

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 Leukozytenantigenensystem (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.