, D. C., Swenson, R. B., Thayer, W. R., & Coutu, J. A. (1987). *Mycobacterium paratuberculosis* Infection in a Colony of Stumptail Macaques (*Macaca arctoides*). *Journal of Infectious Diseases*, 155(5), 1011–1019. <https://doi.org/10.1093/infdis/155.5.1011>

Mora, C., McKenzie, T., Gaw, I. M., Dean, J. M., von Hammerstein, H., Knudson, T. A., Setter, R. O., Smith, C. Z., Webster, K. M., Patz, J. A., & Franklin, E. C. (2022). Over half of known human pathogenic diseases can be aggravated by climate change. *Nature Climate Change*. <https://doi.org/10.1038/s41558-022-01426-1>

Morris, D. L., Taylor, K. E., Fernando, M. M. A., Nititham, J., Alarcón-Riquelme, M. E., Barcellos, L. F., Behrens, T. W., Cotsapas, C., Gaffney, P. M., Graham, R. R., Pons-Estel, B. A., Gregersen, P. K., Harley, J. B., Hauser, S. L., Hom, G., Langefeld, C. D., Noble, J. A., Rioux, J. D., Seldin, M. F., ... Vyse, T. J. (2012). Unraveling multiple MHC gene associations with systemic lupus erythematosus: Model choice indicates a role for HLA alleles and non-HLA genes in europeans. *American Journal of Human Genetics* 91(5), 778-793. <https://doi.org/10.1016/j.ajhg.2012.08.026>

Mühlemann, B., Vinner, L., Margaryan, A., Wilhelmson, H., Castro, C. D. L. F., Allentoft, M. E., Damgaard, P. D. B., Hansen, A. J., Nielsen, S. H., Strand, L. M., Bill, J., Buzhilova, A., Pushkina, T., Falys, C., Khartanovich, V., Moiseyev, V., Jørkov, M. L. S., Sørensen, P. Ø., Magnusson, Y., ... Sikora, M. (2020). Diverse variola virus (smallpox) strains were widespread in northern Europe in the Viking Age. *Science*, 369(6502), eaaw8977. <https://doi.org/10.1126/science.aaw8977>

Nakken, B., Jonsson, R., Brokstad, K. A., Omholt, K., Nerland, A. H., Haga, H. J., & Halse, A. K. (2001). Associations of MHC class II alleles in Norwegian primary Sjögren's syndrome patients: Implications for development of autoantibodies to the Ro52 autoantigen. *Scandinavian Journal of Immunology* 54(4), 428-433. <https://doi.org/10.1046/j.1365-3083.2001.00993.x>

Newfield, T. P. (2017). Malaria and malaria-like disease in the early Middle Ages. *Early Medieval Europe*, 25(3), 251–300. <https://doi.org/10.1111/emed.12212>

Ng, M. H. L., Lau, K., Li, L., Cheng, S., Chan, W. Y., Hui, P. K., Zee, B., Leung, C., & Sung, J. J. Y. (2004). Association of Human-Leukocyte-Antigen Class I (B*0703) and Class II (DRB1*0301) Genotypes with Susceptibility and Resistance to the Development of Severe Acute Respiratory Syndrome. *The Journal of Infectious Diseases*, 190(3), 515–518. <https://doi.org/10.1086/421523>

Nguyen, V. K., Parra-Rojas, C., & Hernandez-Vargas, E. A. (2018). The 2017 plague outbreak in Madagascar: Data descriptions and epidemic modelling. *Epidemics*, 25, 20–25. <https://doi.org/10.1016/j.epidem.2018.05.001>

O'Neil, Y. V. (1993). Diseases of the Middle Ages. In K. Kiple (Ed.), *The Cambridge World History of Human Disease*. Cambridge University Press. <https://doi.org/10.1017/CHOL9780521332866.031>

Penn, D. J., Damjanovich, K., & Potts, W. K. (2002). MHC heterozygosity confers a selective advantage against multiple-strain infections. *Proceedings of the National Academy of Sciences*, 99(17), 11260–11264. <https://doi.org/10.1073/pnas.162006499>

Pérez, V., Marín, J. F. G., & Badiola, J. J. (1996). Description and classification of different types of lesion associated with natural paratuberculosis infection in sheep. *Journal of Comparative Pathology*, 114(2), 107–122. [https://doi.org/10.1016/S0021-9975\(96\)80001-6](https://doi.org/10.1016/S0021-9975(96)80001-6)

Pfizenmaier, S. (2016). *Charterhouse Square: Black Death Cemetery and Carthusian Monastery, Meat Market and Suburb (MOLA/Crossrail Archaeology 7)*. Museum of London Archaeology.

Pfrengele, S., Neukamm, J., Guellil, M., Keller, M., Molak, M., Avanzi, C., Kushniarevich, A., Montes, N., Neumann, G. U., Reiter, E., Tukhbatova, R. I., Berezina, N. Y., Buzhilova, A. P., Korobov, D. S., Suppersberger Hamre, S., Matos, V. M. J., Ferreira, M. T., González-Garrido, L., Wasterlain, S. N., ... Schuenemann, V. J. (2021). Mycobacterium leprae diversity and population dynamics in medieval Europe from novel ancient genomes. *BMC Biology*, 19(1), 220. <https://doi.org/10.1186/s12915-021-01120-2>

Prechel, M. (2002). Eine Lübecker Population von 1350 – Krankheiten und Mangelerscheinungen. In M. Gläser (Ed.), *Lübecker Schriften zu Archäologie und Kulturgeschichte* 26. Dr. Rudolf Habelt.

Prugnotte, F., Manica, A., Charpentier, M., Guégan, J. F., Guernier, V., & Balloux, F. (2005). Pathogen-Driven Selection and Worldwide HLA Class I Diversity. *Current Biology*, 15(11), 1022–1027. <https://doi.org/10.1016/j.cub.2005.04.050>

Raffcliffe, C. (2006). *Leprosy in Medieval England*. Boydell Press.

Rajanna, C., Revazishvili, T., Rashid, M. H., Chubinidze, S., Bakanidze, L., Tsanova, S., Imnadze, P., Bishop-Lilly, K. A., Sozhamannan, S., Gibbons, H. S., Morris, J. G., & Sulakvelidze, A. (2010). Characterization of pPCP1 Plasmids in *Yersinia pestis* Strains Isolated from the Former Soviet Union. *International Journal of Microbiology*, 2010, 1–9. <https://doi.org/10.1155/2010/760819>

Retief, F., & Cilliers, L. (2010). Measles in antiquity and the middle ages. *SAMJ: South African Medical Journal*, 100(4), 216–217.

Rossmann, M. D., Thompson, B., Frederick, M., Maliarik, M., Iannuzzi, M. C., Rybicki, B. A., Pandey, J. P., Newman, L. S., Magira, E., Beznik-Cizman, B., & Monos, D. (2003). HLA-DRB1*1101: A Significant Risk Factor for Sarcoidosis in Blacks and Whites. *The American Journal of Human Genetics*, 73(4), 720–735. <https://doi.org/10.1086/378097>

Sales, A. M., Ponce de Leon, A., Düppre, N. C., Hacker, M. A., Nery, J. A. C., Sarno, E. N., & Penna, M. L. F. (2011). Leprosy among Patient Contacts: A Multilevel Study of Risk Factors. *PLoS Neglected Tropical Diseases*, 5(3), e1013. <https://doi.org/10.1371/journal.pntd.0001013>

Sanchez-Mazas, A. (2020). A review of HLA allele and SNP associations with highly prevalent infectious diseases in human populations. *Swiss Medical Weekly*. <https://doi.org/10.4414/smw.2020.20214>

Sapone, A., de Magistris, L., Pietzak, M., Clemente, M. G., Tripathi, A., Cucca, F., Lampis, R., Kryszak, D., Carteni, M., Generoso, M., Iafusco, D., Prisco, F., Laghi, F., Riegler, G., Carratu, R., Counts, D., & Fasano, A. (2006). Zonulin Upregulation Is Associated With Increased Gut Permeability in Subjects With Type 1 Diabetes and Their Relatives. *Diabetes*, 55(5), 1443–1449. <https://doi.org/10.2337/db05-1593>

- Saporta, L., & Yuksel, A. (1994). Androgenic status in patients with lepromatous leprosy. *British Journal of Urology*, 74(2), 221–224. <https://doi.org/10.1111/j.1464-410X.1994.tb16590.x>
- 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), e1006997. <https://doi.org/10.1371/journal.ppat.1006997>
- Schurr, E., & Gros, P. (2009). A Common Genetic Fingerprint in Leprosy and Crohn's Disease? *New England Journal of Medicine*, 361(27), 2666–2668. <https://doi.org/10.1056/NEJMe0910690>
- Shankarkumar, U. (2003). HLA-DRB1*03 and DQB1*0302 associations in a subset of patients severely affected with systemic lupus erythematosus from western India. *Annals of the Rheumatic Diseases*, 62(1), 92–93. <https://doi.org/10.1136/ard.62.1.92>
- Shankarkumar, U., Ghosh, K., Badakere, S., & Mohanty, D. (2003). Novel HLA Class I Alleles Associated with Indian Leprosy Patients. *Journal of Biomedicine and Biotechnology*, 2003(3), 208–211. <https://doi.org/10.1155/S1110724303210019>
- Shevryev, D., Tereshchenko, V., & Kozlov, V. (2021). Immune Equilibrium Depends on the Interaction Between Recognition and Presentation Landscapes. *Frontiers in Immunology*, 12, 706136. <https://doi.org/10.3389/fimmu.2021.706136>
- Sironi, M., & Clerici, M. (2010). The hygiene hypothesis: an evolutionary perspective. *Microbes and Infection*, 12(6), 421–427. <https://doi.org/10.1016/j.micinf.2010.02.002>
- Schmit, N., Nayagam, S., Thursz, M. R., & Hallett, T. B. (2021). The global burden of chronic hepatitis B virus infection: comparison of country-level prevalence estimates from four research groups. *International Journal of Epidemiology*, 50(2), 560–569. <https://doi.org/10.1093/ije/dyaa253>
- Smiley, S. T., Szaba, F. M., Kummer, L. W., Duso, D. K., & Lin, J.-S. (2019). *Yersinia pestis* Pla Protein Thwarts T Cell Defense against Plague. *Infection and Immunity*, 87(5), e00126–19. <https://doi.org/10.1128/IAI.00126-19>
- Snellgrove, H. (1954). Leprosy in ancient and early medieval times: with especial reference to the Franks. *The Mississippi Quarterly*, 7(4), 1–10.
- Sodeinde, O. A., Subrahmanyam, Y. V. B. K., Stark, K., Quan, T., Bao, Y., & Goguen, J. D. (1992). A Surface Protease and the Invasive Character of Plague. *Science*, 258(5084), 1004–1007. <https://doi.org/10.1126/science.1439793>
- 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>

Stenseth, N. Chr., Samia, N. I., Viljugrein, H., Kausrud, K. L., Begon, M., Davis, S., Leirs, H., Dubyanskiy, V. M., Esper, J., Ageyev, V. S., Klassovskiy, N. L., Pole, S. B., & Chan, K.-S. (2006). Plague dynamics are driven by climate variation. *Proceedings of the National Academy of Sciences*, 103(35), 13110–13115. <https://doi.org/10.1073/pnas.0602447103>

Stoakley, M. (2019). ‘Great fears of the sickness here in the City ... God preserve us all ...’. A Plague Burial Ground in Leith, 1645: An Archaeological Excavation at St Mary’s (Leith) RC Primary School, Leith Links, Edinburgh. *Scottish Archaeological Internet Reports*, 86, 1–82.

van Gerven, N. M. F., de Boer, Y. S., Zwiers, A., Verwer, B. J., Drenth, J. P. H., van Hoek, B., van Erpecum, K. J., Beuers, U., van Buuren, H. R., den Ouden, J. W., Verdonk, R. C., Koek, G. H., Brouwer, J. T., Guichelaar, M. M. J., Vrolijk, J. M., Coenraad, M. J., Kraal, G., Mulder, C. J. J., van Nieuwkerk, C. M. J., ... Verhagen, M. A. M. T. (2015). HLA-DRB1*03:01 and HLA-DRB1*04:01 modify the presentation and outcome in autoimmune hepatitis type-1. *Genes and Immunity*, 16, 247–252. <https://doi.org/10.1038/gene.2014.82>

Vanderborght, P. R., Pacheco, A. G., Moraes, M. E., Antoni, G., Romero, M., Verville, A., Thai, V. H., Huong, N. T., Ba, N. N., Schurr, E., Sarno, E. N., & Moraes, M. O. (2007). HLA-DRB1*04 and DRB1*10 are associated with resistance and susceptibility, respectively, in Brazilian and Vietnamese leprosy patients. *Genes & Immunity*, 8(4), 320–324. <https://doi.org/10.1038/sj.gene.6364390>

Verbrugghe, E., Boyen, F., Gaastra, W., Bekhuis, L., Leyman, B., van Parys, A., Haesebrouck, F., & Pasmans, F. (2012). The complex interplay between stress and bacterial infections in animals. *Veterinary Microbiology*, 155(2–4), 115–127. <https://doi.org/10.1016/j.vetmic.2011.09.012>

Vogler, A. J., Chan, F., Nottingham, R., Andersen, G., Drees, K., Beckstrom-Sternberg, S. M., Wagner, D. M., Chanteau, S., & Keim, P. (2013). A Decade of Plague in Mahajanga, Madagascar: Insights into the Global Maritime Spread of Pandemic Plague. *MBio*, 4(1), e00623–12. <https://doi.org/10.1128/mBio.00623-12>

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](https://doi.org/10.1016/S1473-3099(13)70323-2)

Waldron, T. (1991). Rates for the job. Measures of disease frequency in palaeopathology. *International Journal of Osteoarchaeology*, 1(1), 17–25. <https://doi.org/10.1002/oa.1390010104>

Walker, S. L., & Lockwood, D. N. J. (2007). Leprosy. *Clinics in Dermatology*, 25(2), 165–172. <https://doi.org/10.1016/j.clindermatol.2006.05.012>

Wang, K., Baldassano, R., Zhang, H., Qu, H.-Q., Imielinski, M., Kugathasan, S., Annese, V., Dubinsky, M., Rotter, J. I., Russell, R. K., Bradfield, J. P., Sleiman, P. M. A., Glessner, J. T., Walters, T., Hou, C., Kim, C., Frackelton, E. C., Garriss, M., Doran, J., ... Hakonarson, H. (2010). Comparative genetic analysis of inflammatory bowel disease and type 1 diabetes

implicates multiple loci with opposite effects. *Human Molecular Genetics*, 19(10), 2059–2067. <https://doi.org/10.1093/hmg/ddq078>

Wang, S. F., Chen, K. H., Chen, M., Li, W. Y., Chen, Y. J., Tsao, C. H., Yen, M. Y., Huang, J. C., & Chen, Y. M. A. (2011). Human-leukocyte antigen class I Cw 1502 and Class II DR 0301 genotypes are associated with resistance to severe acute respiratory syndrome (SARS) infection. *Viral Immunology*, 24(5). <https://doi.org/10.1089/vim.2011.0024>

Willmott, H., Townend, P., Swales, D. M., Poinar, H., Eaton, K., & Klunk, J. (2020). A Black Death mass grave at Thornton Abbey: the discovery and examination of a fourteenth-century rural catastrophe. *Antiquity*, 94(373), 179–196. <https://doi.org/10.15184/aqy.2019.213>

Zerva, L., Cizman, B., Mehra, N., Alahari, S., Murali, R., Zmijewski, C., Kamoun, M., & Monos, D. (1996). Arginine at positions 13 or 70–71 in pocket 4 of HLA-DRB1 alleles is associated with leprosy susceptibility to tuberculoid leprosy. *Journal of Experimental Medicine*, 183(3), 829–836.

Zhang, F., Liu, H., Chen, S., Wang, C., Zhu, C., Zhang, L., Chu, T., Liu, D., Yan, X., & Liu, J. (2009). Evidence for an association of HLA-DRB1*15 and DRB1*09 with leprosy and the impact of DRB1*09 on disease onset in a Chinese Han population. *BMC Medical Genetics*, 10(1), 133. <https://doi.org/10.1186/1471-2350-10-133>

Zimble, D. L., Schroeder, J. A., Eddy, J. L., & Lathem, W. W. (2015). Early emergence of *Yersinia pestis* as a severe respiratory pathogen. *Nature Communications*, 6(1), 7487. <https://doi.org/10.1038/ncomms8487>

Declaration

Herewith, I confirm that the submitted dissertation is completely the result of my own work. Apart from the advice of my supervisors, all sources and cooperation partners are listed within the dissertation. This dissertation has not been submitted elsewhere. It has been carried out in strict accordance with the rules of Good Scientific Practice of the *Deutsche Forschungsgesellschaft*. In addition, no academic degree has ever been withdrawn.



Acknowledgements

Where do I even begin? It has been quite a long journey and the road was bumpy, to say the least. When I arrived in Kiel in the middle of September 2018, the nights were long and the days – windy. I still remember my tiny room with just a mattress on the floor. I was fearful, but also excited for the next chapter in my life. Although everything was foreign to me, I felt ready for the challenge. At the beginning of this adventure, I believed that pursuing a PhD, in a field completely out of my expertise at the time, would be the most difficult thing I will have to face in the next three years. Oh, how naïve of me! Life set up a few more obstacles on my way to this day. Although my quest is far from over, I reached an important milestone and it would not have been possible without the support from my family, friends and colleagues.

First and foremost, I want to thank my mother. For almost thirty years, you worked tirelessly to provide me with a fighting chance at a successful and fulfilling life. Thank you for believing in me and supporting all my strange ideas, the good and the not-so-good ones. After all, I survived all of them, which allowed me to learn and grow. I also want to thank my grandparents who have been there for me every step of the way.

I would like to express my appreciation to my first supervisor Prof. Dr. Manuela Dittmar for her support and encouragement.

Special thanks go to Prof. Dr. Almut Nebel, not only for the support and guidance that were essential to my research work, but also for her patience and understanding in times of struggle.

Many thanks to Prof. Dr. Ben Krause-Kyora and Prof. Dr. Tobias Lenz for their scientific advice and support, as well as to Dr. Daniel Unterweger who often helped me see the mountaintop on a cloudy day.

I can say with full confidence that this work would not have been possible without the contributions of my dear colleagues: Magdalena Haller, Dr. Katharina Fuchs, Dr. Alexander Immel, Dr. Janina Dose, Nicolas da Silva and – of course! – Dr. Julian Susat who patiently

taught me everything I needed to know to carry out bioinformatic analyses of bacterial genomes. Thanks a lot!

Jordan Romeyer Dherbey, your love and encouragement unveiled the hidden amounts of motivation within me that were so very needed in the final phase of my PhD. Thank you for believing in me, even when I did not.

Last but not least, I want to thank my friends who were strong for me in scary moments, patient with me on bad days and happy for me in joyful times.

~Joanna♥