Molecular Archaeoparasitology as a novel tool for the study of trading and migration networks through history

Patrik G Flammer
Research Laboratory for Archaeology and the History of Art
Department of Zoology

Merton College

Thesis submitted for the degree of a
Doctor of Philosophy

Trinity 2014

Declaration

I, Patrik Guido Flammer, declare that this thesis has never been submitted, either in the same or a different form, to this or any other university, for a degree or any other qualification. All the work presented in this thesis is the result of my own work unless otherwise stated and was carried out while I was registered as a graduate student at the Research Laboratory of Archaeology and the History of Art and the Department of Zoology, University of Oxford, under the supervision of Prof. A. Mark Pollard and Dr. Adrian L. Smith

Oxford, Trinity Term 2014

Table of Contents

Chapter Four: Medieval Trading Port of Bristol

Chapter Five: Lübeck – Centre of the Hanseatic Trade

Chapter Six: Neolithic parasites from early settlement and grave deposits

Chapter Seven: Slavic - Anglo-Saxon – Viking interactions and considerations on graveyards as an approach for individual based epidemiology

List of Figures

Abstract: Molecular Archaeoparasitology as novel tool for the study of trading and migration networks through history

This project represents the first comprehensive study applying molecular and genetic methods to study historical contexts such as migration and trade based on human parasites. Using specially developed techniques, the study focused on parasites with minor symptoms which allowed the infected person to go about their daily business. The combination of state of the art techniques in archaeology, molecular methods and phylogenetic analysis enabled us to develop a novel powerful tool to study historic events. Diseases have a considerable impact on societies. Various publications indicate that human intestinal parasites are commonly found in a variety of archaeological contexts, including latrines, graves and mummies. These parasites can be detected by microscopy which focuses the work on samples which do close association to humans; widespread prevalence and the possibility for reliable microscopic diagnostics suggest that these parasites are an attractive study system for human activities. Infectious diseases have a much short generation time which offers greater opportunity to track historical events at higher resolution. Looking at a range of human parasites, their different life-cycles allowed insight into various aspects of human culture, comparing different origins of the samples allows an estimation of the epidemiological burden of ancient populations. Application of a parallel sequencing approach (MiSeq) enabled building a comprehensive database of sequences from various archaeological sites dating as far back as 3630 BCE. In-depth phylogenetic analysis reveals patterns in the genetic signatures of both coding and non-coding genetic regions, taking various levels of selective pressure into account. This project has produced the oldest pathogen sequence and the most comprehensive database of ancient pathogen sequences.

Acknowledgements

I was fortunate to be given this great opportunity to work on a unique project which I could shape to make the best of it from the very beginning. It has given me many long days and sleepless nights, but in the end I am grateful and proud of what I could submit as a doctoral thesis. This would not have been possible without the help of many people and I would like to thank them all for their help to make this all possible.

Special thanks goes to my two supervisors, Adrian Smith at the Department of Zoology and Mark Pollard at the Research Laboratory for Archaeology and the History of Art, for giving me the chance – against many odds and sometimes better judgement – for supporting and discussing the ideas on which this unusual project was founded and encouraging me to get through difficult stages of the project.

I thank all current and past members of the Smith Lab in the Department of Zoology for support and patience with me, namely Kirstin Bilham, Stephen Georgiou, Britta Bia, Amy Hong, Amy Boyd, Sylvia Warren, Kirstie McLoughlin, Steven Fiddaman and Stephen Preston, with special thanks to Stephen for his support developing the MiSeq procedure and analysing the resulting data.

On technical issues of my project I could always rely on helpful people. I would like to mention especially Cedric Tan for his help with statistics, Oliver Pybus and Simon Dellicour for their help and support on bioinformatics and phylogenetic analysis.

I would also like to thank all collaborators, especially Dirk Rieger and the team of excavators in Lübeck; Rebecca Nicholson, Vix Hughes and colleagues at Oxford Archaeology; Niels Bleicher and his team at Archaeology in the Amt für Städtebau in Zurich; Joachim Wahl, Helmut Schlichterle and Manuela Fischer at the Landesdenkmalamt Baden-Württemberg; Renáta Přichystalová, Jiri Machacek and Petra Urbanova at the Masaryk University in Brno, Jan Bill and Christian Løchsen Rødsrud at the Cultural History Museum of the University of Oslo; Andrew Jones at the York Archaeological Trust; Annine Moltsen at Natur og Kultur in Copenhagen; and Gunita Zarina at the Latvian University in Riga

The major part of this research was funded by the Swiss National Science Foundation (SNSF) through a Graduate Fellowship. Additional financial support was received from the Boise Fund (Boise Fund Research Award), the Meyerstein Fund and the Graduate Research Award at the School of Archaeology and the Bowman Fund at Merton College.

My family, especially my parents, have been an invaluable support during this thesis and for their support I am deeply grateful.
Ich möchti mich da ganz herzlich bi minere Familie für d Hilf und's Vertraue i mini Arbet bedanke. Es wäri für mich ned immer wiiter gange ohni euri Unterstützig.

My experience of Oxford would not have been the same without the friends I made here who made my time in Oxford very special, and the good times are remembered in the lasting friendship with those who live all over the world. Although I cannot name everybody here I would like to mention some special people who helped me

through difficult times and were always ready to help. First, I would like to thank Stephanie Jones with whom I went through a lot together from matriculation to submission. We faced the same tight deadline for submission, but together we grew stronger and made the stress bearable and the result even more rewarding. We shared many laughs and tears which made our time together unforgettable. Further, I would like to thank these people who supported me through everthing: Juha Huiskonen, Verena Schenzinger, Lutz Warnke, Andreas Düring, Julia Steinberg, David Bitto, Phil Entwistle, Amber and Edmund Hood Highcock, James Moxness, Francesco Fermani, Wolf Rittershofer, Claire Higgins, and Daniel Lavenda.

Chapter 1

Introduction

This doctoral thesis presents an extension of the field of Archaeoparasitology, adding new tools into the archaeologist's toolbox, but also provides insight into the genetic diversity of human parasites over time.

1.1 Impact of diseases on human history – an overview

Diseases have the potential to change human lives and have had a considerable impact on society. Infectious diseases are an integral part of the life and historic records document the recognition of infections as a burden to humans. There are numerous examples throughout history where diseases had devastating effects on mankind. Some outbreaks of diseases resulted in death tolls of millions of people in relatively short times, such as the Black Death, the Spanish Flu, or the recent outbreak of Ebola in Africa. Infections can be asymptomatic, even those causing severe disease such as in the case of Typhoid Mary spreading the pathogen associated with typhoid fever (*Salmonella enterica* subsp. *enterica*, serovar Typhi), and might and have been circulating in the human population for a long time. Different archaeological and historical records showed evidence of tuberculosis (TB), malaria, leprosy and other diseases (Roberts and Manchester 2010; Taylor et al. 2000). Infections studied in this project do not cause rapid health decay and infected individuals might thus not even have recognised the infection.

Diseases, especially fast spreading or ones with devastating effects have been the focus of many studies in archaeology, as there is more historic evidence on diseases which were visible. Although disease causing factors were not commonly identified, counter-measures were taken to limit the spread to and within cities, especially for fast spreading diseases such as the Black Death which came to Europe from the Far East and spread rapidly. Countermeasures such as the introduction of

quarantine (from Italian '*quarantina di giorni*', forty days) did limit but not stop the spread as increased international travel and trade aided the dispersal of diseases (Harrison 2012). This is probably the first example of a disease being spread around the world by the globalisation of trade (Rudbeck et al. 2005).

1.2 Archaeological research – a short introduction to techniques

Archaeology has always been an interdisciplinary field applying methods from other fields to interrogate the evidence found. For the purpose of this overview, archaeological remains will be classified in three categories which do overlap for some remains. All these remains can be used to address specific questions relating to their origin and the culture linked to them. Manmade artefacts offer the most direct insight into the culture of the producer. Human remains give different perspectives on the life and the culture based on diet, diseases and trauma. Environmental samples, such as pollen, seeds or volcanic ash, are less influenced by the human population, thus providing a valuable source for archaeological research.

1.2.1 Artefacts in Archaeology

Every manmade artefact tells a story based on materials used, purpose of the item and craftsmanship. Artefacts are used to get an impression on the excavated site. Some examination techniques are not based on elaborate technology, but are based on collections of similar finds which help to characterise a specific item. Some of the samples in this study relied on artefacts for the identification of the cultural context and the approximation of the dates. Pottery is a commonly used artefact, for example

in this study in context with the Neolithic burials in Bad Mergentheim (DE) which were associated with the Corded Ware Culture named after a characteristic style of pottery (C. Meyer et al. 2009). Architectural styles and features are important cultural artefacts as they give evidence on a permanent settlement and the development of it. Several archaeological remains in this study were dated or located using architectural features, such as the stilt settlements from Zurich (Neolithic stilt settlements) or the Viking houses in Viborg (Fruergaard and Moltsen 2005). These are just some examples on the application of artefact identification and dating. Depending on the nature of the artefacts, other methods such as dating, genetic methods (e.g. on animal bone tools) or metallurgical evaluations can be applied.

1.2.2 Environmental evidence

Environmental evidence may be influenced by the human population and are thus informative on the living conditions and cultural practices of the population. Cultural practices, such as preferences for specific plants for food or medical purposes, would influence the spread of seeds or pollen. The study of pollen (palynology) and remains of seeds and plant material in general can give valuable insight into dietary and medical practise (Meĭen 1987). As an example, a study on *Giardia* and *Cryptosporidium* was partially based on archaeological evidence of intensified production of specific plants (Anthony et al. 2005; Anthony et al. 2007). There is a variety of methods which can be applied on various types of sediment. A special application of sediment analysis is tephrochronology, which makes use of the special nature of volcanic ash (Greek: tephra). By analysing the positions of the tephra layers, dating of these is possible with a considerable accuracy as volcanic eruptions produce tephra only for a limited amount of time (Lowe 2011).

Some common dating techniques can be applied to both sediments and artefacts as they rely on material which can be processed into artefacts, such as wood, clay or other biological material. These methods make use of physical phenomena and statistic evidence in the sediments or artefacts to estimate their age. These techniques can also be combined to increase the accuracy and as they have different time ranges where they perform best a combination offers the most informed insight. The preferred technique depends largely on the nature of the object to be dated. For wooden artefacts, dendrochonological dating can provide very accurate dating based on the tree ring patterns. This technique can, if a reference is available, provide dates with an accuracy of a few years (Pilcher et al. 1984). Radiocarbon dating is also applied on biological material, relying on the radioactive decay of the carbon isotope ^{14}C. Traces of ^{14}C are found in the atmosphere and are steadily incorporated into plant material by photosynthesis. When the plant dies, the incorporation of ^{14}C ceases and the decay of it is measurable as it is longer replenished. The half-life of the isotope is 5,730 years (radioactive beta-decay $^{14}C \rightarrow {}^{14}N + e^-$) which limits the application to approximately 50,000 years (Aitken 1999). There are also dating techniques for non-biological material, such as thermoluminescence dating which measures the accumulated radiation dose from heat or sunlight since the last exposure. This can be applied to sediment (quartz and feldspar, irradiation) or heated artefacts such as ceramics (Aitken 1999; Lian and Roberts 2006).

1.2.3 Archaeological Research on human remains

Bones represent a common resource for archaeological research. Osteology allows the identification of age and sex, as well as information on the health status and potential cause of death. On bone material further techniques can be applied notably radiocarbon dating and stable isotope analysis to evaluate origin and diet and various ancient DNA applications.

1.2.3.1 Osteology

Bones from excavations are evaluated to get information on various issues related to the health state and age of the individuals and the demography of the population. The development of the human skeleton is well understood and allows assessment of age and sex of the individual (Dirkmaat 2012). The accuracy of the analysis is dependent on the preservation state of the bones and the presence of certain key parts of the skeleton. When considering skeletons, cultural aspects also have to be taken into account, such as different burial grounds for different social, gender or age groups. In several excavation a potential sex bias on graveyards has been observed (Roberts and Manchester 2010).

Sex determination should be the easiest factor to identify as the only two options are male or female. The most reliable structure to identify the gender is the pelvis, but other bones and structures, such as the femoral and humeral heads or cranial metrics are frequently used. The choice of the most reliable factors used for gender identity of a skeleton is debated, as in most cases one single factor will not give a conclusive answer (Frias et al. 2013; M. Meyer et al. 2013).

Age at death estimation of a skeletal population is an important factor when studying the population dynamics and demography of individual populations (palaeodemography). Generally, developmental markers (e.g. pubic symphysis, sacro-iliac joint, sternal rib ends) in Juveniles are more reliable as the developmental process is tightly regulated. For adults this is more difficult as the age-related decay is less traceable and depends on the environment (Franklin 2010). There is also some debate concerning biological versus chronological age. The biological age refers to the implied age based on the physiological state, whereas the chronological age is the exact time the individual has lived. The biological age depends on a combination of factors such as genetics, nutrition, environmental factors and health condition. The influences on the biological age are dependent on the individual which makes the correlation with chronological age more difficult. In a collection of skeletons from a single excavation individuals with different apparent biological ages can have the same chronological age (Baccino et al. 1999; Garvin et al. 2012; Milner and Boldsen 2012).

A part of palaeopathological research is focussed on determining diseases and trauma in skeletons based on bone morphology. The main constraint of this research is that only factors which have an impact on the skeleton can be evaluated. Injuries and diseases only affecting soft tissue cannot be diagnosed with a high level of confidence. Major trauma injuries can reveal the cause of death and minor trauma can give evidence on the activities of the individual (e.g. healed battle wounds of soldiers or stress fractures in workers). Bone deformations caused by healed or partially healed injuries or diseases are valuable indices on the life of the deceased's culture. These techniques provide evidence for the reconstruction of historic events. For example, on the important discovery made in 2012 in Leicester, the remains of

Richard III, the last king of England to die in battle, extensive studies on the injuries and the health condition were done. The pattern of the injuries lead to conclusions on how he died and how his body was treated after his death (Appleby et al. 2014).

A wide range of diseases will leave traces of lesions on the bones. Considerable work has been done on infectious diseases such as tuberculosis, leprosy and syphilis by identifying specific markers for these diseases (Mays and Taylor 2003; Mays et al. 2003; Taylor et al. 2000). Analysis of teeth can provide supportive facts on the diet and dental diseases such as caries (Humphrey et al. 2014). Evaluation of the bones provides valuable information on the living conditions, but is restricted to conditions affecting the bones (Roberts and Manchester 2010).

1.2.3.2 Stable isotope analysis

Living organisms have a constant turn-over of biological material by uptake of various nutrients. The differential integration of material from food sources into catabolites allows evaluation of the origin of the source material. Isotope analysis makes use of the fractionation in the integration of different isotopes into biological material. The analysis of the isotope compositions allows insight into diet and geographic origin or migration (Schoeninger and Moore 1992).

The principals that underpin isotope-based analysis are that biologically important chemical elements exist at more than one stable isotope. Of particular interest are carbon ($^{13}C/^{12}C$), nitrogen ($^{15}N/^{14}N$), oxygen ($^{16}O/^{17}O/^{18}O$), sulphur ($^{34}S/^{32}S$), calcium ($^{44}Ca/^{42}Ca/^{40}Ca$) and strontium ($^{87}Sr/^{86}Sr$). Although isotopes have almost the same chemical properties, their difference in weight leads to a different

integration rate (fractionation). The magnitude of the effect is determined by the relative mass difference and the nature of the reactions in which these are integrated. Since there are differences between the ratios of these isotopes between the atmosphere and the ocean or between geographic regions, the analysis of the isotope ratio allows conclusion on the origin of dietary items and changes in dietary patterns (Lee-Thorp et al. 2010).

The main source of carbon for terrestrial plants is the atmosphere, which has lower ^{13}C content than the ocean. The incorporation of carbon into living organisms uses preferably ^{12}C, which leads to an accumulation of ^{12}C in living organisms compared to the source. The carbon ratio ($\delta^{13}C$) of terrestrial plants is determined by the atmospheric carbon ratio and the photosynthetic pathway (C_3, C_4, CAM). Marine organisms have usually different ratios due to the different sources (dissolved atmospheric CO_2, terrestrial detritus) which lead to a higher variability in the $\delta^{13}C$. There are similar mechanisms underlying the integration of other elements (Schoeninger and Moore 1992).

The stable isotope approach has a variety of applications, of which a few examples will be mentioned here. The impact of a marine diet can be estimated, using carbon and nitrogen ratios as reported in a recent study of in victims of the outbreak of the Vesuvius in 79 CE from Herculaneum (Craig et al. 2013). Stable isotopes can be used to assess the influence of diet and source of nutrients on the development (Henderson et al. 2014) or the nutrient sources and dietary changes and adaptations in early hominins.

Several different approaches employing techniques from other disciplines contribute to research in archaeology. Each technique has its advantages and shortcomings, and only combinations of them will contribute to the knowledge in the field (Haak et al. 2008).

1.3 Ancient DNA – methods and applications

The development of PCR created new opportunities to study genomic DNA. The reliable amplification of small amounts of genomic DNA opened new avenues in forensic sciences and archaeology. Even before PCR being available, researchers were able to extract and sequence ancient DNA (aDNA). The first ancient sequence data published was from an extinct member of the horse family (the quagga) by the group of Wilson in 1984 (Higuchi et al. 1984). The first human aDNA was extracted from an Egyptian mummy by Pääbo and colleagues in 1985 (Paabo 1985). Since these initial success stories numerous publications on related issues were published, such as diseases and the evolutionary history of humans and animals (Bos et al. 2011; Dabney et al. 2013; M. Meyer et al. 2013). Reviews on the field summarised commonly accepted issues with aDNA, such as instability and good practice to avoid contamination (Ankarklev et al. 2010; Cooper and Poinar 2000; Hofreiter et al. 2001; Hoss et al. 1996). It was widely accepted that based on experimental evidence and estimations of the thermodynamic instability of aDNA the general upper limit of PCR amplicons is about 200 bp, although it has been shown that well preserved aDNA can produce longer amplications (supported by own data, see chapter 5). Advanced technologies such as the introduction of parallel sequencing techniques (Fornaciari et al. 2010) allowed the generation of longer and more fragmented

sequences by creating contigs (Dabney et al. 2013; M. Meyer et al. 2013). These techniques also allow the closer evaluation of evolution, for example the evolution of the genus *Equus* based on the sequence derived from a Middle Pleistocene horse (Orlando et al. 2013). Preservation of the samples is critical for the recovery of aDNA and it is thus not surprising that the oldest sequences recovered are generally recovered from samples preserved in permafrost (Dabney et al. 2013). Currently, the oldest DNA recovered was extracted from Arctic ice cores containing Holocene and Pleistocene sediments and dated up to 800,000 years old (Willerslev et al. 2003; Willerslev et al. 2007), and the oldest genome sequence was obtained from a Pleistocene horse dated up to 700,000 years old (Orlando et al. 2013). Some research has shown that aDNA can still be extracted from museum specimens which have been removed from its permafrost environment and stored at ambient temperatures for a prolonged period of time (Gilbert et al. 2007).

To evaluate the application of aDNA for disease specimens samples from victims of known or strongly suspected diseases were studied, for example skeletons from Medieval plague pits (Drancourt et al. 1998; Raoult et al. 2000) or skeletons with suspected lepromatous deformations (Taylor et al. 2000). Further studies applied these methods to evaluate the co-existence and co-evolution of diseases with humans. Of particular interest are zoonotic diseases such as tuberculosis (TB) which would provide insight into the evolution of *Mycobacterium* spp., especially the question when the divide between human and bovine TB occurred (*M. tuberculosis* vs. *M. bovis*). Whether the adaptation to humans originated in a time before domestication is a difficult issue due to the limited availability of samples. Ancient samples provide datable ancient evidence which can then be integrated with modern sequence data to provide a dated anchor for the molecular data. The combination of

both ancient and modern sequences can improve the results gained from both perspectives as the relationship between modern and ancient sequences. Although this problem has not yet been fully understood there are hints toward a human adaptation, but these claims remain speculative (Mays and Taylor 2003; Taylor et al. 2007).

The limiting factor for aDNA applications on diseases is the parts of the body affected by the disease. Most diseases affect soft tissue and not the bones, but in the majority of cases only the bones are preserved. This limits the range of diseases detectable. Some blood-borne diseases can be detected in dental pulp as the causing agents were trapped in the fine vascular system within the teeth (Drancourt et al. 1998). A widely discussed example of this limitation of aDNA is the question of the origin (new world or old world) of venereal syphilis which has only minimal effect on bones (Bouwman and Brown 2005). Despite the speculated limitations of the approach for some diseases, the first full genome of *Yersinia pestis* as causing agent of the black death in London (1347-1351 CE) was published in 2011 (Bos et al. 2011). *Y. pestis* reaches high copy numbers in an infected individual which made such a study possible.

Besides aDNA there is a variety of other methods used, such as an adapted ELISA-based fast test for malaria which was used successfully on the remains of Francesco I de Medici, Grand Duke of Tuscany (Fornaciari et al. 2010).

1.4 The biology of the parasites

The parasites in this study have very different life cycles which could be used to address different questions. The differences would influence the perception of the infected individual such as suffering from acute or mild symptoms. Molecular markers such as the expected rate of accumulation of genetic changes by different replication (generation) times gives information on different timescales. Infection routes can provide direct evidence based on presence or absence of parasites, such as for the food-borne tapeworms which do need an intermediate non-human host. This study includes helminth and protozoan parasites. The selection is based on prevalence worldwide and the estimated prevalence in ancient populations, microscopic diagnosis and diversity of life cycles to address various questions.

1.4.1 Helminths

Helminths are a heterogeneous group formed majorly for practical reasons, containing parasites from two phyla: Platyhelminthes and Nematoda. Although grouped they are very diverse in their biology. Based on overall prevalence human intestinal nematodes are the most important helminth group. The two main parasites studied here are the nematodes *Ascaris* sp. and *Trichuris trichiura*, but some work includes *Enterobius vermicularis* and *Anisakis* sp. The cestodes *Diphyllobothrium latum* and *Taenia* spp. were included after microscopic diagnosis in an increasing number of samples.

Helminths are a global problem, according to estimations of the World Health Organisation (WHO) more than a sixth of the human population is affected (Pullan et

al. 2014). Infections are most prevalent in the southern hemisphere, whereas the parasites are nearly extinct in Europe (Figure 1.1, modified from Pullan et al. 2014, most cases in Europe are linked to travelling (Tomaso et al. 2001). The world maps in Figure 1.1 indicate the world-wide prevalence for helminths on the example of *Ascaris* (Figure 1.1 b) and *T. trichiura* (Figure 1.1 d). Comparing data on global numbers of infection from 1990 (in grey, Figure 1.1 a/c) and 2010 are variable and might be slightly decreasing, most notably in Asia (Figure 1.1 a/c), but are nearly stagnant in Latin and Central America (LAC), Sub-Saharan Africa (SSA), North Africa and the Middle East (NAME) and Oceania.

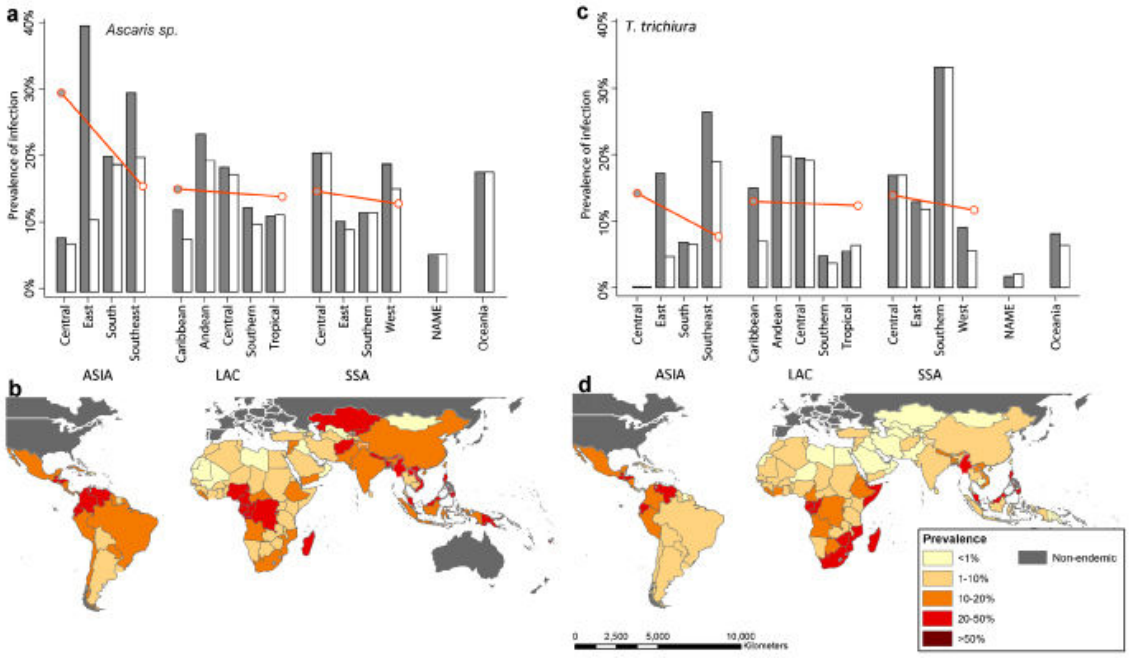


Figure 1.1: Global prevalence of *Ascaris* and *Trichuris*.

Ascaris (a/b) and *T. trichiura* (c/d) remain a global problem with more than one sixth of the world population infected, mostly in the southern hemisphere. By comparing data from 1990 (grey) and 2010 (white) for *Ascaris* sp. (a) and *T. trichiura* (c) trends on global burdens can be estimated. Infections might be decreasing in Asia for both *Ascaris* (a) and *T. trichiura* (c), but as there are only limited data points available and the rates are more stagnant in Latin and Central America (LAC) and Sub-Saharan Africa (SSA) no conclusive result can be seen. Modified from (Pullan et al. 2014)

1.4.1.1 Trichuris trichiura

Trichuris trichiura is a nematode which is in the adult stage between 3 to 5 cm long, with a long pharyngeal region and a wide posterior section. This unusual shape gave it the name whipworm. Infections are caused by ingestion of embryonated eggs. Mature worms produce up to 7,000 eggs per day which require maturation (development of L1 larva) of 10 to 30 days outside the body. After ingestion the larva hatches and develops into an adult worm in approximately 3 months, reaching the caecum where it attaches itself permanently. Light infections are asymptomatic, heavy infections or infections of malnourished individuals are characterised by symptoms such as bloody diarrhoea, anaemia, colic, abdominal pain and in some cases prolapsed rectum. Untreated infections can result in a long-term infection as the life span of the worms is up to eight years. *Trichuris* eggs are easily identified by their characteristic barrel shape with polar plugs on both sides. The eggs measure approximately 30 by 50 µm. Infections with *T. trichiura* are common in developing countries (Figure 1.1) due to a lack in hygiene (Cox 1993; Markell et al. 1999).

There are several *Trichuris* species which are specific to certain hosts, but human infections with other *Trichuris* spp. have been reported before, mainly infections with *T. suis* (Campbell et al. 1992) and *T. vulpis* (Areekul et al. 2010). The species are clearly identifiable by molecular markers (Cutillas et al. 2009).

1.4.1.2 Ascaris spp.

The adult *Ascaris* is the largest nematode to parasitize humans (females reach up to 35 cm and males up to 30 cm). Transmission is largely dependent on human faecal disposal practices, as the eggs are dispersed in the faeces of an infected individual and infection occurs by ingestion of parasite eggs. Adult female worms can produce on average 200,000 eggs per day for several years. The eggs require up to three weeks outside the host to mature. Embryonated eggs hatch in the duodenum and enter the blood or lymphatic stream by penetration of the duodenal wall. The migration eventually reaches the lungs where they moult twice, after about 20 days they ascend the trachea and are swallowed to re-enter the ileum where they remain. Infection with *Ascaris* is in most cases asymptomatic, but among the described symptoms are broncho-pneumonia due to migration larvae to the lungs, obstruction of small intestine, pancreatic duct or other internal organs. Adult worms are generally located in the small intestine. The fertilised egg has is round to oval with a distinct rough outer structure, measuring 45-75 µm by 35-50 µm. (Cox 1993; Markell et al. 1999).

Commonly, there are two species of *Ascaris* described (*A. lumbricoides* and *A. suum*), but there are ongoing debates whether these are in fact two species or just subspecies of one. Both are found in human and porcine hosts (Campbell et al. 1992; Leles et al. 2012).

The differences in the life cycle of *Trichuris* and *Ascaris* are shown in Figure 1.2. The life cycles illustrate the different maturation after infection. Both parasites require the ingestion of embryonated eggs, but the eggs of *Ascaris* sp. require a more complex route of infection compared to *Trichuris*.

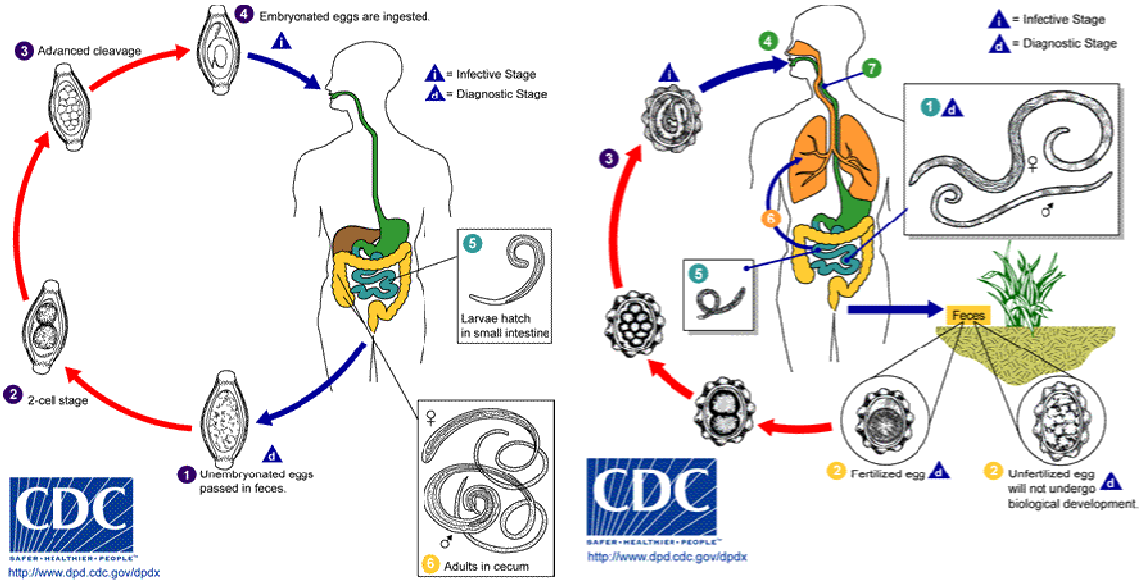


Figure 1.2: Life cycles of *Trichuris trichiura* (left) and *Ascaris* sp. (right)

Both parasites infect humans faecal-oral, the maturation happens outside of the host. Image source Centre for Disease Control (www.dpd.cdc.gov).

1.4.1.3 Enterobius vermicularis

The threadworm (or pinworm) *Enterobius vermicularis* is the most common helminth as the general sanitary measurements required for the reduction of infection with *Ascaris* and *Trichuris* are not sufficient to prevent the infection with *E. vermicularis*. Development to the adult worm after ingestion of the eggs requires about six weeks. The adult worms migrate to the caecum where they remain attached to the mucosa. Male worms are smaller (2-5 mm) than females (8-13 mm). Most infections remain asymptomatic and self-limiting; the commonly described symptom is itching in the perianal region. This itch is caused by gravid female worms crawling out of the anus, depositing eggs on the surrounding skin. Larvae develop within about six hours outside the host and remain infectious for several weeks (Figure 1.3). Infections with *E. vermicularis* are difficult to control as re-infections by ingestion of

eggs from the perianal region are very common due to scratching. The eggs of *E. vermicularis* measure about 30 by 50 µm (Cox 1993; Markell et al. 1999).

1.4.1.4 Anisakis spp.

The genus *Anisakis* consists of several species which infect marine animals. Although humans are a dead-end for *Anisakis* infections they occur relatively frequently (Figure 1.3). The symptoms include abdominal pain, nausea and sometimes diarrhoea occurring one to five days after consumption of raw or undercooked fish or squid (Fukita et al. 2014; Lin et al. 2014).

Infections are common in countries consuming large amounts of raw or undercooked fish, such as Japan (sushi, sashimi), the Netherlands (pickled herring) or Scandinavia (air-dried or pickled fish). The consumption of uncooked fish is considered to be safe when the fish was frozen at -20°C for at least five days, or if the fish has been dried or smoked with the flesh reaching temperatures of 65°C. Although the larvae are usually located in the gut, migration into the muscle tissue is not uncommon (Markell et al. 1999). In ancient samples it is likely that the infection rates were considerably higher than in modern humans as freezing of the fish was not possible and an extensive food safety control was unlikely to be enforced.

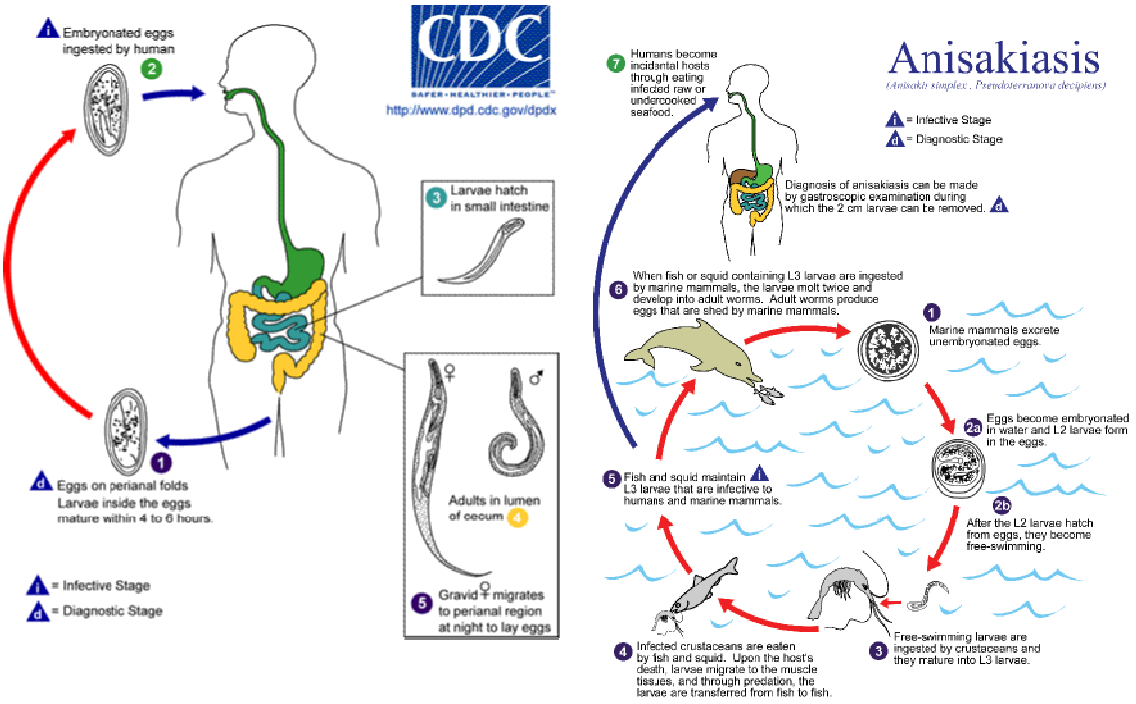


Figure 1.3: Life cycles of *Enterobius vermicularis* (left) and *Anisakis* sp. (right)

The life cycle of *E. vermicularis* (left) shows the self-infection route with a maturation of the infectious stage on the host. *Anisakis* sp. (right) is usually a fish and marine mammal parasite for which the human infection is a dead-end. Image source Centre for Disease Control (www.dpd.cdc.gov).

1.4.2 Cestodes

Tapeworms, or cestodes, are named after their flat and ribbon-like body. Cestodes parasitizing humans do generally have a complex life-cycle which includes generally an intermediate host, which can be, depending on the species, either humans or other species. Tapeworms belong to the largest parasites of humans, with some species (such as *Diphyllobothrium latum*) reaching lengths of up to 20 m (Markell et al. 1999). Three tapeworms are included in this study, indicating consumption of raw or undercooked fish (*D. latum*) or red meat (*Taenia solium* and *Taenia saginata*). This gives direct insight into dietary and cultural practices.

1.4.2.1 Diphyollobothrium latum

The broad tapeworm *Diphyllobothrium latum* occurs in northern temperate regions, such as Scandinavia, Japan or North America (Arizono et al. 2009). The life-cycle of *D. latum* requires two intermediate hosts (Figure 1.4). Eggs excreted through faeces embryonate and hatch in water and are subsequently ingested by the first intermediate host, small crustaceans (such as copepods). After ingestion of infected crustaceans by larger fish the larva continues to grow and develops into a stage called plerocercoid larva. Smaller fish infected by *D. latum* might become prey of larger fish. In those larger the plerocercoid infects the new host, but does not develop further until ingested by a suitable mammal (including humans) in which the larva develops into the adult worm inside the small intestine. Several species of fish have been identified as intermediate host. The adult worms can grow to lengths of 20 m, releasing eggs into the intestine. The eggs are slightly elongated and measure about 50 µm by 70 µm with an operculum on one side. Infections are in most cases asymptomatic, the most frequent symptom described is vitamin B_{12} deficiency. Other symptoms include weight loss, abdominal pain or diarrhoea (Kuchta et al. 2013; Markell et al. 1999; Scholz et al. 2009).

D. latum eggs are less distinct than the nematode eggs and a clear identification is only possible by molecular methods (Guo et al. 2012; Wicht et al. 2010). The eggs are difficult to distinguish in archaeological contexts, but earlier studies have shown their presence in various contexts, such as medieval Lübeck (Paap 1984). The earliest association with humans remains speculative (Le Bailly et al. 2005).

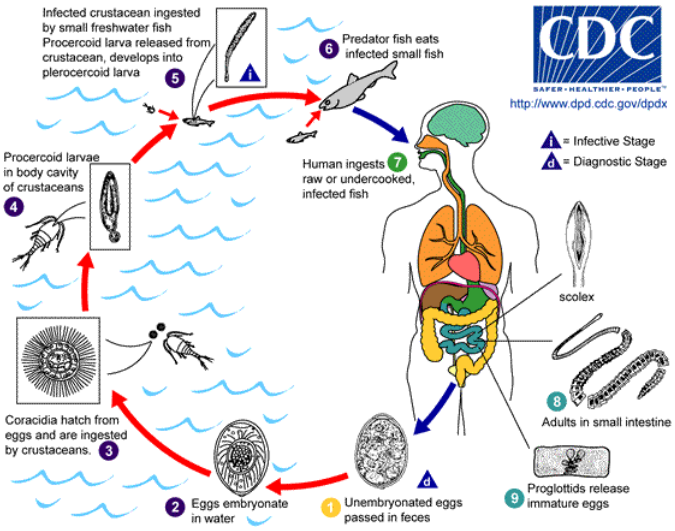


Figure 1.4: Life cycle of *Diphyllobothrium latum*.

Humans are infected by consumption of raw or undercooked infected fish. The ingested larvae develop to adult worms in the human digestive tract producing eggs which are then released into the water. They require two intermediate hosts (crustaceans and fish) to fully develop. Image source Centre for Disease Control (www.dpd.cdc.gov).

1.4.2.2 Taenia solium and Taenia saginata

Similarly to *D. latum*, the *Taenia* tapeworms infect humans by ingestion of raw or undercooked food, in case of *Taenia* red meat (*T. solium* from pork and *T. saginata* from beef). The life-cycle is similar for both *T. solium* and *T. saginata*, with different intermediate hosts (Figure 1.5). The intermediate host gets infected by ingestion of eggs deposited in faecal material. The eggs can remain infective for months after deposition. After migration to the small intestine the hatched larval stage migrates into the striated musculature where they grow and differentiate into the second stage larval state (cysticercus) which is the infectious state for the definitive human host. In the human host they differentiate in the small intestine to their adult form within 12 weeks. The adults are hermaphrodites which can both self-fertilise or cross-fertilise. The adult worm can survive for up to 25 years in humans

and are relatively large as they reach about 5 m for *T. saginata* and 2-7 m for *T. solium*. *T. solium* and *T. saginata* eggs are nearly indistinguishable and molecular methods are required for the species identification. The eggs are round with a diameter of about 30-50 µm and have a characteristic thick, radial striped coat which allows conclusive identification of the genus (Hoberg 2006; Markell et al. 1999; McManus 2006). They can produce more than 10.000 eggs per day (Bobes et al. 2014).

Modern food safety requires the monitoring of infected animals in industrialised countries, reducing the number of infections. A study in France in 2010 found less than 2 % of slaughtered cattle to be infected with *T. saginata*, but various food safety laws restrict human infections (Dupuy et al. 2014). *Taenia* has only been described in few archaeological samples, such as in a natural human mummy from 3rd century BCE Iran (Nezamabadi et al. 2013).

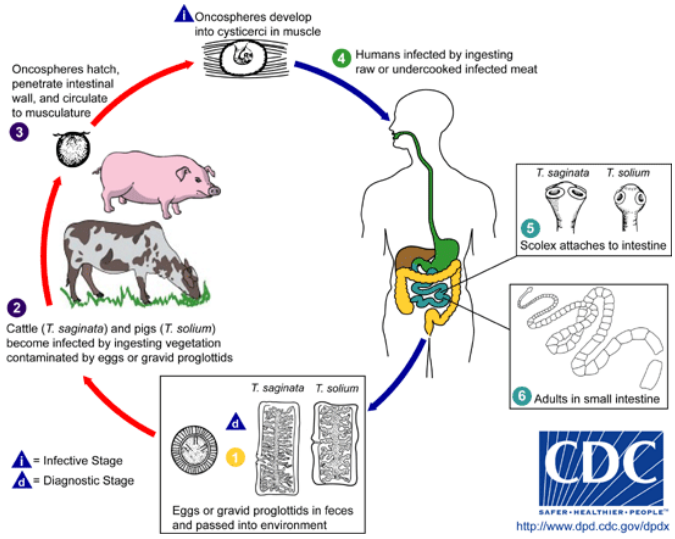


Figure 1.5: Life cycle of *Taenia* sp.

Humans are infected by consumption of raw or undercooked red meat. The ingested larvae develop to adult worms in the human digestive tract producing eggs which are then released through faeces. Image source Centre for Disease Control (www.dpd.cdc.gov).

1.4.3 Protozoan parasites

Protozoans are a diverse group of eurkaryotic microorganisms which have an animal-like behaviour. In this study two protozoan parasites were included: *Cryptosporidium* sp. and *Giardia intestinalis*. Both parasites are still widely prevalent, including European countries, yet very little is known on their ancient distribution.

1.4.3.1 Cryptosporidium sp.

The apicomplexan *Cryptosporidium* sp. usually infects humans through ingestion of contaminated water. They have a worldwide distribution and affect mostly people with a weakened immune system, such as HIV patients (Meisel et al. 1976). *Cryptosporidium* as a human infection was first described in a child in 1976 (Nime et al. 1976). There are two species which infect humans, *C. hominis* and *C. parvum*, although there are debates whether these two are true species or subtypes of the same species (Slapeta 2013). In 2004, two independent groups published the full genome of *C. hominis* and *C. parvum* (Xu et al. 2004).

Cryptosporidium infections are acute, short term infections which include severe diarrhoea. In humans the infection lasts about five weeks and will resolve without intervention in non-immunocompromised patients. Development occurs within epithelial cells where they undergo two asexual generations. The infectious oocycsts are passed in the faeces and can remain infectious for weeks outside of the body (Figure 1.6). *Cryptosporidium* is not affected by chlorine disinfection and due

to its small size (diameter of 4-6 μm) specific filtration techniques have to be applied for efficient removal (Madigan et al. 2006; Markell et al. 1999).

1.4.3.2 Giardia lamblia

Similar to *Cryptosporidium*, *Giardia lamblia* (also called *G. intestinalis* or *G. duodenalis*) infections occur by ingestion of contaminated water. It is a relatively common infection with an estimated 280 million human cases every year (Guy et al. 2003). Symptoms appear 6-15 days after ingestion and include diarrhoea, vomiting, nausea and weight loss (Figure 1.6). The infectious douse is relatively low and *Giardia* sp. can survive for a prolonged time in the environment (Ankarklev et al. 2010; Madigan et al. 2006).

Protozoans are an interesting system to study genetic diversity as they have a significantly faster turnover rate than other parasites or their hosts (Mahbubani et al. 1991). There was further interesting evidence from a 14th century monastery in Scotland for the use of plant remedies. Based on plant traces in latrines and subsequent laboratory experiments it was shown that different plant extracts, most notably blueberry extracts, have an effect on the viability of *Giardia* and *Cryptosporidium* (Anthony et al. 2007; Anthony et al. 2011).

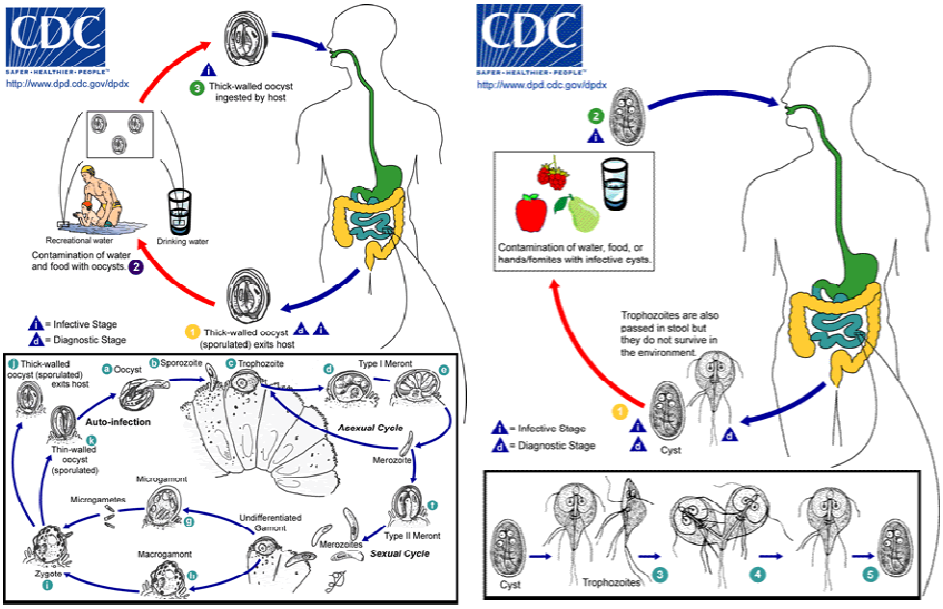


Figure 1.6: Life cycles of *Cryptosporidium* sp. (left) and *Giardia lamblia* (right).

These protozoan parasites infect humans by ingestion of contaminated water or on surfaces which were in contact with such. Image source Centre for Disease Control (www.dpd.cdc.gov).

1.5 Parasites in Archaeology

Various parasitic diseases have been described in archaeological contexts (Anastasiou and Mitchell 2013; Roberts and Manchester 2010)Ectoparasites such as lice, ticks and fleas played an important role as primary infections as well as vectors for the transmission of vector-borne diseases such as the plague (Raoult et al. 2006). In this thesis, only endoparasites were studied and thus the introduction is focused on these parasites in archaeological context.

Microscopic diagnosis of helminth parasites from archaeological specimens has been reported as early as 1910 for *Schistosoma mansoni* (Ruffer 1910). The first report of *Trichuris sp.* and *Ascaris* sp. in a bog body was published in 1944 (Szidat 1944). The identification of the parasite genus by microscopy does not require expensive equipment and is thus reported in various contexts (Araujo et al. 2013;

Aspock et al. 1973; Herrmann and Schulz 1986). Molecular methods were so far limited to short fragments for species identification (Loreille et al. 2001; Oh et al. 2010). The current project aimed at the development of refined molecular techniques for Archaeoparasitology and to explore the potential for using these sequences as an epidemiological tool to interrogate sample sets from different locations and time periods.

Symptoms caused by the intestinal parasites in this study are in most cases mild and non-distinct which makes the identification of parasites based on historic records difficult and unreliable. Prescribed cures in early medical texts such as the Eberts Papyrus (c1550 BCE) or the Corpus Hippocraticum (c5th century BCE) are vague and include general measures of hygiene and nutritional supplements such as garlic (Anthony et al. 2005; Brinkkemper and van Haaster 2012). Hygiene improvements have been shown to be important for the counteraction of infection with helminths in modern populations (Gungoren et al. 2007), but the awareness of the infection rate and the link between the identification of parasites (e.g. expelled worms in faeces) and the appropriate counteractions require further education which is lacking in some modern populations (Phongluxa et al. 2013) and was likely lacking in historic populations.

1.6 Archaeological samples in this study – origin and historic background

During this course of this project archaeological material was obtained from latrines, communal deposits and the abdominal region of skeletal remains from different sites in eight countries in Europe and one site in Lesotho, Africa. The work focuses on sites in Europe as most of the parasites are no longer prevalent in Europe which excludes the danger of contamination with modern samples.

The oldest samples were obtained from cave deposits in Lesotho, dating between 17000 to 70000 BCE. The material was obtained from a temporary human settlement, most of the material from the region of the fireplace in the cave. As these samples were all negative they could not be included in further analysis. Thus, the oldest samples which were diagnosed positively were obtained from eight different Neolithic stilt settlement deposits in Switzerland and southern Germany. These samples were dated between c3920-2900 BCE and were collected from communal deposits in the vicinity of evidence for settlement structures. The oldest samples from a burial context were collected from a Neolithic burial site in Bad Mergentheim, southern Germany. The burials were dated to c2950 BCE and are associated with the corded ware culture and belong to the oldest burials found in central Europe.

The most important sample collections in this project were obtained from excavations of medieval ports in Bristol and Lübeck. From Finzel's Reach in Bristol which was located close to the historic port area a collection of samples from a communal deposit dated between the 12th to 16th century CE was obtained. In Lübeck, the centre of the Hanseatic League, samples from latrines which were associated with town houses in the oldest part of the city were collected. The samples were dated between the mid-12th to mid-17th century CE. Both sample sets span an

important period of growth and increased trade activity in both cities. Other medieval material dated in the same period was obtained from deposits in London and Oxford, as well was graveyards in Ellwangen (DE), Ipswich and Brno (CZ).

Single grave samples were studied from various excavations. The largest set with over 100 samples was obtained from the excavation in Břeclav-Pohansko (CZ). These graves were arranged around a small church within a Slavic stronghold and were dated between 850 - 950 CE. These samples provided information on the frequency of diseases.

Viking samples from Denmark, Norway and England were collected to be linked with Anglo-Saxon samples from England. Latrine and coprolite samples were obtained from Viking settlements in Copenhagen, Viborg (DK) and York, further samples were obtained from settlement deposits in Gokstad (NO) and from graves on Langeland (DK). All Viking samples were dated around 1000 CE. In combination with Anglo-Saxon settlement samples from Ipswich and a latrine sample from York, these samples could provide further insight into the interaction between these groups.

1.7 Overview of thesis

The work presented in this thesis represents a proof of principle for the application of molecular archaeoparasitology on a variety of questions in human history. Recent advances in molecular techniques offer the potential to employ parasite genetics as an archaeological tool. This thesis contains the development of a novel approach based upon analysis of the enteric parasites that have co-existed and co-evolved with humans for millennia which could have generated specific adaptations in the parasite genome. It contains the development of the necessary tools for the diagnosis and specific molecular techniques to study these parasites and their interactions with the human hosts, the acquisition of a sample collection from a variety of geographic and temporal origins, the generation of a sequence database of parasites, and the phylogenetic assessment of these sequences to evaluate their potential for the use of sequences to study human migration and interaction patterns through history.

All samples were screened microscopically and all positive samples and some of the negative samples processed for molecular evaluation and sequence generation. Sequences were generated by extraction of aDNA and PCR amplification with subsequent TA-cloning and Sanger sequencing or bar-coded parallel sequencing. The main impact results from microscopic diagnostics and aDNA analysis. The microscopic diagnosis allows to estimate the load and the different varieties of parasites in a specific sample, whereas the genetic evaluation shows diversity among a species of the parasites which can then be compared to other sites.

1.8 Outlook and further work

During this doctoral dissertation work the validity of the approach was shown. This opens a plethora of possibilities for further research in the area. Former studies focussed on diagnosis and evaluation of prevalence of intestinal parasites, with zoonotic enteric protozoans being neglected in archaeological work. The long term goal of this research is to construct a map of sequence variety and epidemiological prevalence of the parasites through the ages in Europe. A particular interest are trade networks such as the Vikings or of the Middle Ages (e.g. Hanseatic League) and the evolution of the parasites to and within the human population. Further, it will aim at the directional extension of the sample collections to include further sample sets to fit the ongoing research. Extension on the work on protozoans and the potential inclusion of further human parasites (e.g. *Anisakis simplex*) or parasites to evaluate the presence of specific species of livestock (e.g. *Eimeria* spp.) will provide further insight into the history of human migration and trade patterns.

Chapter 2

Development of methods to study enteric parasites in archaeological samples

2.1 Introduction

The research project was initiated to evaluate the use of ancient DNA (aDNA) from human parasites to increase the understanding of human migration and trading patterns through history. The first major aim of this project was to establish standardised techniques for aDNA extraction from environmental samples which were associated with human faeces. Standardised protocols for extraction and amplification of aDNA from bones were available (Cooper and Poinar 2000; Paabo 1989). Environmental samples require additional steps and careful sampling. Intestinal parasites have been recorded as evidence for human diseases as far back as ancient Egypt and based on modern distributions in developing countries it was assumed that a considerable proportion of human populations in the past were affected by parasites. The parasites show a low pathogenicity with only minor clinical symptoms. Further, it has been established for over a century that microscopic diagnosis is possible in archaeological contexts (Ruffer 1910; Szidat 1944). Although intestinal parasites have been described microscopically, the development and application of molecular methods would offer the potential for a new view of parasitological research in archaeology. This chapter will present data underlying the development and refinement of the methodology required for all subsequent work. It is presented as a continuous process of technique development and refinement and is therefore structured with integrated results and discussion.

Reproducible, standardised methods are crucial for the success of the proposed approach. Therefore, a considerable amount of time and effort was invested in the development of methods for the extraction of aDNA and generation of sequences from samples derived from various historical contexts, which included a variety of different soil conditions. Different approaches for the enrichment of

parasites and aDNA extraction were examined. In the course of establishing and refining the methods, several potential problems were identified. A major initial concern which is of special importance with aDNA analysis was the potential for contamination either across the sample set or from other sources, hence the laboratories used in this study never handled modern day samples of the parasites. Although the parasites studied are rare in the industrialised world, which minimised the potential for modern DNA contamination, we still need to take every precaution possible to avoid contamination with modern samples or amplification products of ancient samples, or cross-contamination between different sets of samples.

To aim of identifying genetic signatures which might correlate with the geographical location or the corresponding historic era required the premise that such signatures are possible. An initial basic phylogenetic alignment on sequences available in the GenBank database showed that modern database sequences of *Trichuris trichiura* from different parts of the world where the parasites are still prevalent reveal genetic variation according to geographic origin, yet these initial findings showed only very weak phylogenetic support, even suggesting that some of the sequences labelled *Trichuris trichiura* might not be from the same parasite species. The sequence used in this comparison was the ribosomal internal transcribed spacer 1 (ITS1) located between the 18S and the 5.8S rDNA and commonly used to determine the species of *Trichuris* sp. A part of this ITS1 sequence was subsequently used on the aDNA samples obtained. Further evidence for geographical grouping was found for the related species *T. muris* (parasitizing rodents) from several locations across Europe (Callejon et al. 2010). Since modern samples can be geographically grouped based upon sequence identity it is likely that signatures of temporal and/or geographic origin can be found in ancient samples. Furthermore, in

a recent publication on the molecular epidemiology of *Ascaris* significant genetic differentiation exits between parasites of different countries or even villages based on the mitochondrial cytochrome c oxidase I (CO1) gene (Betson et al. 2014).

2.2 Development of a work-flow

To eliminate sources of contamination, a carefully designed workflow was put in place. It was designed to reflect the necessary precautions to guarantee the sample's fidelity. To avoid cross-contaminations, especially between different steps of the analysis, processing followed a strict uni-directional processing pathway. Samples were handled only in direction of increased processing and aDNA enrichment from original sample to the molecular stages (schematic depiction of the workflow in Figure 2.1). The uni-directional workflow was applied to both samples and the individuals handling them. Samples were processed in different parts of three physically separate laboratories within the Department of Zoology, as represented in the different stages in Figure 2.1. Stage I involved the handling of the original sample from receipt of the samples, sub-sampling, microscopic diagnostics, to aDNA extraction. With the exception of microscopy these steps were performed in a PCR hood with a stainless steel surface and regular UV irradiation (UVP Workbench). For the microscopic diagnosis a 500 µl of the sub-sample was taken from the sample. This microscopy aliquot was discarded after analysis to avoid contamination as the microscope was located in the laboratory where PCR products were handled (laboratory 2 in Figure 2.7). Stage II, the first round PCR set up, was set up on ice in a UV capable laminar flow hood. Stages I and II took place in the same laboratory, the actual PCRs and all subsequent stages were carried out in a different laboratories.

Stage III, the second round PCR, was set up in a similar UV-capable laminar flow hood as stage II, located in laboratory 2. The analysis of the PCR products and further handling of samples took place in laboratory 3. Tools for handling samples were, if not disposable, separate for each step of the process and were cleaned thoroughly after each set of samples (e.g. different sets of pipettes for initial prep, first round PCR, second round PCR, and subsequent steps). The benches used for the all stages were cleaned before and after each set of samples by cleaning the surfaces with 70% ethanol, 1 % Virkon solution and UV irradiation (where available). Intermediate products such as aDNA extracts or PCR products were stored in freezers in the laboratory where they were produced. The original samples were sealed in heavy-duty polythene bags and stored at room temperature in lockable cabinets in laboratory 1. The samples were never brought back into the laboratory where preliminary steps were handled for further processing. The sequences generated by this procedure were different and thus suggest that the procedure was sufficient to avoid contamination.

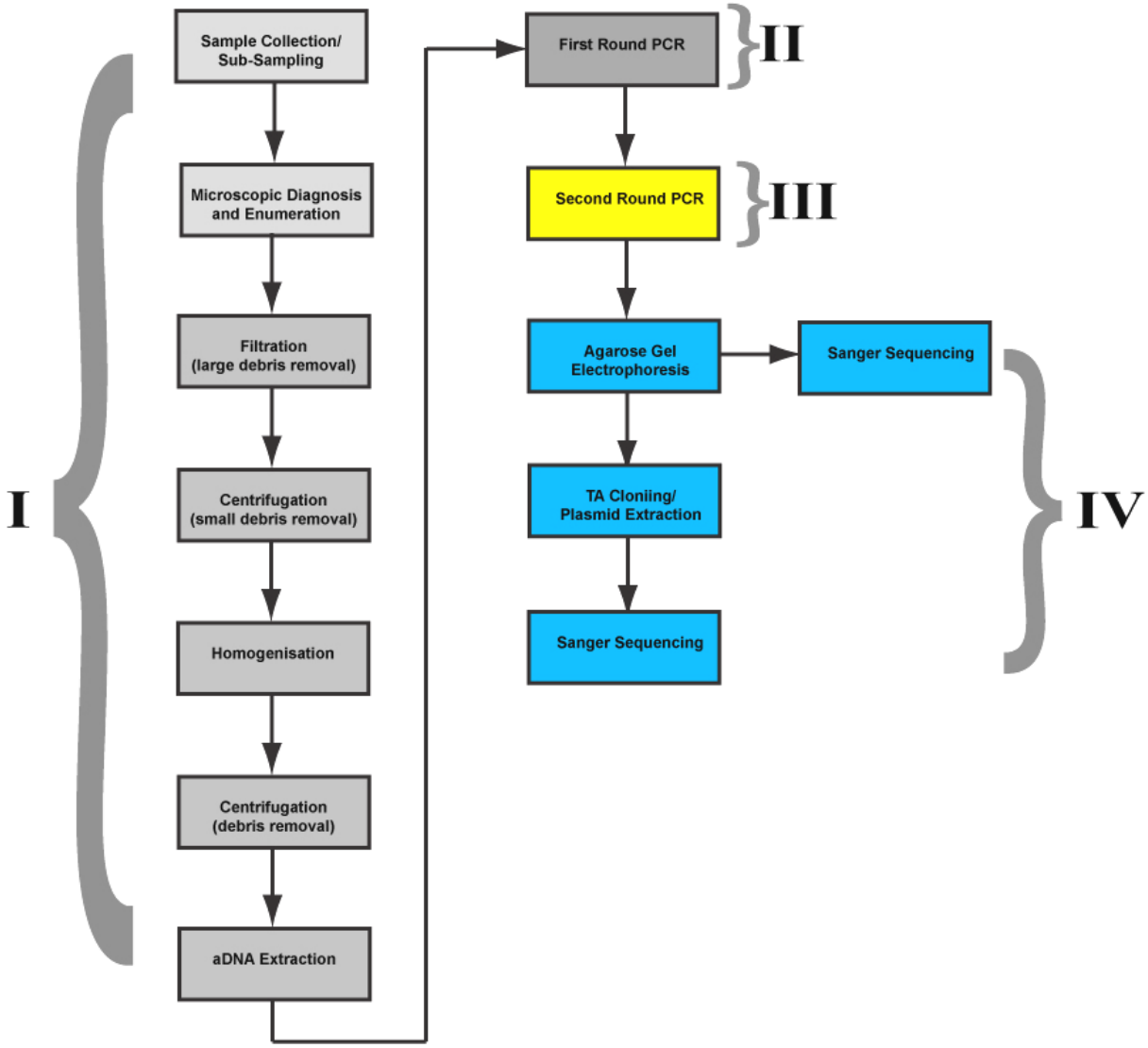


Figure 2.1: Schematic depiction of the workflow for sample handling.

Starting on the top left following the direction of the arrows down are seven steps of stage I from receipt of sample to aDNA extraction. Following the arrows to the top right box, stage II (1st round PCR takes place in the same laboratory as Stage I as indicated by the colour. Stages III and IV take place in laboratory 2 and consist of the subsequent amplification in a 2nd round PCR and processing of the PCR products to sequence generation.

2.3 Sample processing

The following section illustrates the steps of the workflow in more detail. To justify and explain modifications of protocols and procedures, preliminary experiments and their results are described and discussed.

2.3.1 Initial sample collection and preparation

When samples first arrived in the laboratory they were registered in a storage database book and issued a handling code which generally consists of the country of origin and the sample location and/or site and sample description. For example, samples received from the Viking excavation in Viborg, Denmark, were labelled as 'VDKVS' for *V*iking-*D*enmar*k*-*V*iborg *S*øndersø. There was never more than one set from one site handled in one day, eliminating the possibility of contamination with other samples. Cleaning precautions were taken between sets even if they originated from the same site to prevent cross-contamination. Primary samples were handled on a non-absorbing stainless steel surface which was cleaned with 70 % ethanol before and after handling of samples and UV-irradiated after handling samples. Tools like pipettes or tube racks were cleaned and irradiated in the same way.

In a first step a working sample of between 5-10 grams was taken and the remaining sample heat-sealed into a storage bag and stored in a locked storage unit at room temperature before any further processing steps took place. The working sample taken was then re-hydrated in 20 ml of de-ionised (MilliQ) water. If the samples were very dry they were left on a shaker at moderate speed over night (Titertek shaker, level 5). Sediment samples from archaeological excavations contained a variety of different components. During the first step water-soluble components of the sample were washed out by centrifugation of the re-hydrated

sample for 5 min at 400 g (Heraeus Multifuge X3R). The pellet was then re-suspended in 20 ml MilliQ-water.

Of this re-suspended sample, a 500 µl aliquot was taken for microscopy (diagnostic and enumeration). As the microscope was located in a different laboratory, this sub-sample was discarded after the microscopic analysis. Microscopic diagnosis was the first step of the process as it gave an indication of which samples to prioritise. Of the samples from communal deposits, all samples were processed through to aDNA extraction if at least one of the samples from the set was diagnosed positive. This was based on the observation that more than 95 % of the samples within such a set were found to be positive if one was. For samples extracted from burial contexts those diagnosed positively were processed first and only if required other samples were included in the analysis. This was based on the observation that between 15 - 30 % of these samples were positive, which correlates with rates found in modern populations (Betson et al. 2012; Pullan et al. 2014).

Since the size range of the parasites is between 20 to 70 µm, it was possible to concentrate the parasites by removing larger materials by size-exclusion sieving. In preliminary experiments, a sample with relatively high concentration of *Ascaris* and *Trichuris* eggs (Hungate York, Hg) was used to check that the sieving process did not remove the parasite eggs. Subsequently, a series of Nylon sieves with aperture sizes of 1030, 500, and 100 µm, respectively, were used. The sequential use of these sieves was essential as it avoided the clogging of the sieves by larger debris. In initial experiments glass wool filled columns (in 20 ml serological syringes) were used for filtration. Glass wool was chosen for its low cost, easy handling and non-absorbent properties. The columns were prepared directly before use to avoid contamination by previously handled samples. These were less efficient as they had a

low flow-through rate, allowed no control of pore size and were prone to clogging, which reduced the number of samples that could be handled in one day drastically. Filtration through such a glass wool column took generally about one hour (for 20 ml), whereas filtration through the nylon sieves took less than 10 minutes per sample. To reduce the volume, the sieved samples were again centrifuged for 5 min at 400 g (Heraeus Multifuge X3R). The pellets were re-suspended in 2 ml Tris-HCl buffer (pH 7.5) to stabilise the aDNA in the subsequent homogenisation step. Preliminary experiments on an alternative method for parasite egg isolation were evaluated. Extraction of parasite eggs from clinical samples is possible by saturated salt floatation. In a series of experiments between 1-5 g of sample was mixed with up to 20 ml of saturated sodium chloride salt solution. The samples were mixed thoroughly by inversion of the tubes and then centrifuged for 10 min at 400 g (Heraeus Multifuge X3R). After centrifugation 5 % of the added volume was carefully taken off the top and mixed with 10 times the volume of MilliQ water and the centrifugation was repeated. The supernatant was carefully removed and the resulting pellet was re-suspended in 500-1000 µl MilliQ water. Intact eggs should float on saturated salt and then precipitate when the salt concentration drops and parasite eggs should be enriched in the pellet. An experiment prior to the start of this thesis performed by Adrian Smith on a coprolite fragment from a Viking latrine in York enriched intact eggs. However, with other samples these results could not be reproduced and microscopic screening resulted in a very low count of parasites or the complete loss of parasite eggs, Only in one set of samples from a Viking latrine in Copenhagen (VDKKH), three samples (167, 283 and 367) showed low counts of *Ascaris* eggs (1-2 in 10 µl of 1 ml from 5 g sample). Although this procedure would

have the advantage of complete removal of other material, the lack of success on a majority of samples limited its applicability.

2.3.2 Microscopic diagnosis of parasites

Microscopic diagnosis is important step before processing the samples as it gives an indication of which samples are likely to be useful. Several samples could be handled in a relatively short time, providing important information on the variations of parasites in the samples (Nikon Eclipse E400 with Q-Imaging MP5 digital camera). Microscopic detection of the parasite eggs was possible due to the resistant surface structure of the eggs and characteristic size and shapes (Figure 2.2). The microscopic images in Figure 2.2 show both modern pictures (on the left, all from the website of the Centre of Disease Control and Prevention, www.cdc.gov/dpdx) and images from the excavation in Lübeck (on the right). The number of parasite eggs per gram was calculated based on two individual counts of a subsample. The values were calculated in eggs per gram of original sample. Overall, the range of parasites detected varied depending on parasite identity, which could be explained by their different fecundity. The range in the numbers of eggs seen in the samples were as follows: *Ascaris* 19.9 - 1644.9 eggs per gram (Figure 2.2 a/b), *Trichuris* 46.3 - 8859.2 eggs per gram (Figure 2.2 c/d), *Taenia* 133.3 - 8309.6 eggs per gram (Figure 2.2 e/f) and *Diphyllobothrium* 138.6 - 472.4 eggs per gram (Figure 2.2 g/h). The highest numbers of eggs were found in communal deposits, the highest numbers found in single graves were 434.8 eggs/g (CZBP) for *Ascaris* and 330.0 eggs/g (CZBP) for *Trichuris.*

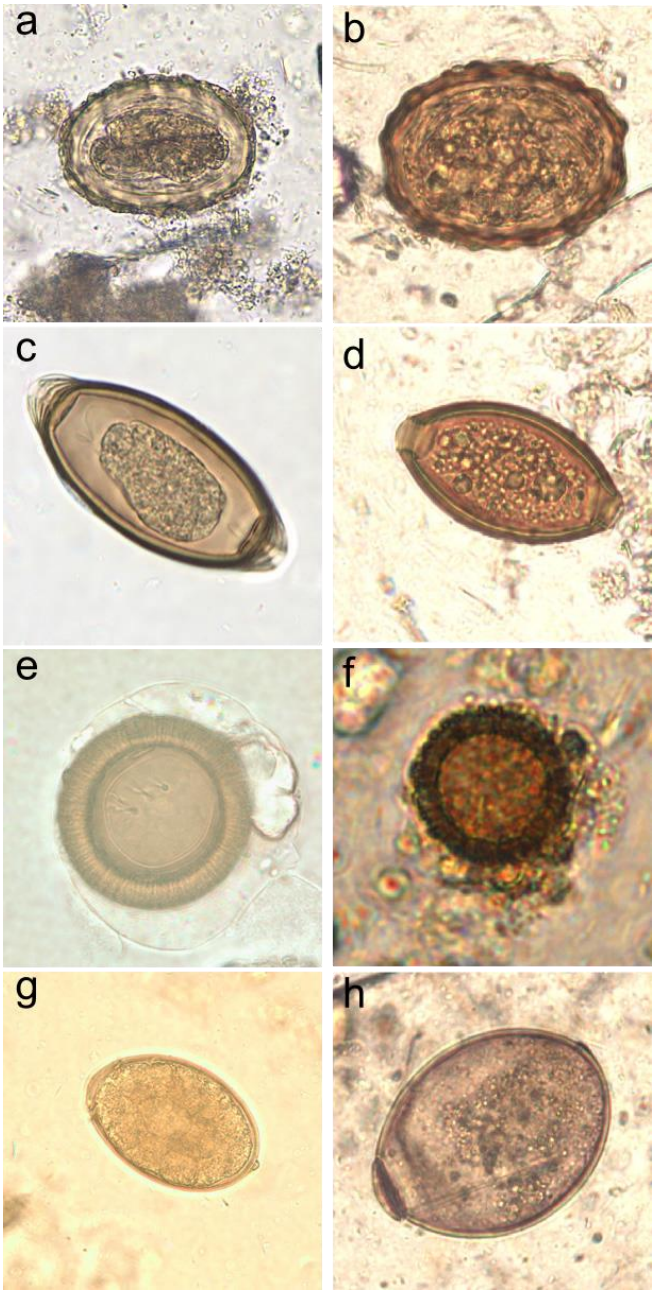


Figure 2.2: Comparison of modern and ancient microscopic pictures of the nematode and cestode parasites in the study.

Ancient parasite eggs show the same characteristic shapes as modern clinical pictures. The images on the left are modern diagnostic pictures (all from the website of the Centre of Disease Control and Prevention, www.cdc.gov/dpdx) and the ones on the right are pictures obtained from samples from Lübeck. The parasite eggs shown are *Ascaris* sp. (a/b), *Trichuris* sp. (c/d), *Taenia* sp. (e/f), and *Diphyllobothrium latum* (g/h)

2.3.3 Homogenisation and aDNA extraction

Preliminary experiments showed that a direct PCR on a purified and enriched sample of (saturated salt floatation) eggs is possible, but very inefficient (Figure 2.3). In the initial experiment the salt-floated parasite eggs from a coprolite fragment from a Viking latrine in York was used directly in a PCR reaction with six specific targets in two parasites (*Ascaris* and *Trichuris*, Figure 2.3). To assess the efficiency of PCR on intact or mechanically disrupted eggs, two sets of PCRs were run. The procedure to use a mechanical disrupter (Bead Beater, BSP Biospec) for parasites was developed for *Eimeria* sp. (Blake et al. 2003). The homogenisation was done using inert glass beads (1 mm glass beads; Hecht Assistent, Germany) in 2 ml screw-top tubes. The eggs were treated for 2 minutes. To evaluate the difference, PCRs were done using primers for *T. trichiura* ITS1 and *Ascaris* sp. CO1 using intact eggs (right) and mechanically disrupted eggs (left) as template (Figure 2.3). The expected band lengths were 189 bp, 201 bp, and 212 bp for *T. trichiura* ITS1 and 246 bp for the *Ascaris* sp. CO1 PCRs. For the pre-treated samples, two PCRs for *T. trichiura* ITS1 and one for *Ascaris* sp. CO1 produced bands of appropriate length, whereas the untreated eggs only produced one band for *T. trichiura* ITS1.

To optimise the homogenisation time for the samples, a set of eight samples from one site with a relatively high count of parasite eggs (latrine from Anglo-Saxon Hungate, York, UK) was subjected to a series of 30 second bursts at maximum speed. Five grams of the sample were prepared as described earlier in 20 ml water. Eight screw-cap tubes were filled to about one third of the volume with 1 mm glass beads and 1 ml of sample. All tubes were inserted into the BeadBeater and after each burst of 30 s one vial was examined for the presence of intact and damaged parasite eggs. Damaged eggs started to appear after 1 min and the last sample containing

intact parasite eggs was after 2.5 min (5 bursts). The parasite eggs were not counted as the presence of intact eggs was deemed sufficient to indicate incomplete homogenisation. As a consequence of this trial all samples handled subsequently were subjected to a minimum of 3 min homogenisation.

The initial optimisation was based on samples with only *Trichuris* and *Ascaris* eggs. In a later stage of the project the experiment was repeated to be applied for other parasite eggs. The samples from the excavation in Lübeck contained comparably high numbers of the cestodes *Taenia* and *Diphyllobothrium*. After an inefficient PCR for *Taenia* on samples where *Taenia* had been diagnosed microscopically, the experiment was repeated with a series of 1 min bursts. *Taenia* eggs were more resistant than the nematode eggs and required 5 min of homogenisation. All relevant samples were re-processed to adjust for the longer homogenisation. After homogenisation tubes were centrifuged for 5 min at 13,300 rpm (Heraeus biofuge primo). This process pellets insoluble debris with the glass beads and leaves the aDNA in buffered solution. The supernatant was processed through a silica column-based DNA extraction based on a commercially available genomic DNA extraction kit (Qiagen DNeasy Blood & Tissue).

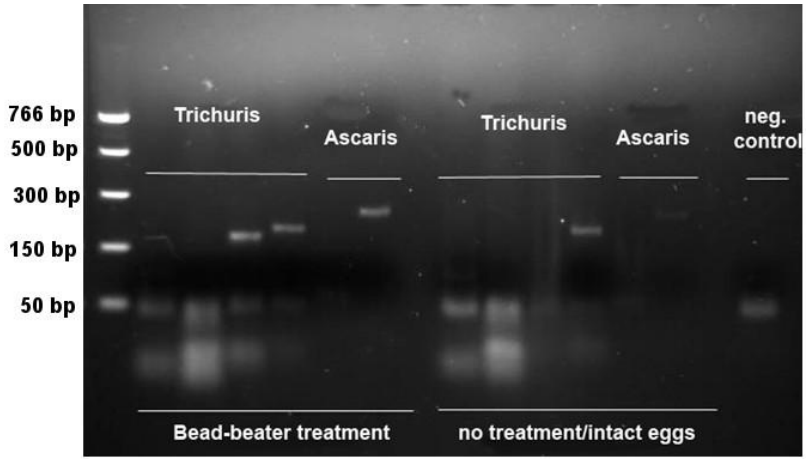


Figure 2.3: PCR amplification on homogenised and non-homogenised salt-floated parasite eggs.

Two sets of PCR reactions for *Trichuris trichiura* (three PCRs) and *Ascaris* spp. (two PCRs) were performed on homogenised eggs (on the left hand side) and non-homogenised eggs (right hand side). The expected band lengths were 189, 201, and 212 bp for *T. trichiura* ITS1, and 243 bp and 242 bp PCRs of *Ascaris* sp. CO1

2.3.4 Isolation of aDNA from homogenised soil samples

The commonly referenced articles on the isolation of aDNA from various substrates recommend an extraction using a phenol-chloroform precipitation protocol (Drancourt and Raoult 2005; Hofreiter et al. 2001; Willerslev and Cooper 2005). A similar protocol was used on intestinal nematode eggs (Oh et al. 2010). There are a number of reasons against employing these methods, including the low efficiency for fragmented DNA and high toxicity of the reagents. The advantage of this approach was the relatively low cost of the reagents and the relatively efficient removal of contaminants soluble in organic solvents. However, a column-based extraction protocol based on a commercially available kit for the extraction of genomic DNA was evaluated (Qiagen DNeasy Blood & Tissue). The unmodified column-based protocol proved to be efficient for the majority of samples. However, as the composition of the soil varied, a disadvantage of the column-based method was the

absence of an efficient method to remove PCR inhibitors such as charcoal. It is thought that charcoal and other charred remains were sometimes used to cover the faeces to reduce the smell or that discarded charred barrels were used to remove the contents of latrines. Some protocols suggested pre-treatment of the samples with 0.5 M NaOH. This approach was evaluated but rejected as it solubilised contaminants which were subsequently eluted into the aDNA extract and gave the extracts a strong colouration (Figure 2.4) and these extract were found to inhibit PCR reactions (Figure 2.5). The introduction of an additional washing step using 6 M Guanidine thiocyanate (Gu-SCN) reduced the amount of visible contaminations (Figure 2.4) and PCR inhibitors (Figure 2.5) in most samples. Even with the adapted protocol some of the extracts contained uncharacterised PCR inhibitors which were not removed with these protocols. Extracts like the ones shown in Figure 2.4 (c/d) contained obvious traces of other materials which inhibited subsequent PCR amplification (Figure 2.5). For samples where the unmodified manufacturer's protocol (Qiagen DNeasy Blood & Tissue Kit) proved to be inefficient a modified protocol was introduced. A commercially available genomic DNA extraction kit developed for extraction of DNA from food items was evaluated (Mericon food kit, Qiagen). This kit was based on a similar silica-based column, but included different washing steps and a chloroform-based removal step for organic solvent soluble substances lowering the DNA yield. Even though this protocol included a step which required chloroform, tests showed that this step was not always required. Smaller amounts of organic materials were efficiently removed by the special extraction buffer as could be observed by the removal of the colour in the supernatant. If the supernatant after the first step was still coloured, the additional chloroform step was required for optimal yield and quality of the aDNA extract. This approach proved to

be efficient on problematic samples, but due to the increased costs and higher toxicity of reagents it was only applied if the extraction with the standard column based kit according to the manufacturer's protocol (Qiagen DNeasy Blood & Tissue, with an additional 6 M Gu-SCN wash step) was not suitable. As an example four extracts are shown in Figure 2.4. These were four replicates obtained from a single sample from the excavation of Finzel's Reach in Bristol (UK) which contained charred remains. Observations of the excavating archaeologists suggested that these samples were from a charred barrel filled with excrement which was disposed of outside of the city walls. Extracts a/b were generated using chloroform-based extraction and extracts c/d using the standard protocol without additional organic compound removal. The removal of the organic colorants in the samples was clearly visible.

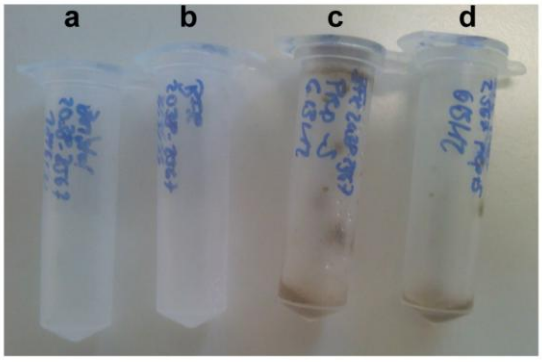


Figure 2.4: aDNA extractions from Bristol Finzel's Reach to evaluate alternative extraction protocols.

The two duplicate samples on the left (a/b) were extracted using a chloroform-based extraction, the duplicate samples on the right (c/d) were extracted using the developed alternative protocol

To demonstrate the presence of PCR inhibitors a series of experiments was designed based on a working PCR reaction for a cloned *Ascaris* CO1 fragment

(Figure 2.5). Lanes 1 and 2 contained two second round PCRs which were known to work. To the same reactions, a small amount (1 µl in a total volume of 25 µl) of a potentially inhibiting extract was added which caused the reaction to stop working (lanes 3 and 4). The same PCRs with only the inhibited extract was also performed and resulted in no product (lanes 5 and 6). These experiments proved that some extracts contain inhibitors since the spiked controls ceased to produce discernible products.

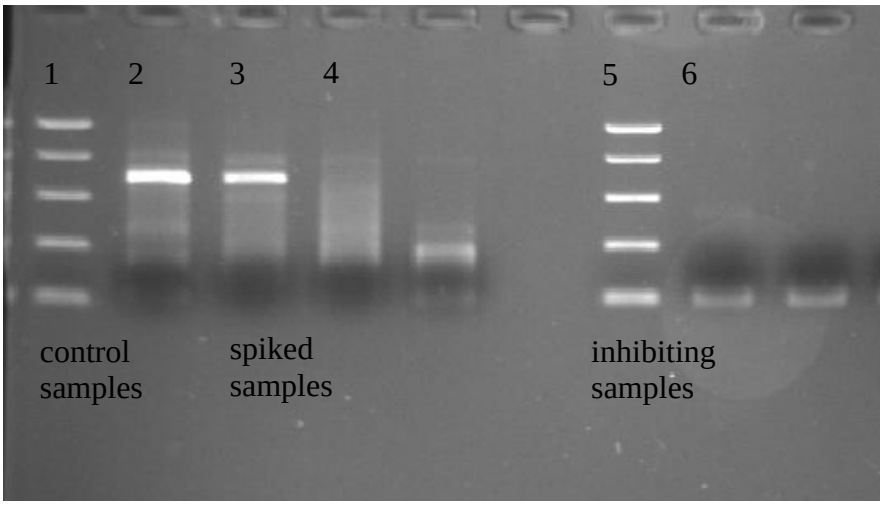


Figure 2.5: PCR for inhibition control of aDNA extracts.

PCR reactions were performed on four samples of which two were known to give PCR products for a cloned *Ascaris* CO1 fragment PCRs (lanes 1/2) and two which were suspected to be inhibiting the PCR (lanes 5/6). The lanes 1 and 2 are the controls which give good bands. The addition (spiking) of these conditions with the potential inhibitor caused the PCR to cease working (lanes 3/4). In lanes 5 and 6 are the two extracts with expected inhibition of PCR.

2.3.5 General PCR protocol

Environmental samples from archaeological contexts contain very small quantities of short fragmented aDNA. Hence, the optimisation of the PCR protocol was crucial for the generation of PCR products and sequences.

Generally, PCR conditions are tested and optimised by using a dilution series of cloned target DNA template and the inclusion of both positive and negative

controls with each set of reactions. However, as the work described here was similar to forensic work where only traces of DNA are available, the inclusion of concentrated modern template would jeopardise the integrity of the results, as the small amounts of aDNA would be masked by higher concentrations and superior quality of modern DNA extracts. Thus, the optimisation was performed in two separate steps. The initial evaluation of primers and PCR conditions was done on selected primary aDNA samples, with selection either based on high concentrations of eggs in the sample or preliminary positive PCR results. The second part of the optimisation was based on dilution series of a cloned PCR product to evaluate the sensitivity of the assay.

To produce a sufficient amount of the PCR amplicon, two rounds of PCR were required, taking a part of the first reaction as template into the second round. To minimise the amplification of inappropriate target sequences the second round PCR primers were nested in one (hemi-nested) or both directions by at least two base pairs. Products of the second round PCR were subjected to electrophoresis with a 2% agarose/ethidium bromide gel. Optimisation was based on the presence or absence and the intensity of appropriately sized bands on the gel. If more than one product or products of non-predicted length were observed, the first step was an in-depth *in silico* analysis of the primers. As the aDNA extracts were derived from environmental samples the DNA of non-target organisms (such as plants, fungi or free-living eukaryotes) may be present in the extracts. As the PCR conditions were optimised for relatively short sequences the selected primers could lead to the amplification of non-target organisms as well as the intended targets. Hence, optimisation of the PCR was very important, especially the choice of suitable targets in each parasite genome and the design of suitable primers for it.

To assess the specificity of the primers and the optimal PCR conditions two experimental optimisation steps were introduced: a temperature gradient and a dilution series. Both optimisation steps require a cloned PCR product generated with appropriate primers and of known concentration to optimise the conditions. As general rule the annealing temperature in the PCR should be slightly below the calculated melting temperature of the primers. To optimise annealing temperature, a series of PCR using the same master-mix of primers, template DNA and reagents was subjected to a series of different annealing temperatures (Applied Biosystems Veriti 96-well Thermal Cycler). Based on the intensity of bands of appropriate size in a gel electrophoresis the optimal temperature for annealing could be estimated. The example shown in Figure 2.6 is based on the optimisation of the *Ascaris* CO1 primer pair F2-R, which generated a fragment of 242 bp length. Two repeats of the PCR based on two different templates (both with verified sequence and concentration) were performed. In the shown example, the optimal temperature was subsequently set to 52°C.The highest temperature of the gradient represents the calculated melting temperature of the reverse primer as the calculated melting temperature of the forward primer was 4°C higher.

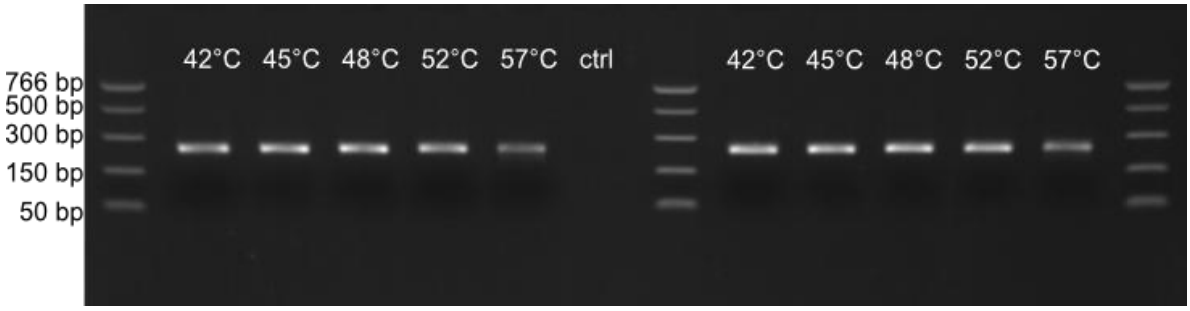


Figure 2.6: Temperature gradient for annealing temperature optimisation.

Based on a cloned product of *Ascaris* CO1 the optimal temperature for the PCR was estimated. The assessment was based on two independent cloned PCR products from different samples.

Optimisation of the annealing temperature was done with a relatively high concentration of target DNA. The PCR produced cloned into the pCRII vector had a length of 4212 bp. For the PCR, 10 ng of DNA was used, which correlates to 3.84 10^{-15} mol or 2.31 10^{9} copies of the vector with insert. Based on these calculations a dilution series of eight 10-fold dilutions was set up, ranging from 10 ng (2.31 10^{9} copies) to 1 fg (2.31 10^{2} copies). The example in Figure 2.7 was based on the same primers and samples as the example in Figure 2.6. The PCRs were performed using two the round protocol (with 5 µl in 25 µl of the first reaction as template for the second round) as used for primary samples. The assay produced weak bands up to a dilution of 10 fp/2310 copies.

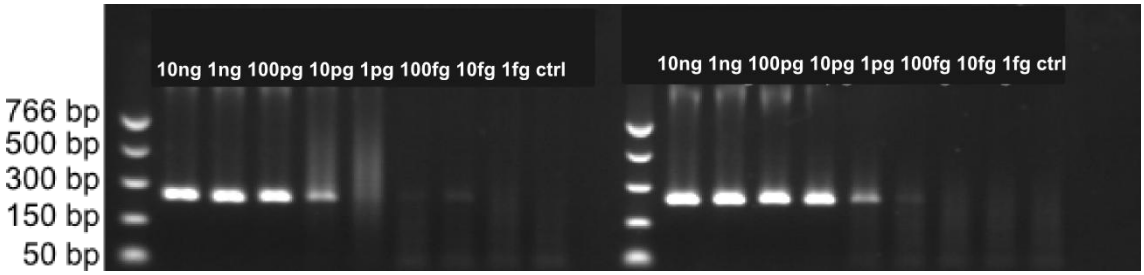


Figure 2.7: Dilution series to evaluate sentitivity of the PCR assay.

Two eight step series of 10-fold dilutions based on two samples were used for a two round PCR. In this gel from the second round some variation between samples was noticed, but the assay was sensitive to amplify a PCR product from 100 fg/2.31 10^{4} copies (right) and 10 fg/2.31 10^{3} copies (left).

2.3.6 Cloning of PCR products and Sanger sequencing setup

Amplified products with lower levels of variation were suitable for direct (Sanger) sequencing. However, most samples contained are mixed populations with more than one sequence present which made direct sequencing not appropriate. Small insertions or deletions lead to unreadable sequence in a mixed population of PCR amplicons. These variations in the amplified sequence necessitate a protocol to

separate these products prior to sequencing. Cloning the fragments generated by PCR into a suitable bacterial vector is the commonly used approach. Therefore, the direct PCR product sequencing was replaced by TA-cloning before sequencing. The principle of this cloning approach is based on the 3'-end single A (adenosine) overhangs of both strands in a PCR using *Taq* DNA polymerase (originally isolated from *Thermus aquaticus*). Standard vectors with a T overhang will bind to the A-overhang of the PCR product which allows the integration (ligation) into an open vector using bacterial DNA ligase enzyme. Commonly, these vectors are cut open within a coding marker gene (usually the *LacZ* cassette) whose function is disrupted if a PCR product is inserted. This allows for a simple selection against bacteria which do contain a vector without an insert. The functional gene metabolises a substrate with which the culture plates are treated. The method used here was the X-Gal/IPTG screening. The *LacZ* gene encodes β-galactosidase which hydrolyses the β-glycosidic bond between galactose and other organic compounds, such as between galactose and glucose in lactose. The enzymatic activity allows for a second level selection based on metabolites which are coloured. The addition of the colourless lactose-analogue X-Gal (5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside) which is hydrolysed into 5-bromo-4-chloro-indoxyl which dimerises and oxidises to form the insoluble blue pigment 5,5'-dibromo-4,4'-dichloro-indigo allows selection based on the colouration of the bacterial colonies, where blue colonies do not contain the PCR insert. IPTG (Isopropyl β-D-1-thiogalactopyranoside) is the strongest known inducer of the reaction. Additionally, the vector contains an antibiotic resistance gene which only allows bacterial cells with a plasmid to grow on the selection media plates (in this case ampicilin). As each vector can only take up a single PCR-generated DNA fragment this technique can be used to separate PCR products which are very similar

by picking a single bacterial colony. On the selection plate, single colonies which originate from single bacterial cells can be selected and the plasmid DNA extracted for sequencing. A further advantage is that the sequencing can start within the plasmid sequence which improves the read quality of the target sequence

In one example presented here, only 1 sequence from 10 initially picked colonies was found to be a true Ascaris CO1 sequence, with the other sequences being either primer concatomers or non-specific targets of environmental DNA which could not be identified by database searches (such as BLAST). Hence, all colonies which could be clearly identified as such (not in a smear) and were clearly not negative by the blue-white selection, were picked. 280 colonies from this one TA-cloning experiment (based on HLGV FS16 1 of 2) were picked and screened using one product primer in combination with vector specific M13 F/R primers. 48 colonies were identified as potentially positive and prepared through MiniPrep (4.5 ml over night culture, Qiagen DNeasy MiniPrep) and sequencing. 48 of 280 is equivalent to 17.14% of the colonies. Sequencing resulted in 42 out of 48 readable sequences. Of these 42, 34 were close to the template sequence (80.95%) and identified as real Ascaris CO1. The remaining 8 sequences group in two clearly distinct groups of 4 sequences each. The source organism of these sequences could not be confirmed since the sequence did not provide sufficiently high homology in a BLAST query of the NCBI database. The alignments showed one predominant product and three other products which showed close to 99 % identity (242 bp length). As a side effect of the extensive screening it was noted that different sequences were found to be represented in different copy numbers in a mixed sample.

To optimise the process, three different commercially available vectors (Invitrogen TA Cloning® Kit pCR™2.1, Qiagen PCR cloning, Promega pGEM®-T Easy Vector System) and two different suppliers of competent cells (DH-5α, Invitrogen and New England Biolabs, NEB) were used. The cloning kits from Invitrogen and Promega performed comparably, whereas the kit obtained from Qiagen did not perform adequately (less than 10 % white colonies in an initial trial). There was no performance difference between the cells from the two evaluated suppliers. As the kits from Invitrogen and Promega performed comparably, the decision on which kit to use for further experiments was based on the simplicity, lower processing time and lower cost of the Promega pGEM®-T Easy Vector System and the lower cost of the NEB cells.

2.3.7 Intestinal parasites and selection of gene targets

The current work includes seven target species: *Ascaris* spp., *Trichuris trichiura*, *Diphyllobothrium latum*, *Taenia* spp., *Cryptosporidium* spp., *Giardia intestinalis* and *Enterobius vermicularis*. Most work has been done on the first two species, and work for the cestode parasites *D. latum* and *Taenia* spp. was largely restricted to the sample set from the Medieval port of Lübeck where these parasites were detected microscopically. For each species, a number of PCR primers for specific targets in the nuclear and mitochondrial DNA were designed.

For *T. trichiura* the targets were the ribosomal ITS1, the cytoskeleton gene β-tubulin, the mitochondrial CO1 and the Cytochrome *b* gene. For *Ascaris.* sp. the targets were ITS1, β-tubulin, the mitochondrial CO1, Cytochrome *b* and NAD1 (NADH dehydrogenase fragment 1). For *Enterobius vermicularis* the used target is the ribosomal 6S rDNA. To date, the PCR for *E. vermicularis* was the only of the

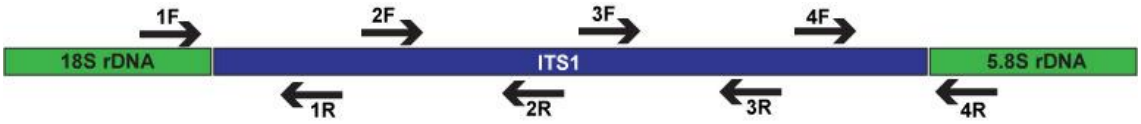
described target organisms which did not show any positive PCR results despite being the most common helminth in the western world. For *D. latum* and *Taenia* spp. the targets were the mitochondrial CO1 and Cytochrome *b*. For *Taenia* spp. different primers for *T. solium* (porcine tapeworm) and *T. saginata* (bovine tapeworm) were designed to distinguish the two species. For *Cryptosporidium* spp. the targets were the 16S ribosomal DNA, HSP70 (70kDa heat-shock protein), gp60 (60 kDa surface glycoprotein), and COWP (*Cryptosporidium* Oocyst Wall Protein). These targets are also commonly used for subtyping of *Cryptosporidium* spp. (L. Xiao et al. 2001; L. H. Xiao and Ryan 2004). The COWP gene was selected because it codes for a protein that is important in maintaining the integrity of the oocyst wall, allowing the parasite to withstand harsh environmental factors (Guy et al. 2003); gp60 and HSP70 (which did not work in initial trials) were chosen because modern studies reported variability and some spatial signatures (Beck et al. 2009; Grinberg et al. 2013; Zintl et al. 2011). For *Giardia intestinalis* the targets were the ribosomal 16S DNA and the β-giardin gene. The β-giardin gene of *Giardia* codes for a structural protein that is a component of the adhesive disk of the parasite. The primers for the β-giardin gene were modified from a publication describing diagnostic procedures (Mahbubani et al. 1991).

For a comprehensive phylogenetic analysis it is useful to focus on those genes with a good number of sequences available in the GenBank database. For all seven parasites (*Ascaris*, *Trichuris*, *Taenia*, *Diphyllobothrium*, *Enterobius*, *Cryptosporidium* and *Giardia*) in this study there were good numbers of sequences available (between 57 for *Ascaris* CO1 to 1660 for *Cryptosporidium* gp60). At the start of the project only the full genome sequences of the protozoan parasites (*Cryptosporidium* and *Giardia*) were available (Abrahamsen et al. 2004; Franzen et

al. 2009). In the course of the project more genomic sequences were published, for *Ascaris* (Jex et al. 2011), *Taenia* (Tsai et al. 2013) and *T. trichiura* (Foth et al. 2014). The genome of *D. latum* has not been published yet and for *E. vermicularis* only a mitochondrial genome is available (Kang et al. 2009). For the choice of the target gene it is useful to identify orthologous in other species. An ideal target would be one which has previously been used for molecular identification or speciation since a larger number of sequences from different geographic origins are available. The main constraint of aDNA is the state of degradation and the reduced length (typically about 200 bp). In the best case these targets would include a variable and a highly conserved region, with the latter being important for the design of the primers. Unfortunately, sometimes these regions exceed the 200-300 bp margin and are less likely to be useful for fragmented aDNA(Cooper and Poinar 2000; Hofreiter et al. 2001; Willerslev and Cooper 2005). In some cases primer sequences were modified from previous published sequences (Guy et al. 2003), in other cases novel primer sequences were tested. For some targets, a series of PCR primers yielding overlapping products of about 200 bp each were designed to amplify longer regions, for instance for the Internal Transcribed Spacer region 1 (ITS1) in *T. trichiura* which is commonly used for species identification, but stretches over more than 800 bp (table in Figure 2.8). In an ideal case, combinations of coding and non-coding sequences from the nuclear and mitochondrial genome would be used as these have different mutation rates. The mitochondria are passed on maternally and exhibit an increased rate of changes compared to the nuclear genome (Brown et al. 1982; Wallace 1994). Moreover, the mitochondrial genome does not contain introns. We hypothesise that a combination of the different properties of these genome sequences might allow us to address questions on different time-scales or over geographical

ranges. Despite the differences between nuclear and mitochondrial DNA, the recovery rates were found to be similar (Binladen et al. 2006).

A forensic study on human bones of different ages (1 to >200 years old) suggested that the total amount of aDNA retrieved from the bone does not vary as much as expected, but the quality is considerably reduced with increasing age (Kaiser et al. 2008). According to these analyses, longer fragments such as *T. trichiura* ITS1 would not be possible to amplify in a single reaction, however, combinations of PCR products could be used for the reconstruction. In the case of *T. trichiura* ITS1, four pairs of primers producing overlapping products (210-258 bp) were designed to span the full sequence and include some of the adjacent sequence (Figure 2.8). The primer design also allowed combination of primers from different pairs to extend the length of the fragment. Interestingly the full 819 bp product was amplified in a single reaction, which included the 620 bp ITS1 and neighbouring ribosomal sequence (primers 1F to 4R, Figure 2.8). This suggests that in at least some samples the aDNA stability was better in the enclosed environment of the parasite eggs than with previously studied bone or tissue material.



	1F	***2F***	***3F***	***4F***
1R	210	-	-	-
2R	410	250	-	-
3R	612	452	252	-
4R	819	659	459	258

Figure 2.8: Design of PCR primers for *T. trichiura* ITS1 and the expected lengths based on published sequences.

Four pairs of primers were designed to give products of 210-258 bp length. The overlap of the products would allow a reconstruction of the full-length of the fragment, as well as combinations of primers from different pairs. The graphical representation of the gene indicates the location of the primers with the ITS1 region indicated in blue and the flanking ribosomal region in green. The table indicates expected fragment lengths for each primer set.

To determine the best combinations of primers for the *T. trichiura* ITS1, all possible combinations of the primers described in Figure 2.8were evaluated in a PCR on two samples from the excavation of Medieval Lübeck. The PCR products from two samples (AS17 1/3 and FS16 1/2) were analysed on a 2% agarose/ethidium bromide gel to visualise lengths (Figure 2.9). Each set of reactions included a negative control (water, no aDNA extract). The largest band of the marker (766 bp) was slightly smaller than the longest predicted fragment (1F-4R, 819 bp).Sequencing analysis confirmed the identity of the region and showed some length variation. Sample 1 performed better than sample 2, however, either one or both samples yielded products in the expected size range (see table in Figure 2.8) for most primer sets. Reactions with primer sets 1F-1R, 1F-3R, 1F-4R, 2F-4R, and 3F-4R were further processed by TA-cloning reaction to verify the sequence. Other primer

combinations were dismissed either due to the complete absence of signal (e.g. 4F-4R, Figure 2.9) or the unclear nature of the signal obtained (e.g. 2F-2R, Figure 2.9).

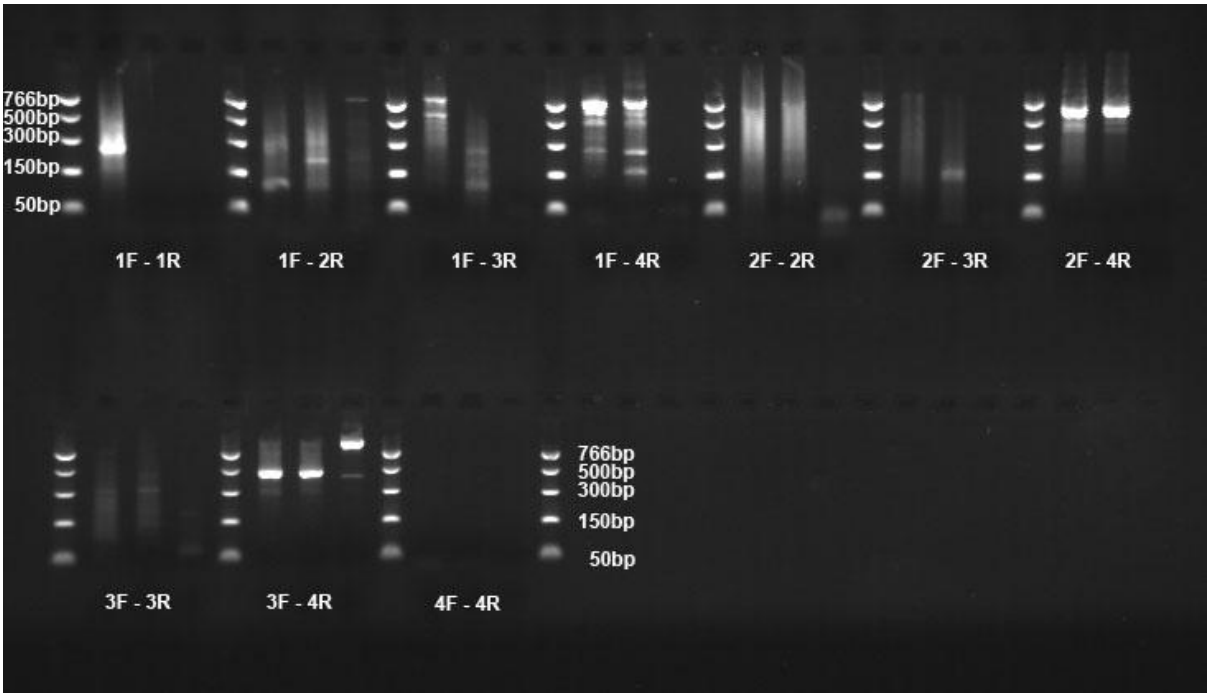


Figure 2.9: Analytic electrophoresis gel for PCR reactions for *T. trichiura* ITS1 on two samples from Lübeck.

The primer pairings are indicated underneath the sample lanes. The ladder used is the NEB QuickLoad with fragment lengths indicated. The layout is for each primer set reaction as follows: Marker, Sample 1, Sample 2, negative control, blank lane.

TA cloning of the fragments revealed that the slight blur at the top of the longer fragments was due to length polymorphisms of the target sequence. These length varied by less than 1 % (811-822 bp) of the amplified sequence. The alignment is included in the appendix.

In almost all samples, irrespective of their age or geographical origin, parasites were detected in the primary sample using microscopic techniques. Overall, the target aDNA fragments were amplified with more than half of the samples, and >100 bp of sequence was obtained with about 80 % of the PCR products.

2.3.8 Sequence evaluation

Preliminary analysis of sequences from modern parasites available in public resources (such as GenBank) showed that various genomic regions can be used to distinguish geographical origin (Callejon et al. 2010) and/or species identity (Figure 2.10). These conclusions were based on sequence alignments and boot-strapped neighbour-joining trees of *Trichuris* spp. which was chosen for the availability of sister species from other hosts and from different geographical origins. Within the *Trichuris* genus there are a number of related species that can be separated by the ITS1 sequence similarity. ITS1 sequences are available for a range of *Trichuris* spp. The largest number of sequences area available for *T. trichiura* which infects humans, *T. suis* (porcines) *T. vulpis* (canines) and *T. muris* (rodents), Although these parasites are relatively host-specific, *T. suis* or *T. vulpis* can occasionally infect humans (Areekul et al. 2010). As an effect of that, some sequences (mostly when isolated from humans) are mis-annotated in the database, which can be seen by the sequence similarity on the neighbour-joining tree (Figure 2.10).

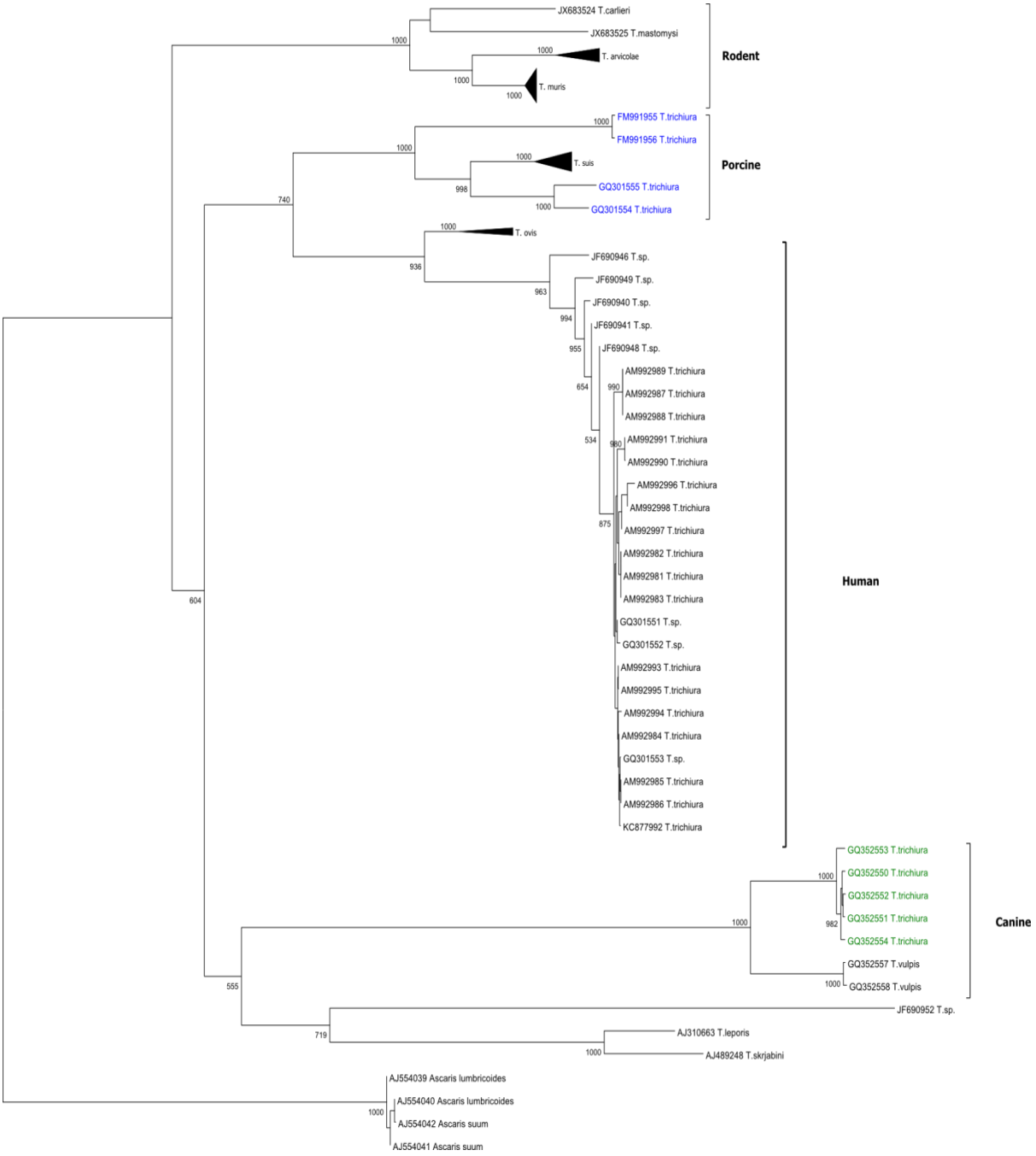


Figure 2.10: Boot-strapped Neighbour-joining tree of *Trichuris* spp. ITS1 sequences for species identification.

Trichuris from different host species, including humans, rodents, dogs, and pigs were analysed. The tree was rooted using *Ascaris* spp. ITS1 sequences. The identification of the host is visible. Sequences were found to be mis-annotated in the database, these were likely sequences where the non-human specialist parasites infected humans (*T. vulpis* in green or *T. suis* in blue) which has been reported in various publications.

PCR amplified target sequences from archaeological samples were either direct sequenced or TA cloned to generate sequence data. The verification of sequence was achieved in several stages. The first step was a coarse alignment of the sequences with each other and the reference sequence. The purpose of this was to exclude sequences which did not fit the alignment due to technical issues such as unreadable sequence or false-positives (e.g. empty TA cloning vectors). The aligned sequences were then examined using the BLAST (Basic Local Alignment Search Tool, Altschul et al. 1990) search algorithm which scores them according to their similarity to database sequences. The next step was a more stringent alignment including multiple modern sequences of the same species and sequences of related and unrelated species. A neighbour-joining tree analysis was then employed to identify that the tested sequences were most closely related to the expected target sequence. As most data was obtained for *Trichuris* ITS1 this will used for the detailed analysis of the approach. Examples of other parasites in the study will be discussed is less detail. A tree was constructed using GenBank entries of the target species *T. trichiura* (green label), the related species *T. suis* (yellow), *T. muris* (red) and *T. vulpis* (magenta) as well as the two other nematode genera *Anisakis* (grey) and *Ascaris* (blue). The sequences of *Anisakis* and *Ascaris* act as out-group reference sequences to demonstrate the identity of the amplified sequence's identity. The resulting phylogenetic tree (Figure 2.11) shows different results. The tree includes some aDNA samples from Lübeck which all fall within the *T. trichiura* clade (GenBank ID GQ35348, GQ35350, GQ35352, GQ35353, GQ35354, and GQ35355). Second, this tree suggests that strains of *T. trichiura* may differ within the *T. trichiura* branch when isolated from Thailand (upper group) or China (lower group). Thirdly, the aDNA sequences also exhibit diversity. This also shows that the aDNA

derived sequences are different and not cross-contaminations as for which we would expect identical sequences. The diversity between two samples from one site is considerable which potentially is very valuable.

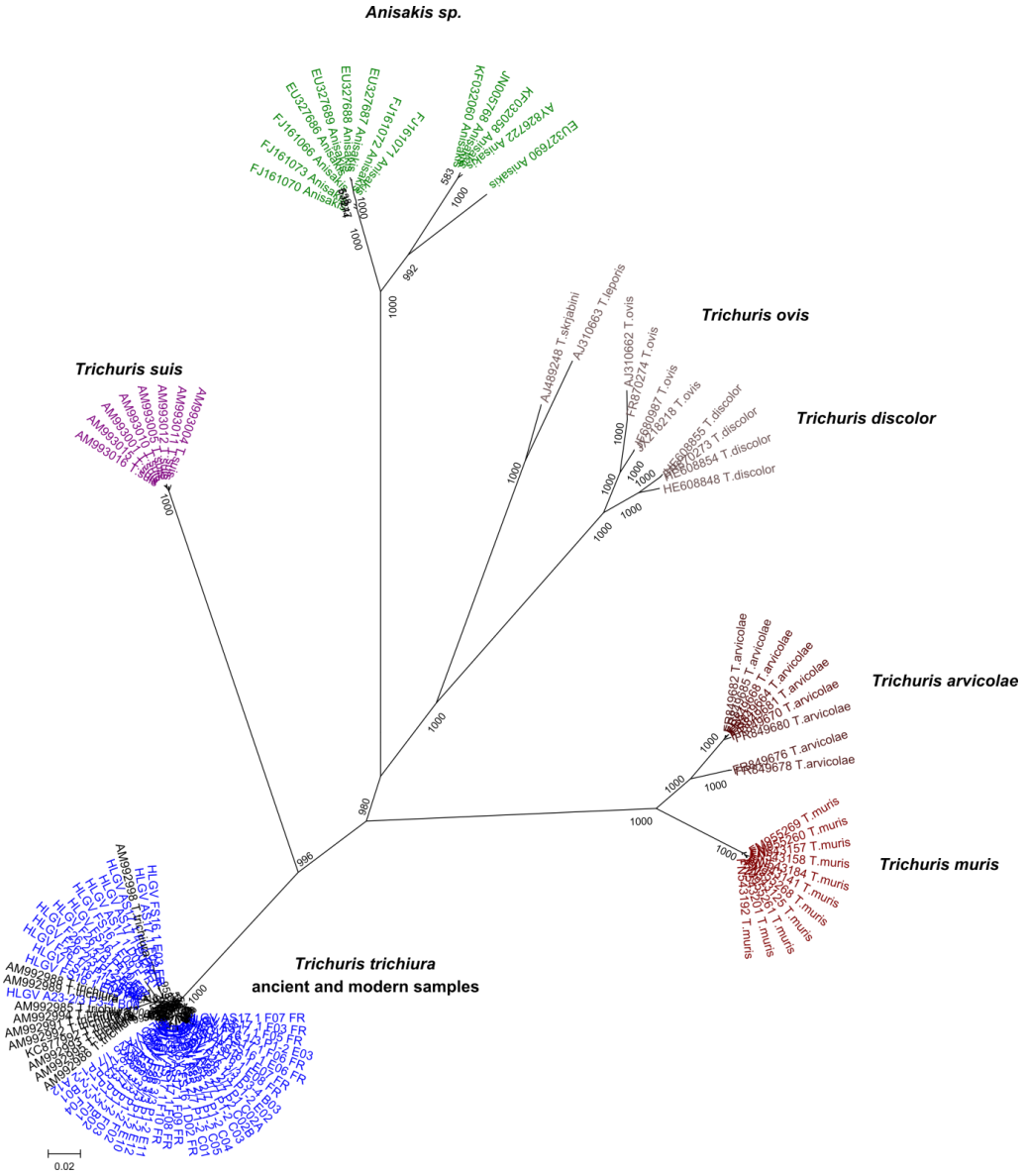


Figure 2.11: Neighbour-joining tree of sequences generated *T. trichiura* ITS1 from Lübeck samples.

The ancient samples (blue) were within the modern *T. trichiura* (black) clade. The ancient samples were clearly distinct from rodent parasites (*T. muris* and *T. arvicolae*, red), porcine hosts (*T. suis*, purple) or sheep and goats (*T. discolour, T. ovis, grey*). The outgroup on the tree is *Anisakis* sp. (green).

The ITS1 sequence is specific enough to determine the species, yet diverse enough to allow for polymorphisms, it offers potential for the application as molecular marker for human interactions and migrations. Although the analysis suggests that this process identifies appropriate sequences it is important to verify the nature of any variation found. A combination of different genetic loci will be used, including loci which are expected to be highly conserved and loci with variable regions. Further, measurements have to be taken to ensure that the sequence variations are genuine and not introduced by PCR polymerase errors. These measurements are based on a three tiered approach. These include (1) the sequencing of multiple products of the same PCR amplification to assess the variety of sequences amplified and the PCR-induced variability, (2) the repeated PCR amplification of a single subsample, and (3) extraction of parasite aDNA from subsamples of our samples to assess the variability through a single sample.

2.4 Analysis of *Ascaris* CO1

Considerable effort was put into the generation of a large number of sequences for *Ascaris* CO1. Although initial results were promising as they showed indications site and age specific variation, there were problems with the specificity of the primers which was addressed by the design of nested primers. These still produced other non-ascaridian products for the aDNA samples. To screen for the *Ascaris* CO1 sequence, a screening PCR based on the cloned PCR product was developed. The screening could be used to determine both the presence of an insert in the vector as well as an identifier for the correct insert. The PCR consists of three primers, two primers bind on a target on the vector (M13) and one that binds on the

insert. There were different possibilities for resulting bands: one low band (212 bp) represents the empty vector, any larger band represents an insert. The expected resulting bands would be at 460 bp for the insert with M13, and at 151 or 169bp (depending on the direction of insertion into the vector). The presence of the double band thus shows the correct insertion (lanes 1, 2, 5, 7, 15, 18, 19, 20, 22 and 25 in Figure 2.12).

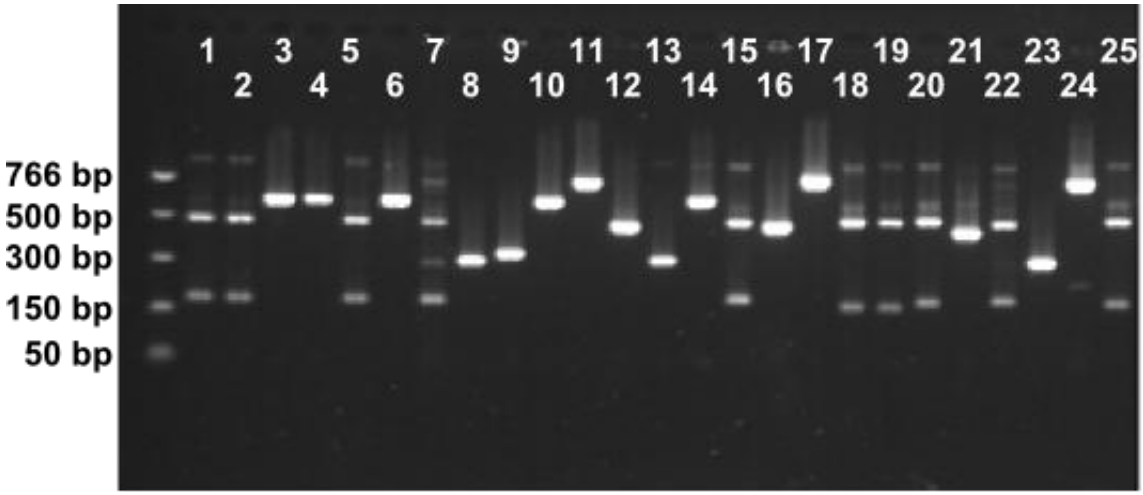


Figure 2.12: PCR gel for screening of colonies from TA cloning of *Ascaris* CO1 PCR products.

The PCR was performed using three primers, of which two bind to the vector and one inside the insert. Expected bands are at 212bp for empty vector, and double bands (151 bp or 169 bp and 460 bp) represent positive results. All other bands represent other inserts.

The PCR-based screening method was not sufficiently discriminating against other products. Although the resulting sequences were found to be closest to modern *Ascaris* CO1 sequences in a BLAST search and were aligning sufficiently with modern sequences, subsequent phylogenetic analysis did show that the majority of the obtained sequences were derived from other eukaryotic species, some of which could not be identified (Figure 2.13). The species of those is not relevant for the project and thus the determination of the exact identity of the source organism was

not regarded as a priority. The neighbour-joining tree in Figure 2.13 showed the relationship between aDNA-derived and modern CO1 sequences from a range of organisms. These results initiated evaluations for a parallel sequencing approach which is introduced and discussed in chapter 3. Parallel sequencing of all sequences present in the sample would show whether the expected *Ascaris* CO1 sequence was amplified from the sample.

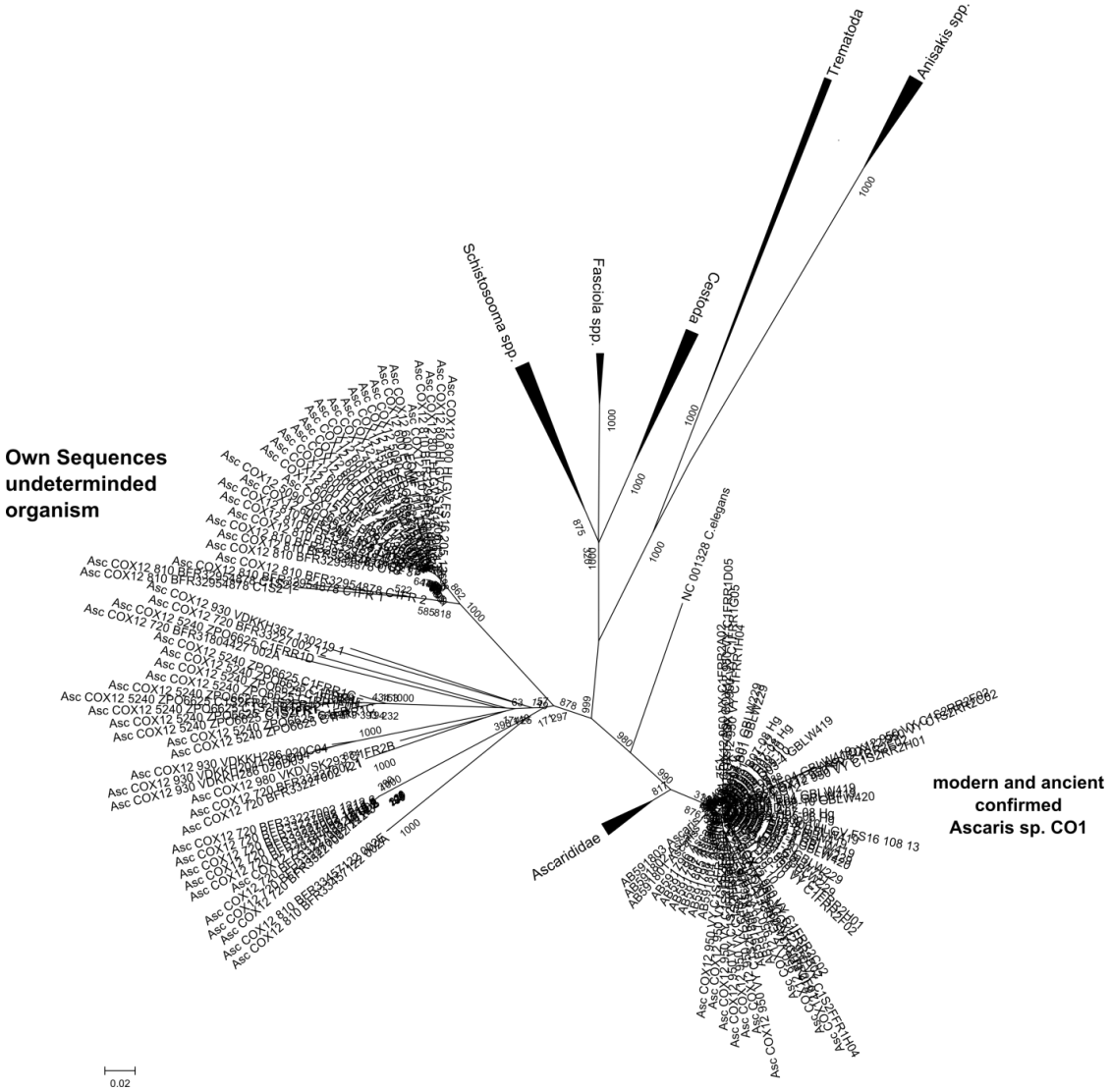


Figure 2.13: Bootstrapped Neighbour-Joining tree of sequences generated with primers for *Ascaris* CO1.

One part of the sequences could be identified as *Ascaris* sp. CO1, but for a sizeable portion of the sequences it remains unclear what species they were. The tree shows various outroups such as *Fasciola, Schistosoma*, *Anisakis* and *Echinococcus* (cestode).

2.4.1 Considerations on target choice

Although the amplification of *Ascaris* CO1 has generated mostly non-target sequences, it is of interest to evaluate the genuine sequences. The sequences generated show only very limited variation and thus it might be an appropriate target for species identification, but not for the evaluation of site differences. In total, 46 confirmed sequences were obtained from five sites. The CO1 gene was chosen as it is an essential gene in the target organism. However, there is very little variation in this gene which makes the distinction between sites very difficult. The variation between the modern consensus and the ancient samples and within the ancient samples is minute, but comparable to the sequences obtained from GenBank. Over the full length of the PCR product there is less than 4 % diversity and some of the sequences are identical. Low sequence variations were not unexpected as it had been published before that the average sequence difference within mitochondrial DNA of *Ascaris* is low (0.3-0.5 %), even compared to other nematodes (Blouin 1998). Therefore, the CO1 gene might be a target suitable for verification of the microscopic diagnosis, but preliminary phylogenetic analysis reveals that there is not enough sequence variation between sites and between ancient and modern sequences to reveal any relationship. The choice of a suitable target is a non-trivial problem as in the ideal case there should be sufficient sequence variation between sites and between modern and ancient samples, but also there should be sufficient conserved sequence to allow for a reliable amplification by PCR. As there are not enough ancient sequences published to predict conserved and variable sites, the prediction is based on modern parasite genomic sequences.

2.5 Ongoing method development

2.5.1 Observations on the effect of sample origin on molecular analysis

Samples received, irrespective of their origin, soil samples which contain a variety of different materials which might influence the preservation of eggs or aDNA, affecting microscopic or molecular diagnosis of parasites. The oldest recovered aDNA (approx. 700,000 years old) was recovered from a permafrost environment (Orlando et al. 2013). The improved conservation in permafrost conditions has been statistically evaluated (Allentoft et al. 2012). The oldest sequence reported outside of permafrost conditions was reported to be the > 300,000 years old full mitochondrial sequence from a Middle Pleistocene cave bear bone (Dabney et al. 2013).

Whereas the temperature conditions can be constant in the case of permafrost, the conditions in other contexts are more variable, especially the chemical properties of the soil. To our knowledge, there is only one study addressing the influence of different burial conditions on the ability to recover parasite eggs. This study by the group of Shin in Korea examined 75 medieval tombs for the potential for microscopic diagnosis of parasites. However, this study could not show any specific factors on favourable or unfavourable conditions for parasite conservation (Seo et al. 2010).

Archaeological publications on soil properties generally describe the conditions from which well preserved bones or artefacts were recovered. This does not necessarily hold for the preservation of parasite material. One set of samples received from a Celtic burial near Herbertingen in Upper Swabia, Germany, the preservation of the bones in the graves was very poor (personal communication

Joachim Wahl, Tubingen) yet there were parasites diagnosed by microscopy. Specific studies on the optimal preservation conditions of parasites are missing. To assess the preservation of parasites and the optimal conditions to extract aDNA we considered also forensic literature although the timescale is different compared to archaeological material. There are studies available on the conservation of various forensic materials, such as blood, bones or teeth, after being buried under controlled conditions (Bogas et al. 2009). Although this study found marshland to be unfavourable for DNA preservation, reports on good preservation of human bodies and artefacts from bogs partially contradict this issue, so for instance the first description of *Ascaris* and *Trichuris* from an archaeological context (Szidat 1944). There are limited long-term studies on buried material which was evaluated with archaeological questions (Crowther 2002; Lawson et al. 2000).

2.6 Sample collection - description of samples

Samples were collected from eight countries and various time periods. Figure 2.14 shows the locations of the samples in Europe. All sites from which samples were obtained are labelled with a black dot. If the samples were obtained from graves, the site is additionally labelled with a cross. A summary table with additional details can be found in the appendix.

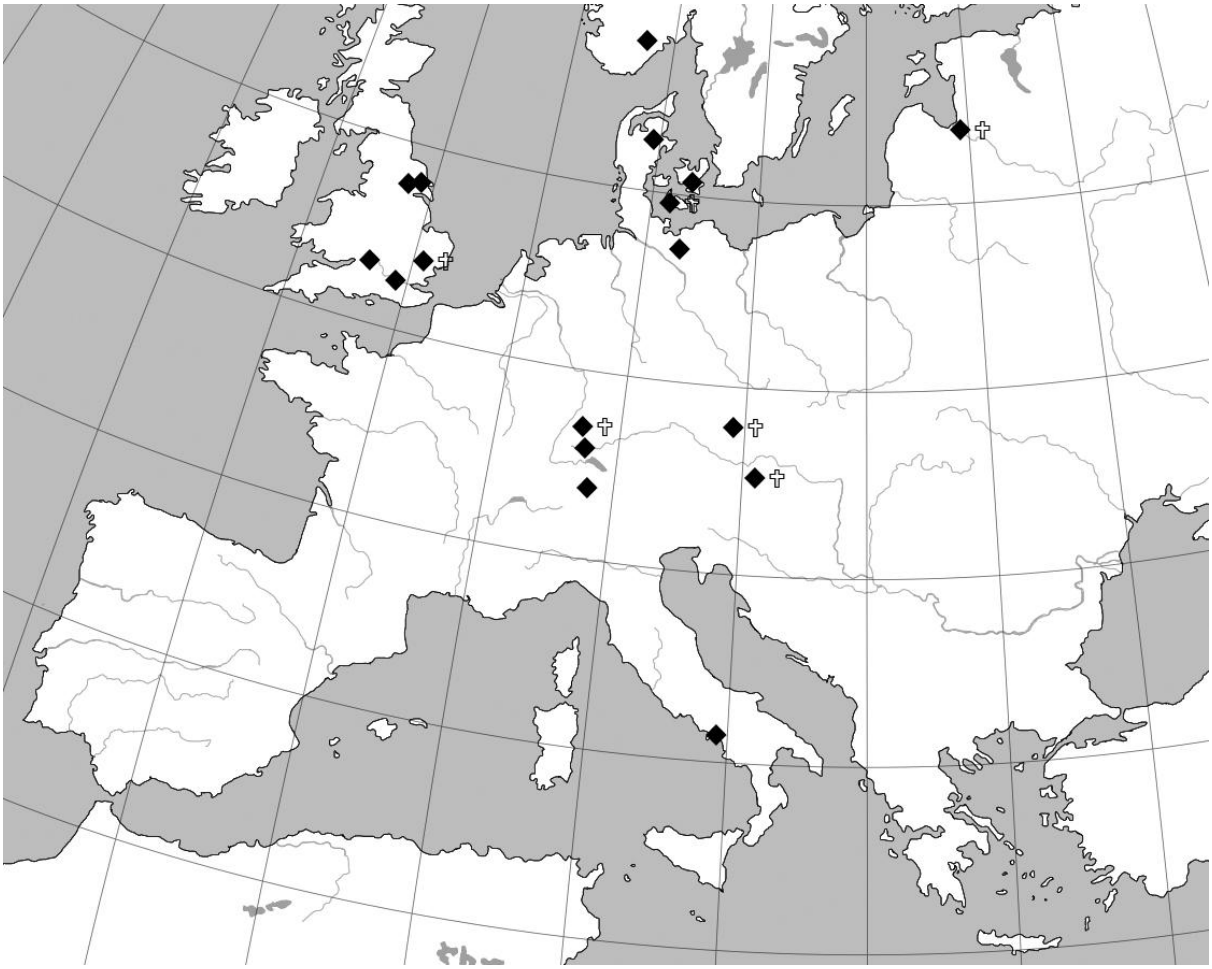


Figure 2.14: European map indicating locations of origin of sample collections

Samples were obtained from sites across Europe; a cross indicates samples collected from graveyards. A detailed description of the sites is found in the appendix.

Chapter 3

Parallel sequencing as an approach analysis of aDNA from ancient parasites

3.1 Introduction

High-throughput sequencing techniques (parallel sequencing / next generation sequencing) allow the generation of larger numbers of sequences and more extensive genome coverage than achieved by conventional cloning and Sanger sequencing. Initially the sequences generated were very short but recent developments enabled many platforms to produce sequences in excess of 200 bp which makes them useable for a broader array of studies. Several studies in archaeology applied these techniques to extend the genome coverage to publish full genomes of ancient organisms (Bos et al. 2011; Orlando et al. 2013). In this chapter a different application for parallel sequencing is described. By sequencing a large number of samples at depths not possible with the standard cloning approach it is possible to address questions relevant to epidemiology of ancient parasites.

Sequencing of short stretches of DNA is commonly done by the chain-termination sequencing technique introduced by Sanger in 1977. This technique involves the inclusion of di-deoxynucleotidetriphosphates (ddNTPs) which stop the extension of the DNA strand in an *in vitro* DNA amplification (Sanger et al. 1977). Chain-termination sequencing is suitable for sequences of approximately 400-700 bp and is relatively easy to handle. However, the number of individual sequences in the output is low compared to parallel sequencing techniques. For this project, an adapted Illumina MiSeq protocol was explored.

Illumina MiSeq parallel sequencing is a different approach to the chain-termination method as it uses sequencing by synthesis (SBS). This approach amplifies numerous fragments in parallel, recording the addition of each nucleotide. All four nucleotides in the SBS technique are reversibly labelled by a fluorescent dye. After each step the fluorescent emission is measured and recorded. In contrast to

chain-termination sequencing the synthesis is not permanently arrested after the addition of the labelled nucleotide allowing for a higher throughput (Fuller et al. 2009).

Parallel sequencing approaches are designed to amplify and read from a very large number of small DNA fragments in parallel. The introduction of labelled (barcoded) smaller fragments can make use of the short fragment read of the technique such as PCR products. The barcode employed a unique seven base-pair sequence which was used for identification of several independent samples. The barcodes were introduced as part of the second round PCR reaction by using barcoded primers. In this project this was applied to *Ascaris* CO1 and *T. trichiura* ITS1 (full primer list in Materials and Methods, chapter 10.2).

Parallel sequencing in archaeoparasitology has the potential to reveal the diversity of a genomic region to evaluate patterns between sample sets and the epidemiology of the target organisms analysed. The approach was successfully adapted for the archaeological parasite samples and two batches of samples were sequenced. The first set prepared in September 2013 consisted of 48 samples of *Ascaris* CO1, the second set prepared in March 2014 consisted of 68 samples of *T. trichiura* ITS1 and a third set of 68 *Ascaris* CO1 samples was submitted in summer 2014, but the data had not yet been analysed at the date of the submission of this thesis. Further runs were planned and the results are expected to be published in peer-reviewed publications in the near future.

3.2 Testing of primers

The barcodes for the primers were generated by Stephen Preston using the principles published by Bystrykh to generate 7 bp additions (barcodes) to the PCR primers. These 7 bp include a 4 bp unique bar code and 3 bp which act as quality control to avoid mis-allocation of a sequence in case of a mutation in the barcode (Bystrykh 2012). As with new primers these had to be tested for efficiency or interference with the PCR. Primer sets were tested on TA-cloned PCR fragments (2.5 ng of plasmid per PCR) with sequences verified by Sanger sequencing. To test the primers each primer was paired with a matching forward (*Ascaris*) or reverse (*Trichuris*) primer. The gel in Figure 3.1 depicts the PCR products for *Ascaris*; the numbers on top indicate the barcode number. Based on this gel, 48 of 50 primers were approved for the application on aDNA sample sets. The gel in Figure 3.2 depicts results for the 75 primers with *T. trichiura* ITS1. The numbers on top indicate the barcode number, with one lane on the left containing the PCR with the template and the negative (water) control on the right. All negative controls did not produce products, showing that there were no primer concatomers. The product was the expected size of 510 bp long. Six primer pairs were eliminated due to no product (018, 049, 058) or nonspecific amplification (021, 032, 050), resulting in 68 approved primers for *Trichuris*.

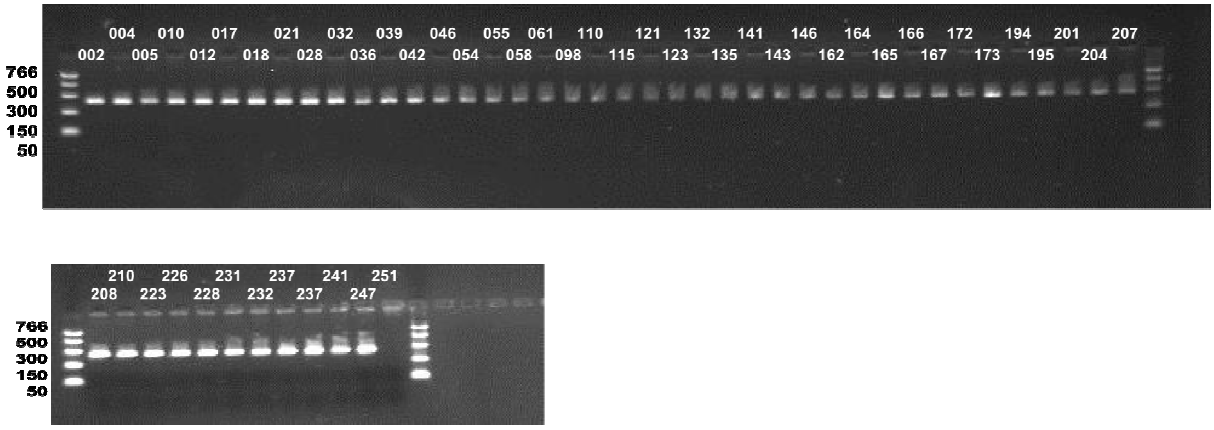


Figure 3.1: PCR gel for primer evaluation of *Ascaris* sp. CO1 barcode primers.

50 sets of primers were evaluated by PCR on a TA-cloning plasmid with known insert (2 ng per PCR reaction) with verified insertion. The lanes are indicated with the number of the barcode primer, with the reaction containing the template on the left and the negative control to the right. The expected product is 282 bp long.

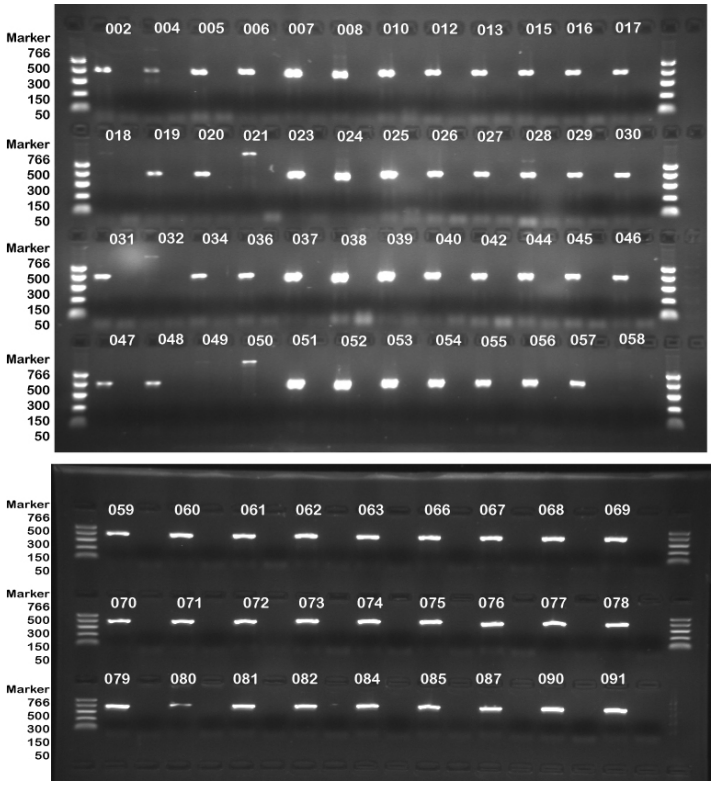


Figure 3.2: PCR gel for primer evaluation of *T. trichiura* ITS1 barcode primers.

75 sets of primers were evaluated by PCR on a plasmid (2.5 ng per PCR reaction) with verified insertion. The lanes are indicated with the number of the barcode primer, with the reaction containing the template on the left and the negative control to the right. The expected product is 510 bp long.

3.3 Preparation of samples for sequencing

Barcoded primers were used on samples which generated a second round PCR product of appropriate length for *T. trichiura* ITS1 or which sequence had been verified (based on cloned ligation mix, *Ascaris* CO1). To minimise PCR errors a high-fidelity DNA polymerase (Phusion, New England Biolabs) was used. The PCR products were processed through a column-based clean-up (Qiagen DNeasy MinElute) and the concentration was measured using Qubit® Fluorometric (invitrogen/Life Technologies). The samples were mixed in equal proportions and added to sequencing libraries.

3.4 Samples sequenced - Ascaris CO1

The first sample set submitted was a collection of 48 samples for *Ascaris* CO1. Previously there were difficulties with mis-priming and primer dimerisation of *Ascaris* CO1 (described in chapter 2.4). We formulated the hypothesis that the parallel sequencing approach should provide a sufficient number of sequences to identify some *Ascaris* CO1 variation.

Samples included in this first sequencing library were taken from different sample collections: Neolithic Bad Mergentheim (SDBM, 3), Anglo-Saxon York (Hg, 1), Viking Copenhagen (VDKKH, 1), Viking York (VY, 2), Medieval Lübeck (HLGV, 7), Medieval Bristol (BFR, 7), Medieval Minchery Farm Oxford (OXMF, 2), Medieval London (GBLW, 9), Medieval Brno Videnska (CZBV, 2) and Viking Gokstad (VNOGNS, 14).

3.4.1 Results Ascaris CO1

Although 58,687 paired reads were obtained from the library containing the *Ascaris* CO1 sequences, none of the sequences contained any barcode. Other unrelated samples contained in the same library did exhibit valid barcodes, indicating that the mistake must have happed earlier in the processing (during the sequencing sample preparation). The sequences were nevertheless extracted and identity clustered using USearch (Edgar 2010). Despite being unable to associate particular sequences with individual samples, it was still possible to obtain important information on the efficiency of the application and the numbers of genuine sequences obtained.

The 58,687 paired read sequences were extracted from the library by software written by in house (Stephen Preston). The search criteria for extraction were that the sequences must contain both primers in a paired read at the expected position (without barcode). These sequences were identity clustered, resulting in 9,678 clusters (Figure 3.3). A considerable number of sequences were found only once (singlets, 7,573, 12.9 %). These sequences were excluded from further analysis as these could be caused by technical issues, such as mis-reads of the MiSeq platform. The published error rate for MiSeq is 0.8 %, which corresponds to one error in a 250 bp read (Quail et al. 2012). There was a decreasing number of clusters with increasing numbers of sequences (Figure 3., in red), however, there were increasing numbers of sequences represented in the clusters with more than 20 identical sequence reads. These 149 clusters represent 41,270 sequences or 70.3 % of the total sequence reads (Figure 3.3, in black).

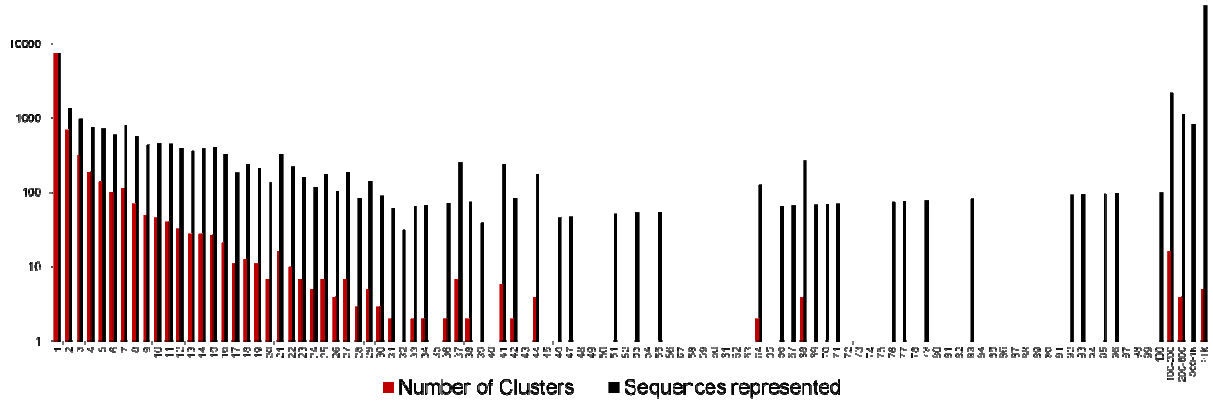


Figure 3.3: Numbers of *Ascaris* CO1 MiSeq sequences in clusters.

Sequences obtained were arranged in clusters with identical sequences. The number of clusters is indicated in red and the cumulative number of reads in these clusters is indicated in black. Both panels are shown with a logarithmic scale on the vertical axis.

To examine the potential for the parallel sequencing approach and to estimate the impact of increased numbers of appropriate sequences compared to the non-*Ascaris* sequences generated, a neighbour-joining tree was constructed using the sequences in clusters with more than 20 sequences (Figure 3.4). Along with the sequences from the parallel sequencing (in blue) the tree shows all sequences from the Sanger sequencing approach (including the non-*Ascaris* sequences, in red) as well as *Ascaris* CO1 sequences from the database (in black) with a nematode outgroup to root the tree (*Anisakis* spp. CO1, in grey). Although the neighbour-joining tree represented in Figure 3.4 contains more than 200 sequences, the original sample identity of the parallel sequencing data could not be determined due to the lack of barcodes. However, the colouration in Figure 3.4 indicates that the group of non-*Ascaris* sequences generated in the Sanger sequencing (red) which appeared in a separate clade outside the target sequence clade was removed by the restrictive filtering of the sequences obtained by parallel sequencing (blue). The filtering checks for the identity of both primers and their position, which would remove the non-

Ascaris sequences generated by Sanger sequencing. The sequences generated by parallel sequencing were subdivided in distinct clades. The results indicated that the approach is suitable to generate a comprehensive sequence collection for further evaluation of the hypothesis.

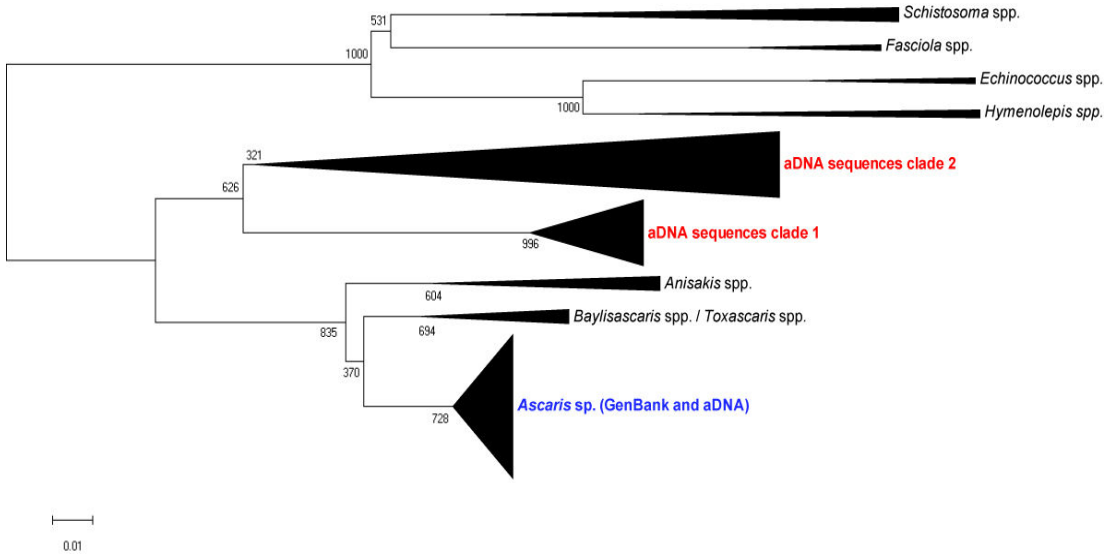


Figure 3.4: Neighbour-joining tree of parallel sequencing cluster sequences against Sanger sequencing and database sequences.

Sequences generated by parallel sequencing all clustered with *Ascaris* sp. sequences obtained from the GenBank database. Some of the sequences generated by Sanger sequencing were found in the same cluster (indicated in blue). A large proportion of aDNA sequences generated by Sanger sequencing did not group with any related species, but formed two clusters (indicated in red). The sequences were compared to related nematodes (*Baylisascaris* spp., *Toxascaris* spp. and *Anisakis* spp.) as well as parasitic cestodes (*Echinococcus* spp. and *Hymenolepis* spp.) and trematodes (*Schistosoma* spp. and *Fasciola* spp.).

3.4.2 Discussion Ascaris CO1

Despite not producing any sequences which could be allocated to a specific sample, the first attempt at parallel sequencing of *Ascaris* CO1 produced a considerable amount of data. The data was analysed in respect to the quality and yield as well as to test the technical tools required for the subsequent analysis. Over

10 % of the sequences generated were found only once, which suggests that these were a result of read error which would match the published error rate for MiSeq at 0.8 % (Quail et al. 2012). Removal of these sequences was technically simple, but these findings did not provide any information on the lower limits of repeated sequences which were not created by technical issues. A conservative approach would be set the cut-off at eight (2^3) repeats as three independent PCR reactions were involved in the library preparation. However, this would only be statistically valid when the capacity of the sequencing technique would be exhausted.

This initial experiment was affected by the failure of the barcoding-PCR which made it impossible to generate useable data for phylogenetic analysis and parasite epidemiology. Analysis of the data still proved valuable to gain experience of handling such large datasets and on the quality estimation and restrictions of the approach. The barcoding was repeated and the samples were re-submitted. By the time of submission of this thesis the results of this sequencing round were not available yet.

3.5 Samples sequenced - Trichuris ITS1

The second batch of samples submitted for MiSeq analysis consisted of 68 *T. trichiura* ITS1 samples produced by second round PCR using barcoded primers. Four samples from Pohansko were based on extracted plasmid DNA from a mixed culture (transformed but not plated bacteria). Samples included in this sequencing library were taken from different sample collections, including Neolithic Zurich (ZPO, 4), Anglo-Saxon York (Hg, 1), early Slavic Pohansko (CZBP, 10), Viking Gokstad (VNOGNS, 17), Viking Copenhagen (VDKKH, 13), Medieval Bristol (BFR, 12) and Medieval Lübeck (HLGV, 11).

3.5.1 Results Trichuris ITS1

A larger number of sequences compared to the *Ascaris* CO1 were generated for *T. trichiura* ITS1. In total 107,047 paired read sequences from 60 of 68 barcodes were retrieved. Clusters were built for each barcode to group all identical sequence reads. The barcoded sequences were sorted into 39,772 clusters, of which 35,741 (89.9 %) consisted of only one single sequence read, which accounted for 33.4 % of retrieved sequences.

The sample with the barcode 082 (Neolithic Zurich, ZPO 8431) contained the largest cluster recorded (7670 sequence reads), the most clusters (11883) and the largest number of sequence reads (29,678). In this sample 1066 clusters (8.97 %) were contained more than one sequence read, adding up to 18,861 (63.6 %) sequences.

Although 60 of 68 barcodes could be retrieved, more than half of those did not contain any repeat sequences. These low-frequency sequences were considered not to be genuine due to the expected error rate which caused the sorting from the original data file to be designed to be very restrictive. However, these samples will have to be considered with care. Of the 27 samples containing larger clusters, 23 contained at least one cluster with at least 10 repeated sequences, and 19 contained at least one cluster with more than 20 repeats. These samples represented 62.9 % of the extracted sequences (67,275) in 4031 clusters (10.2 %, Figure 3.5) which was slightly lower percentage than for *Ascaris* (70.3 %), but the yield of sequences was higher for *Trichuris* ITS1 (58,687 sequences for *Ascaris*).

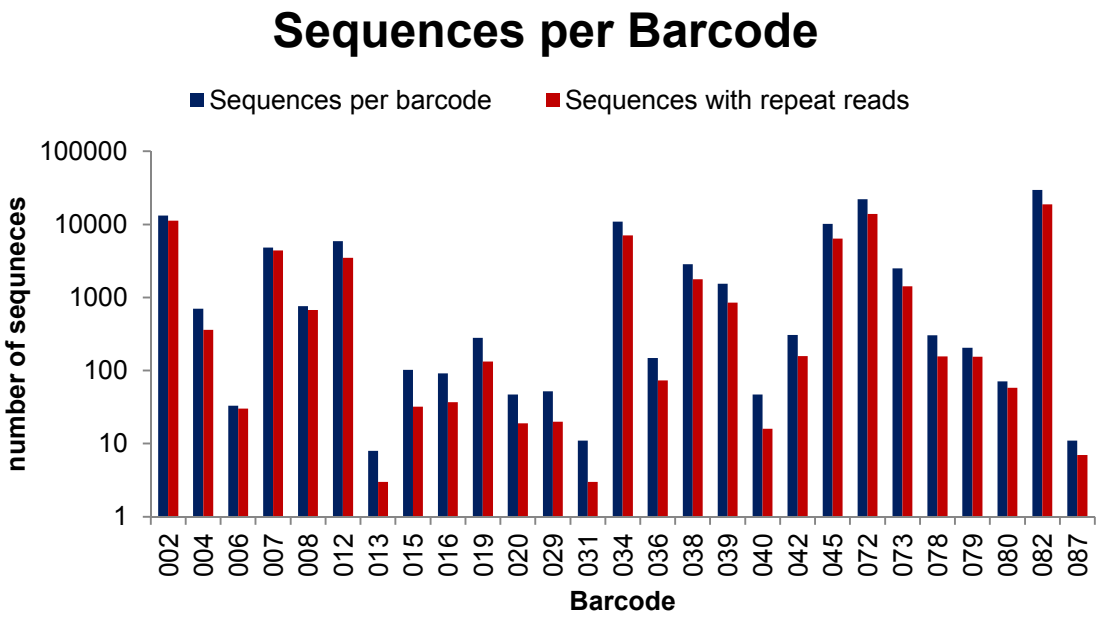


Figure 3.5: Number of sequences per barcode.

The bars in blue indicate the total number of sequences retrieved; the red bars indicate the number of sequences with repeat reads. Samples with barcodes which contained only single reads were not shown.

The main objective of the parallel sequencing approach was the generation of more sequences per sample compared with Sanger sequencing and sequences of low-frequency genotypes. For all samples where Sanger sequencing was available (16 of 27) the sequences generated by Sanger sequencing were also generated by parallel sequencing, generally with large cluster sizes. The parallel sequencing approach added a large number of sequences to the analysis. The parallel sequencing run extended the *T. trichiura* ITS1 aDNA database considerably from 174 Sanger sequences to 744 sequences (540 sequence reads from parallel sequencing).

Six samples in the library were taken from a graveyard site (CZBP, Pohansko, discussed in detail in chapter 7.2). With samples from grave content it was intriguing to ask whether these would contain less diversity than sequences derived from a group of people (e.g. latrine samples). The samples had slightly higher numbers of single read sequences (average 50.3 % against 42.1 % overall). The

diversity was estimated by the distribution of sequence reads in the six largest clusters. Under the assumption that a single individual would have fewer different parasites, similar distributions would be expected. The comparison between grave samples and communal deposits revealed that the distributions of grave samples (red) were more consistent than those from communal deposits (blue, Figure 3.6). Most notably, the large spread of distributions in communal deposits, as indicated by the standard deviation, demonstrated a larger variety in communal deposits.

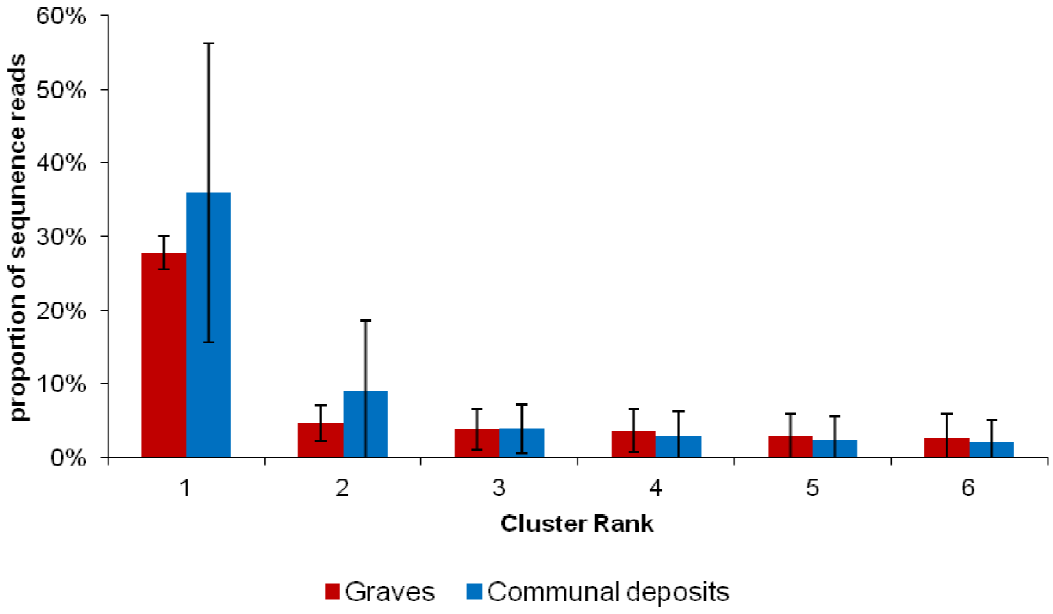


Figure 3.6: Evaluation of graveyard versus communal deposit diversity.

The diversity was estimated by the proportion of sequence reads per cluster for the six clusters with the highest numbers of sequence reads. More variation was observed in the communal deposits (error bar: standard deviation)

3.5.2 Considerations on the reproducibility of the system

To evaluate the reproducibility of the parallel sequencing approach, duplicate samples based on the same soil sample were included in the library. Two individual aDNA extractions of the sample BFR 3194-4472 (Bristol) were used in two individual PCRs (performed on different days). The two samples had barcode

numbers 034 and 042. The sample with the barcode 042 yielded significantly more sequences (10,957 vs. 308). Both samples had a comparable proportion of sequences which were only found once (90.0 % and 89.3 %, respectively) and more than 50 % of the sequences in repeat sequence read clusters (64.8 % and 50.8 %, respectively). The sample with the barcode 034 produced 4285 clusters, of which 429 were repeat sequence read clusters and sample 042 produced 169 clusters of which 17 were repeat sequence read clusters.

When comparing the sequences in the repeat sequence read clusters it was expected that the same sequences would be found in the most frequent clusters with similar relative proportions of sequences in the cluster. For this comparison, the clusters were ranked according to the number of sequences contained. For practical reasons, for the sample with barcode 034 only clusters with more than 10 sequences got a unique identifier (48 clusters). The comparison table (Table 3.1) lists the clusters with the identical sequences from both samples. The comparison showed that eight of the top 10 clusters of sample 034 were also found in sample 042, with the top three clusters being the clusters with the same sequences. Moreover, the relative proportions of the clusters with identical sequences were very similar with about one third of the sequences being in the top cluster (34.94 % in 034 and 32.14 % in 042, respectively), about 4 % in cluster 2 (4.59 % and 3.25 %, respectively) and about 2.5 % in cluster 3 (2.29 % and 2.92 %, respectively).

Barcode 042			Sequence ID	*Barcode 034*		
Frequency Rank	Sequences in Cluster	Relative proportion		Frequency Rank	Sequences in Cluster	Relative proportion
1	99	32.14 %	**a**	1	3828	34.94 %
2	10	3.25 %	**b**	2	503	4.59 %
3	9	2.92 %	**c**	3	251	2.29 %
4	6	1.95 %	**d**	10	49	0.45 %
5	4	1.30 %	**e**	--	--	--
6	3	0.97 %	**f**	6	85	0.78 %
6	3	0.97 %	**g**	16	34	0.31 %
7	2	0.65 %	**h**	4	116	1.06 %
7	2	0.65 %	**i**	5	112	1.02 %
7	2	0.65 %	**j**	7	63	0.57 %
7	2	0.65 %	**k**	11	49	0.45 %
7	2	0.65 %	**l**	15	35	0.32 %
7	2	0.65 %	**m**	23	24	0.22 %
7	2	0.65 %	**n**	27	20	0.18 %
7	2	0.65 %	**o**	16	16	0.15%
7	2	0.65 %	**p**	51	9	0.08 %
7	2	0.65 %	**q**	56	4	0.04 %

Table 3.1: Comparison of ranked clusters of two individual samples from the same original soil sample.

The comparison shows that the top ranked sequences are generated from both samples, despite the significantly lower number of sequences generated in the sample with the barcode 042. The sequence ID is an arbitrary identifier for the sequence found in both barcode sets.

3.6 Discussion

After an initial problem with the barcoding of the PCR products, parallel sequencing has proven a valuable technique to generate large quantities of sequence data. Overall sequence alignment and phylogenetic analysis for *T. trichiura* ITS1 from all samples will be discussed in chapter 8.2.

Results from the first attempt on parallel sequencing were rather disappointing as a technical problem led to the absence of barcodes on the samples, making it impossible to arrange the sequences by sample. Although these sequences could not be integrated by sample the sequence analysis they still provided valuable information on the potential of the technique for the archaeoparasitological research. The numbers of sequences generated indicated how much data could be expected from a parallel sequencing run.

The second attempt was more successful as the PCR produced barcoded products and generated sequence data from 60 samples, of which 27 contained sequence clusters with repeated sequence reads. The issue with the first MiSeq run was likely the use of PCR products which were ligated into a vector. In case of an inefficient PCR, traces of non-ligated and non-barcoded PCR products were sequenced which led to the absence of the barcodes. As observed before, the age of the sample was not an observable issue. The sample producing the largest number of sequences was one of the oldest samples included (Neolithic Zurich). Clustering of the sequences was an important step as it helped to remove sequences which were found without replicates.

The technique was reproducible as shown by the use of two different extracts from the same original sample. For both barcodes, the most frequent sequences were identical and had similar prevalence rates despite considerably different yields of the

total sequence generation. This was a very promising and important result as it confirmed that the samples were treated similarly and that the sub-sampling of the original soil sample was not biasing the sequence output. This analysis could not have been done by Sanger sequencing as this would likely only have picked up the most prevalent sequence which was found to have a nearly ten times higher prevalence than the second most frequent sequence. A large scale cloning and sequencing experiment would be necessary to generate sufficient sequence data which would be infeasible due to high costs and labour-intensive processing.

Overall, this technique represents a valuable addition to the project as it was able to generate large amounts of sequence data which enables the study to address issues of epidemiology and parasite population dynamics.

Chapter 4

The Medieval Trading Port of Bristol

This chapter presents the findings from the first collection of samples obtained from the excavation of Finzel's Reach in Bristol (UK). This medieval port area was an ideal starting point for the establishment of the molecular archaeoparasitology approach. The location just outside of the main city walls and close to the harbour led to the assumption that this area was likely to be populated by a very mixed population of local townspeople and seamen. The Bristol sample set consisted of 26 samples from a waste deposit ditch dating from the late 12th century CE to the early 17th century CE (see maps in Figure 4.1).

4.1 History of the site

Bristol was an important English trade port in the middle ages, trading mainly to the west (Sherborne 1965). The city was founded around 1000 CE, at a location close to the coast and between the rivers Frome and Avon. By about 1020 CE the city had its own mint and at the end of the 11th century CE the city was well fortified to defend itself against invasions; fortifications which were extended to become one of the biggest strongholds and one of the major shipbuilding sites in England (Sherborne 1965). Bristol advanced to one of the biggest trade cities in England, trading extensively with Ireland, Norway, France, Portugal and Spain, as well as distributing products across England. In the 12th century, the most common trade goods were hides, wool, corn and cloth, but also some iron and timber (Carus-Wilson 1967a). Cloth making industry was mentioned in Bristol by 1230 CE, but it's share of the international trade was limited, most of Bristol's cloth production was sold on the home market (Sherborne 1965). Wine became the main trade good for Bristol from the 13th century onwards, mostly imported from France (Carus-Wilson 1967b). Besides wine, woad for the cloth industry was a commodity imported on large scale.

The Hundred Years' War with France led to increased export taxes on wool to benefit the home economy (Sherborne 1965). The effect of the plague (1348-49) resulted in a lack of workers (Boucher 1938) which led to a decline of Bristol's importance as an international trading port while Southampton and London gained more influence.

4.2 The excavation - Bristol Finzel's Reach

Twenty-six samples were obtained from Oxford Archaeology (Rebecca Nicholson). The excavation of the site of Finzel's Reach in Bristol was carried out between April 2007 and April 2008 by Oxford Archaeology. Evidence of industry such as smithing and tanning was found as early as the 13th century, which was replaced by textile production over time. The waterlogged conditions on the site contributed to a good level of preservation of organic artefacts, including three rare wine tuns, as well as a very large and well preserved assemblage of fish bone and samples of waterlogged plant and insect remains. The samples for parasitological analysis were taken from a ditch with evidence of human excrement (extracts from the excavation report and personal communication). The maps in Figure 4.1 depict the location with reference to the modern city and the 16th century CE city map (printed in 1568). While the location of Finzel's Reach in the modern city is fairly central, the earlier map shows Finzel's Reach out of the centre of the city. Initially, the central interest was on the identification of parasites to strengthen the claim that the material was associated with human excrements.

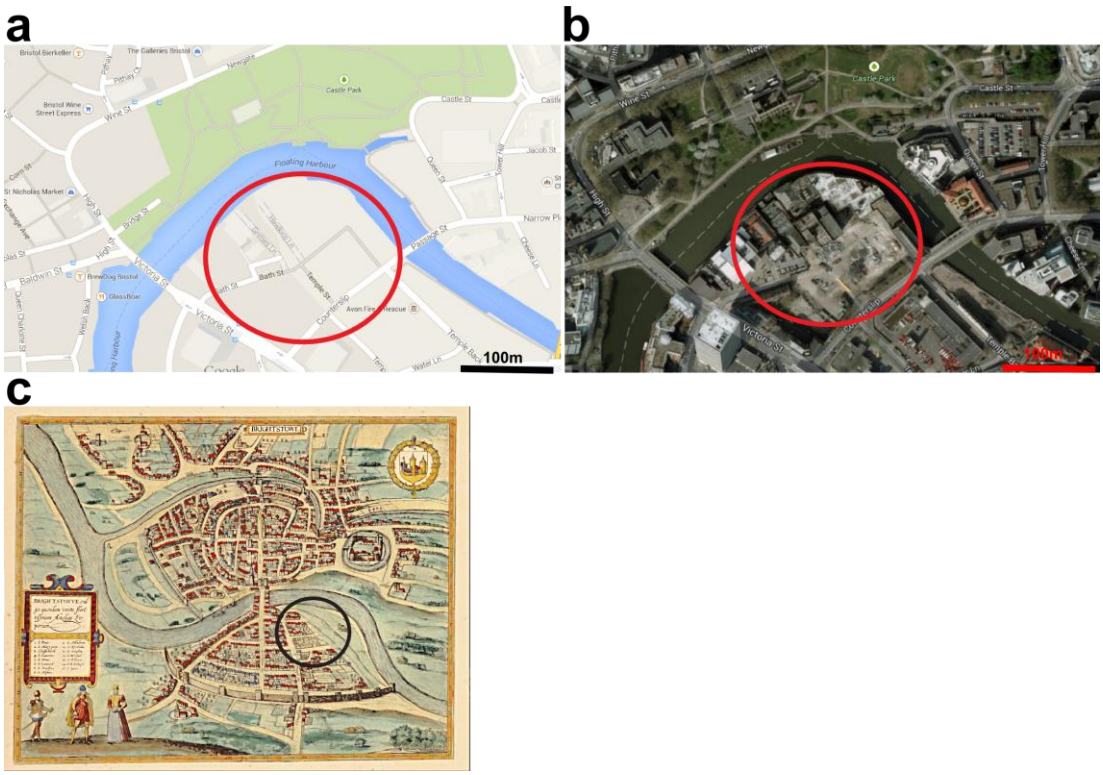


Figure 4.1: Modern map and satellite photograph and historic map of Bristol, indicating the site of the excavation.

Circled on the modern street map (a) and the corresponding satellite photograph during redevelopment construction works (b) is the site of the excavation of Finzel's Reach. The map of 1568 (c) shows the location with respect to the Medieval city walls, locating it out of the city, close to river and port. Map/satellite photograph: maps.google.co.uk; Historic map: maps.bristol.gov.uk

4.3 Samples

Dating of the 26 samples was based upon context finds and stratigraphy. The oldest samples date to the mid-12th century CE and the most recent samples to the 16th to 17th century CE. Most samples were dated in the range of the late 13th to the early 14th century CE. All samples were relatively dry samples, with a predominant colour of beige to dark brown, with the exception of one sample dated to late 15th/early 16th century CE which contained a high proportion of charred organic material. This particular sample was extracted from the remains of a charred barrel (personal communication Rebecca Nicholson, Oxford Archaeology).

4.3.1 Microscopic diagnosis

All samples were analysed microscopically, and most were positive for either one or both of the nematode parasites *Trichuris* and *Ascaris*. Only one sample contained no parasite eggs (no. 3322-7002, mid-13th/14th century CE). To assess the preservation state of the samples, *Trichuris* eggs were separated into those that were intact or visibly damaged (e.g. missing end plugs, see Figure 2.3). This assessment provided information on the sample specific preservation as *Trichuris* eggs were found in 94 % of samples (32 of 34). Overall, 55 % of the counted eggs were intact with some sample to sample variation (Figure 4.2). Three samples only contained damaged eggs, and four samples only contained intact eggs (Figure 4.2).

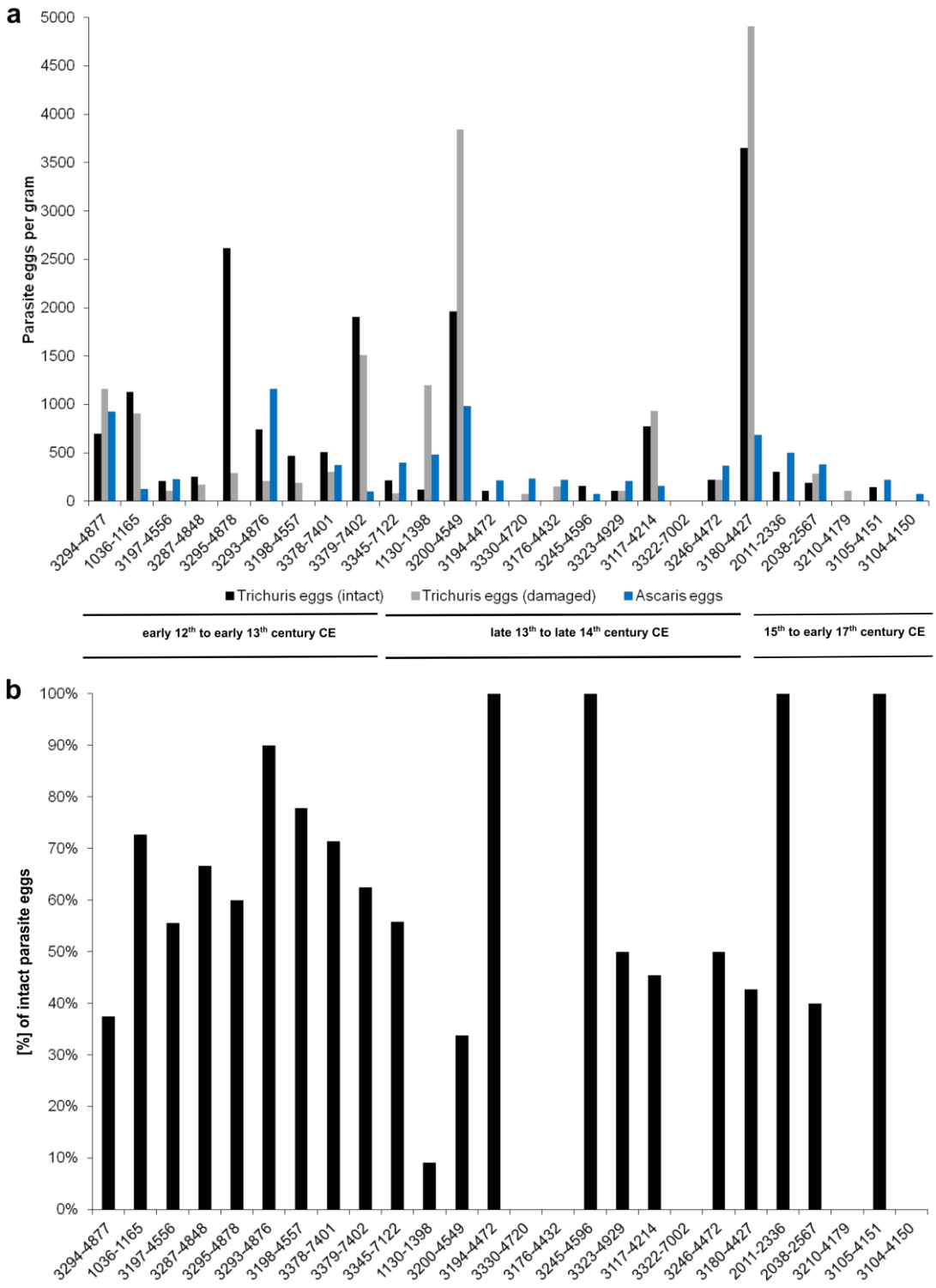


Figure 4.2: Number of counted parasite eggs (a) and percentage of undamaged *Trichuris* eggs (b) from the excavation Finzel's Reach in Bristol.

Varying numbers of *Ascaris* (blue) intact (black) and damaged (grey) *Trichuris* eggs were found in the sample (a). Almost all samples contained parasite eggs. The proportion of intact *Trichuris* eggs varied considerably which provided information on the general preservation state (b). Both panels are arranged in the same order of ascending age.

Statistical evaluation of the distribution of the prevalence of both nematode parasite eggs revealed a linear correlation with a p-value just below 0.05 ($p = 0.044$, $R^2 = 0.398$) between the presence of *Ascaris* and *Trichuris* (Figure 4.3).

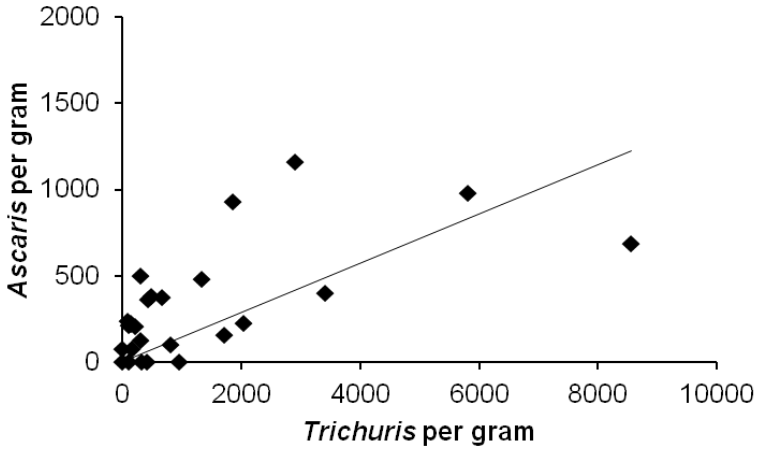


Figure 4.3: Correlation between the numbers of *Trichuris* and *Ascaris* in samples from Bristol.

4.3.2 Molecular analysis

The sample set from Bristol was used to evaluate some aspects of the molecular methods in the project. Extraction of aDNA and PCR was successful for all samples, although some required adaptations to the protocol, particularly those with high charcoal content (as described in chapter 2.3.4). Some the PCR products generated with some PCR primers were non-nematode derived amplicons and most contained the target nematode sequences.

4.3.2.1 Evaluation of sequences

A comprehensive set of sequences for *T. trichiura* ITS1 was generated. Sequences were obtained from 12 samples (46.2 % of those tested) dating between 1200 and 1600 CE, samples dated after 1300 CE were underrepresented (only 3

samples available). Overall, 7,539 sequences were generated by either Sanger sequencing (72 sequences from 10 samples) or MiSeq parallel sequencing (7,467 sequences from 6 samples). Evaluation of the sequences included the identification of the species and an evaluation of the relationship between the sequences from different samples and time periods. Initially the sequence identity was evaluated in a neighbour-joining tree with ITS1 sequences from *Trichuris* spp. (including *T. trichiura*) and a range of other nematode parasites. Figure 4.4 indicates that the sequences generated were *T. trichiura* and not any other *Trichuris* spp. (the species on the tree were *T. muris*, *T. arvicolae*, *T. suis*, *T. discolour*, *T. ovis* and *T. lepronis*) or any other nematode (here: *Anisakis* spp. as outgroup). The identity could be confirmed by the position of the sequences within the clade of the modern *T. trichiura*. Sequences of *T. trichiura* taken from the database are highlighted in red, ancient sequences in blue. Two sets of modern *T. trichiura* sequences obtained from the GenBank database were found to be closer to *T. suis*, yet not inside the *T. suis* clade. These two sequences were derived from wild pigs and based on their position on the tree were most likely mis-annotated *T. suis*. The diversification of the branch is quite clear identifying the generated sequences (in blue, bootstrap value 1000) as being *T. trichiura* ITS1 although in Figure 4.4 the fine-structure topology within the *T. trichiura* ITS1 clade is not clearly visible.

To visualise the sub-structure within *T. trichiura* ITS1 a tree with less subgroups (restricted to *T. muris* and *T. suis*) was generated. The structure of this tree suggests some subdivision into clades within *T. trichiura*. However, most of the observed bootstrap values were small for the majority of branches (<300 at 1000 iterations) on the relationship between samples could not be considered conclusive. The only subsection of the tree supported by high bootstrap values (>900) was that

the samples fall into two clades, of which one (clade 2) consisted of the vast majority of sequences (including all modern sequences) and the other contained of three sequences obtained from two samples (1036-1165 and 3287-4848), both dated to 1300 CE (clade 1). Figure 4.5 depicts the sub-tree with the ancient sequences separated by two reliable bootstrap values. The large branch (clade 2,collapsed) contained >100 sequences *T. trichiura* ITS1 from 12 samples and all included modern sequences for *T. trichiura* ITS1. Both samples with sequences that lie within clade 1 also produced sets that lie within clade 2, sample no. 1036-1165 was represented by 3 (of 5) sequences and sample no. 3287-4848 with 51 (of 52) sequences in clade 2 (Table 4.1).

Table 4.1 summarises the number of sequences generated from the Bristol samples by Sanger and Parallel Sequencing (MiSeq). In 10 of 12 samples sequences were generated using the standard cloning and sequencing (Sanger) approach. By this approach, 72 sequences could be generated. Additionally, 6 samples provided further 7,467 sequences from a MiSeq parallel sequencing approach. For 4 samples (3287-4848, 3246-4472, 3198-4557 and 3117-4214), sequences were generated by both approaches. In samples 3198-4557 and 3117-4214 only one Sanger sequence was generated, which was identical to the most frequent sequence generated by parallel sequencing. All sequences generated using Sanger sequencing from sample 3246-4472 were found in the parallel sequencing data, however, although the most frequent Sanger sequence was identical with the most frequent sequence from parallel sequencing, the other two sequence were found at very low frequencies in parallel sequencing (3 and 1 repeat). Of sample 3287-4848 only 1 of 4 sequences was found in the parallel sequencing data (most frequent sequence in MiSeq data), the other 3 sequences were not found.

Sample number	Date	No of unique Sanger sequences	No of unique MiSeq sequences	Sanger Sequences in MiSeq
3345-7122	1200	0	48 (1)	--
1130-1398	1225	9 (4)	0	--
3176-4432	1300	0	13 (1)	--
3180-4427	1300	11 (5)	0	--
3287-4848	1300	9 (4)	5340 (49)	1/4
3246-4472	1300	5 (3)	109 (2)	3/3
3322-7002	1300	3 (2)	0	--
1136-1165	1300	7 (5)	0	--
3198-4557	1300	1 (1)	652 (7)	1/1
3117-4214	1300	1 (1)	1305 (14)	1/1
2038-2567	1500	3 (1)	0	--
3104-4150	1600	23 (12)	0	--

Table 4.1: Number of sequences generated from the Bristol samples by Sanger and Parallel Sequencing (MiSeq)

Sequences of 12 samples were generated, either by Sanger or Parallel Sequencing. The total number and the number of unique sequences (in brackets) are indicated. Most sequences generated using Sanger sequencing were found in the parallel sequencing approach.

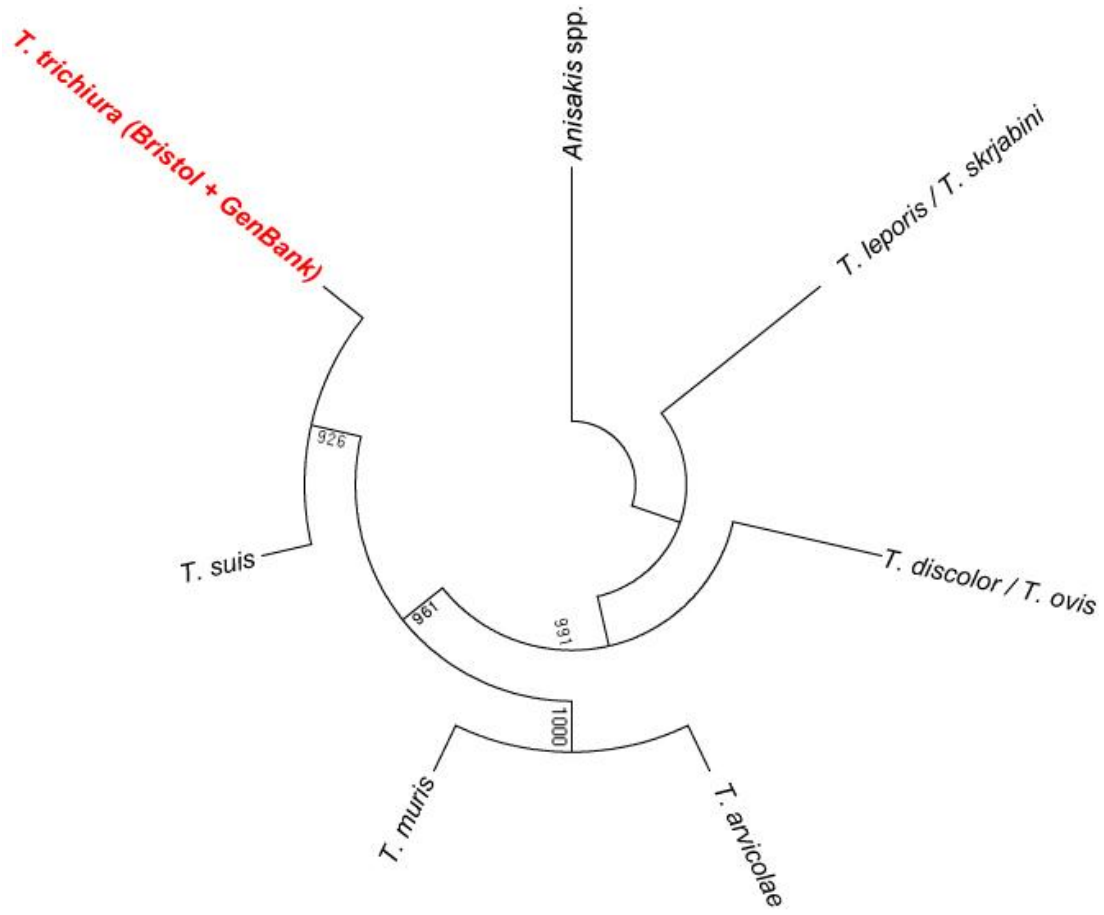


Figure 4.4: Bootstrapped Neighbour-Joining tree for T. trichiura ITS1 sequences from Bristol.

The outgroup on this tree is *Anisakis* spp. and further includes various *Trichuris* spp. (*T. suis*, *T. muris*, *T. arvicolae*, *T. ovis*, *T. discolour*, *T. leporis* and *T. skrjabini*). The obtained ancient sequences show some degree of variation, yet all of them were identified as *T. trichiura*.

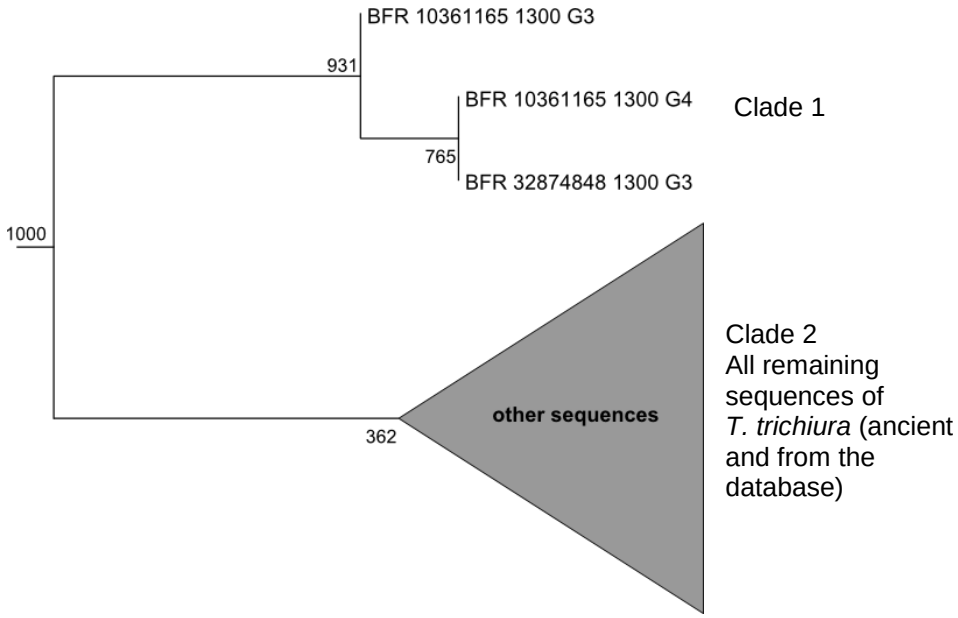


Figure 4.5: Extract from Neighbour-Joining bootstrapped tree for Bristol *T. trichiura* ITS1 sequences samples.

As the bootstrap values were found to be very low for most clades, only the two with the highest values are represented. Most ancient samples and all modern *T. trichiura* ITS1 sequences were represented in one clade, but both clades are clearly separated from other *Trichuris* spp. (full tree in Figure 4.4)

4.3.2.2 Diversity evaluation

The 164 unique *T. trichiura* ITS1 sequences generated were on average 420 bp long, of which 366 bp (87.1 %) were conserved and 15 bp (3.6 %) were different in only one sequence. The 39 bp remaining were evaluated for patterns according to time or sample. The three samples which were found in clade 1 (Figure 4.5) were the most different sequences which showed unique polymorphisms in 6/39 bp. The most frequent transition (T-G, 19.5 % frequency, pos 246) was only found in samples before the 14th century CE. The most frequent transition in the most recent samples (16th century CE) was a C-T transition (position 154) which was found in 6 sequences, which was intriguing as only 12 sequences of this time point were generated.

4.4 Discussion of the Bristol sample set

This sample set obtained from Oxford Archaeology was the first large set in this project. Most optimisations of aDNA extraction and the initial PCRs were performed on this set (as described in chapter 2). Microscopic analysis confirmed earlier publications stating that nematode parasite eggs show outstanding preservation in soil (Loreille et al. 2001; Mitchell et al. 2013). The structure of the *Trichuris* eggs allows assessment of the preservation state since the damaged eggs are distinct from intact ones. Overall, 96 % of the samples contained parasite eggs which suggested a high prevalence of nematode parasites in Medieval England.

The number of intact *Trichuris* eggs did not correlate with the potential to generate readable sequences. ITS1 sequences were generated from 13 samples (38.2 %). For two samples, there were no *Trichuris* eggs (samples 3104-4150 and 3322-7002) or only damaged *Trichuris* eggs detected by microscopy (3176-4432, 148.7 eggs per gram). This reveals the lower limits of microscopic diagnostics. The lowest extrapolated per gram concentrations of eggs found were 106.4 (intact *Trichuris* eggs per gram), 82.0 (damaged *Trichuris* eggs per gram), and 76.2 (*Ascaris* eggs per gram), respectively. Microscopic diagnosis and PCR amplification had different advantages and drawbacks. Whereas PCR amplification was more sensitive in detecting low amounts of parasite eggs, there were issues with contaminations which make the PCR-based approach less reliable as negative results were necessarily correlated with the absence of the parasite in the microscopic diagnosis.

Positive results for PCR and sequencing for at least one sample were generated for *Trichuris* (ITS1 and β-tubulin), *Ascaris* (COX1 and β-tubulin) and *Cryptosporidium* (gp60). For *Giardia* (*β*-giardin) and other targets of *Trichuris*, *Ascaris* and *Cryptosporidium* PCR products were generated.

Molecular work including PCR and sequence generation showed that the techniques developed were appropriate to be applied on archaeological material. The sequences generated by both Sanger sequencing and MiSeq parallel sequencing confirmed that the parasites found were human parasites and that the PCR-based approach was target-specific. The sequences generated using PCR primers for *Ascaris* sp. on the samples from Bristol generated only very few genuine sequences which provided little information. However, they showed that the approach was appropriate to generated sequence data from ancient material. Data generated based on *Trichuris* was very promising; especially the development and application of parallel sequencing increased the depth of the sequence data. The comparison of sequencing methods revealed that in most cases only the most frequent sequences were found by the relatively low depth covered by Sanger sequencing. However, in one sample (3246-4472) the sequences found in Sanger sequencing were not only the most frequent ones in parallel sequencing, but some were found to be at a very low frequency in parallel sequencing.

The sequences of *T. trichiura* ITS1 examined some diversity which is valuable on several levels. The analysis of base pair transitions suggested that some signatures might be related to time or location. The most important finding there was the identification of the polymorphisms which generated the divide of clades 1 and 2. It showed that during the processing sufficient precautions were taken to avoid cross-contaminations between samples, further, ITS1 sequence may exhibit a sufficient level of sequence variation to make it useful to compare sequences from different origins. In the case of Bristol, an important medieval port, comparisons especially with potential trade partners could prove to be very valuable.

Chapter 5

Lübeck

Centre of the Hanseatic Trade

5.1 Introduction – historical background

Lübeck was founded in 1143 and made an Imperial Free City in 1226 which placed it under the direct order of the German emperor giving it freedom to flourish. The wealthy merchant city was the founding place and capital of the Hansatic League which was the most powerful trade union of the middle ages. At its peak the Hanseatic League included about 200 cities (76 core cities and about 130 associated cities) as its members (see map Figure 5.1). It is not surprising that Lübeck has a rich heritage and extensive archives of trade, business, and the people who lived in the town of Lübeck.

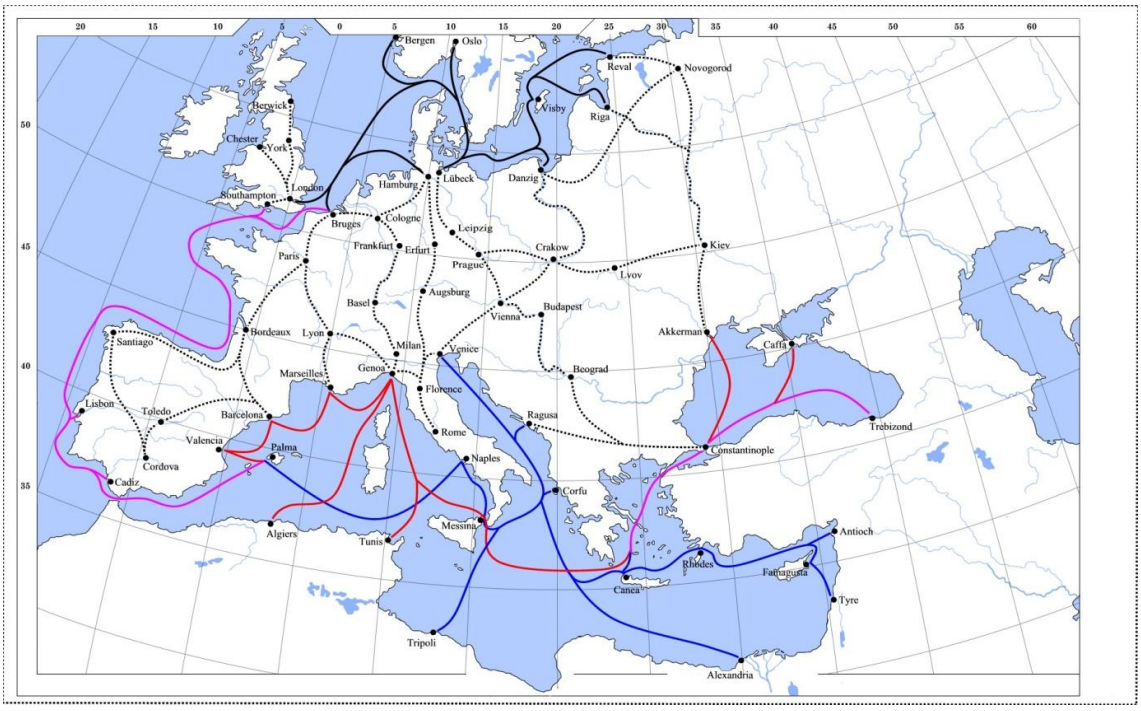


Figure 5.1: Medieval Trade Routes

The black lines show the routes of the Hanseatic League, the blue Venetian and the red Genoese routes. Purple lines are routes used by both the Venetians and the Genoese. Overland and river routes are stippled. (eResources Department of History (University of Warwick 2013), retrieved 10 Oct 2013)

The influence of the Hanseatic League on the trade in Europe cannot be overstated. Trade in the region of the Baltic Sea was nearly exclusively performed by members of the Hanseatic League and trade with Britain was considerable (Figure 5.1). The samples from Lübeck span a relatively long time period in which Lübeck and the Hanseatic League rose to power (until the early 14th century), engaged in warfare (e.g. against Denmark 1360-1370, or the Netherlands 1438-1441), and then slowly lost power and influence (mostly due to increased competition). The last "Hansetag" (congregation of the members) was held in 1669 in Lübeck by which time the powerful trade union had lost most of its members, and only nine cities were left. At the height of the influence of the Hanseatic League on European trade merchandise was traded far beyond the Baltic, and the Kontor in London (Steelyard) played an important role for the trade to the British Isles, especially after being granted tax and customs concessions by Edward I in 1303 (Zimmerling 1976). Tax records and harbour lists provide rich evidence for trade with the British Isles, which was mostly to the western coast, especially London (Kontor at Steelyard), but also with Hull, King's Lynn and Boston, Ipswich and Yarmouth. There are even rare descriptions of Hansatic traders sailing to other cities in England such as Newcastle, York, Norwich, Colchester, Sandwich, Southampton or Bristol (Lloyd 1991; Schulz 1911).

Most of Lübeck's archives have remained intact to the present day and extensive written sources provide valuable background information to life in the medieval city and the trade to and from Lübeck. The tight organisation of all aspects of trade produced unique sources of which considerable amounts have been preserved, including the letters and trade books of the 14th century trader Hildebrand Veckinchusen (Lesnikov and Veckinchusen 1973) and the trade books of merchant

cooperatives such as the ‘Bergenfahrer’ which traded with a specific region, in this case Bergen (Bruns and Brandt 1953). The records on the trade give valuable insight into the changing trading patterns in the Hanseatic League.

The documentation of the houses from which samples were obtained includes information that relates to the inhabitants, their social status, political influence and economic activities, including trade destinations and it is possible to correlate the historic record (city books, ‘Niederstadtbuch’) with archaeological finds (Lutterbeck 2002). This study focuses on one of the oldest parts of Lübeck, initially inhabited by a wealthy group of merchants of which many were members of the council. In the mid-14th century the location became less attractive and many of the houses were sold to less wealthy tradesmen and craftsmen. This could be traced by the deeds of ownership of the houses and the ‘Ratslinie’ (records of members of the city council, (Lutterbeck 2002)). The records of the council members were an important source of information on the social status as the members of the council were elected by the other members of the council and were required to own property in Lübeck, have an excellent reputation and were not allowed to be craftsmen (Lutterbeck 2002). Of the houses from which samples were obtained there are records of them being sold and the former inhabitants moving to a different part of the city. For example the house in Alfstrasse 25 was sold in 1314 by Hinrik van Kampen (member of the council) and the house in Fischstrasse 26 was sold in 1319 by Everhard van Alen, (council member, mayor in 1340) who moved then into the house at Schlüsselbuden 12, in front of the St Mary’s Church (Lutterbeck 2002). The availability of such detailed information was unique and provided valuable additional information for the samples obtained.

5.2 Description of the site

The collection of communal deposit samples originated from one of the oldest parts (Gründungsviertel, German for "Foundation Quarter") of the Hanseatic city of Lübeck. The excavation covered a part of the old city which was destroyed during a Royal Air Force air raid on the night of 28 March 1942. The area was built over in the 1950's without excavation of the underlying structures, which protected them from external disturbance. The excavation is part of the UNESCO World Heritage Site conservational works in Lübeck and extended over 9000 m^2 on several housing plots (Figure 5.2). The current excavation is part of a long-term project which started in the mid-1980s. The samples included in this project are received from the latest excavation between the Alfstrasse and the Fischstrasse, west of the St Mary's Church (Marienkirche). The samples were collected from latrines of individual houses, dating from the early years of Lübeck (12^{th} century CE), through the peak period of Lübeck's influence to the era of dwindling influence of both Lübeck and the Hanseatic League in the early 17^{th} century CE. Extraordinary well preserved foundations of the 12^{th} century city houses allow a clear correlation with contemporary maps (Figure 5.2). The eight plots from which samples were obtained are indicated in Figure 5.2 (Alfstrasse 17, 21, 23, 25, and 27; Fischstrasse 16, 26, and 28).

Figure 5.2: Map of the current excavation site in Lübeck.

The current excavation covers the area of a former school building (Hanseschule, gray overlay). It included 18 plots between Alfstrasse and the Fischstrasse, with plot numbers representing the house numbers of the former town houses. Samples were obtained from cesspits in eight plots (Alfstrasse 17, 21, 23, 25, and 27; Fischstrasse 16, 26, and 28). Image courtesy of Dirk Rieger, Lübeck

Samples were recovered from latrines situated in the back of houses and represent faecal deposits from the dwellers of these houses. The associated finds provided further evidence of the living conditions in these parts of medieval Lübeck and allowed accurate estimation of the usage period for each latrine.

As part of earlier excavations of the site, some parasitological analysis of latrine content was conducted prior to this study and were published in the archaeological reports of the excavation (Herrmann and Schulz 1986; Paap 1984). Both studies applied microscopic techniques (light microscopy and scanning electron microscopy) to describe the presence of *Trichuris*, *Ascaris*, *Fasciola* (liver fluke),

and *Diphyllobothrium*. Although the publication mentioning *Diphyllobothrium* and *Fasciola* did not mention any dating, the authors noticed that the distribution of those two parasites was different from the distribution of *Trichuris* and *Ascaris*. The authors of the 1986 study evaluated samples from a single latrine which was sectioned in roughly 10 cm depth sections (32 samples), in which they noted that the top seven samples (more modern samples) did not contain *Diphyllobothrium*, but two of them contained *Fasciola* (Herrmann and Schulz 1986).

5.3 Description of samples

The collection provided for this study composed of 31 samples from 8 plots and two samples from the transitional horizon to the pre-Christian settlement (before 1143 CE, samples not dated, Table 5.1). Some of the plots excavated contain more than one latrine (indicated in Figure 5.2) representing different periods. Two of five samples from the plot on Fischstrasse 28 (two latrines) show clear evidence of re-using of older structures. While one of the latrines was most likely built in the 14th century CE, the contents show artefact evidence from the 16th century CE. These older structures were completely emptied before reuse. The samples were taken from the middle of these structures to avoid potential leftover contamination of older contents (personal communication Dirk Rieger, Lübeck). On some plots there were also other types of waste dump such as pit fills or well fills. Between two and seven samples were provided for each site, representing different stratigraphic layers of the excavation. Dated stratigraphic samples allow a comparison of the parasites vertically though time (late 11th century to 17th century) and across the plots. Each stratigraphic layer was dated by artefacts contained in the layer, or in adjacent layers. Each sample comprised of about 50-100 grams of soil.

Plot	Sample number	Depth	Estimated age
Alfstrasse 17	1/3	2 m	1275 CE
Alfstrasse 17	2/3	3 m	1250 CE
Alfstrasse 17	3/3	4 m	1225 CE
Alfstrasse 21	1/3	1.5 m	1150 CE
Alfstrasse 21	2/3	1.8 m	1100 CE
Alfstrasse 21	3/3	3 m	1150 CE
Alfstrasse 23	1/3	1 m	1500 CE
Alfstrasse 23	2/3	2 m	1250 CE
Alfstrasse 23	3/3	2.5 m	1200 CE
Alfstrasse 25	1/7	1 m	1600 CE
Alfstrasse 25	2/7	1.5 m	1350 CE
Alfstrasse 25	3/7	2.1 m	1500 CE
Alfstrasse 25	4/7	2.5 m	1325 CE
Alfstrasse 25	5/7	3 m	1350 CE
Alfstrasse 25	6/7	3.5 m	1225 CE
Alfstrasse 25	7/7	4 m	1225 CE
Alfstrasse 27	1/5	1 m	1550 CE
Alfstrasse 27	2/5	1.2 m	1425 CE
Alfstrasse 27	3/5	2.5 m	1400 CE
Alfstrasse 27	4/5	3.5 m	1300 CE
Alfstrasse 27	5/5	4 m	1275 CE
Fischstrasse 16	1/2	1 m	1250 CE
Fischstrasse 16	2/2	1 m	1250 CE

Fischstrasse 26	1/3	2 m	1625 CE
Fischstrasse 26	2/3	3 m	1500 CE
Fischstrasse 26	3/3	4 m	1350 CE
Fischstrasse 28	1/5	1 m	1325 CE
Fischstrasse 28	2/5	1.6 m	1275 CE
Fischstrasse 28	3/5	1.8 m	1200 CE
Fischstrasse 28	4/5	2.8 m	1525 CE
Fischstrasse 28	5/5	3.5 m	1550 CE
culture layer	1301-369	--	before 1100 CE
culture layer	1480-457	--	before 1100 CE

Table 5.1: List of samples obtained from the excavation in Lübeck, indicating the plot they were retrieved from, the depth and the estimated age.

5.3.1 Microscopic diagnosis

All 33 received samples were assessed by microscopy. The parasite densities were determined for *Trichuris*, *Ascaris*, *Taenia* and *Diphyllobothrium* (photomicrographs in Figure 5.3). At least one of these four parasites was detected in each sample except for the two samples dating before the establishment of the city.

In 26 of the plots *Trichuris* was detected at an average density of 1005.9 eggs per gram (range 0-4934.6 eggs/g). *Ascaris* was detected diagnosed in 26 samples with an average density of 449.7 eggs per gram (range: 0-1644.9 eggs/g). *Taenia* was detected in 19 plots at an average density of 1149.6 eggs per gram (range: 0-8309.6 eggs/g). *Diphyllobothrium* was detected in 14 plots at an average density of 165.5 eggs per gram (range: 0-1414.2 eggs/g). The microscopic detection was based on the

characteristic shapes and sizes of the parasites; Figure 5.3 depicts photomicrographs of each parasite.

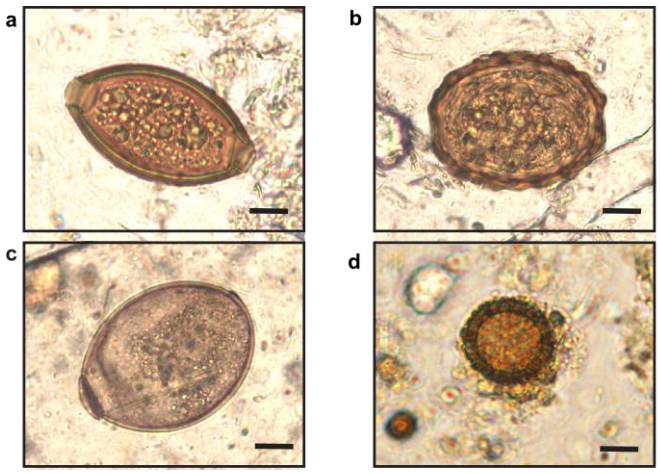


Figure 5.3: Photomicrographs of parasite eggs from the samples from Lübeck

The clearly distinguishable shape allows microscopic diagnosis of the four parasites. The pictures show one example of each (a) *Trichuris* spp., (b) *Ascaris* spp., (c) *Diphyllobothrium* spp. and (d) *Taenia* spp. Scale bar: 10 µm.

The distribution of the parasites over time is shown in Figure 5.4 (a-d). The prevalence of the nematodes *Trichuris* and *Ascaris* was independent of time and space (binomial distribution with z-values of 0.90 and -0.10, respectively; Figure 5.4, a and b). Similar distributions were found with a contemporary sample set from an excavation of the medieval port area of Bristol (Figure 4.3). In our experience the two nematode parasites are commonly found in archaeological excavations, the cestode parasites are uncommon and of particular interest. These parasites require an intermediate host and are associated with consumption of raw or insufficiently cooked fish (*D. latum*) or red meat (*T. solium* in pork and *T. saginata* in beef) which implies comments on dietary or cultural habits. The presence of *Diphyllobothrium* changed significantly over time (z-value 2.42 / p-value 0.008), but was independent

of the location (z-value -1.02). The prevalence of the fish-associated parasite *Diphyllobothrium* dropped significantly after 1300 CE to near complete absence (Figure 5.4 c). The decreasing numbers of *Diphyllobothrium* correlated with an increasing number of *Taenia* eggs. Despite *Taenia* spp. being present before 1300 CE, there was a clear trend to increased proportions of positive samples (25 % before 1300 CE and 93.3% after 1300 CE; Figure 5.4 d and e), a trend that was statistically significant (z-value -2.22 / p-value 0.013), but independent of location (z-value 1.48). For statistical analysis of the counted data, as compared to incidence, a Quasi-Poisson model was required to account for over dispersion of counts.

Most samples contained both nematode parasites and there was a significant correlation between the numbers of *Ascaris* and *Trichuris* in Lübeck (Figure 5.4 f, linear model p-value <0.001). There is a weak negative correlation between *Taenia* and *D. latum*, but as most samples contain either one or the other and only five contain both (3 of 16 before 1300 CE, 2 of 15 after 1300 CE) the correlation was less apparent. There was no correlation between the number or presence of *Trichuris* or *Ascaris* to either of the cestodes. Sampling was slightly biased to the western end of the two streets due to the ongoing excavation, but this had no significant impact on the parasite distribution (Figure 5.4 g).

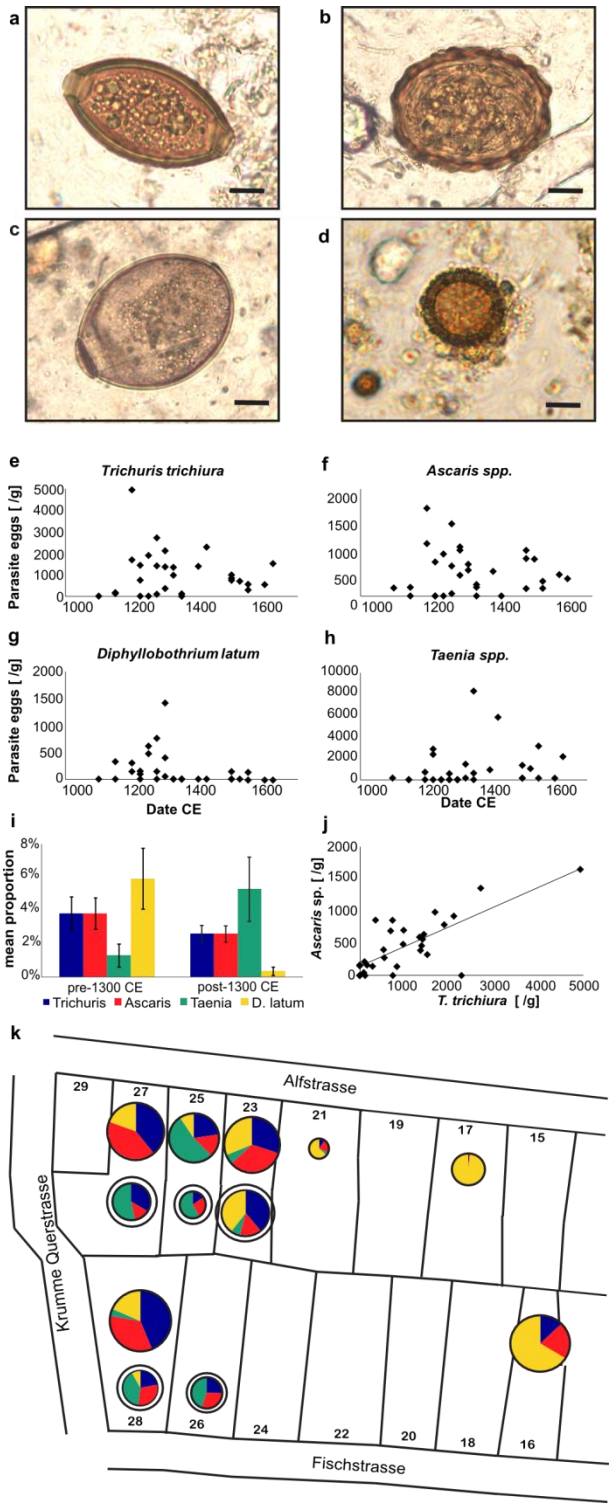


Figure 5.4: Evaluation of parasite counts of the samples from Lübeck.

The numbers of parasites found per gram of soil sample are displayed against time for *Trichuris* (a), *Ascaris* (b), *Diphyllobothrium* (c), and *Taenia* (d). The change of prevalence of *Diphyllobothrium* to *Taenia* around 1300 CE is shown in the mean proportion of the parasites per sample (e). The presence of *Ascaris* or *Trichuris* was indicative for the presence of the other parasite (binomial model, f). The map (g) represents the proportion of the four parasites per plot before and after 1300 CE, where the size of the charts represents the proportion of the overall parasite count in the sample.

5.3.2 Molecular data

Using appropriate primers to amplify specific targets for PCR and sequencing it was possible to determine the species identity for each parasite. For *Ascaris* sp. CO1 was used, but as described in chapter 2.4, the majority of the sequences had to be excluded as non-nematode sequences, however, there were still genuine sequences confirming the microscopic diagnosis (Figure 5.5).

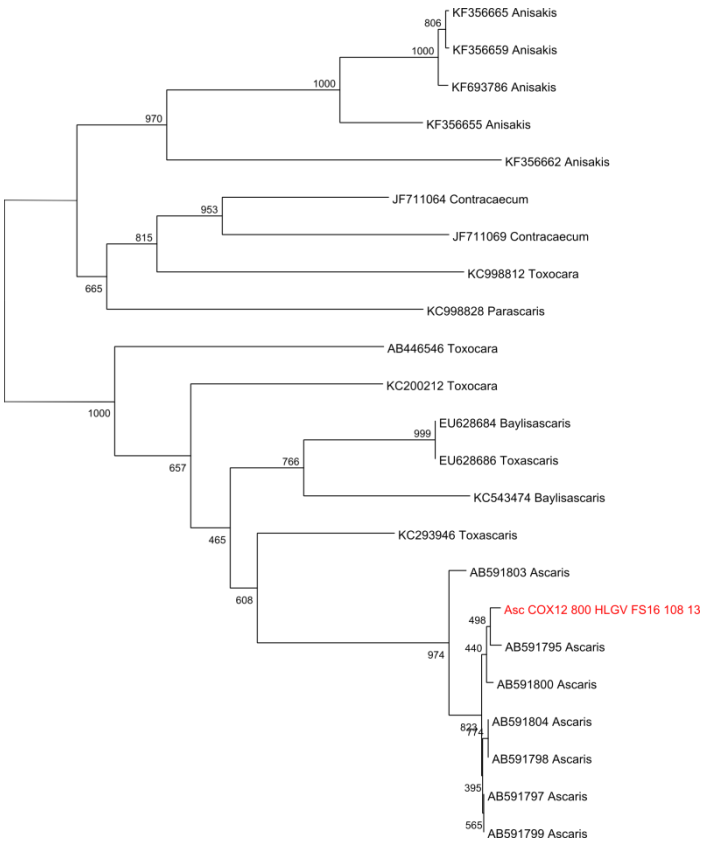


Figure 5.5: Bootstrapped Neighbour-Joining tree for *Ascaris* sp. CO1

The sequence generated from a sample from Lübeck (highlighted in red) was located inside the modern *Ascaris* sp. clade. The tree contained other Ascaridians and *Anisakis* sp. as distant outgroup.

CO1 was also used to confirm the identification of *Diphyllobothrium latum*. For the other described species the genus identification based on the morphology of the eggs was reliable, but for *Diphyllobothrium* spp. this was more difficult as the eggs exhibit less characteristic features. The primers for the PCR on CO1 were

adapted from previously published primers (Wicht et al. 2010). Sequences were obtained from two samples and were used to verify the species identity on a bootstrapped neighbour joining tree. The tree was constructed using a nematode outgroup (*Caenorhabditis elegans*) and various other *Diphyllobothrium* spp. as comparison . All sequences from Lübeck (in blue) were found within the *D. latum* clade (Figure 5.6).

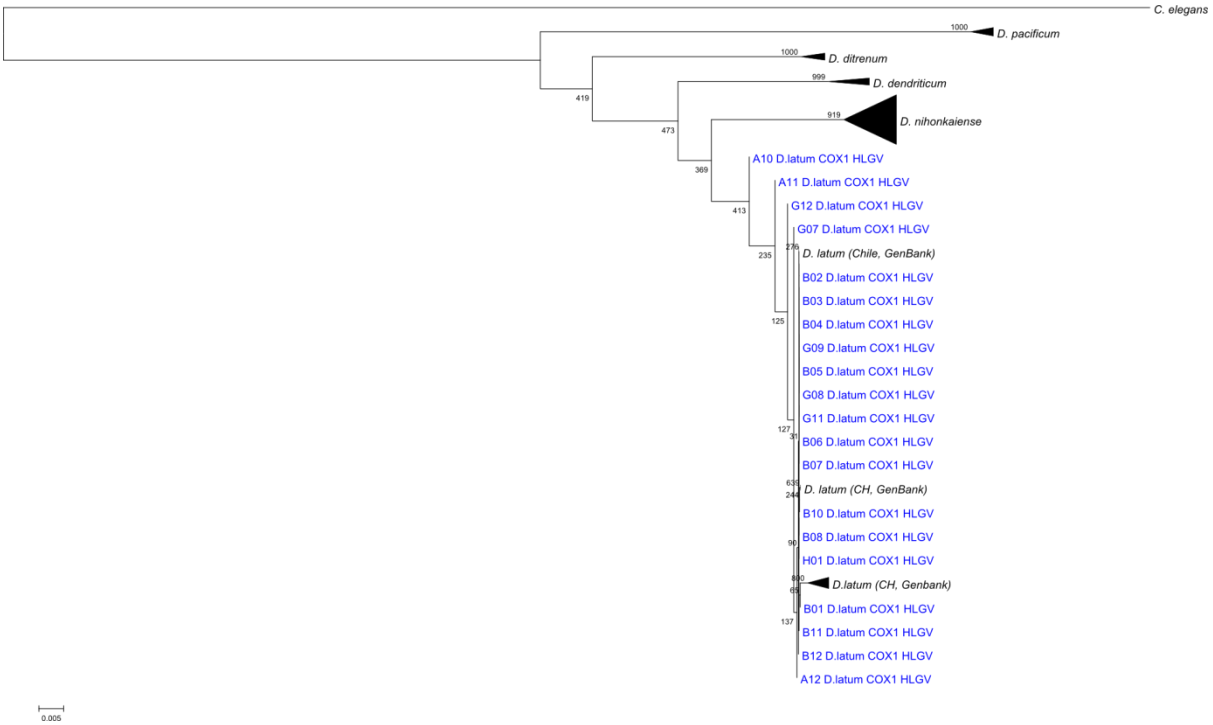


Figure 5.6: Bootstrapped Neighbour-Joining tree to assess the species identity of *D. latum*.

The nematode *Caenorhabditis elegans* was used as outgroup. The obtained sequences from Lübeck (highlighted in blue) were compared to various *Diphyllobothrium* spp. obtained from the GenBank database (*D. pacificum*, *D. ditrenum*, *D. dendriticum*, *D. nihonkaiense* and *D. latum*) showing the closest relationship to *D. latum*.

For the molecular identification of *Taenia* spp. the mitochondrial cytochrome *b* gene was used. A series of primers were designed to distinguish between the two species *T. solium* (pork tapeworm) and *T. saginata* (beef tapeworm), as the microscopic diagnosis could only determine the genus. A bootstrapped neighbour-joining tree was constructed using a combination of nematodes (*C. elegans*, *Ascaris* sp. and *Trichuris* spp.) as outgroup (Figure 5.7). The bootstrap values show a clear separation between the cestodes *D. latum* and *Taenia* spp., as well as between *T. solium* and *T. saginata* (991/1000). The sequences derived from the Lübeck samples all grouped within the *T. saginata* clade (highlighted in blue, Figure 5.7). However, as these sequences were generated from only one sample it could not be excluded that some of the samples contained *T. solium*.

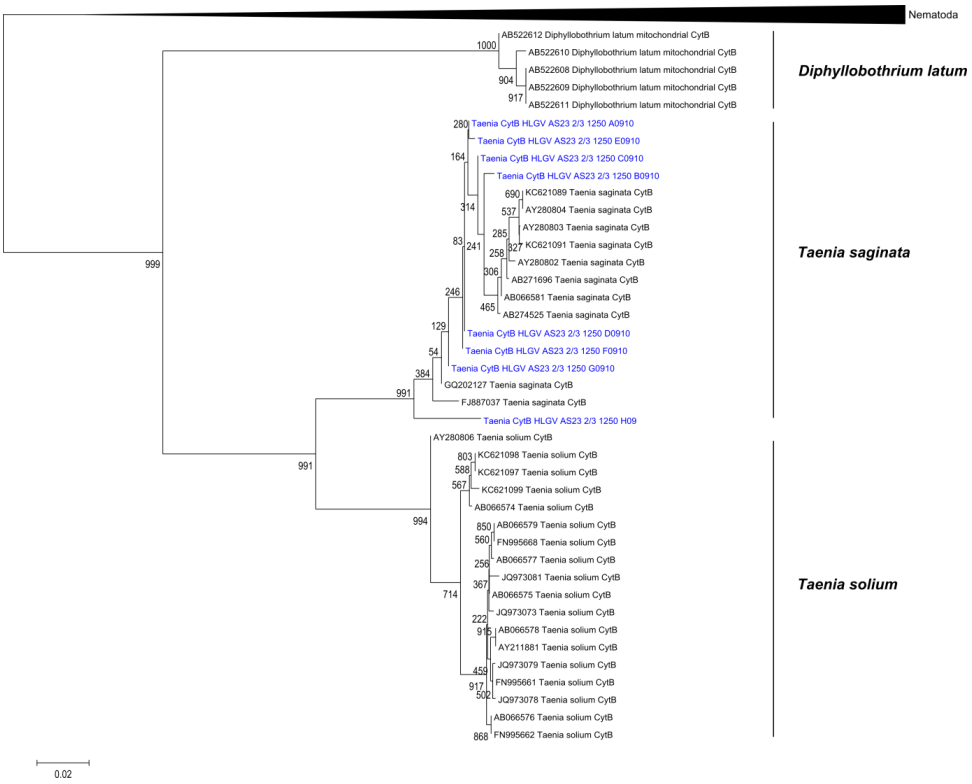


Figure 5.7: Bootstrapped Neighbour-Joining tree for species identification of *Taenia* spp. for the samples from Lübeck.

Sequence comparison identified the *Taenia* spp. as *T. saginiata* (beef tapeworm) in at least one ancient sample (highlighted in blue). Nematode outgroup: *C. elegans*, *Anisakis* sp and *Ascaris* sp

The most comprehensive set of sequences was generated from the ribosomal internal transcribed spacer 1 (ITS1) for *Trichuris* spp. which allowed further insight into distributions within and between sites. Initial assessment of the sequences on a neighbour-joining tree confirmed that the sequences obtained were *T. trichiura* (Figure 5.8 and Figure 4.3).

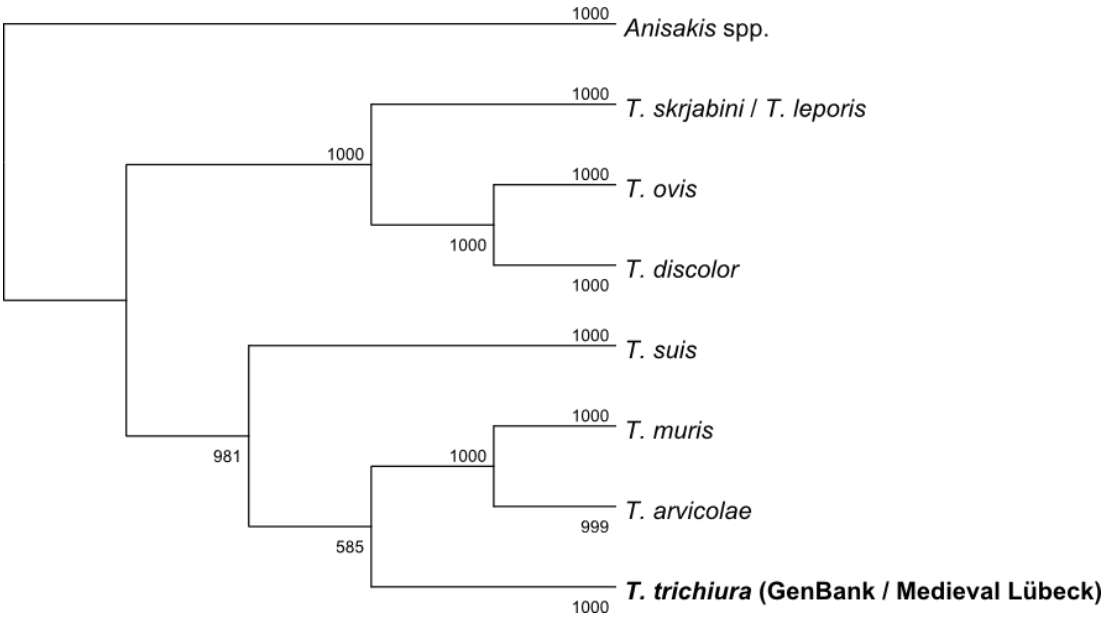


Figure 5.8: Bootstrapped Neighbour-Joining tree to identify the species of the samples from Lübeck.

All sequences obtained from the site were clearly identified as *T. trichiura*.

As displayed in Figure 5.8, the sequences generated using *T. trichiura* ITS1 PCR primers showed the highest similarity to *T. trichiura* ITS1 obtained from the GenBank database. Within the *T. trichiura* clade considerable variation was found (Figure 5.9), which made a further evaluation comparing the obtained sequences with those obtained from other sites necessary. The tree in Figure 5.9 shows the *T. trichiura* branch from Figure 5.8 in more detail.

Figure 5.9: Bootstrapped Neighbour-Joining tree comparing the *T. trichiura* ITS1 sequences from Lübeck.

Modern sequences obtained from the GenBank database were collapsed in groups. The root correlates to the end of the branch in Figure 5.8. The variation between the sequences was found to be considerable.

To investigate the potential to use these parasites to trace human activities, the sequences generated from two other medieval sites (Bristol and Ellwangen) were

added to the analysis (Figure 5.10). Phylogenetic analysis on *T. trichiura* ITS1 suggests two distinct groups, one being nearly exclusively found in Lübeck, independent of time which also contained three sequences from Bristol as depicted in Figure 4.4 (indicated in red, clade 1), and one mixed clade that included samples from samples from the ports of Lübeck and Bristol, the land-locked market town of Ellwangen in Southern Germany, and all modern sequences (indicated in grey, clade 2) which indicated exchange between. The samples from Ellwangen were received from a graveyard associated with a church. A more detailed analysis of the site of Ellwangen will be in chapter 7.

A smaller collection of sequences was acquired from *T. trichiura* β-tubulin. The *β*–tubulin gene was considered relatively stable. The primers were designed based on published primers generating a 163 bp product (Diawara et al. 2009). Unfortunately, the number of generated products was limited due to mispriming on non-target sequences and primer dimerisation. The tree in Figure 5.11 depicted a neighbour-joining tree of medieval sequences from three sites (Lübeck, Bristol and Oxford). The tree showed a similar structure to the *T. trichiura* ITS1 tree in Figure 5.10, except that for β-tubulin the sequences from Lübeck associated with the modern database entries (clade 1) and the sequences from Bristol formed an external clade. This further supported the evidence of the existence of two separate sub-clades of *T. trichiura.*

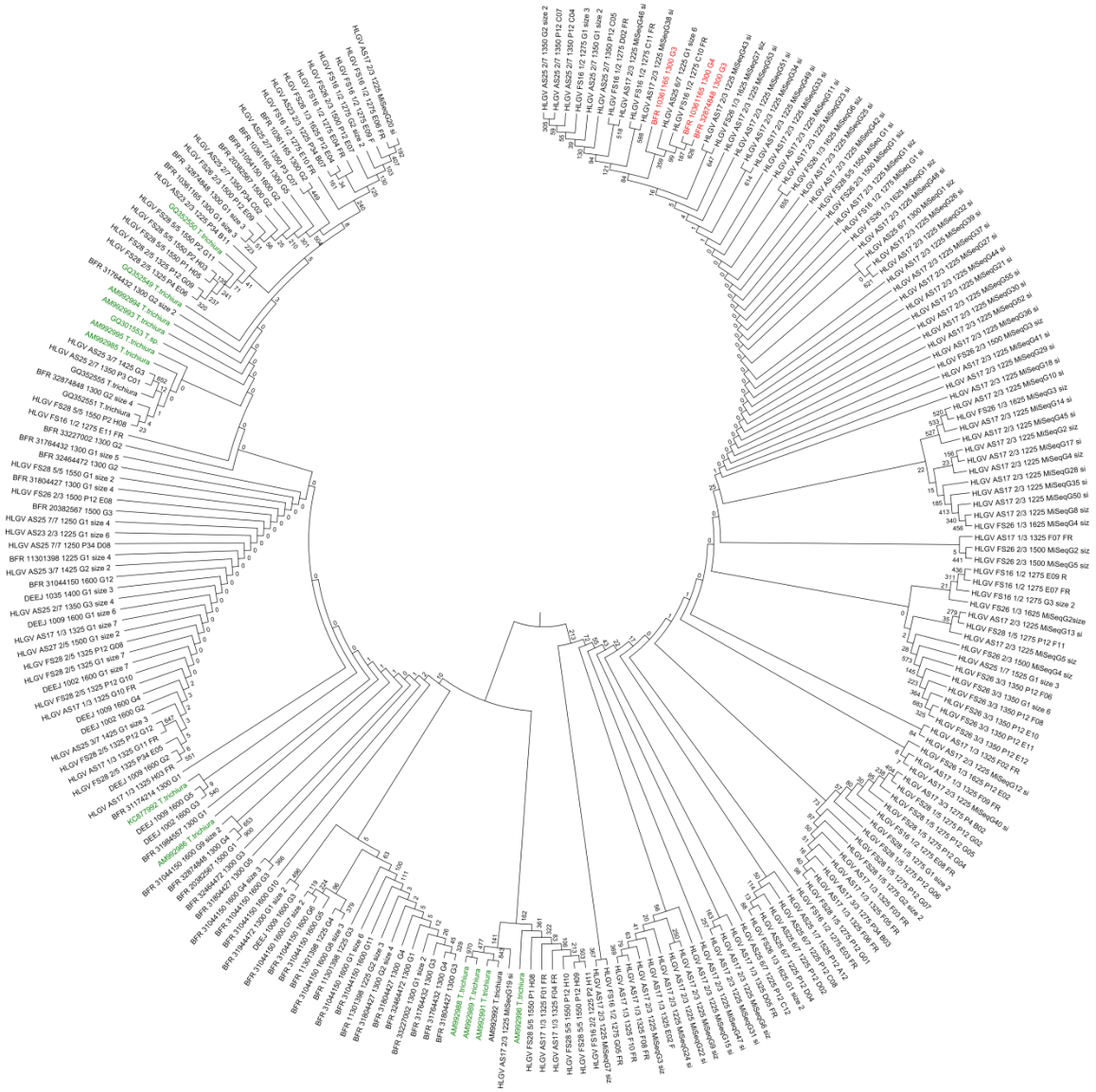


Figure 5.10: Cut-out from Bootstrapped Neighbour-Joining Tree of *T. trichiura* ITS1 from Lübeck (HLGV), Bristol (BFR) and Ellwangen (DEEJ).

The samples fall in two clades, with clade 1 (right) being dominated by samples from Lübeck, whereas the clade 2 (left) was intermixed, including all modern sequences obtained from GenBank (highlighted in green). In the clade dominated by the samples from Lübeck there were only three samples (highlighted in red) from a different site (Bristol).

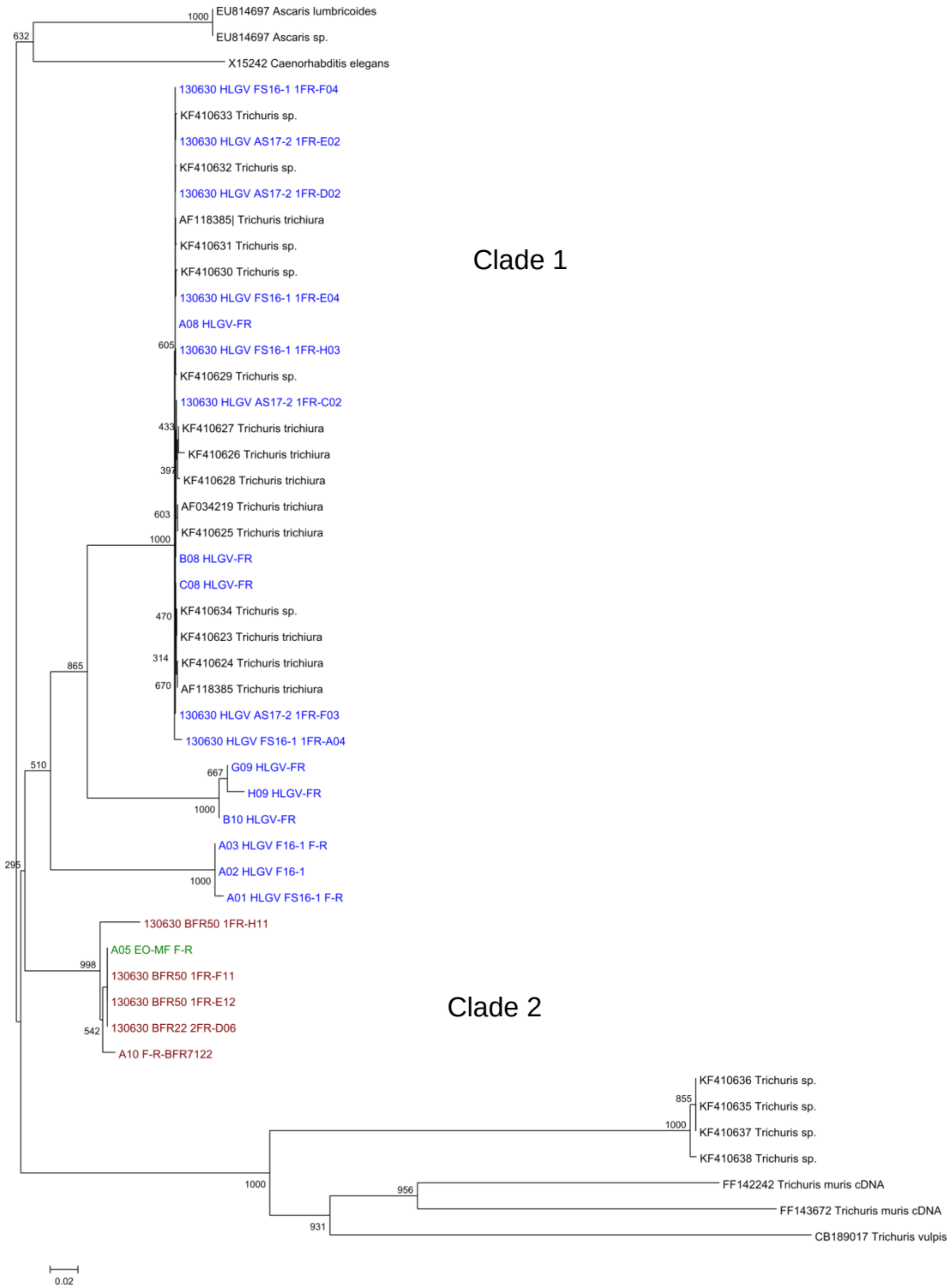


Figure 5.11: Bootstrapped neighbour-joining tree for *T. trichiura* β-tubulin.

Sequences from three the medieval European sites in Lübeck (blue), Bristol (red) and Oxford (green) were analysed on this tree against modern *T. trichiura*, *T. muris*, *T. vulpis*, using the nematodes *Ascaris* sp. and *C. elegans* as outgroups. The samples fall into the *T. trichiura* clade where the sequences from Lübeck associate into a separate clade (clade 1) compared to the other sequences.

5.4 Discussion

This study represents the first comprehensive study employing microscopic and molecular analysis of human parasites in any archaeological context. The extensive sample set from Lübeck allowed a variety of conclusions to be made, relating to the life of this thriving Medieval European port city. Parasites represent a unique source of information which can be interpreted in relation to human activities. Other sources of information can be actively biased (by behaviour).

The fluctuations in number of *T. trichiura* and *Ascaris* spp. were comparable with observations in other sites in Europe (for example Bristol, as described in Chapter 4) and since infection rates vary in modern developing world populations (Pullan et al. 2014) it seems to be reasonable to assume that this will have been the case in Medieval Europe. Comparing the Lübeck data with previously published material showed interesting similarities and differences. Paap (1984) analysed 12 samples from a latrine dated to the 17^{th}-18^{th} century from the house in the Gerade Querstrasse 1 between the plots at Fischstrasse 28 and Alfstrasse 29 (see map in Figure 5.2, (Paap 1984)). The samples were dated to the 17^{th}-18^{th} century slightly later than those examined in the current project. The average densities of *Ascaris* and *Trichuris* obtained by Paap were 324 Ascaris eggs per gram, and 317 Trichuris eggs per gram was comparable to the youngest samples from the current study (441.3 *Ascaris* eggs per gram, and 765.1 *Trichuris* eggs per gram). Herrmann and Schultz (1986) studied a core through a single latrine in a different part of the city (Fronerei am Schrangen, East of the St Mary's Church, approx. 0.5 km from the site) (Herrmann and Schulz 1986). Unfortunately, the publication does not contain dates for the core, just depths relative to mean sea level. The core studied was 3.54 m deep and was sliced in 32 layers. The fluctuations in egg numbers were comparable with

those seen in our samples, yet the absolute numbers were higher with 2976.9 *Ascaris* eggs per gram (ranging from 80 to 11620 eggs per gram) than the 449.7 eggs per gram in the current study and 10498.1 *Trichuris* eggs per gram (ranging from 1300 to 22480 eggs per gram) compared to the 449.7 eggs per gram in the current study. The differences may be due to the difference in location and setting. The core was taken from a deposit in the Fronerei (prison) which implies that not only the number of people contributing to the deposit may be higher but it is also likely that their general health condition was worse.

The dramatic temporal shift from high numbers of *D. latum* to *Taenia* (likely *T. saginata*) was an intriguing find as this was not seen before in any other sample sets presented in this thesis and to our knowledge published elsewhere. The observed change in the cestode population is an independent indicator for a change in cultural or dietary practices. Molecular analysis confirmed the identity of *D. latum* (Figure 5.6) and *T. saginata* (Figure 5.7). Although the identification of *T. saginata* was based on one sample and although the presence of *T. solium* cannot be refuted, it provided strong supporting evidence for the increasing beef consumption. Studies on the bone content of latrines from excavations in Lübeck in 1990 indicated that beef was more popular than pork (Paul 1990; Pudek 1990). The excavation at the Heilig-Geist-Hospital (Holy Spirit Hospital) reported 62.1 % beef bones and only 22.5 % pork bones in layers dated to between the 13th-15th century CE (Pudek 1990). Similar results were found in latrines from the excavations in the Königsstrasse where 70.8 % of the bones were beef bones, 15.1 % pork bones and 21.8 % either sheep or goat bones (Paul 1990).

The evidence found for *D. latum* was interesting as the detection of fish bones in archaeological contexts is generally limited to larger bones, as fishbone is less durable and would require washing of the entire content to extract the small bones. Bigger bone remains were recorded, but not quantified. In most cases the focus of the fish remains was on stockfish-related remains (mainly cod), but various other fish species have been either reported from excavations or historic records, including sturgeon, pike, carp, and bream ((Fehring and Gläser 1991) and personal communication Dirk Rieger, Lübeck). Infection of humans with *D. latum* requires the consumption of raw or undercooked fish, it is possible that the air-dried stockfish was a source of infection, although it cannot be excluded that incautious preparation of fresh fish (i.e. not properly gutted fish or contamination of cooked fish with uncooked fish remains) contributed to the infections. Herrmann and Schulz (1986) reported *Fasciola* sp. and *Diphyllobothrium* sp. eggs. The publication did not provide numbers, but based on the publication the numbers for *D. latum* could be estimated to be in the range of <1000 per gram. The current study showed *D. latum* in 16 of 33 samples with a bias towards the earlier samples. Herrmann and Schulz (1986) found similar evidence for the presence of *D. latum* (25 of 32) and they did not report *D. latum* eggs in later samples (Herrmann and Schulz 1986). The disappearing of one particular species of parasites may be linked to cultural changes or the availability of Stockfish. In the early years of Hanseatic trade, Norwegian Stockfish (air-dried fish, mostly cod) was a commonly traded commodity which was traded in large volumes through Lübeck from where it got re-distributed to other cities (Gläser et al. 1999). The trade in the Hanseatic League changed over time from a spoke and hub distribution where all good were traded via Lübeck to more direct trade between satellite cities. On the example of Stockfish, it is worth to consider the changes in

trade from Bergen in Norway. Stockfish was the main export good from Bergen in Norway, which was the centre of the Norwegian foreign trade in the middle ages. Bergen was ideally situated as trading port with good access for ships from the British Isles and north-eastern Europe. Trade in Bergen was dominated by foreigners, in the late 12th century CE mainly English traders (Øye 1996). Stockfish accounted to 80-90 % of all exported goods (Nedkvitne 2014). Trade of Stockfish to Germany has been reported as early as 1186 when German traders bought fish and butter and sold wine (Gläser et al. 1999; Øye 1996). Stockfish was such an important commodity that the cooperation of Hanseatic traders with Bergen ('Bergenfahrer') had a crowned stockfish in their crest (Bruns and Brandt 1953). Although the German traders were initially not very welcome, they gained more importance when with the growing population in Europe the demand for Stockfish rose and the trade volume required the import of fresh fish from all over Norway (mostly the Lofoten and Vesterålen) to produce more Stockfish which was sold in exchange for grain, flour and malt from Lübeck to feed the growing population of Bergen (Øye 1996). Although political tensions were growing, the city of Bergen became more dependent on the grain imports that in 1250 the Norwegian king Håkon IV signed a free trade agreement with Lübeck. This agreement made it possible that the Baltic rye sold by Lübeck traders was better available than the English wheat and that some German traders stayed in Bergen for the winter (Nedkvitne 2014). The German settlement in Bergen (in Bryggen, sometimes also called Tyskebryggen which is Norwegian for German quay) grew as Hanseatic League traders and craftsmen settled in Bergen. King Magnus VI grated foreigners the same rights and duties as the Norwegians (bylov, 1276 CE) and later the same trade privileges (1294 CE, except for the trade with the territories north of Bergen which were still only

privilege of Norwegian traders) which led to the establishment of the fourth Hanseatic Kontor in 1343 CE (Bruns and Brandt 1953; Nedkvitne 2014; Øye 1996). With the growing German population and the growing influence of the Hanseatic League on the trade the English traders were gradually replaced, even on the trade routes to England. Customs documents show that the first Hanseatic traders brought Stockfish to England in 1280 CE and by the mid-14th century CE the Stockfish trade to England was dominated by the Hanseatic League (Bergenfahrer) which operated through King's Lynn and from 1303 CE through Boston which became the centre of the Stockfish trade in England (Nedkvitne 2014). The change of practice to no longer import the Stockfish to Lübeck and to distribute it from Bergen made it a less common commodity in Lübeck. Further, after the Black Death (1349-50 CE) the demand and prices for Stockfish were rising (Gläser et al. 1999; Øye 1996) which might have led to a further replacement of Stockfish on the dining tables in Lübeck and thus a reduction of *D. latum* in their latrines. These events may have contributed to the reduced levels of *Diphyllobothrium* at around 1300 CE.

However, the biology of *D. latum* suggests that the marine fish which were used to produce Stockfish were unlikely the source of the infection, as *D. latum* is predominantly found in freshwater fish. Although the hypothesis of Stockfish being the source of the infection cannot be refuted completely due to the predatory nature of the fish used, it is more likely that environmental or cultural change led to the lower infection rates. In communication with the archaeologists in Lübeck it was speculated that the growth of the city led an increased demand in food. Thus, the source of the fish could have changed. Further, the increased demand in other goods, especially meat and leather, led to increased numbers of production facilities to produce these goods. Most of the butcheries and tanneries in Lübeck were found on

the western part of the city. Waste from these facilities might have polluted the Wakenitz river, which is not tidal like the Trave on the eastern side of the city. Wastewater, especially from tanneries, could potentially interrupt the lifecycle of *D. latum* (see Figure 1.4) by affecting copepods in the water (Cooman et al. 2003; Lafferty 1997).

The *T. trichiura* ITS1 sequences provided insight into the structure of the parasite population potential links with another contemporary port in Bristol and the land-locked market town of Ellwangen in Southern Germany. There were two main clades of *T. trichiura* ITS1 sequences (Figure 5.10). One of the two clades contained sequences from Lübeck and three sequences from 1300 CE Bristol (clade 1), whereas the other clade represented a mixed population containing sequences from all three sites and all the modern reference sequences (clade 2). Clade 1 might represent an extinct Northern European variant of *T. trichiura*. The sequences from Bristol in this clade could represent some form of link (e.g. trade) between Lübeck and Bristol. Supporting evidence for the two clade hypothesis was found in the independent *T. trichiura* β-tubulin sequence (Figure 5.11).

As the samples were all taken from latrines of individual houses (see Figure 5.2), the sampled population might be more restricted and re-infection of one predominant type would be facilitated by only a small group of users for one latrine. The sequences from Lübeck samples which were not in clade 1 would then represent *T. trichiura* variants which were acquired outside of Lübeck. This closed circle of re-infection could then represent a "Lübeck variant" of *T. trichiura*. The single grave sequences from Ellwangen in Southern Germany were found in the mixed clade 2 which suggested that the clade 1 variants were not prevalent in Ellwangen. However, as sequences were only obtained from three graves in Ellwangen this must be

considered speculative. The value of the Ellwangen graves will be strengthened by the evaluation of the 92 further samples which were obtained from Ellwangen just before submission of this thesis.

The difference in the sampling could also provide insight into the distribution of the sequences. The samples from Bristol were obtained from a ditch outside the city walls. These deposits were likely to contain contributions from the poorer population of Bristol, whereas the samples from Lübeck were obtained from merchant's latrines. These had likely considerably fewer users than contributed to the deposits in Bristol. This would suggest a more mixed population in the port of Bristol and a relatively contained parasite population within the merchants in Lübeck. This would be further supported by a note by Sherborne that Bristol merchants frequently embarked with their cargo on foreign ships (Sherborne 1965).

This sample set showed the potential of the approach to apply parasitological data on archaeological samples. Further correlation with historic records where available will strengthen evidence to elucidate historic events.

Chapter 6

Neolithic parasites from early settlement and grave deposits

One critical element of aDNA work is how far back aDNA can be extracted to produce reproducible sequences. This chapter describes the oldest samples in this study which generated the oldest available pathogen sequences. The sample collections consisted of material from stilt dwellings in Switzerland and Germany and a collection of some of the oldest excavated burials in Germany.

6.1 The sites and their history

A collection of stilt settlement samples from excavations in Switzerland and Germany represented the oldest European samples in the study. This settlement type has been found in various places around the Alps and a collection of 111 small individual sites has been listed on the UNESCO World Heritage List in 2011. As the sites were generally close to water the preservation of the archaeological material has been found to be exceptional (Schlichtherle et al. 1997).

The Neolithic settlement samples from Switzerland were obtained from the excavation in front of the Opera house in Zurich in 2010 ('Kleiner Hafner'). During building works for underground parking the archaeologists found traces of a stilt settlement which was in use between the Pfyn culture (c3900-3500 BCE) and the Horgen culture (c3500-2800 BCE). Samples were obtained from two building sites, one dating to the Pfyn culture (samples dated c3630 BCE) and one to the Horgen culture (samples dated c3060 BCE). In total, 15 samples were obtained from the archaeological service of the city of Zurich (Niels Bleicher).

The state archaeology of Baden-Württemberg (Landesdenkmalamt Baden-Württemberg, Helmut Schlichterle) provided a collection of 19 samples from six sites in the region of Lake Constance (Bad Buchau-Bachwiesen, Bad Buchau-Torwiesen II, Olzreute-Enzisholz, Alleshausen-Hartöschle, Alleshausen-

Täschenwiesen and Sipplingen-Osthafen). The sites showed cultural associations of the Schussenrieder Gruppe (c4200-3700 BCE), the Pfyn culture or the Horgen culture and were dated between c3800 BCE and c2800 BCE.

The third set of samples was obtained from seven graves in Bad Mergentheim associated with the Corded Ware culture (provided by Joachim Wahl, Landesdenkmalamt Baden-Württemberg and University of Tubingen). This culture is sometimes also referred to as the single grave culture as burials are one of the signature inventions. There is little known on the culture or economy during this period. These samples from Bad Mergentheim are special in the collection as they are among the oldest ever recovered single burials and were the oldest burials evaluated for parasites. The samples were dated by the inclusion of the culture's namesake corded ware pottery into the burials to c2900 BCE (Furholt 2004). Despite the underrepresentation of settlement finds compared to burial finds it is widely accepted that the corded ware culture was farming based rather than a nomadic hunter gatherer culture. Burials in this period are generally orientated east-west (in central Europe) with the deceased lying on the side with tucked up legs. Generally, women were buried lying on their left and men lying on their right (personal communication Joachim Wahl, Landesdenkmalamt Baden-Württemberg and in Furholt 2004). In the later phase of the culture an increased amount of burial objects are found (personal communication on excavation, Joachim Wahl).

Parasites were found microscopically in all three collections. For the samples from Zurich there was also sequence data obtained from two parasites.

6.2 Results – Neolithic burial ground Bad Mergentheim

Seven samples from the burial ground were obtained, representing five single graves and one double grave (33-379 A/B). The double grave was most likely the burial of a mother and child, but detailed analysis of the relationship have not been conducted. This sample set was unusual compared to other single grave sites where overall prevalence rates of 15-30 % were found as all seven samples were diagnosed positively for *Ascaris* sp. and one for *Trichuris* (discussed in detail in chapter 7). The concentrations were relatively low, comparable to the lower end of finds of other sites. The concentration range of *Ascaris* was between 19.9-121.1 eggs per gram and for *Trichuris* the detected concentration was 78.9 eggs per gram (Figure 6.1).

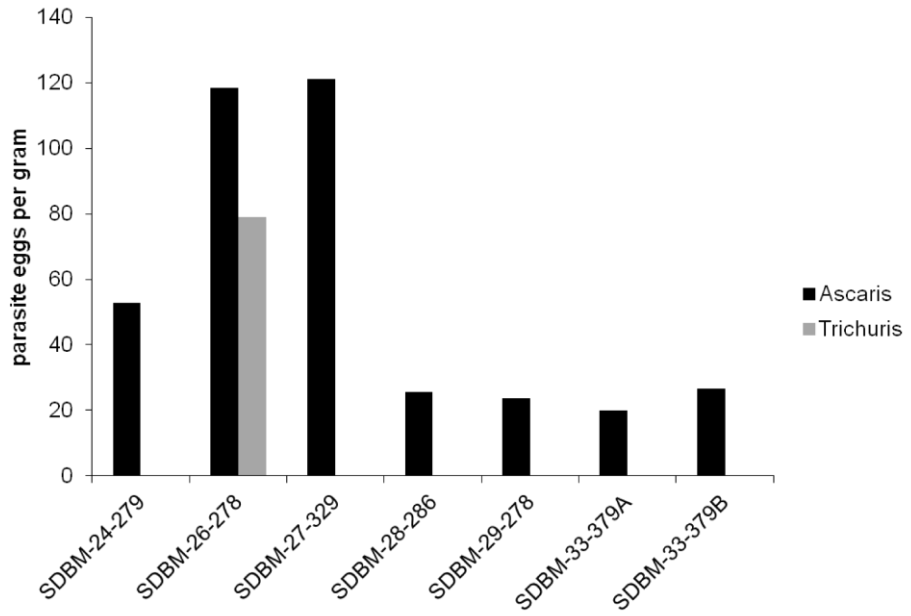


Figure 6.1: Parasite counts in the burials at Bad Mergentheim

Number of parasite eggs per gram of soil sample. The two last samples were excavated from one grave and were an adult woman (A) and a child (B) were buried together.

6.3 Results – Stilt settlements in the Region of Lake Constance

The samples of this collection were excavated between 1993 (Alleshausen-Hartöschle) and 2011 (Sipplingen-Osthafen). Parasite eggs were detected in nine of the nineteen samples with the most prevalent parasites being of the genera *Trichuris* and *Ascaris*. In one sample there was a single structure which was tentatively identified as *Taenia* sp. (Sipplingen-Osthafen).

Overall, the 19 samples were collected at seven excavations at six sites. Eight samples in five sites were positive for parasite eggs with no positive samples in Alleshausen-Hartöschle (three samples). Parasites of the genera *Ascaris* and *Trichuris* were found in the sediments, eight of the samples were positive for one or both of the parasites with densities of 200-650 eggs per gram for *Trichuris* and 150-450 eggs per gram for *Ascaris* in dry sediment.

6.4 Results – Kleiner Hafner – Excavation Opéra Zurich

The excavation in front of the Opera house in Zurich was part of a rescue excavation. The conditions for the preservation of organic material were very good even large wooden objects were found. Of the 15 samples from the excavation in Zurich, 13 were positive for nematodes, of which eight were positive for *Trichuris,* ten for *Ascaris* and 5 were positive for both nematodes. In the samples diagnosed positively, the numbers of parasite eggs ranged from 200-1240 parasite eggs per gram for *Ascaris* and 509-1649 parasite eggs per gram for *Trichuris* (Figure 6.2). Interestingly, the highest and the lowest number of parasite eggs for both parasites were found in the same two samples.

The samples from Zurich were the oldest samples to produce sequences. Sequences were obtained for *T. trichiura* ITS1, *T. trichiura* β-tubulin and *Ascaris* sp.

cytochrome *b*. Sequences for *Ascaris* were obtained from six samples, by Sanger sequencing and MiSeq parallel sequencing (for *T. trichiura* ITS1 only). Sequences for *T. trichiura* ITS1 were obtained from two samples by Sanger sequencing and for four by parallel sequencing, including two which were not diagnosed positive by microscopy.

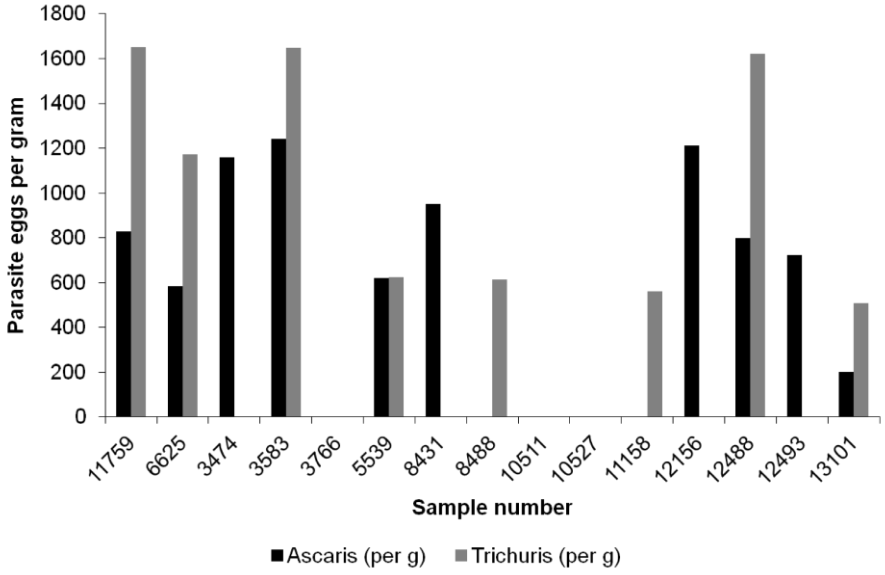


Figure 6.2: Parasite counts in the excavation of Zurich.

Most samples contained at least *Trichuris* or *Ascaris*, the counts per gram were indicated for both parasites.

In samples from Bristol (chapter 4) and Lübeck (chapter 5), a positive correlation between the numbers of *Ascaris* and *Trichuris* were observed. For the samples from Neolithic Zurich this could not be verified. A weak correlation was observed (Figure 6.3), but the on the r^2 value (0.08) suggested an insufficient statistical fit, probably due to a considerable number of samples which only contained one or the other.

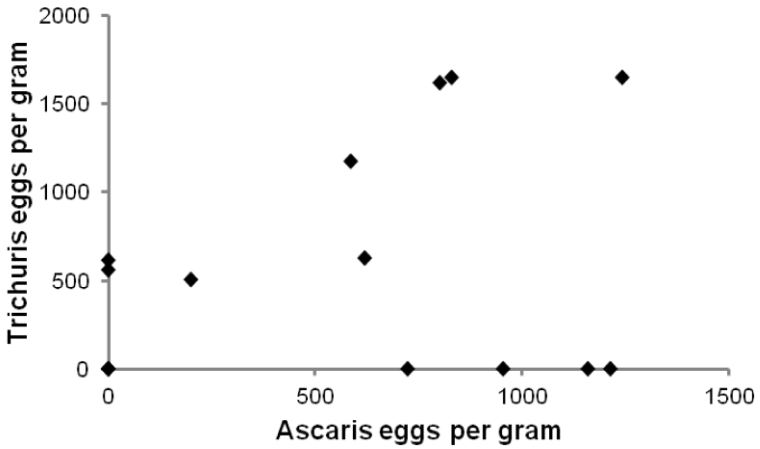


Figure 6.3: Correlation between the numbers of *Ascaris* and *Trichuris*

Number of *Ascaris* and *Trichuris* were displayed against each other. No clear correlation can be seen.

6.4.1 Sequence evaluation for Zurich samples

Sequences were obtained from six samples (10 sequences, Sanger sequencing) for *Ascaris* sp. cytochrome *b*, 349 unique sequences (of which 336 from MiSeq parallel sequencing) from four samples for *T. trichiura* ITS1, and from two samples (two sequences, Sanger sequencing) for *T. trichiura* β-tubulin.

The neighbour-joining tree for *Ascaris* Cytochrome *b* was constructed with GenBank sequences from *T. trichiura*, and the Ascarididae *Toxascaris* spp., *Parascaris* spp., *Baylisascaris* spp. and *Ascaris* sp. The 10 aDNA sequences generated from the excavation of Zurich (highlighted in blue, Figure 6.4) were located within the *Ascaris* clade, separate from all other Ascarididae and other nematodes (*Trichuris*). The structure of the clade suggested a subdivision into three parts which contained only modern sequences, only ancient sequences and a mixture of ancient and modern sequences, respectively. The subdivision indicated separation of modern sequences within the *Ascaris* spp. clade, but this separation did not did not follow the annotated divide (*A. suum* vs. *A. lumbricoides*). The sequence diversity

against the sequences in the database was visualised on a logo plot (Crooks et al. 2004; Schneider and Stephens 1990; Schneider 1996), indicating a relatively low variation (9/194 bp, 4.64 %; Figure 6.5, created using weblogo.berkeley.com). The variable sites are indicated on the logo plot in Figure 6.5 the 9 variable sites were distributed across the sequence (in positions 15, 30, 32, 64, 73, 96, 97, 125 and 128) without larger clusters.

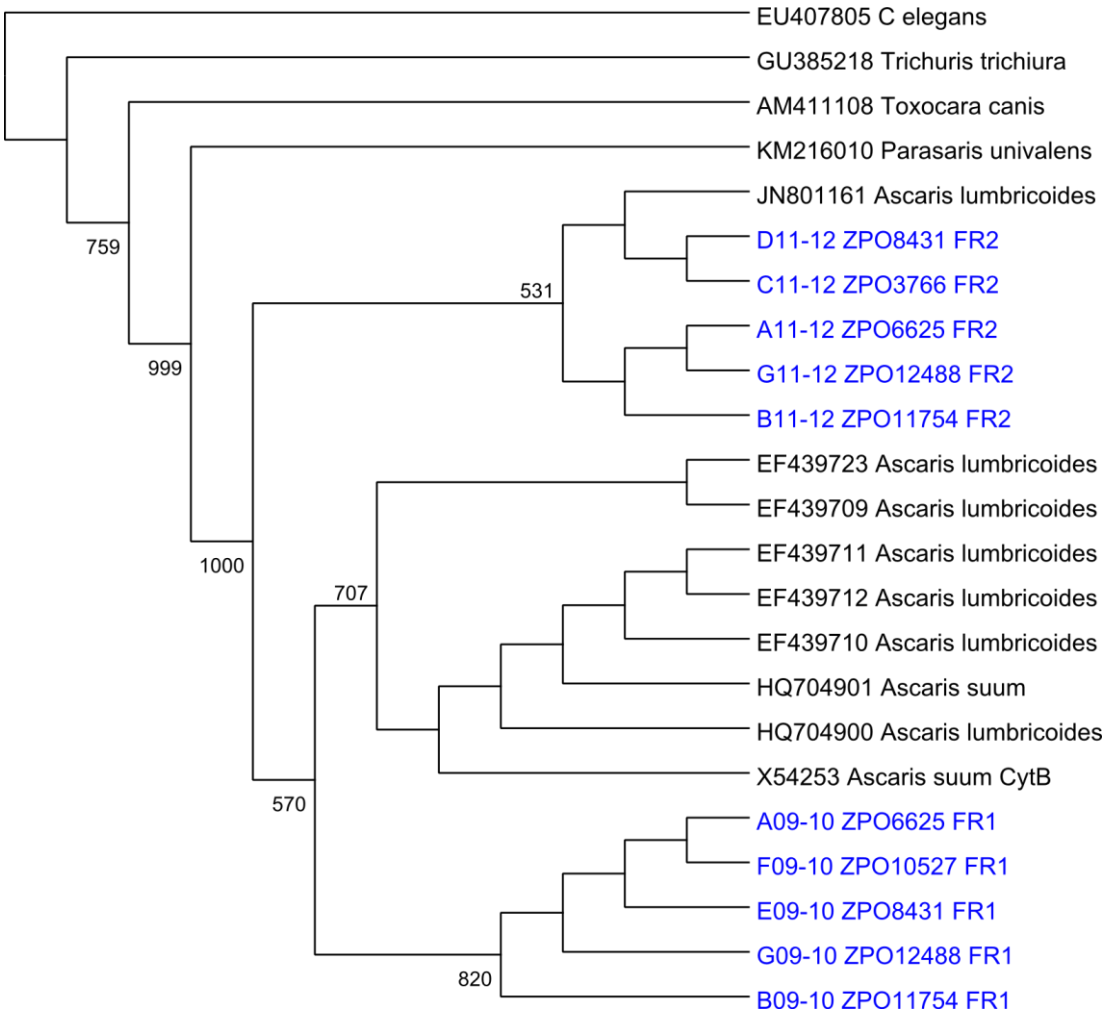


Figure 6.4: Bootstrapped Neighbour-Joining tree of *Ascaris* sp. cytochrome *b* sequences from samples of Neolithic Zurich and modern database sequences.

The sequences were compared to modern *A. suum* and *A. lumbricoides*, as well as *Toxascaris*, *Parascaris*, *Baylisascaris* and *Trichuris* sequences. Ancient samples (highlighted in blue) were found to be located within the *Ascaris* spp. clade.

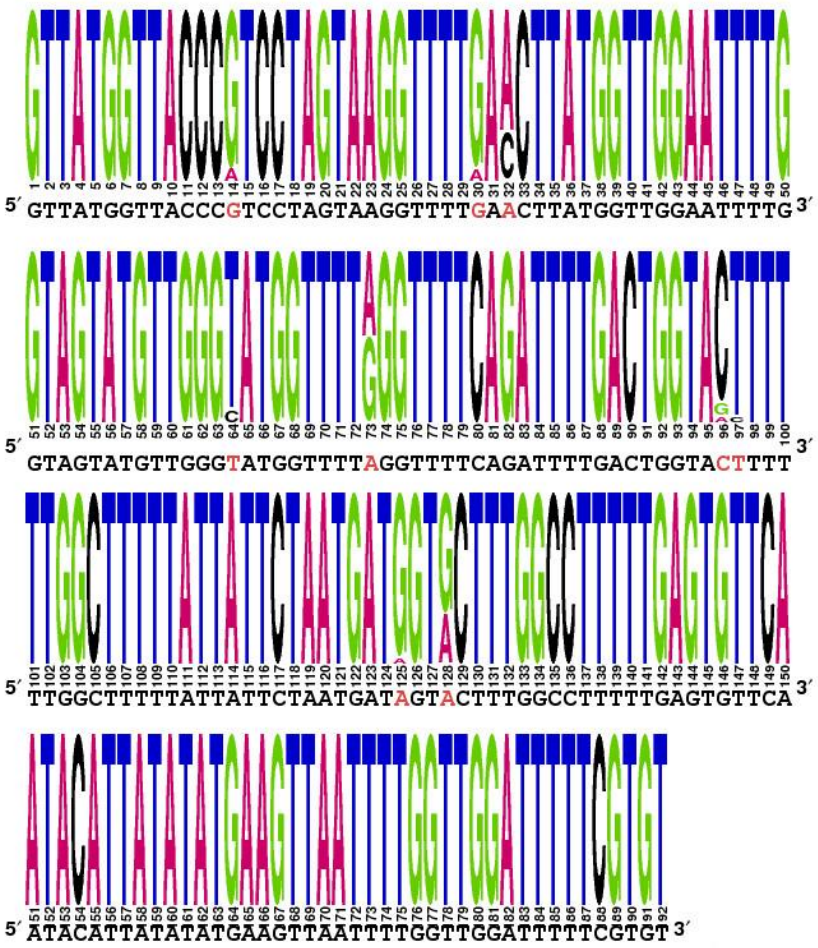


Figure 6.5: Sequence logo plot of the sequence of *Ascaris* sp. cytochrome *b*.

The size of the letters represents the frequencies or proportion of the bases in the sequence at each position; polymorphic sites were indicated on the consensus sequence. Generated using WebLogo (weblogo.berkeley.org)

A considerable amount of the sequence data for *T. trichiura* ITS1 was generated by parallel sequencing (336 of 349 sequences). The MiSeq parallel sequencing approach (as described in chapter 3) generated 54597 sequences for the Zurich samples and the largest collection of any single barcode (082, 29678 sequences) in the parallel sequencing run targeting *T. trichiura* ITS1. Two of the samples generated about equal amounts of sequence (22121 and 29678 sequences, respectively) with the two other samples generating considerably fewer sequences

(2495 and 303 sequences, respectively). Analysis of the sequence collection of a neighbour-joining tree identified all sequences obtained to be *T. trichiura* ITS1. However, within the *T. trichiura* clade there was no subdivision visible and the sequences will require comparison with those created from other sites (Figure 6.6).

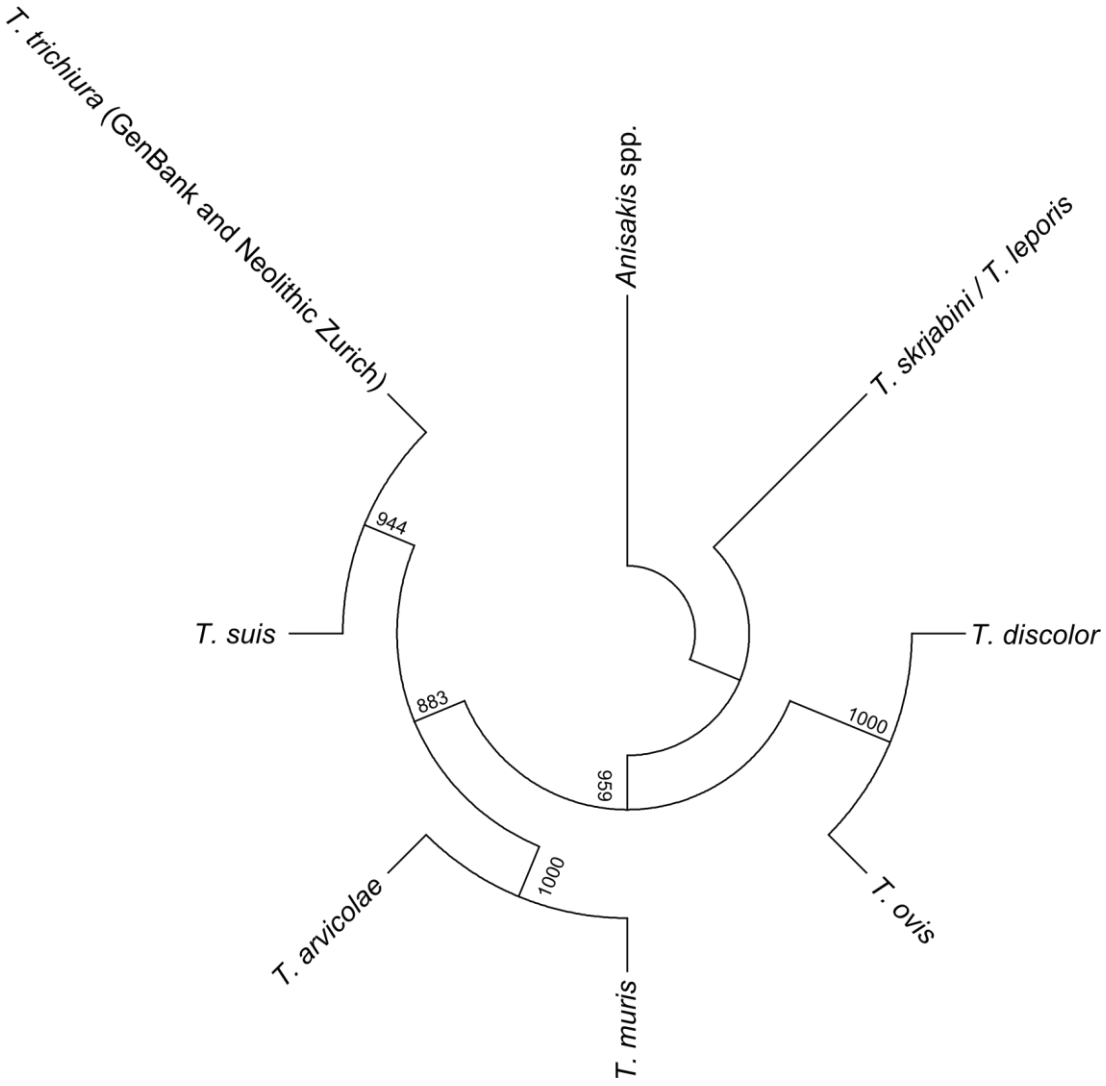


Figure 6.6: Bootstrapped Neighbour-Joining tree of *Trichuris* sp. ITS1 sequences from samples of Neolithic Zurich and modern database sequences. The sequences were compared to modern *Trichuris* spp. (using *Anisakis* spp. as outgroup) to confirm the identity of the sequences. All ancient sequences were located in the *T. trichiura* clade. This clade did not exhibit a well-defined sub-structure.

6.5 Discussion

The Neolithic samples contained a wide range of interesting samples, including the oldest helminth sequence available. Some of the oldest burials excavated in Germany revealed an unusually high prevalence rate for *Ascaris* (7/7 samples), but *Trichuris* was only found in one grave. However, the incidence rate of 14.3 % (1/7) was not far off other samples of later graveyards were prevalence rates of about 20 % were found (see chapter 7). The unusually high prevalence of *Ascaris* could be a coincidence caused by a non-random subset of the population. Due to the small sample set this cannot be evaluated, but evidence from the sample collection from Pohansko could explain these findings. The 102 samples were received in batches of about 20; in the first batch there were no samples positive for *Ascaris*, but 11 were positive for *Trichuris,* this pattern was then reversed in the third batch where 7 samples were positive for *Ascaris* but only one for *Trichuris* (detailed analysis see chapter 7).

The oldest sequences described in this project were dated to 3630 BCE, which are currently the oldest available sequences of any pathogen. Conditions which were favourable for the preservation of biological material in archaeology such as waterlogged sites have been shown to be suitable conditions for the preservation of parasite remains. The stilt settlement site samples from Switzerland and Germany were obtained from fully or partially waterlogged contexts.

Sequence analysis of the samples from Zurich confirmed the microscopic diagnosis of the parasite species. For *Ascaris* sp. cytochrome *b* the variation in the sequence was evaluated and visualised in a logo graph. The sequence was of sufficient length to indentify the species, but for speculations on species identity it was probably not sufficient. Further, only a limited number of samples and sites

produced this sequence and further work will be necessary to generate more sequence to investigate the possibilities.

Sequences generated for *T. trichiura* ITS1 confirmed the species identity, but as the sequences were generated from four of thirteen samples the variations between them were small (based on a neighbour-joining tree). To elucidate the variations of the *T. trichiura* ITS1 fragment further sequence analysis of a larger collection of samples is required.

Chapter 7

Slavic - Anglo-Saxon – Viking interactions

and considerations on graveyards as an

approach for individual based epidemiology

7.1 Introduction

The work presented in this chapter focused on the collection of samples dated between the 8th and 11th century CE, with considerations on additional analysis of samples derived from single graves.

Northern Europe between the 8th and 11th century CE was influenced heavily by the Vikings. Although commonly associated with their raids, they founded many cities in Scandinavia and the British Isles (Richards 2005). This project contains samples from five Viking sites: Viborg Søndersø (DK, four latrine samples, received from Annie Moltsen, (Fruergaard and Moltsen 2005; Hjermind et al. 1998)), Copenhagen (DK, 9 latrine samples, received from Peter Nejsum, Copenhagen), Langeland Bogøvej (DK, 29 graveyard samples, received from Pia Bennike (Grøn et al. 1994)), Gokstad Nedre Sandesfjord (NO, 28 communal deposit samples, received from Christian Løchsen Rødsrud and Jan Bill, Oslo) and York (UK, one coprolite sample). No parasites were diagnosed by microscopy in the samples from Langeland Bogøvej, therefore these were excluded from further analysis.

On the British Isles the population consisted of different groups, which are generally referred to as Anglo-Saxons and sometimes Anglo-Scandinavians. The sample collection described here included further samples from communal deposits of Anglo-Saxon settlement in Ipswich (16 settlement samples, received from Rebecca Nicholson, Oxford Archaeology) and Anglo-Scandinavian York-Hungate (one cesspit sample, received from Andrew Jones, York Archaeological Trust).

The third group of samples described here consisted of a large collection (102 samples) from the Slavic stronghold of Břeclav-Pohansko in the Czech Republic. This was a collection of grave sediment samples. The practice of human burials is documented for several millennia. Why humans started to bury their dead

is unknown, but with the establishment of stable settlements with farming instead of hunter-gatherer civilisations the number of burials increases. Graveyards are frequently excavated and the health state of the population is generally assessed using biological factors such as osteology and stable isotope analysis. Generally, the number of samples available from graveyards is considerably larger than for samples of other origin, especially when the graveyards are located in an obvious place (e.g. in a churchyard). This allows the evaluation of a large number of samples at the level of individuals.

The samples from the UK could provide evidence on links between Anglo-Saxons and Vikings on the British Isles. The collection of Viking samples could reveal interaction between the different tribes and might give insight into the origin of the Vikings in Britain which are called 'þa Deniscan' (the Danes) in Anglo-Saxon chronicles.

7.2 Břeclav-Pohansko – a Great Moravian stronghold

The largest collection of single grave samples (102 samples) was obtained from an excavation in Břeclav-Pohansko in the Czech Republic. This elite settlement was established around 850 CE and in use until the destruction around 950 CE (Macháček and Ungerman 2011). The excavations on this site have been ongoing for over 60 years and were conducted by the Masaryk University in Brno (CZ). The samples were taken from the graves in and surrounding the smaller church ‘rotunda’ of the settlement (excavation 2008-2012, see Figure 7.1, modified from (Macháček and Ungerman 2011)). Of the 152 graves excavated, samples from 102 graves were received. The buried population most likely belonged to the court of a bailiff (economic trustee) of the stronghold, with the smaller church as part of his court. The

graves had less grave goods than other graves within the stronghold and the majority of the graves found were children (94/152, personal communication on unpublished material from Renáta Přichystalová, Masaryk University Brno, Czech Republic).

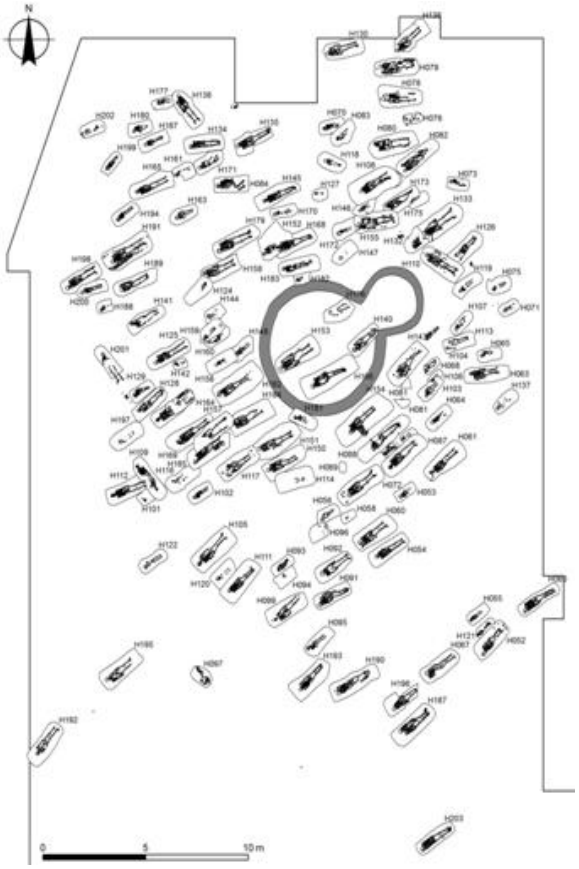


Figure 7.1: Rotunda church with graves arranged around.

152 graves were excavated around the church (indicated in grey), each graves carried a unique identifier number. Image courtesy of Renáta Přichystalová, Brno

7.2.1 Microscopic analysis

Sixty-one were processed for microscopic and molecular analysis with 24 (40 %) positive for *Ascaris* sp. and 7 (11.7 %) for *Trichuris* sp., with 6 positive for other parasite-like structures (possibly *Fasciola hepatica*). Overall 29 (48.3 %) samples were positive for at least one parasite and 6 with more than one parasite.

The graveyard from Břeclav-Pohansko was the largest sample set analysed in this study. Large single grave sets like Břeclav-Pohansko confirmed the prevalence rate of the parasites. The infection rate in Pohansko for *Ascaris* was comparable to a study in Uganda where 42.5 % of mothers and 30.4 % of their children were found to be infected (Betson et al. 2012). The incidence of *Ascaris* in South America were reported to be 15.6 % and for *Trichuris* the prevalence was 12.5 % (Chammartin et al. 2013). Higher prevalence rates of 95.6 % for *Trichuris* and 47.8 % for *Ascaris*, with overall prevalence rates of 98.4 % in Malaysian schoolchildren were published recently (Al-Delaimy et al. 2014). Although the possibility of false-negatives cannot be excluded in the ancient samples, the observed prevalence rates are comparable with a modern population without access to modern pharmaceutical treatment or a fully functional sanitation system. The impact of medical treatment and/or a functional sanitation system is tremendous and can reduce infection and prevalence close to zero (Ziegelbauer et al. 2012).

In a modern population increased levels of helminth infection is generally associated with poverty, underdeveloped infrastructure or the lack of access to modern medication. To assess whether social status may affect the helminth load in ancient samples, the level of infection was correlated with social status as judged by grave good and positions within the graveyard. In collaboration with the excavators at the Masaryk University in Brno a comparison map was built based on 42 randomly chosen graves (Figure 7.2). The map in Figure 7.2 (a) indicates the social status of the buried individual by position of the grave and grave goods. The samples were from high status individuals buried inside the church (red), close to the church in preferred areas with grave goods (green) and without grave goods (blue), in non-preferred areas with grave goods (yellow) and without grave goods (grey). The map

in Figure 7.2 (b) indicates the presence or absence of *Ascaris* (green), *Trichuris* (red), and other unidentified parasites (yellow). Parasites were found in all groups. Together with the finds from communal deposits in the wealthy parts of Lübeck (see Chapter 5) this strengthens the argumentation that infections are independent of social status or wealth and were unlikely to have caused social stigmatisation.

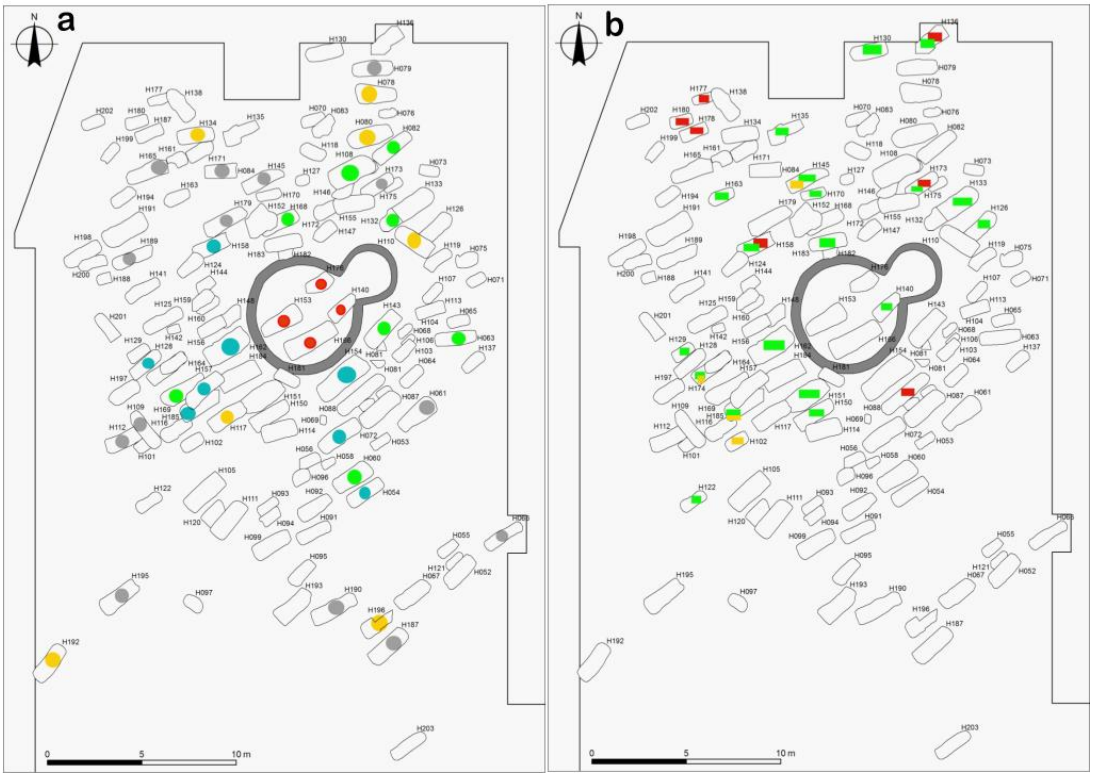


Figure 7.2: Map of location of graves with indicator of social status (a) and parasite infection (b).

The social status (a) was evaluated by the position of the grave relative to the church and the presence of grave goods. The parasite infections (b) were mapped according to species. Image courtesy of Renáta Přichystalová, Brno

7.3 Comparison of microscopic analysis – burials and settlements

The number of burials positive for any parasite was as expected lower than the number of samples from other deposits. The number of parasites found was also lower in burial samples compared to deposits or latrines. To estimate the differences between the different deposit types (latrines, other communal deposits and graves), the highest values of parasite counts were compared. The highest values for latrines and communal deposits were comparable (12,034 *Trichuris* eggs per gram and 1,645 *Ascaris* eggs per gram in latrines to 9,209 *Trichuris* eggs per gram and 1,162 *Ascaris* eggs per gram in a waste deposit), but the values found in graves was considerably lower (435 *Trichuris* eggs per gram and 303 *Ascaris* eggs per gram, Figure 7.3).

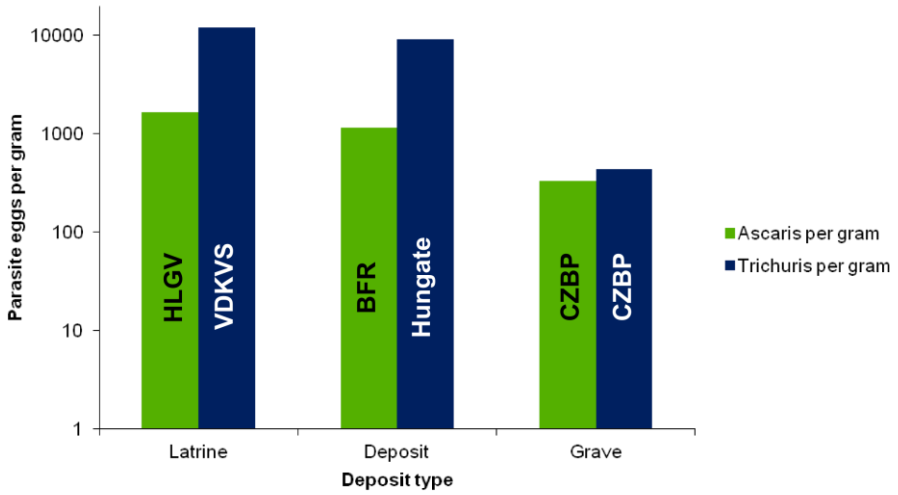


Figure 7.3: Comparison of highest parasite counts from different archaeological deposits.

The differences between the highest values in latrines and communal deposits were not very different, yet the graves were considerably lower. The highest values were found in Lübeck (HLGV, Medieval port), Viborg (VDKVS, Viking latrine), Bristol (BFR, Medieval port), Hungate (Anglo-Saxon settlement deposit), and Pohansko (CZBP, early Slavic graveyard). The y-axis is shown on a logarithmic scale.

7.4 Molecular analysis of sample collection

A range of sequences were used to analyse the samples. Sequences were produced from *T. trichiura* ITS1, *Ascaris* cytochrome *b* and CO1 and *Cryptosporidium* 60 kDa glycoprotein (gp60) gene. The results were summarised in Table 7.1.

The results for *Trichuris* ITS1 will be discussed in detail in context with all sequences acquired from all sample collections (chapter 8.2). The results for *Ascaris* CO1 were discussed in chapter 2 (Figure 2.12).

Site code/ Sequence	CZBP	VDKKH	VDKVS	VNOGNS	VY	Hg	UKIPS
T. trichiura beta-tubulin	yes	yes	yes	no	no	no	no
T. trichiura ITS1	yes	yes	yes	yes	yes	no	no
Ascaris sp. Cytochrome *b*	yes	no	no	no	no	no	no
Ascaris sp. CO1	no	yes	no	no	yes	yes	no
Cryptosporidium sp. gp60	yes	no	no	no	no	no	no

Table 7.1: Summary of sequences obtained from sample collection

Neither *Cryptosporidium* gp60 nor *Ascaris* cytochrome *b* allowed comparison within the contemporary group presented here. Sequences for *Cryptosporidium* gp60 were only obtained from one grave in Pohansko and one of the Neolithic Zurich deposit samples. The sequencing trials resulted in a large proportion of environmental (algae) sequences which made the acquisition of a comprehensive

sequence collection very challenging. The development of a parallel sequencing approach (as described in chapter 3) has been initiated to assess the potential to overcome the limitations of the conventional approach and to enable the acquisition of more sequences in the near future.

Ascaris cytochrome *b* sequences were obtained from four graves in Pohansko (grave numbers 2730, 3343, 3219 and 2777) and six samples from Zurich (11 sequences). Analysis of the sequences on the bootstrapped neighbour-joining tree (Figure 7.4) did not reveal any connection, but confirmed the species identity.

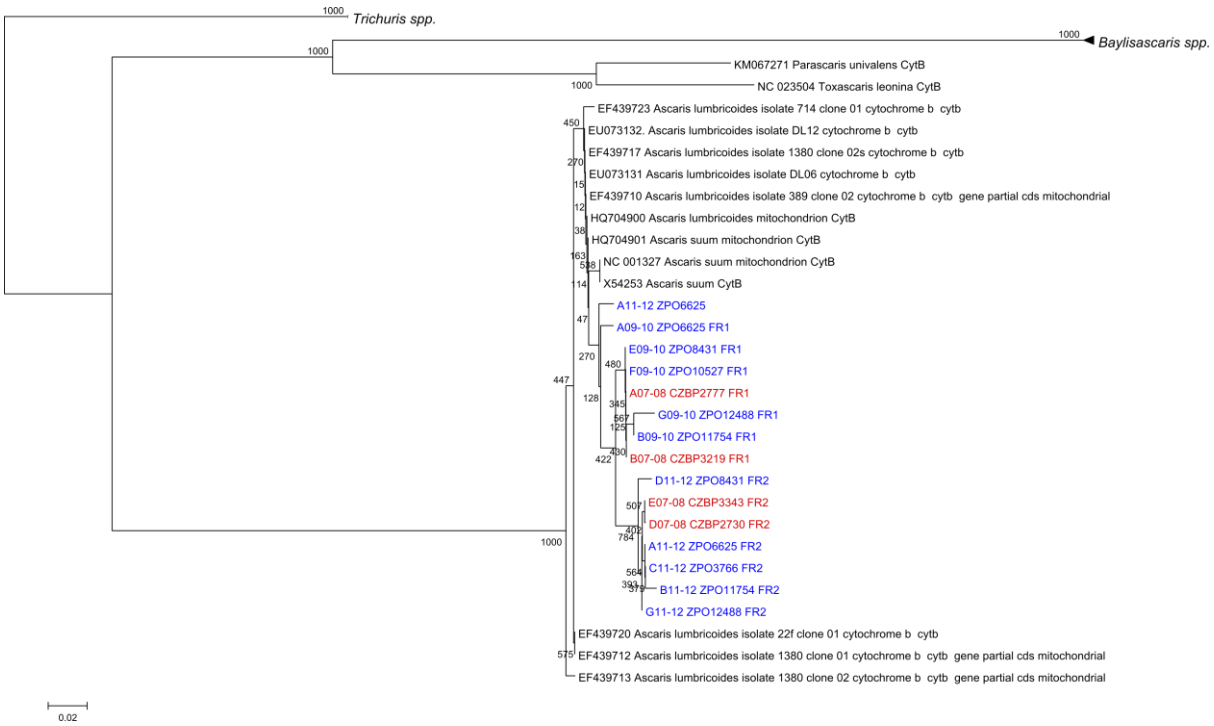


Figure 7.4: Bootstrapped Neighbour-Joining tree for *Ascaris* sp. cytochrome *b* sequences from Pohansko and Zurich.

Sequences highlighted were obtained from Pohansko (red) and Zurich (blue). The position within the *Ascaris* spp. clade confirms the species of the sequences. Other species on the tree were other Ascaridians (*Baylisascaris* spp, *Parascaris univalens* and *Toxascaris leonia*) and *Trichuris* sp.

7.5 Trichuris β-tubulin – indications of a link?

The *β*–tubulin gene is considered a relatively stable genomic region as it has an essential function in the cytoskeleton structure. The primers were designed based on published primers generating a 163 bp product (Diawara et al. 2009). Unfortunately, the number of generated products was limited due to mispriming on non-target sequences and primer dimerisation. The *β*–tubulin sequences generated were mainly obtained from the samples from Pohansko and samples from Viking settlements in Denmark. The comparison of the sequences from five graves from Pohansko (red) and the five samples from the Viking settlements of Viborg (three sequences) and Copenhagen (two sequences, blue) on a neighbour-joining tree suggested an interaction of the Danish Viking population with the Slavic settlement in Pohansko (Figure 7.5). The linkage is speculative and based on the sequence identity, but the majority of the samples fall into one clade, which is most likely *T. trichiura*. It is quite likely that the Danish Vikings were travelling by rivers to Pohansko (via the Elbe to the South-East) or that trade was established via the Rus who where reigning extensive territories in today's Russia (named after the Rus), Belarus and Ukraine. Both theories would explain the unilateral exchange from Denmark to Pohansko. Further, of two similar Viking latrines, only one shows the linkage to Pohansko, suggesting that these were separate tribes with different interactions to other places.

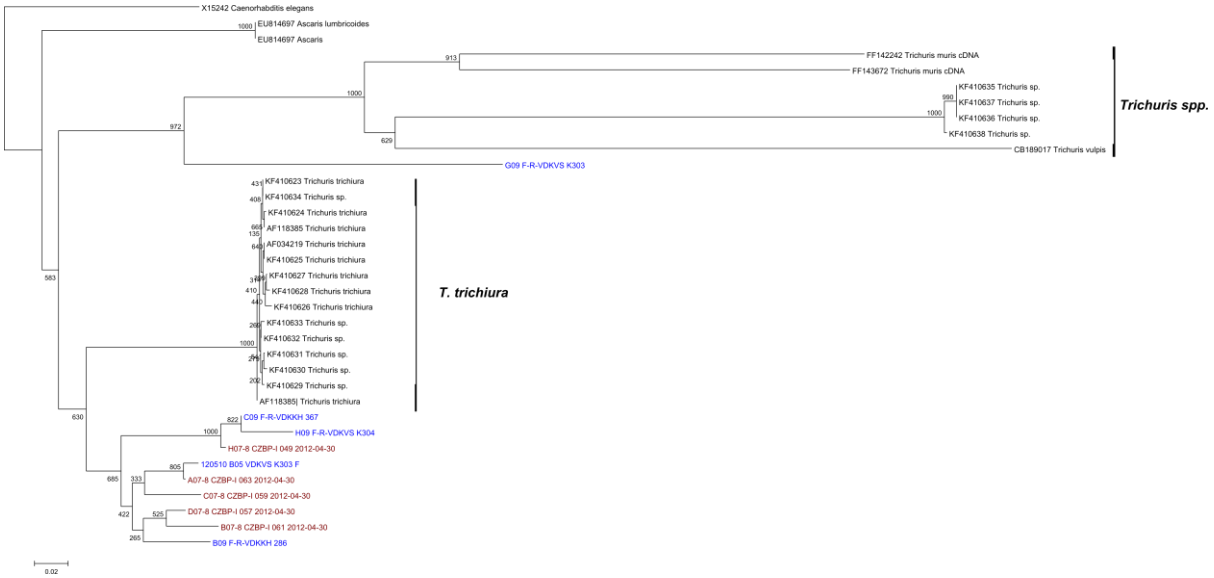


Figure 7.5: Bootstrapped Neighbour-Joining tree showing *T. trichiura beta*-tubulin sequences from Viking settlements in Denmark and the graveyard of Pohansko.

Sequences from the Slavic settlement of Pohansko (red) were compared with sequences from the Viking settlements (blue) in Viborg and Copenhagen, Denmark, against modern sequences from GenBank. Besides *T. trichiura*, *T. muris*, and *T. vulpis* for the identification of the species, two nematode outgroups (*Ascaris* spp. and *C. elegans*) were added to the tree.

7.6 Parallel sequencing – insight into kinship or epidemiology?

Parallel sequencing techniques were used to generate larger quantities of sequences from a variety of targets. A first insight into the matter was gained based on the *T. trichiura* ITS1 parallel sequencing data which contained sequences from six samples from Pohansko (CZBP 3421, 3795, 2857, 3869, and two independent PCR products of 3752). In total, 7003 sequences were obtained, of which 4035 (57.6 %) sequences were present in more than a single readout. The sequence clusters were evaluated to determine whether there were groups of clusters across different samples, which would suggest a link between these two. The parallel sequencing technique detects both high and low-frequency sequences through a high sequence

output. Sequences from five samples from Pohansko were obtained using parallel sequencing techniques. The graves from which sequences could be generated are highlighted in Figure 7.6 (sample numbers 3752, 3795, 3421, 3867, and 2857). For this preliminary analysis, the most frequent sequence from each sample from the MiSeq parallel sequencing run was compared to the sequences from the other samples. The most frequent sequence of samples 3795, 3752 and 3421 were identical, as were the most frequent ones from samples 2857 and 3869. These suggested that the distribution of the most frequent parasite sequence was not random, but seem to follow an east-west divide.

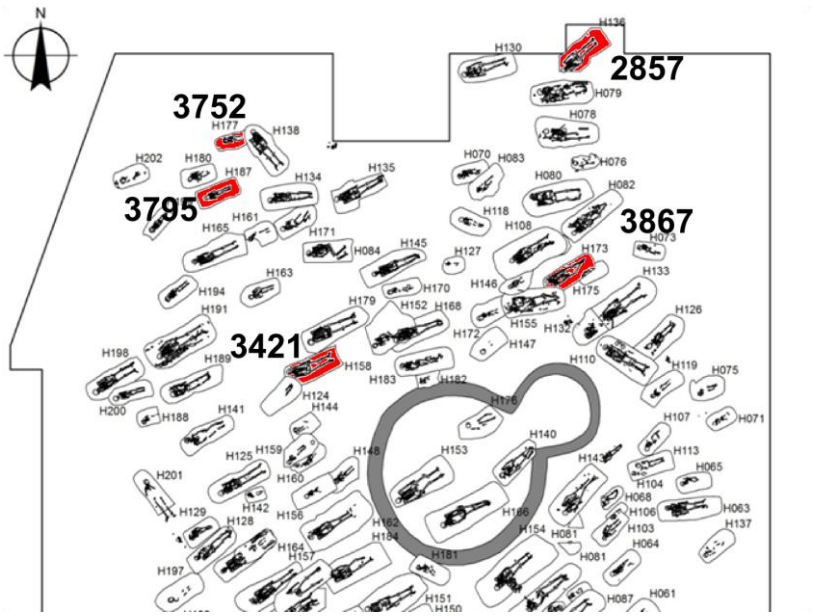


Figure 7.6: North section of the excavation map with graves highlighted from which sequences for *T. trichiura* ITS1 were generated by MiSeq parallel sequencing.

All graves were found north of the chapel (in grey). Image courtesy of Renáta Přichystalová, Brno

7.7 Medieval cemeteries in Europe

Samples from three medieval cemeteries were obtained and analysed. The first set was obtained from Brno-Vídeňská in the Czech Republic (CZVB). Samples from 14 graves (two per grave) were obtained dated to the 14th century CE. The context of this graveyard was not entirely clear as it was a rescue excavation, it might have belonged to a smaller church, but this could not be confirmed (personal communication Petra Urbanova, Masaryk University Brno). A second, larger sample set was received from Ellwangen-Jagst in Southern Germany (DEEJ). This excavation was part of a larger excavation of the graveyard surrounding the church of the market town. Of this excavation, 28 samples have been evaluated, and further 92 samples were obtained at date of submission of this thesis for parasitological research. The third and largest set was obtained from the excavation of Stoke Quay in Ipswich, England (UKIPS). In contrast to the other two land-locked cities, Ipswich is connected to the sea by the river Orwell. This is of particular interest to compare ports to land-locked market towns. The samples from Ipswich were, like the ones from Ellwangen and probably those from Brno, connected to a church.

The incidence rates in these three sites were similar to the ones found in Pohansko, although these samples were several centuries later than the ones from Pohansko. The incidence rates are depicted in Figure 7.7, indicating the rates for infections with *Ascaris* and *Trichuris*, as well was the incidence of at least one nematode or both nematodes. The incidence rates for *Ascaris* were similar in all three sets (28.6 % in CZVB, 21.8 % in UKIPS, 21.4 % in DEEJ), but considerably lower than in the samples from Pohansko (40 %). On the other hand, the rates for *Trichuris* were higher in Brno and Ellwangen (42.9 % and 32.4 %, respectively) than in Pohansko (11.7 %), whereas the observed rate of *Trichuris* in Ipswich was very

low (1.8 %, one grave). The incidence rate for at least one nematode parasite detected was slightly higher in Brno (50 %) than in Pohansko (48.7 %), but lower in both Ellwangen (35.7 %) and Ipswich (21.8 %). Variation of incidence rates has been reported in modern populations, even without pharmaceutical intervention (Adams et al. 1994; Olsen et al. 2009; Shapiro et al. 2005; Woodburn et al. 2009).

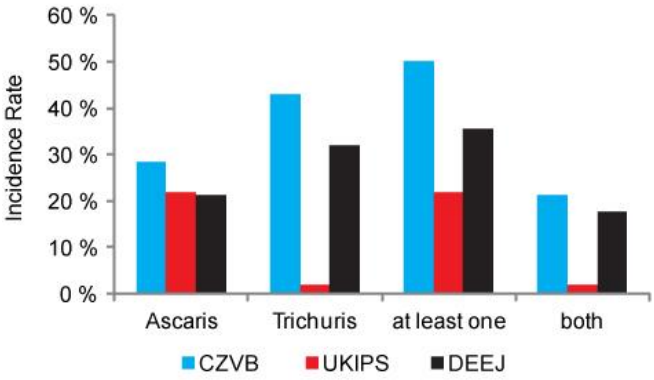


Figure 7.7: Incidence rates for parasites on medieval graveyards in Europe. Samples were obtained from Brno-Vídeňská in the Czech Republic (CZVB, blue), Ipswich in England (UKIPS, red) and Ellwangen-Jagst in Germany (DEEJ, black), indicating the incidence rates for *Ascaris* and *Trichuris*, as well as the overall incidence rate (either *Trichuris* or *Ascaris*) and the rate for infection with both nematode parasites.

7.8 Discussion

This chapter presented the diverse nature of samples which were found useful to the approach to use parasite genomic material to gain insight into connections between human populations, especially when written sources are scarce such as for the interactions between the Vikings and other contemporary societies such as the Anglo-Saxons or the Slaves, or even between Vikings of different origin.

The major part of this chapter discussed the possibilities and limitations of samples which were derived from graveyards. Such samples were generally obtained

in considerably large sample collections as by-product of excavations of graveyards. The major issue with these samples was the rate of negative samples (only about 20-30 % positive) and the relatively low quantities of parasite eggs per sample (Figure 7.3). The differences between graveyards seemed to be, as with the communal deposit samples, mostly dependent on the location and the sampling. The sampling of soil for parasitological analysis from a graveyard is more challenging as the parasites are expected to remain in a relatively restricted area between the pelvic bone and the rib cage (remains of the inner organs).

The sequences obtained from *Trichuris* β-tubulin from five graves from Pohansko and five settlement samples from Viking Denmark suggested a link between the two. As the number of sequences is limited this will be discussed further in chapter 9 to put the molecular data into a broader context.

However, especially the successful application of parallel sequencing on samples derived from graves revealed their full potential for in-depth analysis on an individual-based sample collection. This might provide insight into palaeoepidemiology or kinship between individuals on a burial ground. The analysis of the parallel sequencing data of the graves from Pohansko (Figure 7.6) was just the beginning of this kind of analysis. Due to the very limited sample size in this analysis the validity of the conclusions remain speculative, but provide some intriguing insight into what will be possible based on a large-scale analysis of grave-derived samples.

Results presented in this chapter are very promising and demonstrated the potential for the application of archaeoparasitological research on various fields, such as links between populations, individual-based palaeoepidemiology or even kinship evaluations.

Chapter 8

Global analysis and discussion of molecular data

8.1 Introduction – issues and expected outcome of a global analysis

To evaluate the power of the approach all sequences generated for one target were aligned and larger neighbour-joining trees were constructed. This allowed for a broader view on the context of all the sites in the study. Despite having a large collection of samples it clearly was not a complete set to represent all of European history. Therefore, the nature of the samples could influence the conclusions drawn based on the dataset collected. This chapter provides an insight into the potential of the approach based entirely on the aDNA sequences. At this point of the analysis the data is interrogated based solely on the genetic variation, the age of the parasite sequence had not been taken into account when building the alignments and the neighbour-joining trees. This evaluation provides a first glimpse onto the relationship between the samples and the sites, but a more comprehensive analysis will be required in future work.

In this chapter, the collected data for two sequence collections will be presented, the sequences of *T. trichiura* ITS1 and *β*-tubulin. These sets consist of the most extensive sample collection and showed thus the largest potential for the analysis.

8.2 Trichuris trichiura ITS1

The largest collection of sequences was generated from *T. trichiura* ITS1. In modern infections, ITS1 has been used extensively for species identification and thus a considerable number of modern sequences were available in public databases. As ITS1 is a variable non-coding sequence, it has the potential to be most informative as considerable variation could be expected between sites. This is the largest collection of ancient parasite sequences and it offers a unique opportunity to interrogate the evolution of the parasite as these samples were dated. At this point, the evaluation will be based on the observations of the topology of the tree and the diversity of the sequences when compared between sets.

The sequences were generated from eight sites, spanning about 5000 years, 744 sequences (of which 23.4 % were produced by Sanger sequencing) were generated. The largest set of samples was acquired from the Neolithic settlement in Zurich (ZPO), which consisted of 349 (46.9 %) of sequences. The sets from the medieval ports of Lübeck (HLGV) and Bristol (BFR) consisted of 170 (22.8 %) and 164 (22.0 %) sequences, respectively. Further, sequences from Pohansko (CZBP, 48 sequences/6.4 %), Ellwangen (DEEJ, 9 sequences/1.2 %) and Viking York (VY, 4 sequences/0.5 %) were included.

Based on the tree topology (Figure 8.1) there is a clear division of the *T. trichiura* clade into two separate clades, one of which consists nearly exclusively of samples from Lübeck, including only three sequences from two samples from Bristol (clade 1, circled in Figure 8.1) whereas the other clade contains all other sequences irrespective of origin or age, including the modern reference sequences (clade 2), although all sequences were found to be closer to *T. Trichiura* than to other species (Figure 8.1 a). There are several possible explanations for such a pattern. The in our opinion most likely theory is that the Lübeck sequences fall into a *T. trichiura* variant which was prevalent at the time, but is no longer prevalent in modern samples. This could be due to geographical signature where these sequences represent a 'Northern European' variant or to the absence of matching modern sequences. The argument in favour of this theory is the restricted distribution of the genetic signature. The setting in Lübeck, with latrines of individual houses with a restricted group of users, would maintain such as population as it would facilitate re-infections between users of the same facility.

The more diverse clade 2 contained 645 (of which 28 were modern database sequences) sequences, which corresponds to 83.2 % of all *T. trichiura* and 86.4 % of ancient sequences. The part of the phylogenetic tree containing all *T. trichiura* sequences is displayed enlarged in Figure 8.1 b.

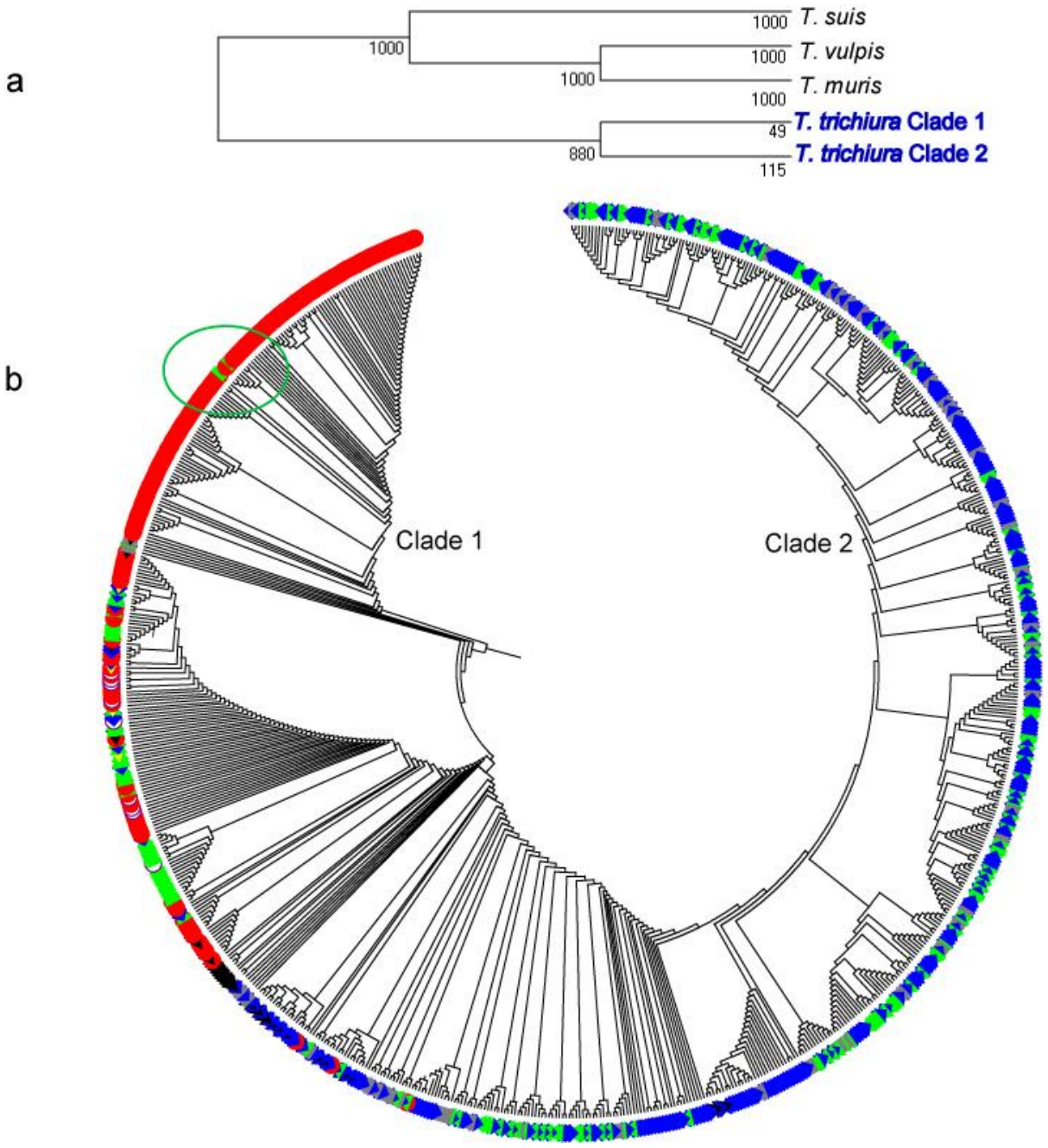


Figure 8.1: Bootstrapped neighbour-joining tree for *Trichuris* ITS1.

All ancient sequences were found to be within the *T. trichiura* clade (a). The clade containing *T. trichiura* is displayed in more detail (b). All sites are colour-coded: Lübeck (Medieval, red), Bristol (Medieval, green), Ellwangen (Medieval, purple),York (Viking, yellow), Pohansko (early Slavic, grey) Zurich (Neolithic, blue), GenBank reference sequences (black). The divide between the two clades is clearly visible. The three sequences from Bristol are circled.

To evaluate the nucleotide frequency, a stretch of 424 bp which was covered in the parallel sequencing run was analysed. 225 bp (53.1 %) were conserved and 93 bp (21.9 %) were different in only one sequence (Singleton), which left 106 bp (25.0 %) for the analysis, which were extracted to analyse the sequence variations. Most polymorphisms were relatively rare with a <10 % frequency (100/106) and five of the more frequent polymorphisms (in 16 % of the sequences) were only found in sequences in clade 1. The sixth one of these more frequent polymorphisms was found in various sequences from all sites where MiSeq data was available (all except Ellwangen and York), suggesting it was a ubiquitous polymorphism which was found at a relatively low frequency in all populations (Figure 8.1).

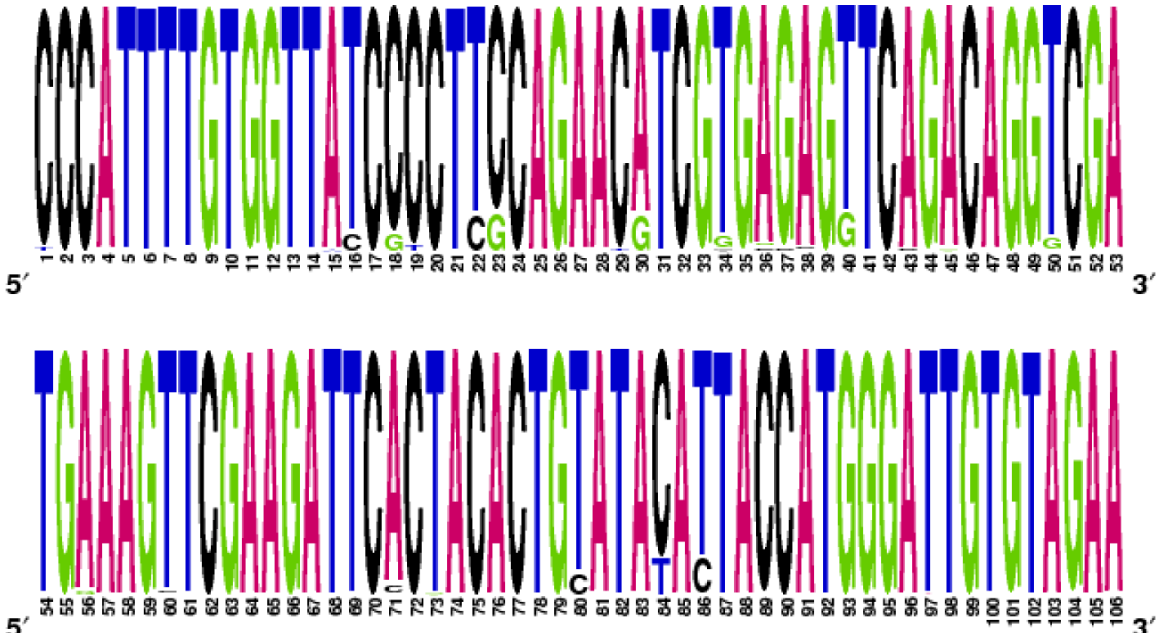


Figure 8.1: Sequence logo plot visualising the SNPs from the sequence.

Only the bases which were found to be different in more than one occasion were displayed. Figure generated using WebLogo (weblogo.berkeley.edu)

It has been published the most frequent exchange in aDNA due to age is the deamination of cytosine into uracil (C->U), which leads then to mis-priming in the PCR to be a U-A annealing (Hoss et al. 1996; Paabo 1989; Willerslev and Cooper 2005). In the

compared sequences the C-T change is relatively frequent (15.1 %), but not the most frequent (Figure 8.2). This suggested that the degradation of aDNA has an effect, but as it was not the most frequent it could be assumed that the sequences represent true different genotypes. However, for most cases (15/16) it was assumed that the deamination of cytosine was a random age-related effect as it was found in less than 3 % (up to 20/744) of the sequences, whereas the other incidence was very frequent and associated with clade 1 as described. This shows that the unspecific deamination of cytosine accounts only for a few of the changes in the sequence, and it provided further evidence for the divide of the samples into two clades.

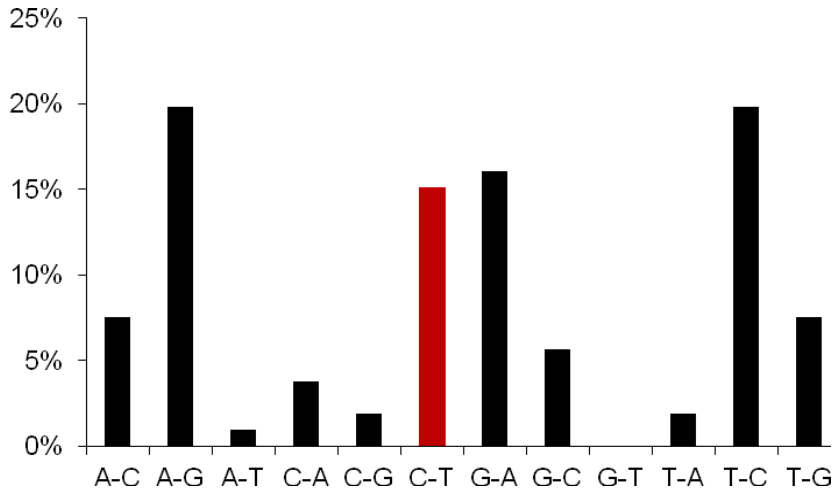


Figure 8.2: Frequencies of base exchanges in the *T. trichiura* ITS1 ancient sequences.
The deamination of cytosine described in literature is highlighted in red. Other changes are more frequent which suggested that these represent different genotypes.

To evaluate whether the frequency of C-T transitions increased over time, the frequencies were shown against age of sample on a graph (Figure 8.). No apparent correlation between age and an increased frequency of C-T transitions could be found. Notably, the graveyard samples (in Pohansko, CZBP and Ellwangen, DEEJ) as well as in the coprolite from York (VY) the rates of C-T transitions were notably lower (on average 9.2 %, 3.1 % and 5.2 %, respectively), which would indicate that at least a proportion of the C-T transitions were not associated with random decay.

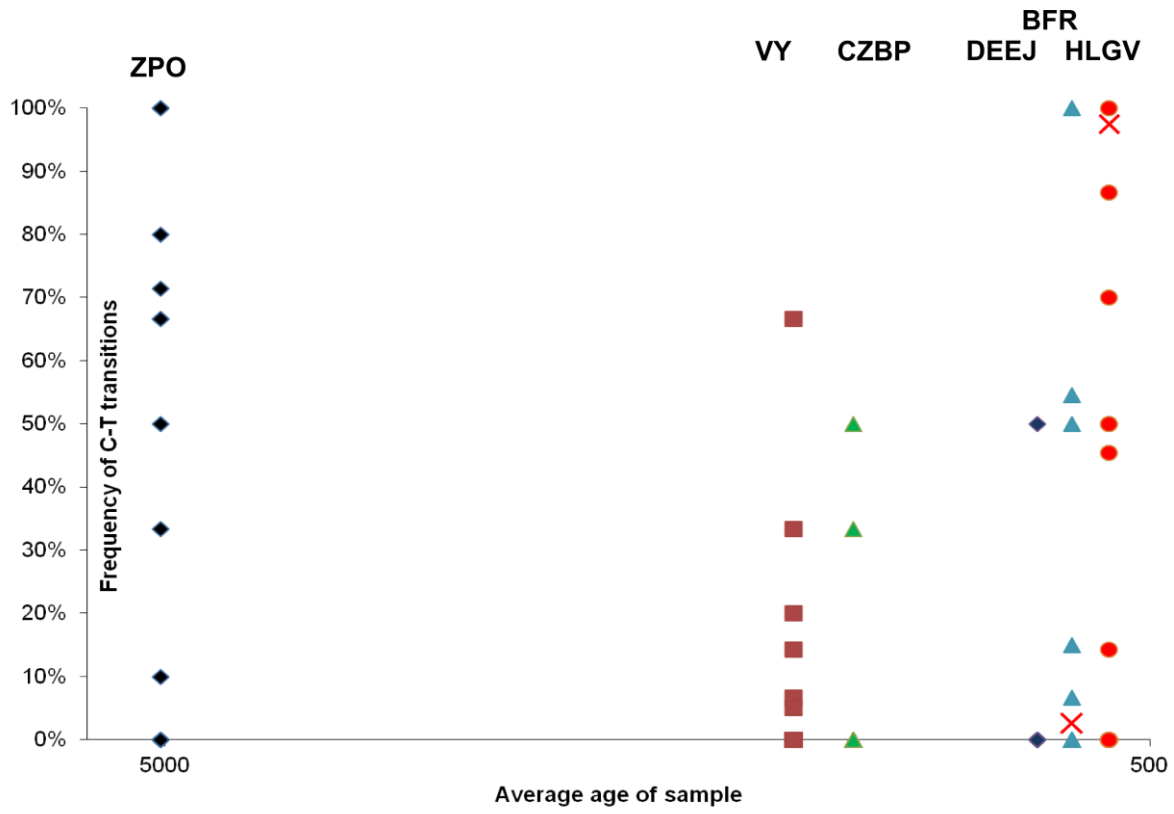


Figure 8.4: Frequencies of C-T transitions correlated with age.

The frequencies were taken from the frequencies on each of the C-T transition sites. The two values indicated with a cross (in Lübeck and Bristol) were defining for their positions in clade 1 of the tree.

Besides the changes defining clade 1, no other site-specific changes were found. This was not unexpected as more than a single change would be necessary to define a sequence to site. The evaluation showed that the differentiation on the tree could be traced to sequence changes. Further, the rate of deamination of cytosine causing a C-T transition might not only be due to age, but also due to the chemistry of the sediment. To fully evaluate this, more samples from more different contexts and ages would be required.

An influencing factor on the sequence diversity could also be the region of the polymorphisms on the amplicon. Figure 8.5 indicates the position of polymorphisms relative to the sequence length. Only polymorphic positions were indicated. The pattern indicates that the central section of the sequence exhibited higher levels of polymorphism and that these

were not randomly distributed. The high level of polymorphisms would also explain why one of the designed primers (ITS 4F, Figure 2.7) did not produce PCR products.

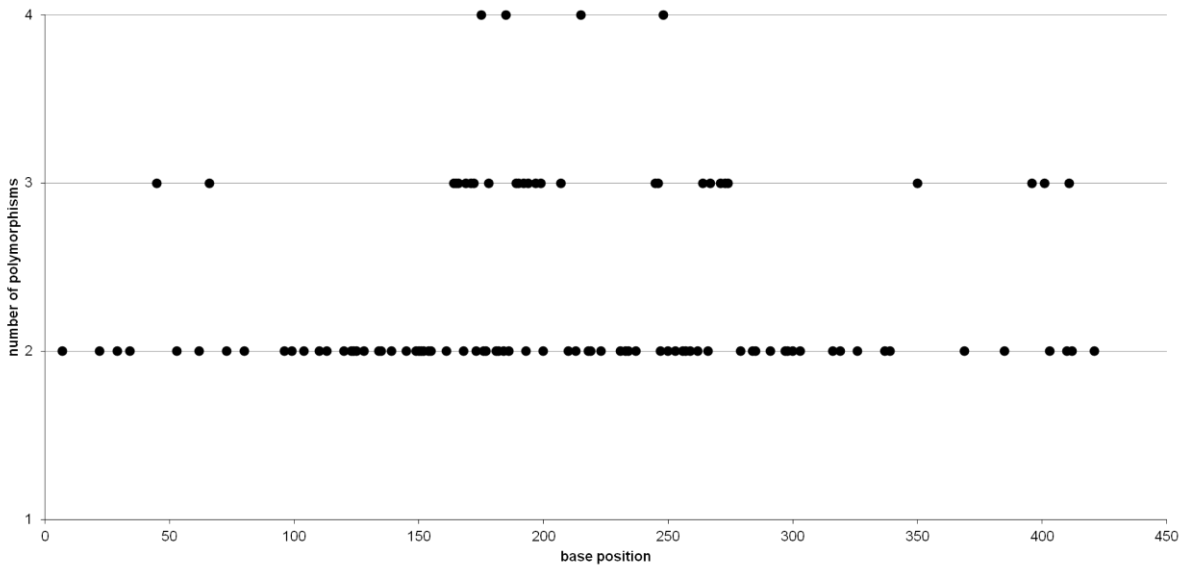


Figure 8.5: Number of polymorphisms relative to the base position.

The figure indicates the position of the polymorphism within the amplified sequence. Only sites with polymorphisms were indicated. The polymorphisms were not randomly distributed, but were focused in the centre section of the sequence.

8.3 Trichuris trichiura β-tubulin

A smaller collection of sequences was obtained from *T. trichiura β*-tubulin (Figure 8.5). Sequences were obtained from single graves in Pohansko, Viking settlements in Denmark, Bristol, Oxford, Zurich and Lübeck. The structure on the tree generated two clades, of which clade 1 consisted solely of samples from Lübeck and the modern *T. trichiura* sequences, whereas clade 2 consisted of samples from Lübeck and all other ancient sequences. In chapter 7, we speculated about a link between Viking Denmark and Pohansko in the Czech Republic. While this is still a valid explanation, the tree in Figure 8.6 strengthens the hypothesis of the two clades of which one is mainly represented in Lübeck. The tree in Figure 8.5 showed a mirrored image of the situation in *T. trichiura* ITS1 (Figure 8.1) regarding the modern sequences, but this is still consistent with the hypothesis.

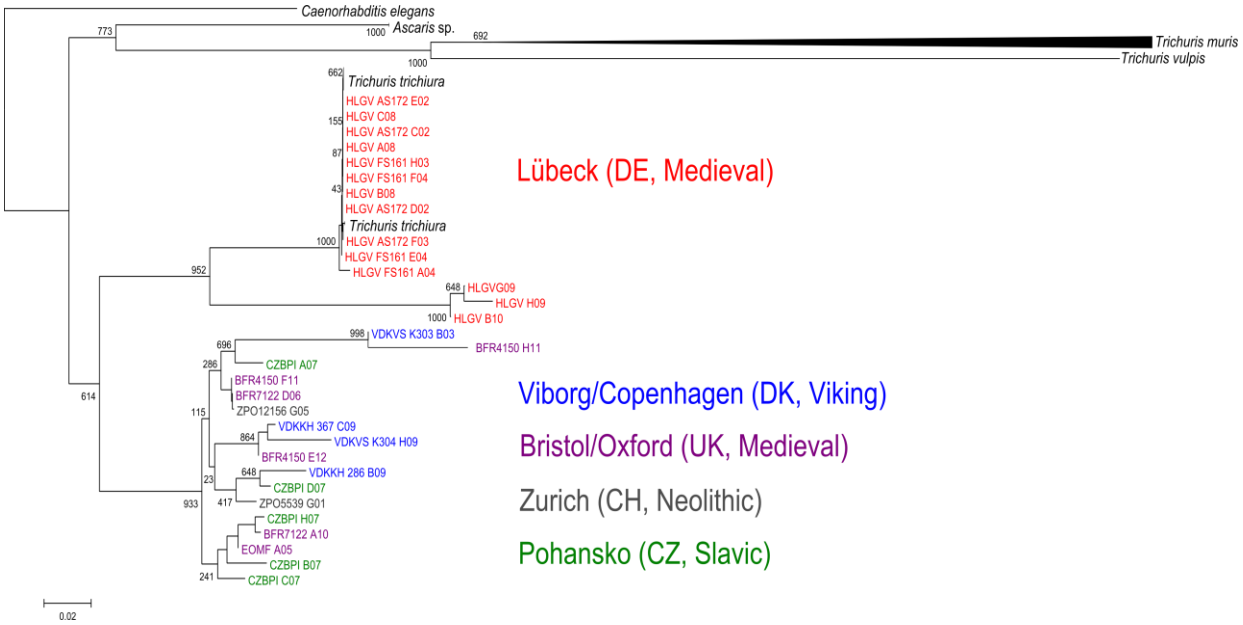


Figure 8.6: Bootstrapped Neighbour-joining tree of *T. trichiura* β-tubulin.

The generated sequences from Lübeck (red), the Danish Viking sites at Viborg and Copenhagen (blue), the Neolithic site in Zurich (grey), Medieval Bristol and Oxford (purple), as well as the single grave samples from Pohansko (green) were compared with modern *T. trichiura*, *T. muris*, and *T. suis* with *Ascaris* sp. and *C. elegans* as outgroups. The ancient samples associated with *T. trichiura*, forming two clades. Clade 1 consisting of the modern *T. trichiura* sequences and all sequences from Lübeck (red) and the second clade consisting of all remaining ancient samples

8.4 Alternative methods of phylogenetic analysis

The phylogenetic analysis of the ancient sequences was done using the Neighbour-Joining method (Saitou and Nei 1987). This is the most commonly used algorithm due to its simplicity and low computational requirements. The method calculates the divergence from all other species to evaluate which pair has the lowest distance from all other species to create a new node. This method was used to evaluate the data as it requires relatively low computational resources and still provides reliable data by bootstrapping the resulting trees. The bootstrap value provides a statistical measure of the likelihood of a node being accurate by re-sampling the available data (Lemey et al. 2009). However, the Neighbour-Joining algorithm might not produce the best possible model for the relationship between the

sequence species in the dataset. Therefore, one sub-set of samples was used to construct trees with other algorithms. The dataset chosen contained sequences of *T. trichiura* ITS1 from the medieval period and its neighbour-joining tree was constructed for Figure 5.10. The set was deemed to be suitable as it contained a good number of sequences which show a divergence into two clades. To confirm that this was not an artefact created by the neighbour-joining algorithm, the sequences were evaluated using the Maximum Parsimony and Maximum Likelihood algorithms available in the software MEGA5 (Tamura et al. 2011). The principle of Maximum Parsimony is to construct a tree which requires the minimal number of evolutionary events to explain diversity (Lemey et al. 2009). Maximum Likelihood is similar to Maximum Parsimony as it evaluates the number of changes necessary to explain the data, but it employs an estimation of probability of these events happening to refine the model (Lemey et al. 2009).

To evaluate the three methods, Neighbour-Joining, Maximum Parsimony and Maximum Likelihood, the same dataset was imported into the software. All trees were computed with a bootstrap value of 100. The sequences were generated from samples from Lübeck (red), Bristol (blue) and Ellwangen (green). For the tree building, modern *T. trichiura* sequences from the GenBank (black) and the three closest sister species (*T. muris*, *T. suis* and *T. vulpis*; three sequences each) were included. The Neighbour-Joining tree depicted in Figure 8.7 shows the same pattern as the one in Figure 5.10 with the divide into two clades of which one contains nearly exclusively samples from Lübeck. The support for the trees is relatively weak, but shows slight variation between the different tree-building algorithms. The strongest support was found using the Maximum-Likelihood algorithm (Figure 8.8), whereas the Maximum-Parsimony (Figure 8.9) showed the weakest support for the separation of the two clades.

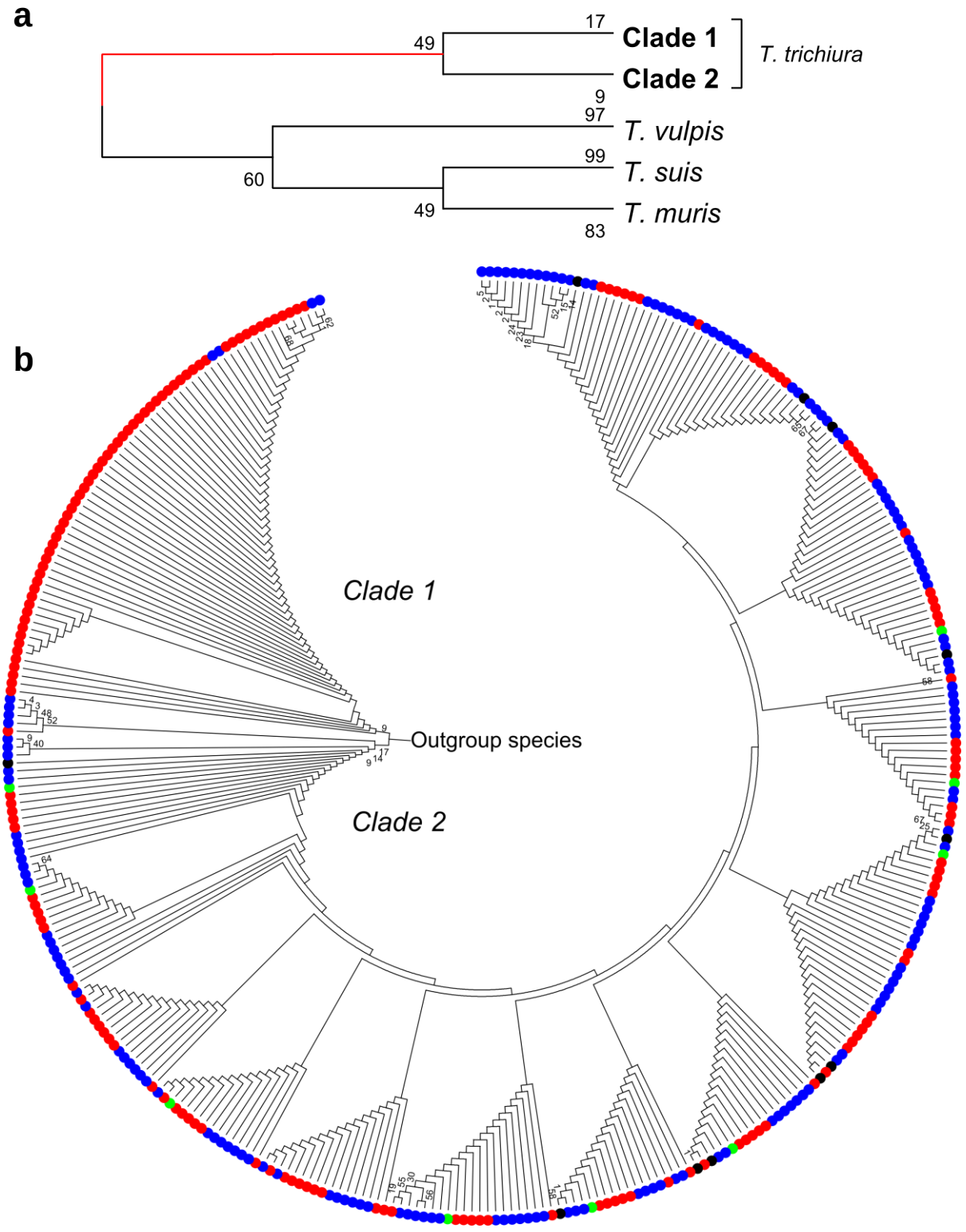


Figure 8.7: Bootstrapped Neighbour-Joining tree for Medieval *T. trichiura* ITS1 samples. Samples from three sites were included and are colour coded (Lübeck in red, Bristol in blue, Ellwangen in green and GenBank in black). The tree a is reduced to species level with the exception of *T. trichiura* where the two clades are shown, and tree b is rooted as tree a and cropped at the highlighted branch and is displayed showing topology only with bootstrap values of 0 removed for clarity.

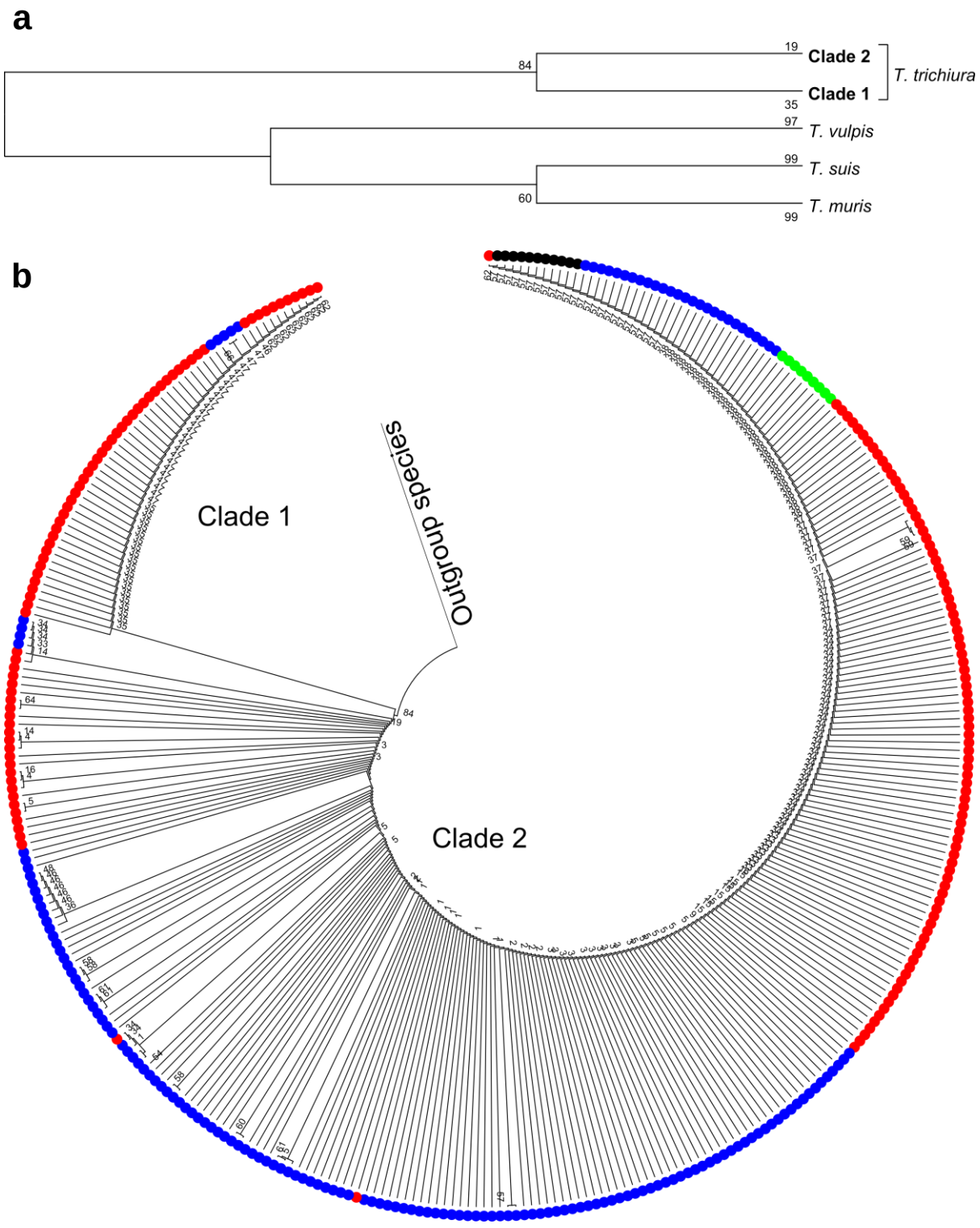


Figure 8.8: Bootstrapped Maximum Likelihood tree for Medieval *T. trichiura* ITS1 samples. Samples from three sites were included and are colour coded (Lübeck in red, Bristol in blue, Ellwangen in green and GenBank in black). The tree a is reduced to species level with the exception of *T. trichiura* where the two clades are shown. Tree b is rooted as tree a with the outgroup species collapsed to one branch as indicated. The tree is displayed showing topology only with bootstrap values of 0 removed for clarity.

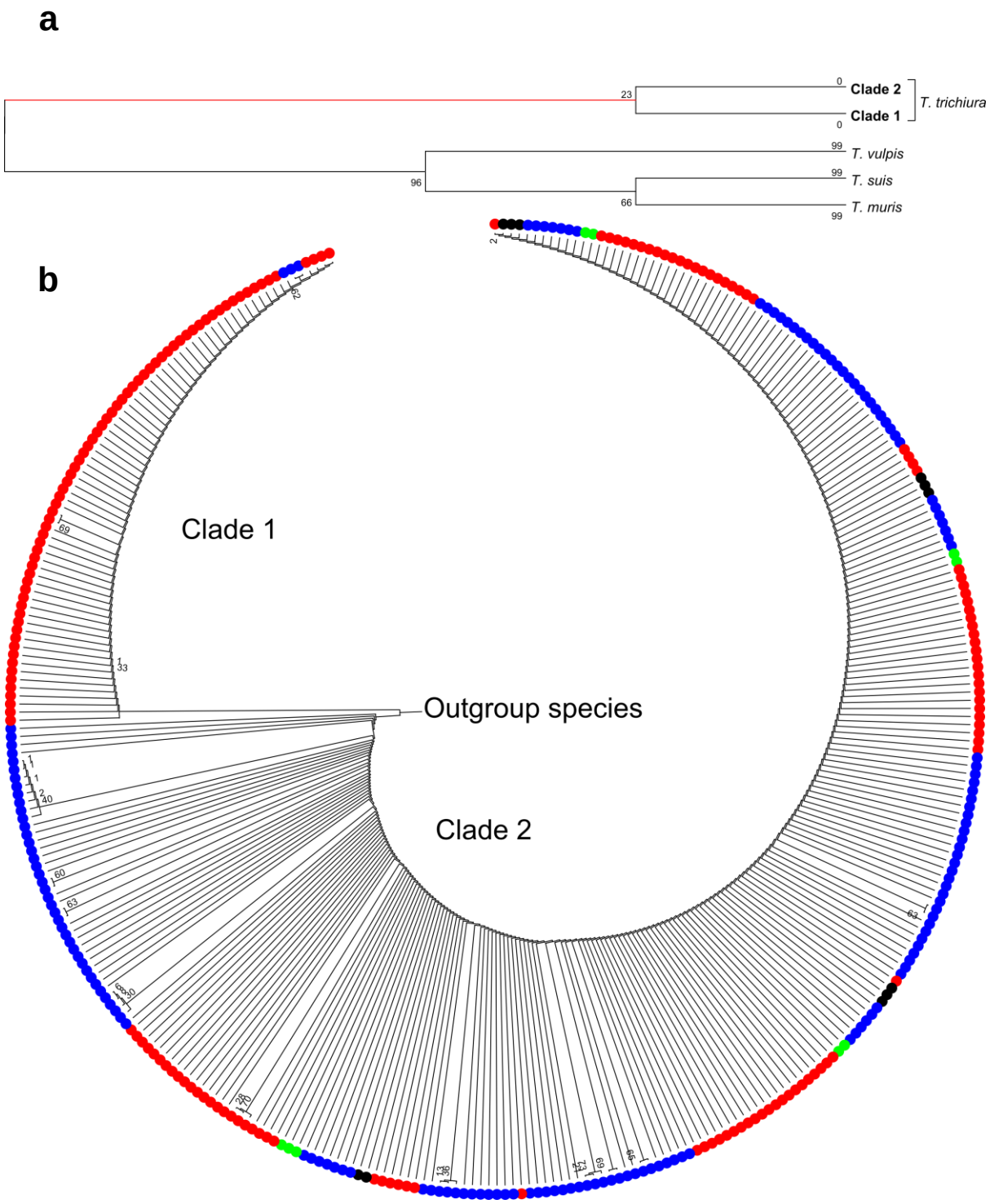


Figure 8.9: Bootstrapped Maximum Parsimony tree for Medieval *T. trichiura* ITS1 samples. Samples from three sites were included and are colour coded (Lübeck in red, Bristol in blue, Ellwangen in green and GenBank in black). The tree a is reduced to species level with the exception of *T. trichiura* where the two clades are shown, and tree b is rooted as tree a and cropped at the highlighted branch and is displayed showing topology only with bootstrap values of 0 removed for clarity.

Evaluation of the Maximum Likelihood (Figure 8.8) and Maximum Parsimony (Figure 8.9) trees revealed similar patterns, although the support for the split between the clades varies between the different algorithms used and is thus not strong enough for a decisive argument. The clearest divide and the highest support was found the Maximum-Likelihood algorithm (Figure 8.8). However, further work will be required to confirm the divide into two clades, including more sophisticated algorithms such as the Bayesian Inference of phylogeny (Huelsenbeck et al. 2001; Huelsenbeck and Ronquist 2001).

Chapter 9

Global discussion and conclusion

This doctoral thesis is an example of interdisciplinary research which links Biology with Archaeology and History. The approach does not claim to provide all new answers, but is conceived as an additional tool to put together the big picture of human history by providing novel insight.

Diseases have been described through history as challenges which need to be overcome. The description and the development of countermeasures, i.e. medical treatment, is one of the oldest scientific disciplines. The most apparent diseases have always been the ones with devastating symptoms or high fatality rates. The infections of humans with parasites in this study have been described as far back as ancient Egypt (medical texts) and the first description of finds in an archaeological context was 70 years ago (Szidat 1944). There were several descriptions of helminth parasite eggs in various contexts, but the majority of the studies did not employ any molecular methods or did indeed go further than the mere mention of their existence in the sample.

9.1. Acquisition of samples and planning of the project

Archaeological samples were obtained from a variety of different sites from eight countries in Europe and one site in southern Africa. Samples evaluated were obtained from a variety of different contexts, including different time periods and different geographic locations. The oldest samples diagnosed positive for parasites were dated to 3630 BCE and were obtained from one of the UNESCO listed Neolithic stilt settlements in Switzerland (excavation Opéra at the Kleiner Hafner in Zurich). Presence of nematode parasites in these samples demonstrated a long standing relationship between humans and their parasites. Samples were, where possible, acquired to be larger sets of a single location and fitting in with other

previously acquired sets of samples to maximise the information gained by the evaluation. The focus areas of the project were thus defined largely by the availability of suitable samples. Over the course of the project, a sample collection of considerable size was obtained and studied, although some very central samples were still missing. Three focal eras formed during the process. The medieval trade though the samples from the trade ports of Bristol, Lübeck and Ipswich; with other supporting samples from contemporary sites was the major topic. The medieval trade offered the opportunity to correlate the findings with extensive historic records (mainly from Lübeck and the Hanseatic League), which supported the formulated hypotheses and helped to evaluate the potential of the approach. The plethora of available sources is very fortunate, but for the other interest areas such extensive sources did not exist. The interaction between different societies between the 8th and the 11th century CE was the second focus of the project. Many different populations and tribes inhabited Europe at the time – the Anglo-Saxons and Vikings the British Isles, the Vikings Scandinavia, and the Slaves central Eastern Europe. A collection of Viking and Anglo-Saxon settlement samples combined with a large collection of grave sediments from a Slavic stronghold was studied to provide some insight into the complex interactions. The third area of interest was the question of the origin of the parasites. For this purpose, collections of the oldest possible human remains were acquired. The collection of samples includes settlement samples from UNESCO-listed waterlogged stilt settlement samples from Switzerland and Germany, but also samples from some of the oldest excavated graves in Europe (corded ware culture). Besides the question on how old the parasitic infestations of humans were, the possibility of studying samples from geographically distinct, but not very distant places was an opportunity to evaluate human migration in the Neolithic.

The work focused on sites in Europe as most of the parasites were no longer prevalent in Europe which excluded the danger of contamination with modern samples and added safeguard in the reconstruction of the ancient parasite population. The European sampled range from the Mesolithic to the 18th century CE, the African samples from Lesotho were considerably older (c17 000 BCE). For the dating we relied on the accuracy of the methods applied by the excavating archaeologists. Samples were collected from latrines, communal deposits (e.g. waste pits) or from the abdominal region of skeletons. All samples were initially screened microscopically to assess the parasite content and all positive samples and some of the negative samples processed for molecular evaluation and sequence generation. The majority of samples from communal deposits (>90 %) were found positive for at least one of the described parasites, most frequently *Trichuris* or *Ascaris*. In undisturbed burials the parasites were contained in the abdominal region (Oh et al. 2010). The analysis of individuals revealed the prevalence of any parasite as well as the prevalence of co-infections. The rates of infection were estimations as we were evaluating one particular time-point in the life of an individual, which might not represent the situation in the population. Based on examinations of graves and comparison with studies from regions where helminths are still prevalent it was possible to reasonably assume that in ancient populations the prevalence rates of people infected were about 25%, but varied between sites and with pharmaceutical interventions (Betson et al. 2012; Pullan et al. 2014; Schmidlin et al. 2013). These findings were consistent both with modern studies and with the wide prevalence in deposit and latrine samples. Although the parasite eggs have extraordinarily sturdy shells, it cannot be excluded that different soil conditions have an influence on the preservation of the parasite eggs in the sediment. These hardy shells and the

relatively high proportion made these infections especially suitable for studies of larger populations in archaeology.

9.2. Methods development

At the onset of the project no suitable methods were available and all procedures necessary had to be developed. The development was a long process which is still ongoing as the refinement and addition of new approaches can provide possibilities which open new avenues for research. An example for this was the introduction of a parallel sequencing approach. Although the generation of sequences by extraction of aDNA and PCR amplification with subsequent TA-cloning and Sanger sequencing was successful, the bar-coded parallel sequencing approach added the possibility to generate more sequences than possible by traditional Sanger sequencing.

Results achieved in this project were on several levels which combine to the greater picture of the research. Most stages of the processing generated data which had a value on its own but also contributed to deeper understanding. The main interest was focussed on microscopic diagnostics and genomic analysis. Both stages gave information on the parasite population and of their diversity. The microscopic diagnosis allows to estimate the load and the different varieties (to a genus level) of parasites in a specific sample, whereas the genetic evaluation showed diversity within a species of the parasites which can then be compared to other sites, which could be demonstrated most notably using *Trichuris trichiura* ITS1(Figure 8.1).

The process of developing suitable methods is described and discussed extensively in chapter 2 (general methods) and chapter 3 (parallel sequencing),

which described all the issues und challenges of the process of methods development.

9.3. What can we learn from samples diagnosed negatively for parasites?

Samples with no parasites potentially contain more information than initially perceived, depending on the reason for the absence of the parasites. One of the key advantages of the parasites in this study is that they used to be prevalent in Europe, but were nearly completely eradicated since. This is a major advantage as it drastically reduces the risk of contamination with modern material which would be more abundant than the ancient DNA.

Depending on the environmental conditions there might be issues with the microscopic or molecular detection of parasites. The absence of the parasites did not allow the conclusion of the complete absence of the parasites from the population. In two cases it is likely that on deposition there were parasites present, but when the samples were received no parasites could be detected. These samples were both received from Mark Robinson (Environmental Archaeology, University of Oxford) dating to the Roman period. These samples were taken for environmental studies, such as palynological studies (pollen analysis). One was a latrine sample from Silchester (UK) and the other one from the sewage system of Herculaneum (Italy). Both samples were most likely associated with human excrements. The sample from Silchester was stored for a several years; sieved and dried for the study of pollen and seeds. These conditions were clearly not favourable for the detection of parasites. The sample from Herculaneum was probably severely changed after the outbreak of the Vesuvius in 79 CE. It is likely that these sediments were covered with hot

volcanic ashes and therefore most organic material was destroyed. In this particular sample seeds were often found as imprints or mineralised cores with the organic shell missing (personal communication Mark Robinson).

A slightly different case where no parasites were found were samples received from a Navy graveyard in southern England (Roberts and Mays 2010). The obtained samples were clumps of soil associated with the excavated bones from the pelvic region. The skeletons were dated between the 18th and the early 19th century CE. There are different possibilities why these samples did not contain parasites and it remains speculative which one is the most likely explanation. The first explanation would be that these seamen lived in an era where the prevalence of infections in Europe was very low. However, as they were travelling, likely to the British Colonies, this is not an entirely suitable explanation. As the parasites are persistent the seamen could still show traces of them even after a few years back home. A second explanation could be the disposal of human waste material on sea. As it is likely that excrement was released directly to the sea, the direct contact required for the faecal-oral transmission is not given. It is thus likely that the parasites would not spread the same as in a similar community on land. Thirdly, it could also be that the deposition conditions are unfavourable for the preservation or detection of the parasites. The clay-rich soil makes the microscopic detection very difficult as it is very dense. Other clay-rich samples have also shown difficulties in subsequent preparation steps such as the filtration (clogging of filters) and aDNA extraction (clogging of silica membranes and extraction of PCR inhibitors). For these samples it is likely that the combination of the environmental conditions and the habits of the individuals have to be taken into account. The relatively small sample collection

might not represent the entire population and thus a larger set of samples has to be taken into account.

In modern-day Europe, there are hardly any cases of helminth infections diagnosed, and if they are diagnosed there is frequently a history of travelling to countries where these parasites are still prevalent (Tomaso et al. 2001). The analysis of other sample sets has shown that parasites were prevalent in Europe in the late middle ages. Based on observations of graveyard samples it could be assumed that the prevalence of these samples before the introduction of sanitation was similar to the prevalence in modern developing countries. It is unclear when exactly these parasites disappeared, but with increased sanitation and a new understanding of hygiene and food safety in the late 18th century CE it is likely that these intestinal parasites gradually disappeared. Assessing an exact date is difficult as the absence traces of the parasites do not necessarily mean the absence in the living population. The analysis of the graveyard of the Radcliffe Infirmary in Oxford gave insight into the transition period where parasites were made extinct in the population of Western Europe. The Radcliffe Infirmary was a University hospital established in 1770 and in use until 2007. The attached graveyard was used between 1770 and 1855 with records of the buried individuals, but as none of the grave markings (i.e. tombstones) were found it was not possible to indentify the individuals. The general health state of the skeletons was good (personal communication Vix Hughes, Oxford Archaeology), which indicates that the individuals received good care in the hospital. From the excavation, an extensive sample set of over 100 samples was received, of which 50 were analysed. None of these samples showed trace of any parasites, which would indicate incidence rates below 2 % based on the detection rate in similar samples (graveyard samples, for instance the ones from Pohansko). The soil

conditions did not show any obvious restrictions on the preservation of the parasites or limitations on the microscopic or molecular diagnosis. The condition of the bones was reported to be very good, thus, there seem not be have been restrictions on this side. This most likely indicates that at the end of the 18th century CE the sanitation and health care system was established well enough to significantly reduce the incidence rates of the parasites in the population. This does not allow conclusions on the incidence rates in other places in Europe, England or even Oxfordshire, as the samples were from a hospital graveyard where patients were treated for a variety of illnesses and were in a very clean environment and the graveyard is located in a relatively wealthy city which does not allow conclusions on the surrounding countryside. These findings will require further evidence, but are a good indicator that parasites were getting less common with increased urbanisation and industrialisation - even without specific counter measurements.

Other sample sets established that parasites were prevalent in Viking settlements in Denmark, Norway and the UK. However, when samples from the Viking graveyard on Langeland, Denmark (Grøn et al. 1994) were analysed, none of the 27 samples were tested positive. These samples were collected in 2013 from the archaeological archives at the University of Copenhagen, Denmark (courtesy of Pia Bennike). The excavations on Langeland were carried out between 1982 and 1989. Therefore, the samples were stored dry for about thirty years with only fragmented documentation on the sampling procedures. The long storage period, an optically dense clay-rich soil and potential sub-optimal sampling cumulated in the negative diagnosis of all 27 samples. Prolonged storage should not be a problem in general, as the samples have been *in situ* for longer periods, but the complete drying of these samples might have accelerated the decay of the parasite traces. A similar issue was

encountered with one of the subsets of the sample collection from Neolithic sites in Southern Germany (Alleshausen-Hartöschle) where none of the three samples were diagnosed positively. These samples were excavated in 1993 and stored in a dry storage since. It is likely that this had an effect on the parasite traces, but further work, such as an experiment to artificially dry a formerly positively diagnosed sample to demonstrate the decay of the parasite eggs under these circumstances.

The oldest samples cave deposits from Lesotho received from Charles Arthur and Peter Mitchell (School of Archaeology, University of Oxford) dated between 17.000-70.000 BCE and are thus considerably older than any other samples in this study. Unfortunately, none of the samples showed any trace of parasites. Although it can be assumed that the parasites have been associated with humans for a considerable time during the human evolution, the origin of the parasitic lifestyle remains speculative. Although unlikely, it has been taken into account that these samples might pre-date human infections. As these were not permanent settlements it is possible that sampling is an issue. Samples were usually taken from places where clear traces of human occupation were found (such as remains of hearths). Whether the samples were taken of sediments which are associated with faeces cannot be verified. These caves were generally used as temporary dwellings and no permanent structures were erected. Thus, dilution of the material has to be taken into account. It could also be that the faecal material was deposited outside of the caves or transported out of the cave after deposition. As hearths are a common sign for human use of these caves, the sediments close to these were frequently collected. As it cannot be reconstructed how close these were to the actual fire upon deposition it is possible that the organic material was destroyed when deposited. Further samples have to be evaluated to assess interesting questions such as the origin of the parasites.

9.4 On the origin of the parasite – the oldest samples

Limitations of a new approach are important to be probed. Age was the most obvious limitation to tackle. Compared to the oldest aDNA recovered (>700,000 y) the oldest parasite in the study, which provided the oldest nematode parasite sequence available, these samples were not comparable (Dabney et al. 2013; Orlando et al. 2013). However, as it is still unclear when the nematode parasites infected humans (or hominins) first, the data obtained is very valuable. Further, as all the parasites were recovered from sediments which were clearly associated with humans and their excrements, it will be difficult to find suitable sediments which predate the Neolithic period and the start of the establishment of permanent settlements. The samples obtained from Lesotho represent an earlier sample source in the form of a natural shelter (cave) which was used temporarily. The next issue would be to identify within these sediments which part would be most useful for the parasitological analysis as the parasite remains, unlike bones, are not visible to the naked eye and large-scale microscopic screening of large amounts of soil is impracticable.

9.5 Outlook on further work

The tools developed as part of the work on this thesis have been demonstrated to have the potential to produce more a large amount of genetic data from ancient parasites. The global analysis of the sequences as presented in chapter 8 allowed us a glimpse into the connections which existed between the sites in the study. However, although the collection of samples was extensive, it can never be complete. Further work will be required at several parts of the project to enable a more complete view on the connections and interactions between the sites in Europe.

The most important first step will be to publish the data obtained in this thesis in peer-reviewed journals to advertise the research which will most likely facilitate the acquisition of further sample sets and collaborations. Many aspects of the research might also appeal to the general public which might require the publication of some parts of this work in journals accessible to the wider public or in public lectures associated with excavations or organised by local archaeology or heritage conservation organisations.

An important part of the work to establish methods is the establishment of collaborations and thus the extension of the sample set. This chapter will give an overview on the samples collected or expected which have not been analysed yet. The constant growth of the sample collection is an important part to link the samples within the project. The success of the first parts of the project raises expectations for further applications, such as the further extension of the sample collection and the mapping of parasite infections across Europe through time.

Samples described in the first parts of this chapter concerning the medieval trade through Europe from Bristol (chapter 4), Lübeck and Ellwangen (chapter 5) showed intriguing insight into the trading patterns. However, direct links are so far difficult to see as there are only few cities involved which were historically linked, but to which degree remains unclear. The extension of the sample set could give valuable insight into the issue. Intriguing findings such as the change in the fish consumption in Lübeck linked to the decreased prevalence of *D. latum* lead to the hypothesis, that this might be linked to the trade with Norwegian stockfish. Lübeck was the main trading centre for stockfish until the establishment of direct trade from Bergen to other cities. It is therefore important to verify a potential link to Bergen. The University of Bergen and the associated University Museum conducted several

excavations in and around Bergen. A set of settlement and graveyard samples will be collected and provided for further research. Trading towards the east, Lübeck and the Hanseatic League had strong links to the cities on the coast of the Baltic Sea. Samples from one of the big cities on the eastern coast might give valuable insight into the trading patterns. A formal collaboration with the University of Latvia in Riga was established and access to an extensive sample collection was given. The samples were collected in different parts of medieval Riga and from rural areas in the Hinterlands of Latvia which might reveal differences between the urban and the rural areas. The publication of the first results from Lübeck and the very active collaboration with the excavators (namely Dirk Rieger) will likely attract other samples from excavations in the other Hanseatic League cities. The trade network of the Hanseatic League offers unique opportunities for the integration of the bio-archaeological research with historical materials. Several cities of the region have extensive archives with trade records, ownership deeds of ships, houses and goods, city and parish records which allow reconstruction of the social structure and in some cases even extensive written correspondences between traders.

A collection of Mesolithic burial and settlement samples were received from the University of Latvia in Riga. These samples might push the oldest sequences back establishing the origin of the human parasites. Burials in general will become an important source for future work into the assessment of palaeoepidemiology. The results so far, based on graveyards from the Neolithic to the early modern period provided an insight into the matter, but as considerable variations were found, such as the fact that all samples from the Neolithic burial ground of Bad Mergentheim were diagnosed positive for nematode parasites, the study of more graves will be

necessary to allow reliable conclusions. So far, we can only assume that the variations observed correlate with the prevalent parasite load in the population.

The vast majority of the research so far has been done based on four species of parasites, the nematodes *T. trichiura* and *Ascaris* sp. and the cestodes *D. latum* and *Taenia* spp. Although preliminary work has been done to include other parasites such as *Enterobius vermicularis*, *Anisakis simplex*, *Giardia intestinalis* and *Cryptosporidium* sp. the results obtained from these parasites were limited as the work had to remain focussed on the central aspects to generate a meaningful data set. The different life cycles and infection routes of helminth parasites allow diversification of the approach, which was demonstrated with the inclusion of the cestodes in the analysis of the samples from Lübeck which allowed conclusions on a cultural and dietary change around 1300 CE.

To assess some of these parasites the developed current protocol might not be entirely adequate, as for instance the protozoan parasites *Giardia* and *Cryptosporidium* are not expected to be detectable in microscopic diagnosis. In clinical parasitology there are a variety of techniques available to identify parasite species. Some of these techniques can potentially be applied to archaeological samples. Some of the structures used in clinical parasitology cannot be observed in samples derived from archaeological contexts; however, there are possibilities to apply methods based on the outer structure of the oocysts or eggs of the parasites. In particular, we intend to identify two species of protozoan parasites using fluorescence microscopy techniques. The protozoan parasites *Giardia intestinalis* and *Cryptosporidium* spp. are common infections of humans across the world. There are two main techniques which are used in clinical parasitology. One involves a chemical stain (auramine phenol) which stains *Cryptosporidium* spp., but also some

yeasts. This very inexpensive chemical compound would be a simple test for the potential of fluorescence microscopy based diagnostics. There are also more specific identification kits based on antibody-linked fluorophores available. Antibodies are also used for ELISA-based approaches, which might not be very useful in our investigation. The evaluation of these methods will be a pre-requisite for the proceeding of the research.

To assess the power of the sample collection, more elaborate bioinformatics tools will be required to enable the research to take further variables into account. The analysis so far did not take one of the most important factors into account – the dating. Although all sequences were obtained with an archaeological date, this data was not yet used to its full potential. The dates provide an additional source of information which could be used in an in-depth phylogenetic analysis, as sequences from samples which were dated earlier will have to be at a more basal level of a tree. This will provide a more comprehensive view onto the connections between the sites by their parasites.

9.6 Closing remarks

In conclusion, this doctoral thesis presents the tools and first results of a novel approach for the study of human actives through time by the analysis of non-human genomic material. This thesis has set the foundations of the approach and provides thus a starting step for further investigations. Further work will require the acquisition of more samples and the establishment of collaborations in various fields and the application and development of more tools to interrogate the data. The main focus of this thesis was to develop the starting point an approach which will have some impact onto research in archaeology. To move from this starting point onward, further techniques will be required, especially on the level of data analysis as there is more potential in the data collected that has not yet been made use of. The challenge and opportunity for future work will also include the continuous development and refinement of the methods used to maximise the efficiency of sample processing and the usability of the produced data for research.

Chapter 10

Materials and Methods

As a major part of this thesis consisted of the development of microscopic and molecular methods, these were describe in detail in chapter 2 (general method development) and chapter 3 (parallel sequencing).

10.1 Materials

Chemicals were, unless otherwise stated, obtained from Sigma or Fisher Scientific. Plastic consumables, including pipettes and Flacon tubes (15 ml and 50 ml) were obtained from Corning Inc. Microcentrifuge tubes were obtained from Eppendorf, Axygen (Corning Life Sciences) and Sarstedt (screw-top tubes). PCR tubes were obtained from Axygen and PCR plates from Web Scientific. Pipette tips were obtained from StarLabs. Petri Dishes were obtained from Greiner and VWR. Nuclease Free Water was obtained from Qiagen, if not specifically stated, MilliQ water was used.

10.1.1 Reagents

LB broth and LB-Agar plates

Obtained in powdered form (Sigma)

Agarose

Molecular grade (Fisher/Sigma)

DNA Loading Dye

30 % v/v Glycerol (Sigma)

0.25 % w/v Bromophenol blue (Sigma)

Water (MilliQ)

TBE (10x)

55 g Tris (2-Amino-2-hydroxymethyl-propane-1,3-diol, Sigma)

27.5 g Bornoic Acid (Fisher/Sigma)

4.65 g EDTA (Ethylenediaminetetraacetic acid, Sigma)

Up to 500 ml water

Autoclaved (15 min at 121°C)

Tris-HCl pH 7.5

121.1 g g Tris (2-Amino-2-hydroxymethyl-propane-1,3-diol, Sigma)

Up to 1000 ml water

Adjust pH to 7.5 using HCl

Amicilin

100 mg/ml, 0.2 μm filtered (Sigma)

Kanamycin (VWR)

Powdered, prepared to 100 mg/ml, sterile filtered 0.22 μm in Nuclease Free Water

X-Gal (5-Bromo-4-chloro-3-indolyl β-D-galactopyranoside)

Prepared to 40mg/ml (Sigma)

IPTG (Isopropyl β-D-1-thiogalactopyranoside)

Prepared to 40mg/ml (Sigma)

Salter buffer (6M Guanidine thiocyanate)

5.9g Guanidine thiocyanate (Sigma)

50 ml water (MilliQ)

10.1.2 PCR Reagents

AmpliTaq Gold 360 and AmpliTaq Gold LD (Life Technologies) were used for all PCRs, except for the barcoding reactions for MiSeq parallel sequencing where Phusion Hot Start (New England Biolabs) was used.

10.1.3 TA-Cloning

Invitrogen pCRII dual promoter (Life Technologies)

Promega pGEM t-easy vector system

Competent cells: One Shot TOP10 Chemically Competent *E. coli* (Invitrogen) and NEB 5-alpha Competent *E. coli* (High Efficiency, New England Biolabs)

10.1.4 DNA extraction and clean-up kits

The modified protocols are described in chapter 2, basis for the extractions were:

DNeasy Blood & Tissue (Qiagen)

Mericon food kit (Qiagen)

DNeasy and MinElute Gel extraction kits (Qiagen)

DNeasy and MinElute PCR purification kit (Qiagen)

DNeasy MiniPrep (single tube and 96 well, Qiagen)

10.2 Primers

Primers used for this study (table xx) were organised by target organism and synthesised by Invitrogen and Integrated DNA technologies (IDT). The lyophilised primers were diluted in Nuclease Free Water (Qiagen) under sterile conditions to obtain a concentration of 100 mM. Of all primers a working stock of 10 mM was produced (to be stored at 4°C) and the original stocks were stored at -20°C.

10.2.1 Primers for General PCR

Trichuris trichiura

Name	Genetic Region	Sequence [5' – 3']
Ttri-COX1-1F	CO1	GGAGGAGGAGACCCCATCCT
Ttri-COX1-1R	CO1	TATAGCCAAAAACTACTCCTAT
Ttri-NAD1-F	NAD1	GTTTACATAGGAGTTTCCCAACG
Ttri-NAD1-R	NAD1	TTCAACATTGTAGCCTCTAACTAATTC
Ttri-NAD1-inF	NAD1	TGAGCATCCAATAACAAATTTGCC
Ttri-NAD1-inR	NAD1	CTATTATACCCAAGAGTACAATTA
***T. trichiura* β-tubulin F**	β-tubulin	AGGTTTCAGATACAGTTGTAG
***T. trichiura* β-tubulin R**	β-tubulin	CAAATGATTTAAGTCTCCG
Tri_bT_F_Jan14	β-tubulin	AGGTTTCAGATACAGTTGTAGAAC
Tri_bT_R_Jan14	β-tubulin	ATGATTTAAGTCTCCGTAAGTTGGTG
Ttri_CO1_F	CO1	TTTTGGTTCTTCGGACACCC
Ttri_CO1_R	CO1	TGGTGTATATGTGGTGTCCTC
Ttri_CO1_Fn	CO1	GTTCTTCGGACACCCAGAAG
Ttri_CO1_Rn	CO1	GTGTATATGTGGTGTCCTCA
Tri_ITS1_1F	ITS1	GAGGCCTTTGGACCGATGCC
Tri_ITS1_1F_N	ITS1	AGGCCTTTGGACCGATGCCGC
Tri_ITS1_1R	ITS1	GCAGGCCAGCGGAGCAGCGC
Tri_ITS1_2F	ITS1	CAAAGCGGCCGTAGCTGCCG
Tri_ITS1_2F_N	ITS1	AAAGCGGCCGTAGCTGCCGTG
Tri_ITS1_2R	ITS1	TGCGCGCAGGCGGCGAGCAC
Tri_ITS1_3F	ITS1	GCTCGATCAGGCAGCAGCGG
Tri_ITS1_3F_N	ITS1	GCTCGATCAGGCAGCAGCGGTT
Tri_ITS1_3R	ITS1	TTTTCCAACCGGCGACGAGCCCG
Tri_ITS1_4F	ITS1	AACAGCCGTTCGACCTCGAG
Tri_ITS1_4R_N	ITS1	AATTCACGTCAATTCTCGAGTG
Tri_ITS1_4R	ITS1	AATTCACGTCAATTCTCGAG
Tri_CytB_F-A	Cytochrome *b*	TGCGGGAAGTGAAATATGGCT
Tri_CytB_R-A	Cytochrome *b*	AACAGTTGCGGCTCAAAACC
Tri_CytB_F-A	Cytochrome *b*	TAGGGTTTTGAGCCGCAACT
Tri_CytB_R-A	Cytochrome *b*	AGGAGAATGGGATCCAAGAGGA

Ascaris sp.

Name	Genetic Region	Sequence [5' – 3']
***Asaris β*-tubulin F**	*β*-tubulin	AGGTTTCTGATGTGGTGTTGGA
***Asaris β*-tubulin R**	*β*-tubulin	TATGTGGGATTTGTAAGCTTCAG
Asc A2 F2	CO1	ACTCTCGGGCTTATTTTACTGC
Asc A2 R	CO1	ACATAATGAAAATGACTAACAAC
Asc_CO1-F2n_p2	CO1	CTCTCGGGCTTATTTTACTGCTG
Asc_CO1-Rn_p2	CO1	ACATAATGAAAATGACTAACAACTA
Asc_CO1-Rn_Neo	CO1	ACATAATGAAAATGACTAACAACTA
Asc_CO1_F	CO1	TTTTTTGGTCATCCTGAGGTTTAT
Asc_CO1_R2	CO1	CAGTAGGAACAGCAATAACC
Asc COX1 F2-mod	CO1	CACCAGTAGGAACAGCAATAAC
Asc COX1 F2-mod RC_seq	CO1	GTTATTGCTGTTCCTACTGGTG
Asc-ITS1-1F	ITS1	TTTTTGGCGGACAATTGCATGC
Asc-ITS1-1R	ITS1	TTTTTTGAGTTTTGGCGGCC
Asc-ITS1-2F	ITS1	GTCGAGTAGACTCAATGAGCC
Asc-ITS1-2R	ITS1	CCCGATGGCGCAATGTGCGTTCG
Asc-ITS1-3F	ITS1	TAAATAGTGCGAATTGCAGACAC
Asc-ITS1-3R	ITS1	TGATGTATATATTCTCTCTCGC
Asc_CytB_F	Cytochrome *b*	GTTAGGTTACCGTCTAGTAAGG
Asc_CytB_R	Cytochrome *b*	CACTCAAAAAGGCCAAAGCAC
Asc_CytB_R2	Cytochrome *b*	ACACGAAAAATCCAACCAAAAT
Asc_CytB_R3	Cytochrome *b*	CAGATACCCAAACCTTCTTC

Enterobius vermicularis

Name	Genetic Region	Sequence [5' – 3']
***E. vermicularis* 6S-rRNA F-ex**	6S-rRNA	CACTTGCTATACCAACAACAC
***E. vermicularis* 6S-rRNA R-ex**	6S-rRNA	GCGCTACTAAACCATAGAG
***E. vermicularis* 6S-rRNA F-nt**	6S-rRNA	ACAACACTTGCACGTCTC
***E. vermicularis* 6S-rRNA R-nt**	6S-rRNA	GAATTGCTCGTTTGC

Cryptosporidium sp.

Name	Genetic Region	Sequence [5' – 3']
***Cryptosporidium* sp. 18S-rRNA F**	18S-rDNA	GGGGAATTAGGGTTCGATTC
***Cryptosporidium* sp. 18S-rRNA R**	18S-rDNA	ACCAGACTTGCCCTCCAAT
***Cryptosporidium*_18S_F**	18S-rDNA	GAACTGTACAGATGCTTGGGAGAAT
***Cryptosporidium*_18S_R**	18S-rDNA	CCTTCGTTAGTTGAATCCTCTTTCCA
Crypto_COWP_F	COWP	CAAATTGATACCGTTTGTCCTTCTG
Crypto_COWP_R	COWP	GGCATGTCGATTCTAATTCAGCT
Crypto_Hsp70_F	Hsp70	ACCAAGAGGTGTACCACAAA
Crypto_Hsp70_R	Hsp70	CTCCAAAGAGTTCTTAGCCT
Crypto_Hsp70_R_long	Hsp70	GTTCTCCAAAGAGTTCTTAGCCT
Crypto_Hsp70_F_N	Hsp70	ACCAAGAGGTGTACCACAAATTG
Crypto_Hsp70_R_s_N	Hsp70	CTCCAAAGAGTTCTTAGCCTC

Crypto_Hsp70_R_l_N	Hsp70	TCTCCAAAGAGTTCTTAGCCTC
Cryptosporidium_gp60_F1_N	gp60	TATTCTCAGCCCCAGCCGTTCCACTC
Cryptosporidium_gp60_R1_N	gp60	CAGTACCACTAGAATCTTGACTGC
Cryptosporidium_gp60_F1	gp60	TATTCTCAGCCCCAGCCGTTCC
Cryptosporidium_gp60_R1	gp60	CAGTACCACTAGAATCTTGAC

Giardia lamblia

Name	Genetic Region	Sequence [5' – 3']
***Giardia* 16S-rRNA F**	16S-rDNA	ATCCGGTCGATCCTGCCG
***Giardia* 16S-rRNA R**	16S-rDNA	GGGGTGCAACCGTTGTCCT
***Giardia β*-Giardin P200 F**	*β*-Giardin	CATCCGCGAGGAGGTCAA
***Giardia β*-Giardin P200 R**	*β*-Giardin	GCAGCCATGGTGTCGATCT
***Giardia β*-Giardin P400 F**	*β*-Giardin	CAAGAGCCTGAACGATCTC
***Giardia β*-Giardin P400 R**	*β*-Giardin	GGTCGTACATCTTCTTCCTT

Anisakis simplex

Name	Genetic Region	Sequence [5' – 3']
***A.simplex*_ITS1_F**	ITS1	AACTCGTACACACGTGGTGGC
***A.simplex*_ITS1_Rs**	ITS1	GTCTGCTCAACTCTGCCTT
***A.simplex*_ITS1_Rl**	ITS1	GGCTCATTAAGTCTGCTCAA
***A.simplex*_COX1_F**	CO1	GATTCTTGTTTCGTTGATATGGGT
***A.simplex*_COX1_R**	CO1	TATGCTCCAAAGAAATAGACCTC

Taenia spp.

Name	Genetic Region	Sequence[5'–3']
***T.solium*_Tso31_F**	Tso31	GGTGTCCAACTCATTATACGCTGTG
***T.solium*_Tso31_R**	Tso31	GCACTAATGCTAGGCGTCCAGAG
***T.solium*_Tso31_F_n**	Tso31	GGTGTCCAACTCATTATACGCTGTGAC
***T.solium*_Tso31_R_n**	Tso31	GCACTAATGCTAGGCGTCCAGAGGTG
F1_test_Tso	cytochrome *b*	CTTTTTGTTCATATGGCCCGTG
R1_test_Tso	cytochrome *b*	CTTTTTGTTCACATGGCCCGTG
F2_test_Tsg	cytochrome *b*	ATGGGTGAGGCTTTTACTGG
R2_Tsg	cytochrome *b*	GCATATTTGTTTGGGTTTTGTT
Taenia_HDP2_F_long	HDP2	CAGTGGCATAGCAGAGGAGGAA
Taenia_HDP2_F_short	HDP2	CTTCTCAATTCTAGTCGCTGTGGT
Taenia_HDP2_F_long_r	HDP2	GGACGAAGAATGGAGTTGAAGGT
T.solium_COX1_F1	CO1	TTTATACTTTGTTGGGTTTGTG
T.solium_COX1_R1	CO1	AATGGAATTAAGTATTTACCA
T.solium_COX1_F2	CO1	TTATCGGATTTGAATTTACC
T.solium_COX1_R2	CO1	TAAAAATACTTGATACACCAG
T.saginata_COX1_R2	CO1	ATACTTGATACACCAGCTAAATGC
T.saginata_COX1_R2_n	CO1	TAGACGCACCCGCCAAATGCAACG
T.saginata_COX1_F2	CO1	TTATCTGATTTGAATTTACCCCG
T.saginata_COX1_May14_F	CO1	TGCGTTTTTTGATCCATTGGGTGG
T.saginata_COX1_May14_FN	CO1	GTTTTTTGATCCATTGGGTGGTGG
T.saginata_COX1_May14_R	CO1	CATATGATGACCCCACACACTTCTCC
T.saginata_COX1_May14_RN	CO1	ATGACCCCACACACTTCTCCCCAAAC

Diphyllobothrium latum

Name	Genetic Region	Sequence[5'–3']
Diphyllobothrium_COX1_F1	CO1	TTGATCCTTTGGGTGGTGG
Diphyllobothrium_COX1_R1	CO1	ACATAGCAAATAACAACCCG
Diphyllobothrium_COX1_F2	CO1	GGGGTGTTACGGGTATTATACTC
Diphyllobothrium_COX1_R2	CO1	ATACTTATTCAATCTCACACCTG

10.4.2 Primers for parallel sequencing

Ascaris

Name	Genetic Region	Sequence[5'–3']
002.Asc_CO1-F2-1	CO1	TTCTCCAGGTATGGATCTTGACTCTCGGGC
004.Asc_CO1-F2-1	CO1	AACACCTGGTATGGATCTTGACTCTCGGGC
005.Asc_CO1-F2-1	CO1	CTCTCACGGTATGGATCTTGACTCTCGGGC
010.Asc_CO1-F2-1	CO1	TACACGAGGTATGGATCTTGACTCTCGGGC
012.Asc_CO1-F2-1	CO1	ATCTCGTGGTATGGATCTTGACTCTCGGGC
017.Asc_CO1-F2-1	CO1	TCCTACCGGTATGGATCTTGACTCTCGGGC
018.Asc_CO1-F2-1	CO1	GTCGACAGGTATGGATCTTGACTCTCGGGC
021.Asc_CO1-F2-1	CO1	TTCGAACGGTATGGATCTTGACTCTCGGGC
028.Asc_CO1-F2-1	CO1	CTCGAGTGGTATGGATCTTGACTCTCGGGC
032.Asc_CO1-F2-1	CO1	CGCAATTGGTATGGATCTTGACTCTCGGGC
036.Asc_CO1-F2-1	CO1	TACTGCTGGTATGGATCTTGACTCTCGGGC
039.Asc_CO1-F2-1	CO1	CACTGAGGGTATGGATCTTGACTCTCGGGC
042.Asc_CO1-F2-1	CO1	AACTGGAGGTATGGATCTTGACTCTCGGGC
046.Asc_CO1-F2-1	CO1	ACCGGTAGGTATGGATCTTGACTCTCGGGC
054.Asc_CO1-F2-1	CO1	CGCTTAAGGTATGGATCTTGACTCTCGGGC
055.Asc_CO1-F2-1	CO1	TACGTAGGGTATGGATCTTGACTCTCGGGC
058.Asc_CO1-F2-1	CO1	CACGTGAGGTATGGATCTTGACTCTCGGGC
061.Asc_CO1-F2-1	CO1	AACGTTCGGTATGGATCTTGACTCTCGGGC
098.Asc_CO1-F2-1	CO1	CGAAGCAGGTATGGATCTTGACTCTCGGGC
110.Asc_CO1-F2-1	CO1	CTAGGTAGGTATGGATCTTGACTCTCGGGC
115.Asc_CO1-F2-1	CO1	GAATTCGGGTATGGATCTTGACTCTCGGGC
121.Asc_CO1-F2-1	CO1	CAATTGCGGTATGGATCTTGACTCTCGGGC
123.Asc_CO1-F2-1	CO1	GTAATGGGGTATGGATCTTGACTCTCGGGC
132.Asc_CO1-F2-1	CO1	TTGACCTGGTATGGATCTTGACTCTCGGGC
135.Asc_CO1-F2-1	CO1	CTGACAGGGTATGGATCTTGACTCTCGGGC
141.Asc_CO1-F2-1	CO1	GTGACTCGGTATGGATCTTGACTCTCGGGC
143.Asc_CO1-F2-1	CO1	CAGTCTGGGTATGGATCTTGACTCTCGGGC
146.Asc_CO1-F2-1	CO1	CAGGACAGGTATGGATCTTGACTCTCGGGC
162.Asc_CO1-F2-1	CO1	TAGAGCAGGTATGGATCTTGACTCTCGGGC
164.Asc_CO1-F2-1	CO1	ATGTGCTGGTATGGATCTTGACTCTCGGGC
165.Asc_CO1-F2-1	CO1	CAGAGACGGTATGGATCTTGACTCTCGGGC

167.Asc_CO1-F2-1	CO1	GTGTGAGGGTATGGATCTTGACTCTCGGGC
172.Asc_CO1-F2-1	CO1	AAGAGGTGGTATGGATCTTGACTCTCGGGC
173.Asc_CO1-F2-1	CO1	CTGTGTCGGTATGGATCTTGACTCTCGGGC
194.Asc_CO1-F2-1	CO1	CCTTCCAGGTATGGATCTTGACTCTCGGGC
195.Asc_CO1-F2-1	CO1	TTTGCCGGGTATGGATCTTGACTCTCGGGC
201.Asc_CO1-F2-1	CO1	ATTGCGCGGTATGGATCTTGACTCTCGGGC
204.Asc_CO1-F2-1	CO1	GCTTCGTGGTATGGATCTTGACTCTCGGGC
207.Asc_CO1-F2-1	CO1	TCTTCTGGGTATGGATCTTGACTCTCGGGC
208.Asc_CO1-F2-1	CO1	GTTGCTTGGTATGGATCTTGACTCTCGGGC
210.Asc_CO1-F2-1	CO1	TCTGACAGGTATGGATCTTGACTCTCGGGC
223.Asc_CO1-F2-1	CO1	GCTGATGGGTATGGATCTTGACTCTCGGGC
226.Asc_CO1-F2-1	CO1	GCTAGCAGGTATGGATCTTGACTCTCGGGC
228.Asc_CO1-F2-1	CO1	CGTTGCTGGTATGGATCTTGACTCTCGGGC
231.Asc_CO1-F2-1	CO1	AGTTGAGGGTATGGATCTTGACTCTCGGGC
234.Asc_CO1-F2-1	CO1	GGTTGGAGGTATGGATCTTGACTCTCGGGC
237.Asc_CO1-F2-1	CO1	TGTTGTCGGTATGGATCTTGACTCTCGGGC
241.Asc_CO1-F2-1	CO1	GATATCCGGTATGGATCTTGACTCTCGGGC
247.Asc_CO1-F2-1	CO1	CGTGTAGGGTATGGATCTTGACTCTCGGGC
251.Asc_CO1-F2-1	CO1	CATATGGGGTATGGATCTTGACTCTCGGGC

Trichuris trichiura

Name	**Genetic Region**	**Sequence[5'–3']**
002_Tr_ITS1_4Rn	ITS1	TTCTCCAATTCACGTCAATTCTCGAGTG
004_Tr_ITS1_4Rn	ITS1	AACACCAATTCACGTCAATTCTCGAGTG
005_Tr_ITS1_4Rn	ITS1	CTCTCAAATTCACGTCAATTCTCGAGTG
006_Tr_ITS1_4Rn	ITS1	TGCGCAAATTCACGTCAATTCTCGAGTG
007_Tr_ITS1_4Rn	ITS1	GACACAAATTCACGTCAATTCTCGAGTG
008_Tr_ITS1_4Rn	ITS1	ACCCCAAATTCACGTCAATTCTCGAGTG
010_Tr_ITS1_4Rn	ITS1	TACACGAATTCACGTCAATTCTCGAGTG
012_Tr_ITS1_4Rn	ITS1	ATCTCGAATTCACGTCAATTCTCGAGTG
013_Tr_ITS1_4Rn	ITS1	CACACTAATTCACGTCAATTCTCGAGTG
015_Tr_ITS1_4Rn	ITS1	GTCTCTAATTCACGTCAATTCTCGAGTG
016_Tr_ITS1_4Rn	ITS1	AGCGCTAATTCACGTCAATTCTCGAGTG
017_Tr_ITS1_4Rn	ITS1	TCCTACAATTCACGTCAATTCTCGAGTG
018_Tr_ITS1_4Rn	ITS1	GTCGACAATTCACGTCAATTCTCGAGTG
019_Tr_ITS1_4Rn	ITS1	AGCAACAATTCACGTCAATTCTCGAGTG
020_Tr_ITS1_4Rn	ITS1	CACCACAATTCACGTCAATTCTCGAGTG
021_Tr_ITS1_4Rn	ITS1	TTCGAAAATTCACGTCAATTCTCGAGTG
023_Tr_ITS1_4Rn	ITS1	AACCAAAATTCACGTCAATTCTCGAGTG
024_Tr_ITS1_4Rn	ITS1	CCCTAAAATTCACGTCAATTCTCGAGTG
025_Tr_ITS1_4Rn	ITS1	TGCAAGAATTCACGTCAATTCTCGAGTG
026_Tr_ITS1_4Rn	ITS1	GACCAGAATTCACGTCAATTCTCGAGTG

027_Tr_ITS1_4Rn	ITS1	ACCTAGAATTCACGTCAATTCTCGAGTG
028_Tr_ITS1_4Rn	ITS1	CTCGAGAATTCACGTCAATTCTCGAGTG
029_Tr_ITS1_4Rn	ITS1	TACCATAATTCACGTCAATTCTCGAGTG
030_Tr_ITS1_4Rn	ITS1	GCCTATAATTCACGTCAATTCTCGAGTG
031_Tr_ITS1_4Rn	ITS1	ATCGATAATTCACGTCAATTCTCGAGTG
032_Tr_ITS1_4Rn	ITS1	CGCAATAATTCACGTCAATTCTCGAGTG
034_Tr_ITS1_4Rn	ITS1	ATCAGCAATTCACGTCAATTCTCGAGTG
036_Tr_ITS1_4Rn	ITS1	TACTGCAATTCACGTCAATTCTCGAGTG
037_Tr_ITS1_4Rn	ITS1	GTCAGAAATTCACGTCAATTCTCGAGTG
038_Tr_ITS1_4Rn	ITS1	AGCCGAAATTCACGTCAATTCTCGAGTG
039_Tr_ITS1_4Rn	ITS1	CACTGAAATTCACGTCAATTCTCGAGTG
040_Tr_ITS1_4Rn	ITS1	TCCGGAAATTCACGTCAATTCTCGAGTG
042_Tr_ITS1_4Rn	ITS1	AACTGGAATTCACGTCAATTCTCGAGTG
044_Tr_ITS1_4Rn	ITS1	TTCAGGAATTCACGTCAATTCTCGAGTG
045_Tr_ITS1_4Rn	ITS1	GACTGTAATTCACGTCAATTCTCGAGTG
046_Tr_ITS1_4Rn	ITS1	ACCGGTAATTCACGTCAATTCTCGAGTG
047_Tr_ITS1_4Rn	ITS1	CTCAGTAATTCACGTCAATTCTCGAGTG
048_Tr_ITS1_4Rn	ITS1	TGCCGTAATTCACGTCAATTCTCGAGTG
049_Tr_ITS1_4Rn	ITS1	ACCATCAATTCACGTCAATTCTCGAGTG
050_Tr_ITS1_4Rn	ITS1	CTCCTCAATTCACGTCAATTCTCGAGTG
051_Tr_ITS1_4Rn	ITS1	TGCTTCAATTCACGTCAATTCTCGAGTG
052_Tr_ITS1_4Rn	ITS1	GACGTCAATTCACGTCAATTCTCGAGTG
053_Tr_ITS1_4Rn	ITS1	ATCCTAAATTCACGTCAATTCTCGAGTG
054_Tr_ITS1_4Rn	ITS1	CGCTTAAATTCACGTCAATTCTCGAGTG
055_Tr_ITS1_4Rn	ITS1	TACGTAAATTCACGTCAATTCTCGAGTG
056_Tr_ITS1_4Rn	ITS1	GCCATAAATTCACGTCAATTCTCGAGTG
057_Tr_ITS1_4Rn	ITS1	AGCTTGAATTCACGTCAATTCTCGAGTG
058_Tr_ITS1_4Rn	ITS1	CACGTGAATTCACGTCAATTCTCGAGTG
059_Tr_ITS1_4Rn	ITS1	TCCATGAATTCACGTCAATTCTCGAGTG
060_Tr_ITS1_4Rn	ITS1	GTCCTGAATTCACGTCAATTCTCGAGTG
061_Tr_ITS1_4Rn	ITS1	AACGTTAATTCACGTCAATTCTCGAGTG
062_Tr_ITS1_4Rn	ITS1	CCCATTAATTCACGTCAATTCTCGAGTG
063_Tr_ITS1_4Rn	ITS1	TTCCTTAATTCACGTCAATTCTCGAGTG
066_Tr_ITS1_4Rn	ITS1	GGATCCAATTCACGTCAATTCTCGAGTG
067_Tr_ITS1_4Rn	ITS1	AAAGCCAATTCACGTCAATTCTCGAGTG
068_Tr_ITS1_4Rn	ITS1	CCAACCAATTCACGTCAATTCTCGAGTG
069_Tr_ITS1_4Rn	ITS1	TGATCAAATTCACGTCAATTCTCGAGTG
070_Tr_ITS1_4Rn	ITS1	GAAGCAAATTCACGTCAATTCTCGAGTG
071_Tr_ITS1_4Rn	ITS1	ACAACAAATTCACGTCAATTCTCGAGTG
072_Tr_ITS1_4Rn	ITS1	CTACCAAATTCACGTCAATTCTCGAGTG
073_Tr_ITS1_4Rn	ITS1	TAAGCGAATTCACGTCAATTCTCGAGTG
074_Tr_ITS1_4Rn	ITS1	GCAACGAATTCACGTCAATTCTCGAGTG
075_Tr_ITS1_4Rn	ITS1	ATACCGAATTCACGTCAATTCTCGAGTG

076_Tr_ITS1_4Rn	ITS1	CGATCGAATTCACGTCAATTCTCGAGTG
077_Tr_ITS1_4Rn	ITS1	TCAACTAATTCACGTCAATTCTCGAGTG
078_Tr_ITS1_4Rn	ITS1	GTACCTAATTCACGTCAATTCTCGAGTG
079_Tr_ITS1_4Rn	ITS1	AGATCTAATTCACGTCAATTCTCGAGTG
080_Tr_ITS1_4Rn	ITS1	CAAGCTAATTCACGTCAATTCTCGAGTG
081_Tr_ITS1_4Rn	ITS1	GTATACAATTCACGTCAATTCTCGAGTG
082_Tr_ITS1_4Rn	ITS1	AGAGACAATTCACGTCAATTCTCGAGTG
084_Tr_ITS1_4Rn	ITS1	TCACACAATTCACGTCAATTCTCGAGTG
085_Tr_ITS1_4Rn	ITS1	GGAGAAAATTCACGTCAATTCTCGAGTG
087_Tr_ITS1_4Rn	ITS1	CCACAAAATTCACGTCAATTCTCGAGTG
090_Tr_ITS1_4Rn	ITS1	ACACAGAATTCACGTCAATTCTCGAGTG
091_Tr_ITS1_4Rn	ITS1	CTATAGAATTCACGTCAATTCTCGAGTG

Bibliography

Abrahamsen, M. S., et al. (2004), 'Complete genome sequence of the apicomplexan, Cryptosporidium parvum', *Science,* 304 (5669), 441-5.

Adams, E. J., et al. (1994), 'Physical-Activity and Growth of Kenyan School-Children with Hookworm, Trichuris-Trichiura and Ascaris-Lumbricoides Infections Are Improved after Treatment with Albendazole', Journal of Nutrition, 124 (8), 1199-206.

Al-Delaimy, A. K., et al. (2014), 'Epidemiology of Intestinal Polyparasitism among Orang Asli School Children in Rural Malaysia', *PLoS Negl Trop Dis,* 8 (8), e3074.

Allentoft, M. E., et al. (2012), 'The half-life of DNA in bone: measuring decay kinetics in 158 dated fossils', *Proc Biol Sci,* 279 (1748), 4724-33.

Aitken, M. J. (1999), 'Archaeological dating using physical phenomena', *Reports on Progress in Physics,* 62 (9), 1333-76.

Altschul, S. F., et al. (1990), 'Basic local alignment search tool', *J Mol Biol,* 215 (3), 403-10.

Anastasiou, E. and Mitchell, P. D. (2013), 'Palaeopathology and genes: Investigating the genetics of infectious diseases in excavated human skeletal remains and mummies from past populations', *Gene,* 528 (1), 33-40.

Ankarklev, J., et al. (2010), 'Behind the smile: cell biology and disease mechanisms of Giardia species', *Nat Rev Microbiol,* 8 (6), 413-22.

Anthony, J. P., Fyfe, L., and Smith, H. (2005), 'Plant active components - a resource for antiparasitic agents?', *Trends Parasitol,* 21 (10), 462-8.

Anthony, J. P., et al. (2011), 'Differential effectiveness of berry polyphenols as anti-giardial agents', *Parasitology,* 138 (9), 1110-6.

Anthony, J. P., et al. (2007), 'The effect of blueberry extracts on Giardia duodenalis viability and spontaneous excystation of Cryptosporidium parvum oocysts, in vitro', *Methods,* 42 (4), 339-48.

Appleby, J., et al. (2014), 'Perimortem trauma in King Richard III: a skeletal analysis', *Lancet*.

Araujo, A., et al. (2013), 'Paleoparasitology: the origin of human parasites', *Arquivos De Neuro-Psiquiatria,* 71 (9B), 722-26.

Areekul, P., et al. (2010), '*Trichuris vulpis* and *T. trichiura* infections among schoolchildren of a rural community in northwestern Thailand: the possible role of dogs in disease transmission', *Asian Biomedicine,* 4 (1), 49-60.

Arizono, N., et al. (2009), 'Diphyllobothriasis Associated with Eating Raw Pacific Salmon', *Emerging Infectious Diseases,* 15 (6), 866-70.

Aspock, H., Flamm, H., and Picher, O. (1973), '[Intestinal parasites in human excrements from prehistoric salt-mines of the Hallstatt period (800-350 B.C.)]', *Zentralbl Bakteriol Orig A,* 223 (4), 549-58.

Baccino, E., et al. (1999), 'Evaluation of seven methods of estimating age at death from mature human skeletal remains', *J Forensic Sci,* 44 (5), 931-6.

Beck, H. P., et al. (2009), 'Molecular approaches to diversity of populations of apicomplexan parasites', *Int J Parasitol,* 39 (2), 175-89.

Betson, M., et al. (2014), 'Molecular epidemiology of ascariasis: a global perspective on the transmission dynamics of ascaris in people and pigs', *J Infect Dis,* 210 (6), 932-41.

Betson, M., et al. (2012), 'Genetic diversity of Ascaris in southwestern Uganda', *Trans R Soc Trop Med Hyg,* 106 (2), 75-83.

Binladen, J., et al. (2006), 'Assessing the fidelity of ancient DNA sequences amplified from nuclear genes', *Genetics,* 172 (2), 733-41.

Blake, D. P., Smith, A. L., and Shirley, M. W. (2003), 'Amplified fragment length polymorphism analyses of Eimeria spp.: an improved process for genetic studies of recombinant parasites', *Parasitol Res,* 90 (6), 473-5.

Blouin, M. S. (1998), 'Mitochondrial DNA diversity in nematodes', *J Helminthol,* 72 (4), 285-9.

Brinkkemper, O. and van Haaster, H. (2012), 'Eggs of intestinal parasites whipworm (Trichuris) and mawworm (Ascaris): Non-pollen palynomorphs in archaeological samples', *Review of Palaeobotany and Palynology,* 186, 16-21.

Bruns, Friedrich and Brandt, Ahasver von (1953), *Das Frachtherrenbuch der Lübecker Bergenfahrer* (Bergen: J. Griegs Boktrykkeri) 79 p.

Bobes, R. J., et al. (2014), 'Evolution, molecular epidemiology and perspectives on the research of taeniid parasites with special emphasis on Taenia solium', *Infection Genetics and Evolution,* 23, 150-60.

Boucher, C. E. (1938), 'The Black Death in Bristol', *Transactions of the Bristol and Gloucestershire Archaeological Society,* 60, 31-46.

Bos, K. I., et al. (2011), 'A draft genome of Yersinia pestis from victims of the Black Death', *Nature,* 478 (7370), 506-10.

Bouwman, A. S. and Brown, T. A. (2005), 'The limits of biomolecular palaeopathology: ancient DNA cannot be used to study venereal syphilis', *Journal of Archaeological Science,* 32 (5), 703-13.

Brown, W. M., et al. (1982), 'Mitochondrial DNA sequences of primates: tempo and mode of evolution', *J Mol Evol,* 18 (4), 225-39.

Bystrykh, L. V. (2012), 'Generalized DNA barcode design based on Hamming codes', *PLoS One,* 7 (5), e36852.

Callejon, R., et al. (2010), 'Molecular evolution of Trichuris muris isolated from different Muridae hosts in Europe', *Parasitology Research,* 107 (3), 631-41.

Campbell, A. T., Robertson, L. J., and Smith, H. V. (1992), 'Viability of Cryptosporidium parvum oocysts: correlation of in vitro excystation with inclusion or exclusion of fluorogenic vital dyes', *Appl Environ Microbiol,* 58 (11), 3488-93.

Carus-Wilson, E. M. (1967a), *Medieval merchant venturers : collected studies* (2nd edn., University paperbacks; London: Methuen) xxxvi, 314 p.

Carus-Wilson, E. M. (1967b), *The overseas trade of Bristol in the later Middle Ages* (New York: Barnes & Noble) xii, 338 p.

Chammartin, F., et al. (2013), 'Soil-transmitted helminth infection in South America: a systematic review and geostatistical meta-analysis', *Lancet Infect Dis,* 13 (6), 507-18.

Cooman, K., et al. (2003), 'Tannery wastewater characterization and toxicity effects on Daphnia spp.', *Environmental Toxicology,* 18 (1), 45-51.

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

Cox, Francis E. G. (1993), *Modern parasitology : a textbook of parasitology* (2nd edn.; Oxford: Blackwell Scientific) xii, 276 p.

Craig, O. E., et al. (2013), 'Evaluating marine diets through radiocarbon dating and stable isotope analysis of victims of the AD79 eruption of Vesuvius', *Am J Phys Anthropol,* 152 (3), 345-52.

Crowther, J. (2002), 'The experimental earthwork at Wareham, Dorset after 33 years: Retention and leaching of phosphate released in the decomposition of buried bone', *Journal of Archaeological Science,* 29 (4), 405-11.

Crooks, G. E., et al. (2004), 'WebLogo: a sequence logo generator', *Genome Res,* 14 (6), 1188-90.

Cutillas, C., et al. (2009), 'Trichuris suis and Trichuris trichiura are different nematode species', *Acta Trop,* 111 (3), 299-307.

Dabney, J., et al. (2013), 'Complete mitochondrial genome sequence of a Middle Pleistocene cave bear reconstructed from ultrashort DNA fragments', *Proc Natl Acad Sci U S A,* 110 (39), 15758-63.

Diawara, A., et al. (2009), 'Assays to detect beta-tubulin codon 200 polymorphism in Trichuris trichiura and Ascaris lumbricoides', *PLoS Negl Trop Dis,* 3 (3), e397.

Dirkmaat, Dennis (2012) *A companion to forensic anthropology* [online], Wiley-Blackwell, http://www.oxford.eblib.com/patron/FullRecord.aspx?p=877782

Drancourt, M., et al. (1998), 'Detection of 400-year-old Yersinia pestis DNA in human dental pulp: an approach to the diagnosis of ancient septicemia', *Proc Natl Acad Sci U S A,* 95 (21), 12637-40.

Dupuy, C., et al. (2014), 'Prevalence of Taenia saginata cysticercosis in French cattle in 2010', *Veterinary Parasitology,* 203 (1-2), 65-72.

Edgar, R. C. (2010), 'Search and clustering orders of magnitude faster than BLAST', *Bioinformatics,* 26 (19), 2460-1.

Fehring, Günter P. and Gläser, Manfred (1991), 'Beiträge zur Bevölkerung Lübecks, zu ihrer Ernährung und Umwelt', *Lübecker Schriften zur Archäologie und Kulturgeschichte,* 21, 183-202.

Fornaciari, G., et al. (2010), 'Malaria was "the killer" of Francesco I de' Medici (1531-1587)', *Am J Med,* 123 (6), 568-9.

Foth, B. J., et al. (2014), 'Whipworm genome and dual-species transcriptome analyses provide molecular insights into an intimate host-parasite interaction', *Nature Genetics,* 46 (7), 693-700.

Franzen, O., et al. (2009), 'Draft genome sequencing of giardia intestinalis assemblage B isolate GS: is human giardiasis caused by two different species?', *PLoS Pathog,* 5 (8), e1000560.

Franklin, D. (2010), 'Forensic age estimation in human skeletal remains: Current concepts and future directions', *Legal Medicine,* 12 (1), 1-7.

Frias, L., Leles, D., and Araujo, A. (2013), 'Studies on protozoa in ancient remains--a review', *Mem Inst Oswaldo Cruz,* 108 (1), 1-12.

Fruergaard, J. and Moltsen, A.S.A. (2005), 'Latrinen. I Viborg Søndersø 1018-1030. ', in M et al. Iversen (ed.), *Viborg Søndersø 1018-1030, Arkæologi og naturvidenskab i et værkstedsområde fra vikingetid* (Højbjerg), 117-27.

Fukita, Y., Asaki, T., and Katakura, Y. (2014), 'Some Like It Raw: An Unwanted Result of a Sushi Meal', *Gastroenterology,* 146 (5), E8-E9.

Fuller, C. W., et al. (2009), 'The challenges of sequencing by synthesis', *Nat Biotechnol,* 27 (11), 1013-23.

Furholt, Martin (2004), 'Entstehungsprozesse der Schnurkeramik und das Konzept eines Einheitshorizontes', in Römisch-Germanisches Zentralmuseum Mainz., Nordwestdeutscher Verband für Altertumsforschung., and West- und Süddeutscher Verband für Altertumsforschung. (eds.), *Archäologisches Korrespondenzblatt* (4; Mainz am Rhein,: Verlag P. von Zabern.), v.

Garvin, Heather M., et al. (2012), 'Developments in Forensic Anthropology: Age-at-Death Estimation ', in Dennis Dirkmaat (ed.), *A companion to forensic anthropology. Blackwell companions to anthropology* (Blackwell Publisher), 202-23.

Gilbert, M. T., et al. (2007), 'Whole-genome shotgun sequencing of mitochondria from ancient hair shafts', *Science,* 317 (5846), 1927-30.

Gläser, Manfred, et al. (1999), *Lübecker Kolloquium zur Stadtarchäologie im Hanseraum II : der Handel* (Lübeck: Schmidt-R©œmhild) 624 p.

Grinberg, A., et al. (2013), 'Extensive intra-host genetic diversity uncovered in Cryptosporidium parvum using Next Generation Sequencing', *Infect Genet Evol,* 15, 18-24.

Grøn, Ole, Krag, Anne Hedeager, and Bennike, Pia (1994), *Vikingetidsgravpladser på Langeland* (Langelands Museum, Rudkøbing).

Gungoren, B., et al. (2007), 'Effect of hygiene promotion on the risk of reinfection rate of intestinal parasites in children in rural Uzbekistan', *Trans R Soc Trop Med Hyg,* 101 (6), 564-9.

Guo, A. J., et al. (2012), 'Molecular identification of Diphyllobothrium latum and a brief review of diphyllobothriosis in China', *Acta Parasitologica,* 57 (3), 293-96.

Guy, R. A., et al. (2003), 'Real-time PCR for quantification of Giardia and Cryptosporidium in environmental water samples and sewage', *Appl Environ Microbiol,* 69 (9), 5178-85.

Haak, W., et al. (2008), 'Ancient DNA, Strontium isotopes, and osteological analyses shed light on social and kinship organization of the Later Stone Age', *Proc Natl Acad Sci U S A,* 105 (47), 18226-31.

Harrison, Mark (2012), *Contagion : how commerce has spread disease* (New Haven, Conn. ; London: Yale University Press) xviii, 376 p. [30] p. of plates.

Henderson, R. C., Lee-Thorp, J., and Loe, L. (2014), 'Early life histories of the London poor using delta13C and delta15N stable isotope incremental dentine sampling', *Am J Phys Anthropol,* 154 (4), 585-93.

Herrmann, Bernd and Schulz, Ursula (1986), 'Parasitologische Untersuchungen eines spätmittelalterlich-frühneuzeitlichen Kloakeninhaltes aus der Fronerei auf dem Schrangen in Lübeck', in Günter P. Fehring (ed.), *Lübecker Schriften zur Archäologie und Kulturgeschichte 12, 1986.* (12; Bonn: Verlag Dr. Rudolf Habelt).

Higuchi, R., et al. (1984), 'DNA-Sequences from the Quagga, an Extinct Member of the Horse Family', *Nature,* 312 (5991), 282-84.

Hjermind, Jesper, Kristensen, Hans Krongaard, and Iversen, Mette (1998), *Viborg Søndersø 1000-1300 : byarkæologiske undersøgelser 1981 og 1984-85* (Jutland Archaeological Society publications; Højbjerg: Jysk arkæologisk selskab, i kommission hos Aarhus Universitetsforlag) 372 p.

Hoberg, E. P. (2006), 'Phylogeny of Taenia: Species definitions and origins of human parasites', *Parasitol Int,* 55 Suppl, S23-30.

Hofreiter, M., et al. (2001), 'Ancient DNA', *Nat Rev Genet,* 2 (5), 353-9.

Hoss, M., et al. (1996), 'DNA damage and DNA sequence retrieval from ancient tissues', *Nucleic Acids Res,* 24 (7), 1304-7.

Humphrey, L. T., et al. (2014), 'Earliest evidence for caries and exploitation of starchy plant foods in Pleistocene hunter-gatherers from Morocco', *Proc Natl Acad Sci U S A,* 111 (3), 954-9.

Jex, A. R., et al. (2011), 'Ascaris suum draft genome', *Nature,* 479 (7374), 529-33.

Kaiser, C., et al. (2008), 'Molecular study of time dependent changes in DNA stability in soil buried skeletal residues', *Forensic Sci Int,* 177 (1), 32-6.

Kang, S., et al. (2009), 'The mitochondrial genome sequence of Enterobius vermicularis (Nematoda: Oxyurida)--an idiosyncratic gene order and phylogenetic information for chromadorean nematodes', *Gene,* 429 (1-2), 87-97.

Kuchta, R., et al. (2013), 'Tapeworm Diphyllobothrium dendriticum (Cestoda)-Neglected or Emerging Human Parasite?', *Plos Neglected Tropical Diseases,* 7 (12).

Lafferty, K. D. (1997), 'Environmental parasitology: What can parasites tell us about human impacts on the environment?', *Parasitology Today,* 13 (7), 251-55.

Lawson, T., et al. (2000), 'The experimental earthwork at Wareham, Dorset after 33 years, 3, Interaction of soil organisms with buried materials', *Journal of Archaeological Science,* 27 (4), 273-85.

Le Bailly, M., et al. (2005), 'Diphyllobothrium: Neolithic parasite?', *J Parasitol,* 91 (4), 957-9.

Lee-Thorp, J. A., et al. (2010), 'Stable isotopes in fossil hominin tooth enamel suggest a fundamental dietary shift in the Pliocene', *Philos Trans R Soc Lond B Biol Sci,* 365 (1556), 3389-96.

Leles, D., et al. (2012), 'Are Ascaris lumbricoides and Ascaris suum a single species?', *Parasit Vectors,* 5, 42.

Lesnikov, Michail P. and Veckinchusen, Hildebrand (1973), *Die Handelsbücher des Hansischen Kaufmannes Veckinchusen* (Forschungen zur mittelalterlichen Geschichte,; Berlin: Akademie-Verlag) xl, 560 p., plates (facsims.).

Lian, O. B. and Roberts, R. G. (2006), 'Dating the Quaternary: progress in luminescence dating of sediments', *Quaternary Science Reviews,* 25 (19-20), 2449-68.

Lin, A. H., et al. (2014), 'An Extended Study of Seroprevalence of Anti-Anisakis simplex IgE Antibodies in Norwegian Blood Donors', *Scandinavian Journal of Immunology,* 79 (1), 61-67.

Lloyd, T. H. (1991), *England and the German Hanse, 1157-1611 : a study of their trade and commercial diplomacy* (Cambridge: Cambridge University Press) ix, 401 p.

Loreille, O., et al. (2001), 'Ancient DNA from Ascaris: extraction amplification and sequences from eggs collected in coprolites', *Int J Parasitol,* 31 (10), 1101-6.

Lowe, D. J. (2011), 'Tephrochronology and its application: A review', *Quaternary Geochronology,* 6 (2), 107-53.

Lutterbeck, Michael (2002), *Der Rat der Stadt Lübeck im 13. und 14. Jahrhundert : politische, personale und wirtschaftliche Zusammenhänge in einer*

städtischen Führungsgruppe (Veröffentlichungen zur Geschichte der Hansestadt Lübeck. Reihe B; Lübeck: Schmidt-Röpmhild) x, 491 p.

Macháček, Jiří and Ungerman, Šimon (2011), *Frühgeschichtliche Zentralorte in Mitteleuropa* (Studien zur Archäologie Europas; Bonn: Habelt Verlag) 689 pages.

Madigan, Michael T., Martinko, John M., and Brock, Thomas D. (2006), *Brock Biology of microorganisms* (11th edn.; Upper Saddle River, NJ: Pearson Prentice Hall) xxv, 992 p.

Mahbubani, M. H., et al. (1991), 'Detection of Giardia cysts by using the polymerase chain reaction and distinguishing live from dead cysts', *Appl Environ Microbiol,* 57 (12), 3456-61.

Markell, Edward K., John, David T., and Krotoski, Wojciech A. (1999), *Markell and Voge's medical parasitology* (8th edn.; Philadelphia ; London: Saunders) viii, 501 p.

Mays, S. and Taylor, G. M. (2003), 'First prehistoric case of tuberculosis from Britain', *International Journal of Osteoarchaeology,* 13 (4), 189-96.

Mays, S., Crane-Kramer, G., and Bayliss, A. (2003), 'Two probable cases of treponemal disease of Medieval date from England', *Am J Phys Anthropol,* 120 (2), 133-43.

McManus, D. P. (2006), 'Molecular discrimination of taeniid cestodes', *Parasitology International,* 55, S31-S37.

Meĭen, Sergeĭ Viktorovich (1987), *Fundamentals of palaeobotany* (London: Chapman and Hall) xxi, 432 p.

Meisel, J. L., et al. (1976), 'Overwhelming watery diarrhea associated with a cryptosporidium in an immunosuppressed patient', *Gastroenterology,* 70 (6), 1156-60.

Meyer, C., et al. (2009), 'The Eulau eulogy: Bioarchaeological interpretation of lethal violence in Corded Ware multiple burials from Saxony-Anhalt, Germany', *Journal of Anthropological Archaeology,* 28 (4), 412-23.

Meyer, M., et al. (2013), 'A mitochondrial genome sequence of a hominin from Sima de los Huesos', *Nature*.

Milner, George R. and Boldsen, Jesper L. (2012), 'Skeletal Age Estimation: Where We Are and Where We Should Go ', in Dennis Dirkmaat (ed.), *A companion*

to forensic anthropology. Blackwell companions to anthropology (Blackwell Publisher), 224-38.

Mitchell, P. D., et al. (2013), 'The intestinal parasites of King Richard III', *Lancet,* 382 (9895), 888.

Nedkvitne, Arnved (2014), *The German Hansa and Bergen 1100-1600* (Quellen und Darstellungen zur hansischen Geschichte) 785 pages.

Nezamabadi, M., et al. (2013), 'Identification of Taenia sp. in a natural human mummy (third century BC) from the Chehrabad salt mine in Iran', *J Parasitol,* 99 (3), 570-2.

Nime, F. A., et al. (1976), 'Acute Enterocolitis in a Human Being Infected with Protozoan Cryptosporidium', *Gastroenterology,* 70 (4), 592-98.

Oh, C. S., et al. (2010), 'Amplification and sequencing of Trichuris trichiura ancient DNA extracted from archaeological sediments', *Journal of Archaeological Science,* 37 (6), 1269-73.

Olsen, A., et al. (2009), 'Albendazole and mebendazole have low efficacy against Trichuris trichiura in school-age children in Kabale District, Uganda', *Transactions of the Royal Society of Tropical Medicine and Hygiene,* 103 (5), 443-46.

Orlando, L., et al. (2013), 'Recalibrating Equus evolution using the genome sequence of an early Middle Pleistocene horse', *Nature,* 499 (7456), 74-8.

Øye, Ingvild (1996), *Bergen und die deutsche Hanse* (Bergen: Bryggens Museum) 86 p.

Paabo, S. (1985), 'Preservation of DNA in Ancient Egyptian Mummies', *Journal of Archaeological Science,* 12 (6), 411-17.

Paap, Norbert (1984), 'Untersuchungen zu Darmparasiten aus der Kloake Gerade Querstrasse 1 zu Lübeck', in Günter P. Fehring (ed.), *Lübecker Schriften zur Archäologie und Kulturgeschichte 8, 1984.* (8; Bonn: Verlag Dr. Rudolf Habelt).

Paul, A. (1990), 'Untersuchungen an Tierknochen aus dem mittelalterlichen Lübeck (Grabung Königsstrasse 59-63)', *Lübecker Schriften zur Archäologie und Kulturgeschichte,* 2, 7-104.

Pudek, N (1990), 'Untersuchungen an Tierknochen des 13. bis 20. Jahrhunderts aus dem Heilig-Geist-Hospital in Lübeck', *Lübecker Schriften zur Archäologie und Kulturgeschichte,* 2, 107-201.

Phongluxa, K., et al. (2013), 'Helminth infection in southern Laos: high prevalence and low awareness', *Parasit Vectors,* 6 (1), 328.

Pilcher, J. R., et al. (1984), 'A 7,272-Year Tree-Ring Chronology for Western-Europe', *Nature,* 312 (5990), 150-52.

Pullan, R. L., et al. (2014), 'Global numbers of infection and disease burden of soil transmitted helminth infections in 2010', *Parasit Vectors,* 7, 37.

Quail, M. A., et al. (2012), 'A tale of three next generation sequencing platforms: comparison of Ion Torrent, Pacific Biosciences and Illumina MiSeq sequencers', *BMC Genomics,* 13, 341.

Raoult, D., et al. (2000), 'Molecular identification by "suicide PCR" of Yersinia pestis as the agent of medieval black death', *Proc Natl Acad Sci U S A,* 97 (23), 12800-3.

Raoult, D., et al. (2006), 'Evidence for louse-transmitted diseases in soldiers of Napoleon's Grand Army in Vilnius', Journal of Infectious Diseases, 193 (1), 112-20.

Richards, J. D. (2005), *The Vikings : a very short introduction* (Oxford: Oxford University Press) 152 p.

Roberts, C. and Manchester, K. (2010), *The Archaeology of Disease* (3 edn.: The History Press) 338.

Roberts, C. and Mays, S. (2010), 'Study and restudy of curated skeletal collections in bioarchaeology: A perspective on the UK and the implications for future curation of human remains', *International Journal of Osteoarchaeology*,

Rudbeck, L., et al. (2005), 'mtDNA analysis of human remains from an early Danish Christian cemetery', *Am J Phys Anthropol,* 128 (2), 424-9.

Ruffer, M. A. (1910), 'Note on the Presence of "Bilharzia Haematobia" in Egyptian Mummies of the Twentieth Dynasty [1250-1000 B.C.]', *Br Med J,* 1 (2557), 16.

Sanger, F., Nicklen, S., and Coulson, A. R. (1977), 'DNA sequencing with chain-terminating inhibitors', *Proc Natl Acad Sci U S A,* 74 (12), 5463-7.

Schlichtherle, Helmut, Aspes, Alessandra, and Verband der Landesarchäologen in der Bundesrepublik Deutschland. (1997), 'Pfahlbauten rund um die Alpen', *Archäologie in Deutschland. Sonderheft,* ([Stuttgart]: Theiss), 131 p.

Schmidlin, T., et al. (2013), 'Effects of hygiene and defecation behavior on helminths and intestinal protozoa infections in Taabo, Cote d'Ivoire', *PLoS One,* 8 (6), e65722

Schneider, T. D. (1996), 'Reading of DNA sequence logos: prediction of major groove binding by information theory', *Methods Enzymol,* 274, 445-55.

Schneider, T. D. and Stephens, R. M. (1990), 'Sequence logos: a new way to display consensus sequences', *Nucleic Acids Res,* 18 (20), 6097-100.

Schoeninger, Margaret J. and Moore, Katherine (1992), 'Bone Stable Isotope Studies In Archaeology', *Journal of World Prehistory,* 6 (2), 50.

Scholz, T., et al. (2009), 'Update on the Human Broad Tapeworm Genus Diphyllobothrium), Including Clinical Relevance', *Clinical Microbiology Reviews,* 22 (1), 146-+.

Schulz, Friedrich N. (1911), *Die Hanse und England, von Eduards iii bis auf Heinrichs viii Zeit* (Abhandl. zur Verkehrs- u. Seegesch.; Berl.).

Seo, M., et al. (2010), 'The influence of differential burial preservation on the recovery of parasite eggs in soil samples from Korean medieval tombs', *J Parasitol,* 96 (2), 366-70.

Sherborne, J. W. (1965), *The port of Bristol in the Middle Ages* (Port of Bristol series; Bristol: Bristol Branch of the Historical Association) 30 p

Shapiro, A. E., et al. (2005), 'Epidemiology of helminth infections and their relationship to clinical malaria in southwest Uganda', *Transactions of the Royal Society of Tropical Medicine and Hygiene,* 99 (1), 18-24.

Slapeta, J. (2013), 'Cryptosporidiosis and Cryptosporidium species in animals and humans: a thirty colour rainbow?', *Int J Parasitol,* 43 (12-13), 957-70.

Szidat, Lothar (1944), 'ÜBER DIE ERHALTUNGSFÄHIGKEIT VON HELMINTHENEIERN IN VOR- UND FRÜHGESCHICHTLICHEN MOORLEICHEN.', *Zeitschrift für Parasitenkunde,* 13 (3), 265-74.

Taylor, G. M., et al. (2000), 'A Mediaeval Case of Lepromatous Leprosy from 13-14th Century Orkney, Scotland', *Journal of Archaeological Science,* 27, 1133-38.

Taylor, G. M., et al. (2007), 'First report of Mycobacterium bovis DNA in human remains from the Iron Age', *Microbiology-Sgm,* 153, 1243-49.

Tsai, I. J., et al. (2013), 'The genomes of four tapeworm species reveal adaptations to parasitism', *Nature,* 496 (7443), 57-63.

Tomaso, H., Dierich, M. P., and Allerberger, F. (2001), 'Helminthic infestations in the Tyrol, Austria', *Clinical Microbiology and Infection,* 7 (11), 639-41.

University of Warwick, Department of History 'Urban society in the Medieval World',

Wicht, B., et al. (2010), 'Multiplex PCR for Differential Identification of Broad Tapeworms (Cestoda: Diphyllobothrium) Infecting Humans', *Journal of Clinical Microbiology,* 48 (9), 3111-16.

Wallace, D. C. (1994), 'Mitochondrial DNA sequence variation in human evolution and disease', *Proc Natl Acad Sci U S A,* 91 (19), 8739-46.

Willerslev, E. and Cooper, A. (2005), 'Ancient DNA', *Proc Biol Sci,* 272 (1558), 3-16.

Willerslev, E., et al. (2003), 'Diverse plant and animal genetic records from Holocene and Pleistocene sediments', *Science,* 300 (5620), 791-5.

Willerslev, E., et al. (2007), 'Ancient biomolecules from deep ice cores reveal a forested southern Greenland', *Science,* 317 (5834), 111-4.

Woodburn, P. W., et al. (2009), 'Risk Factors for Helminth, Malaria, and HIV Infection in Pregnancy in Entebbe, Uganda', *Plos Neglected Tropical Diseases,* 3 (6).

Xiao, L., et al. (2001), 'Identification of 5 types of Cryptosporidium parasites in children in Lima, Peru', *J Infect Dis,* 183 (3), 492-7.

Xiao, L. H. and Ryan, U. M. (2004), 'Cryptosporidiosis: an update in molecular epidemiology', *Current Opinion in Infectious Diseases,* 17 (5), 483-90.

Xu, P., et al. (2004), 'The genome of Cryptosporidium hominis', *Nature,* 431 (7012), 1107-12.

Ziegelbauer, K., et al. (2012), 'Effect of sanitation on soil-transmitted helminth infection: systematic review and meta-analysis', *PLoS Med,* 9 (1), e1001162.

Zimmerling, Dieter (1976), *Die Hanse : Handelsmacht im Zeichen der Kogge* (1. Aufl. edn.; Düsseldorf ; Wien: Econ-Verlag) 405 p.

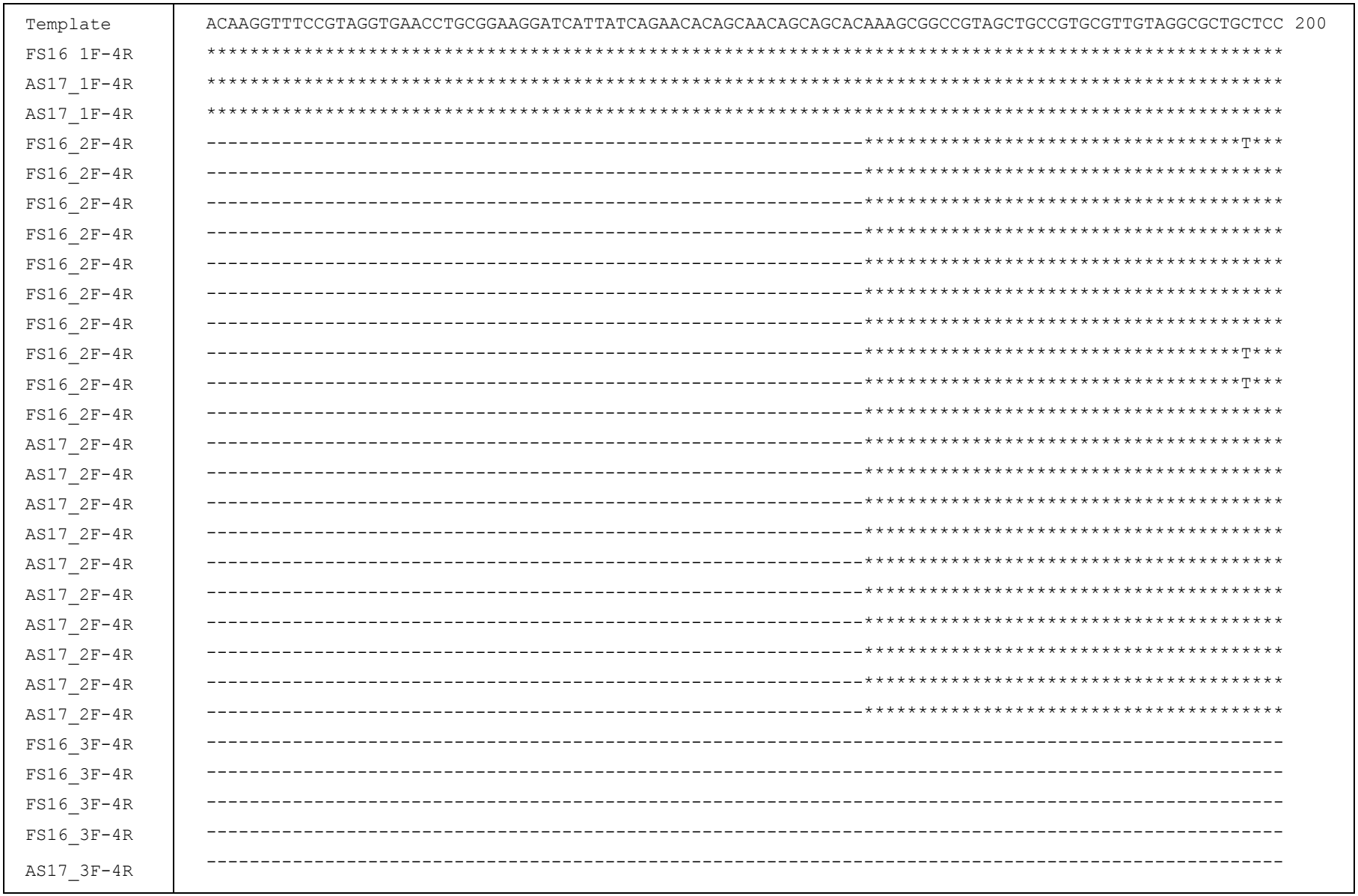
Zintl, A., et al. (2011), 'Longitudinal and spatial distribution of GP60 subtypes in human cryptosporidiosis cases in Ireland', *Epidemiol Infect,* 139 (12), 1945-55.

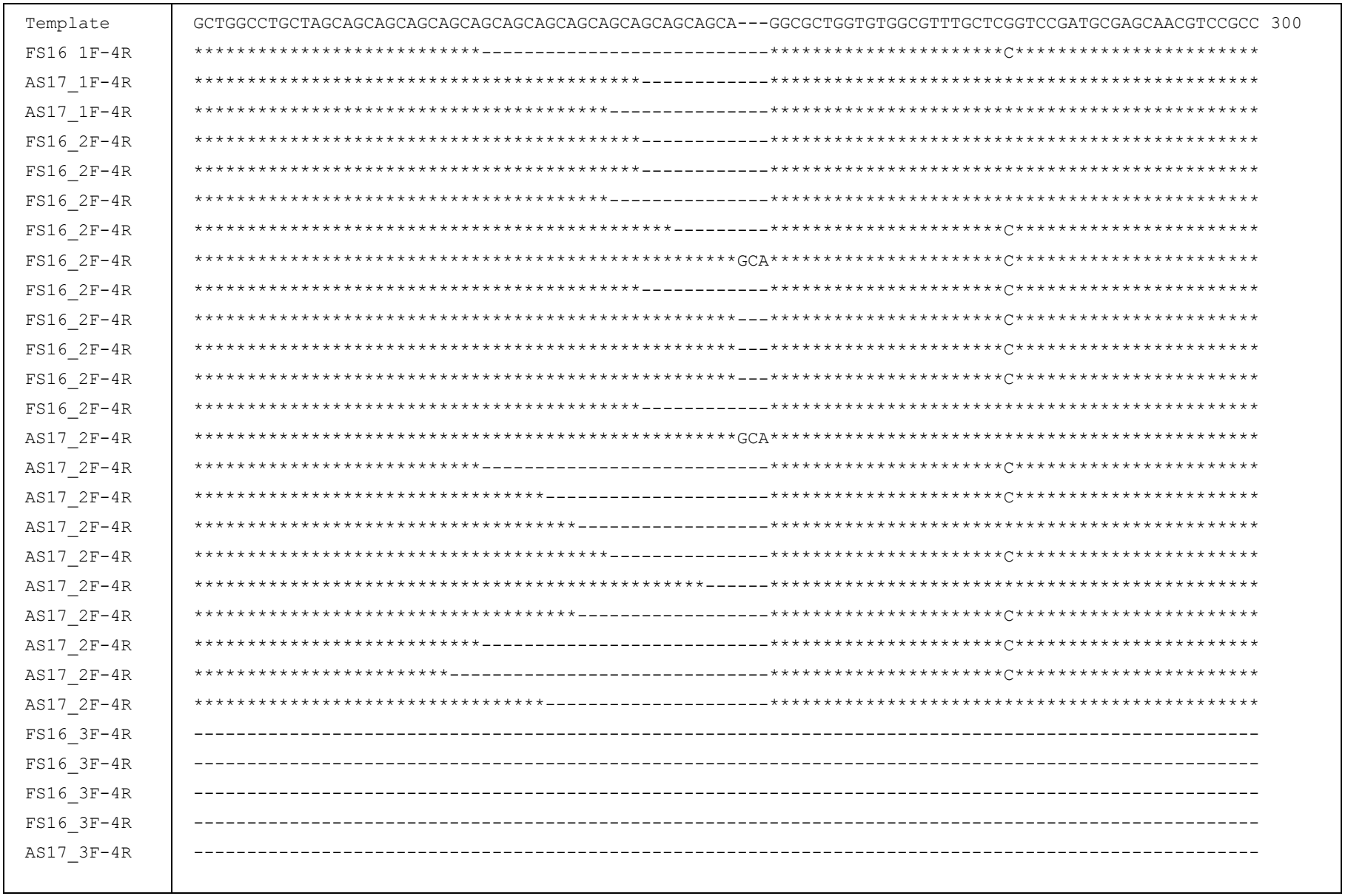
Appendix

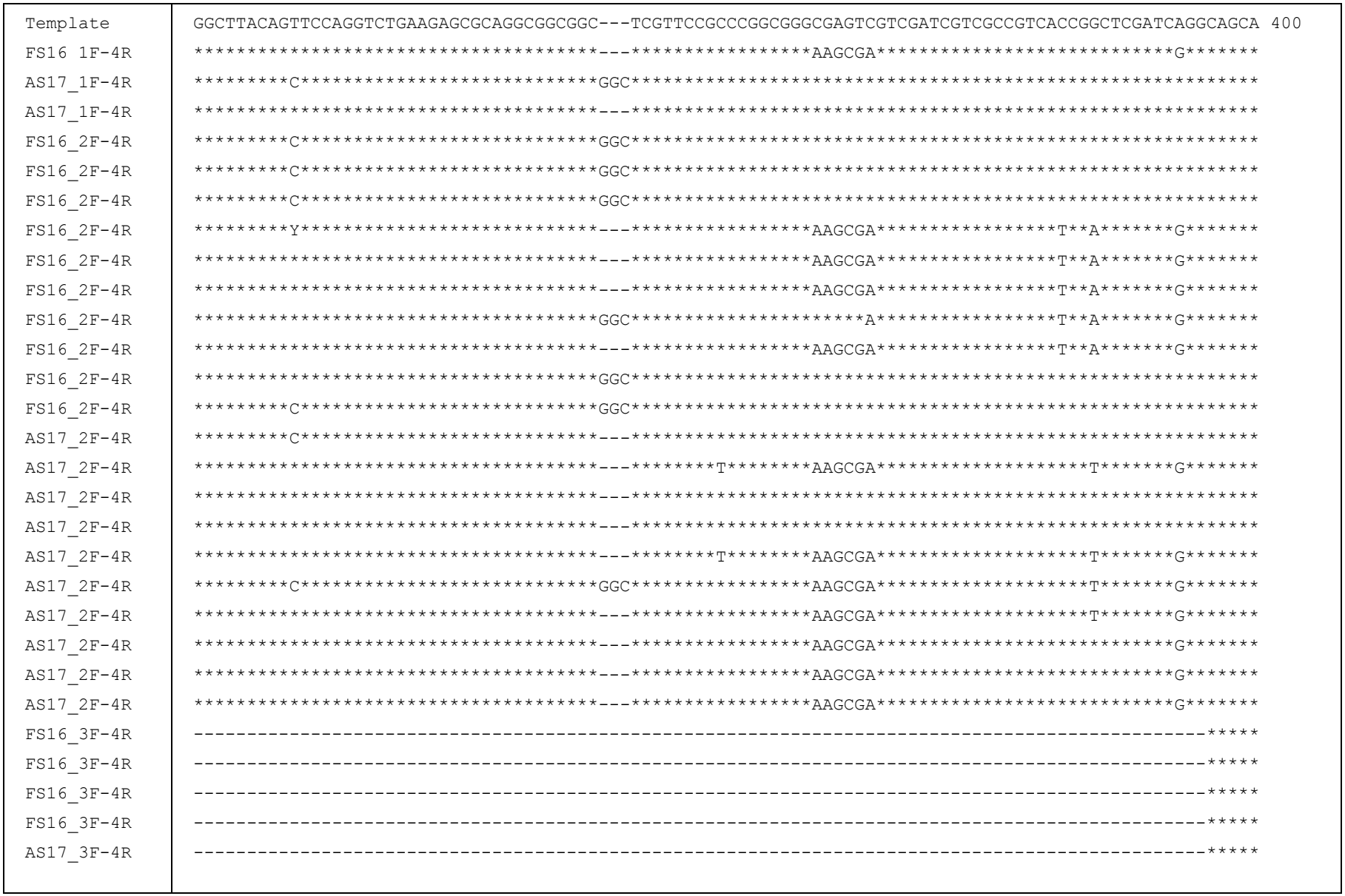
Appendix 1

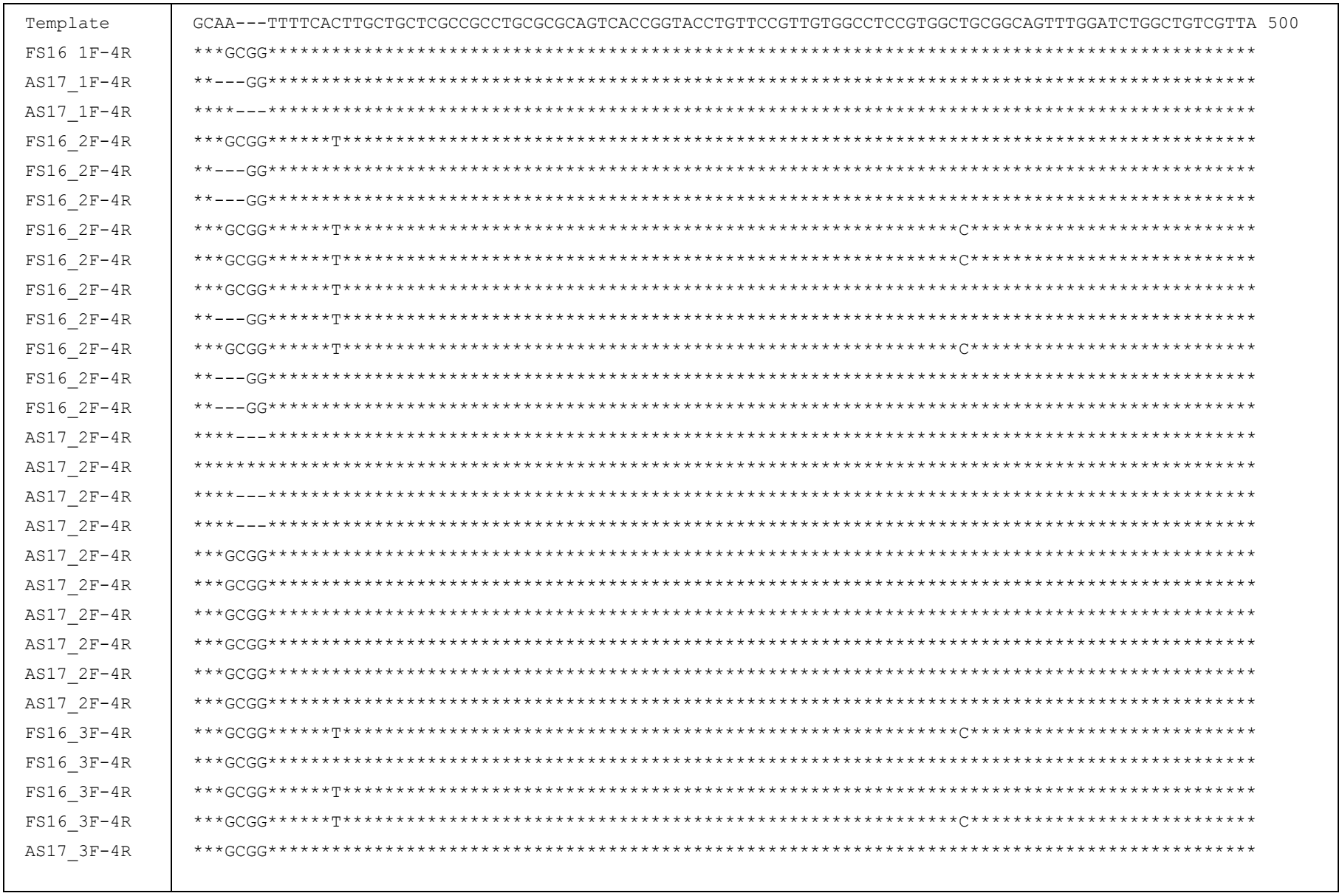
Full alignment of T. trichiura ITS1 from aDNA Lübeck against modern template

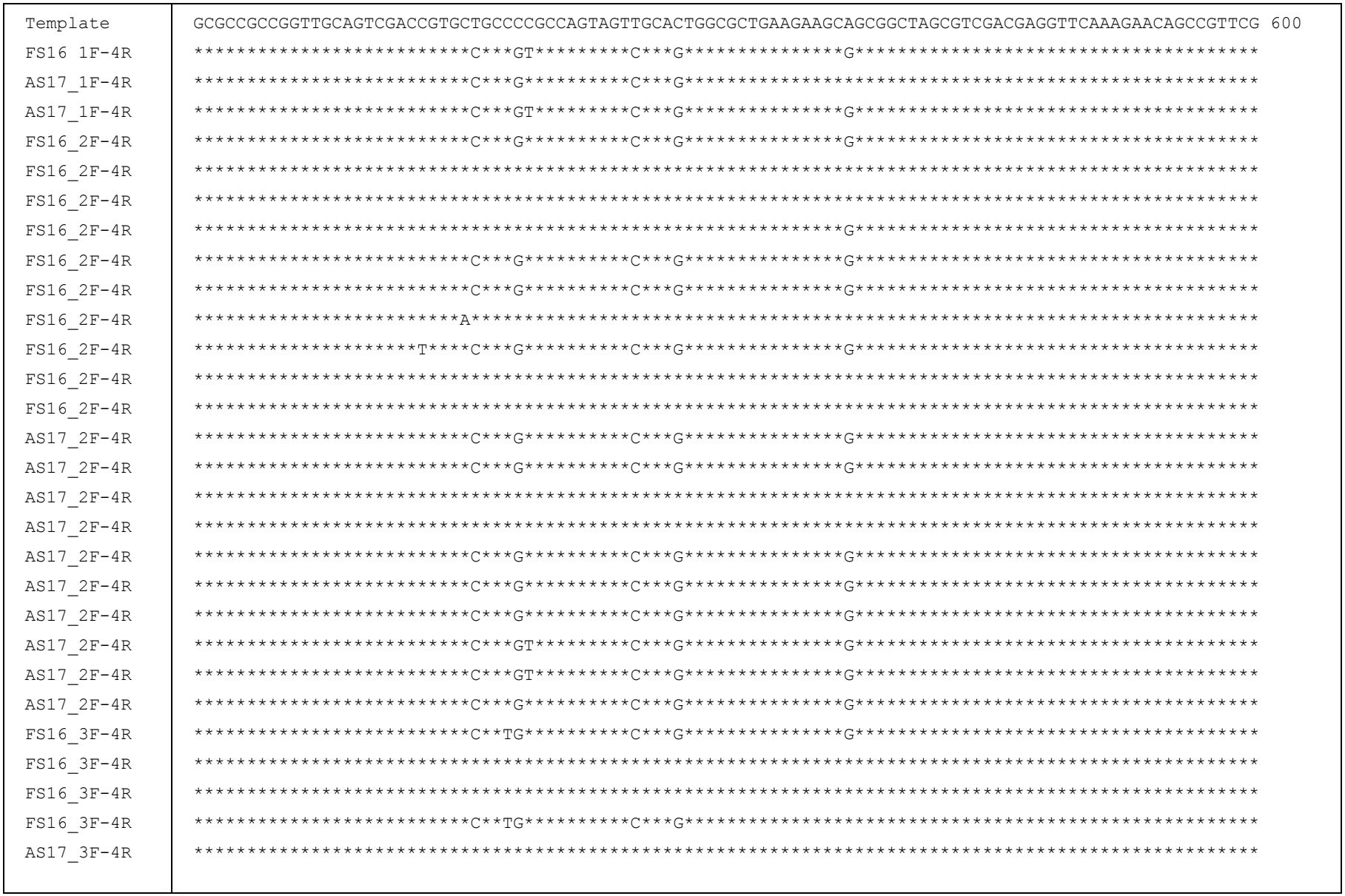
Template	GAGGCCTTTGGACCGATGCCGCCGGTGTTCGGCCGTCAGGTTGAACCACGGCGTTGGAAAACTGGCCGAACCAAGCCATCTAGAGGAAGTAAAAGTCGTA 100
FS16_1F-4R	-**
AS17_1F-4R	-**
AS17_1F-4R	-**
FS16_2F-4R	----------
FS16_2F-4R	----------
FS16_2F-4R	----------
FS16_2F-4R	----------
FS16_2F-4R	----------
FS16_2F-4R	----------
FS16_2F-4R	----------
FS16_2F-4R	----------
FS16_2F-4R	----------
FS16_2F-4R	----------
AS17_2F-4R	----------
AS17_2F-4R	----------
AS17_2F-4R	----------
AS17_2F-4R	----------
AS17_2F-4R	----------
AS17_2F-4R	----------
AS17_2F-4R	----------
AS17_2F-4R	----------
AS17_2F-4R	----------
FS16_3F-4R	----------
FS16_3F-4R	----------
FS16_3F-4R	----------
FS16_3F-4R	----------
AS17_3F-4R	----------


Template
ACAAGGTTTCCGTAGGTGAACCTGCGGAAGGATCATTATCAGAACACAGCAACAGCAGCACAAAGCGGCCGTAGCTGCCGTGCGTTGTAGGCGCTGCTCC 200
FS16_1F-4R
AS17_1F-4R
AS17_1F-4R
FS16_2F-4R
FS16_2F-4R
FS16_2F-4R
FS16_2F-4R
FS16_2F-4R
FS16_2F-4R
FS16_2F-4R
FS16_2F-4R
FS16_2F-4R
FS16_2F-4R
AS17_2F-4R
AS17_2F-4R
AS17_2F-4R
AS17_2F-4R
AS17_2F-4R
AS17_2F-4R
AS17_2F-4R
AS17_2F-4R
AS17_2F-4R
FS16_3F-4R
FS16_3F-4R
FS16_3F-4R
FS16_3F-4R
AS17_3F-4R


Template GCTGGCCTGCTAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCA---GGCGCTGGTGTGGCGTTTGCTCGGTCCGATGCGAGCAACGTCCGCC 300
FS16 1F-4R
AS17_1F-4R
AS17_1F-4R
FS16_2F-4R
FS16_2F-4R
FS16_2F-4R
FS16_2F-4R
FS16_2F-4R
FS16_2F-4R
FS16_2F-4R
FS16_2F-4R
FS16_2F-4R
FS16_2F-4R
AS17_2F-4R
AS17_2F-4R
AS17_2F-4R
AS17_2F-4R
AS17_2F-4R
AS17_2F-4R
AS17_2F-4R
AS17_2F-4R
AS17_2F-4R
FS16_3F-4R
FS16_3F-4R
FS16_3F-4R
FS16_3F-4R
AS17_3F-4R


Template GGCTTACAGTTCCAGGTCTGAAGAGCGCAGGCGGCGGC---TCGTTCCGCCCGGCGGGCGAGTCGTCGATCGTCGCCGTCACCGGCTCGATCAGGCAGCA 400
FS16 1F-4R **********AAGCGA**********G*******
AS17_1F-4R *********C**********GGC**********
AS17_1F-4R **********---**********
FS16_2F-4R *********C**********GGC**********
FS16_2F-4R *********C**********GGC**********
FS16_2F-4R *********C**********GGC**********
FS16_2F-4R *********Y**********---**********AAGCGA**********T**A*******G*******
FS16_2F-4R **********---**********AAGCGA**********T**A*******G*******
FS16_2F-4R **********---**********AAGCGA**********T**A*******G*******
FS16_2F-4R **********GGC**********A**********T**A*******G*******
FS16_2F-4R **********---**********AAGCGA**********T**A*******G*******
FS16_2F-4R **********GGC**********
FS16_2F-4R *********C**********GGC**********
AS17_2F-4R *********C**********---**********
AS17_2F-4R **********---*********T********AAGCGA**********T*******G*******
AS17_2F-4R **********---**********
AS17_2F-4R **********---**********
AS17_2F-4R **********---*********T********AAGCGA**********T*******G*******
AS17_2F-4R *********C**********GGC**********AAGCGA**********T*******G*******
AS17_2F-4R **********---**********AAGCGA**********T*******G*******
AS17_2F-4R **********---**********AAGCGA**********G*******
AS17_2F-4R **********---**********AAGCGA**********G*******
AS17_2F-4R **********---**********AAGCGA**********G*******
FS16_3F-4R ----------*****
FS16_3F-4R ----------*****
FS16_3F-4R ----------*****
FS16_3F-4R ----------*****
AS17_3F-4R ----------*****


Template GCAA---TTTTCACTTGCTGCTCGCCGCCTGCGCGCAGTCACCGGTACCTGTTCCGTTGTGGCCTCCGTGGCTGCGGCAGTTTGGATCTGGCTGTCGTTA 500
FS16 1F-4R ***GCGG**********
AS17_1F-4R **---GG**********
AS17_1F-4R ****---**********
FS16_2F-4R ***GCGG******T**********
FS16_2F-4R **---GG**********
FS16_2F-4R **---GG**********
FS16_2F-4R ***GCGG******T**********C**********
FS16_2F-4R ***GCGG******T**********C**********
FS16_2F-4R ***GCGG******T**********
FS16_2F-4R **---GG******T**********
FS16_2F-4R ***GCGG******T**********C**********
FS16_2F-4R **---GG**********
FS16_2F-4R **---GG**********
AS17_2F-4R ****---**********
AS17_2F-4R **********
AS17_2F-4R ****---**********
AS17_2F-4R ****---**********
AS17_2F-4R ***GCGG**********
AS17_2F-4R ***GCGG**********
AS17_2F-4R ***GCGG**********
AS17_2F-4R ***GCGG**********
AS17_2F-4R ***GCGG**********
AS17_2F-4R ***GCGG**********
FS16_3F-4R ***GCGG******T**********C**********
FS16_3F-4R ***GCGG**********
FS16_3F-4R ***GCGG******T**********
FS16_3F-4R ***GCGG******T**********C**********
AS17_3F-4R ***GCGG**********


Template GCGCCGCCGGTTGCAGTCGACCGTGCTGCCCCGCCAGTAGTTGCACTGGCGCTGAAGAAGCAGCGGCTAGCGTCGACGAGGTTCAAAGAACAGCCGTTCG 600
FS16 1F-4R **********C***GT*********C***G**********G**********
AS17_1F-4R **********C***G*********C***G**********
AS17_1F-4R **********C***GT*********C***G**********G**********
FS16_2F-4R **********C***G*********C***G**********G**********
FS16_2F-4R **********
FS16_2F-4R **********
FS16_2F-4R **********G**********
FS16_2F-4R **********C***G*********C***G**********G**********
FS16_2F-4R **********C***G*********C***G**********G**********
FS16_2F-4R **********A**********
FS16_2F-4R **********T****C***G*********C***G**********G**********
FS16_2F-4R **********
FS16_2F-4R **********
AS17_2F-4R **********C***G*********C***G**********G**********
AS17_2F-4R **********C***G*********C***G**********G**********
AS17_2F-4R **********
AS17_2F-4R **********
AS17_2F-4R **********C***G*********C***G**********G**********
AS17_2F-4R **********C***G*********C***G**********G**********
AS17_2F-4R **********C***G*********C***G**********G**********
AS17_2F-4R **********C***GT*********C***G**********G**********
AS17_2F-4R **********C***GT*********C***G**********G**********
AS17_2F-4R **********C***G*********C***G**********G**********
FS16_3F-4R **********C**TG*********C***G**********G**********
FS16_3F-4R **********
FS16_3F-4R **********
FS16_3F-4R **********C**TG*********C***G**********
AS17_3F-4R **********

Template	ACCTCGAGCATCGAGGACGGCTCGTCGCCGGTTGGAAAAAGAAAAAGCAGCTTCACCAGTGCACAAGGCGTCGTCCTACTACGCACGCCGATGTGCCTTC 700
FS16 1F-4R	**********C**********T*****C********
AS17_1F-4R	**********T*****C********
AS17_1F-4R	**********
FS16_2F-4R	**********T*****C********
FS16_2F-4R	**********T*****C********
FS16_2F-4R	**********T*****C********
FS16_2F-4R	**********T*****C********
FS16_2F-4R	**********T*****C********
FS16_2F-4R	**********T*****C********
FS16_2F-4R	**********T*****C********
FS16_2F-4R	**********T*****C********
FS16_2F-4R	**********T*****C********
FS16_2F-4R	**********
AS17_2F-4R	**********T*****C********
AS17_2F-4R	**********T*****C********
AS17_2F-4R	**********
AS17_2F-4R	**********
AS17_2F-4R	**********T*****C********
AS17_2F-4R	**********T*****C********
AS17_2F-4R	**********T*****C********
AS17_2F-4R	**********
AS17_2F-4R	**********T*****C********
AS17_2F-4R	**********
FS16_3F-4R	**********T*****C********
FS16_3F-4R	**********
FS16_3F-4R	**********
FS16_3F-4R	**********
AS17_3F-4R	**********

Template	GACACGCAGACAC----TGCTGTCGGCGCAATGAGCGCTAGTAGCATTCGAACTGCTGCTGTACGCGCAGGAAAGTACATTCCGAACGGTGGATCACTTG 800
FS16 1F-4R	GACACGCAGACAC----TGCTGTCGGCGCAATGAGCGCTAGTAGCATTCGAACTGCTGCTGTACGCGCAGGAAAGTACATTCCGAACGGTGGATCACTTG
AS17_1F-4R	**********----**********
AS17_1F-4R	**********----**********
FS16_2F-4R	**********----**********
FS16_2F-4R	**********ACAC**********
FS16_2F-4R	**********ACAC**********
FS16_2F-4R	**********ACAC**********
FS16_2F-4R	**********ACAC**********
FS16_2F-4R	**********ACAC**********
FS16_2F-4R	**********ACAC**********
FS16_2F-4R	**********ACAC**********
FS16_2F-4R	**********ACAC**********
FS16_2F-4R	**********ACAC**********
AS17_2F-4R	**********----**********
AS17_2F-4R	**********----**********
AS17_2F-4R	**********----**********
AS17_2F-4R	**********----**********
AS17_2F-4R	**********----**********
AS17_2F-4R	**********----**********
AS17_2F-4R	**********----**********
AS17_2F-4R	**********----**********
AS17_2F-4R	**********----**********
FS16_3F-4R	**********ACAC**********
FS16_3F-4R	**********----**********
FS16_3F-4R	**********----**********
FS16_3F-4R	**********----**********
AS17_3F-4R	**********----**********

Template	GCTCGTAGATCGTTGAAGAACGACGTGACACTCGAGAATTGACGTGAATT	850
FS16 1F-4R	**********	
AS17_1F-4R	**********	
AS17_1F-4R	**********	
FS16_2F-4R	**********	
FS16_2F-4R	**********	
FS16_2F-4R	**********	
FS16_2F-4R	**********	
FS16_2F-4R	**********	
FS16_2F-4R	**********	
FS16_2F-4R	**********	
FS16_2F-4R	**********	
FS16_2F-4R	**********	
FS16_2F-4R	**********	
FS16_2F-4R	**********	
AS17_2F-4R	**********	
AS17_2F-4R	**********	
AS17_2F-4R	**********	
AS17_2F-4R	**********	
AS17_2F-4R	**********	
AS17_2F-4R	**********	
AS17_2F-4R	**********	
AS17_2F-4R	**********	
AS17_2F-4R	**********	
AS17_2F-4R	**********	
FS16_3F-4R	**********	
FS16_3F-4R	**********	
FS16_3F-4R	**********	
FS16_3F-4R	**********	
AS17_3F-4R	**********	

Appendix 2

Detailed list of samples included in thesis

United Kingdom

York

- VY: Viking York (Jorvik), coprolite fragment, c1080 CE, 1 sample
  Sample received from Adrian Smith
- Hg: Anglo-Saxon/Anglo-Scandinavian York (Hungate), cesspit sample, c6th - 8th c CE, 1 sample
  Sample received from Andrew Jones, YAT

Bristol

- BFR: Excavation medieval Finzel's Reach, waste deposit/ditch, c 12th-16th c CE, 26 samples
  Samples received from Rebecca Nicholson, Oxford Archaeology

London

- GBLW: Excavation "The Wheatsheaf", midden deposit/latrine, Medieval deposits, 14 samples
  Samples received from Rebecca Nicholson, Oxford Archaeology

Oxford

- OXMF: East Oxford Minchery Farm, medieval midden deposit, 13th-15th c CE, 1 sample
  Sample received via Rebecca Nicholson, Oxford Archaeology
- OXRI: Oxford Radcliffe Infirmary Graveyard, grave deposits, c1770-1850 CE, 51 samples
  Samples received from Vix Hughes, Oxford Archaeology

Portsmouth

- RNP: Royal Navy graveyard, grave deposits/bone associated soil samples, c16th c CE, 18 samples
  Samples received from Ceri Boston, School of Archaeology, University of Oxford

Ipswich

- UKISP-A: middle Saxon deposits Ipswich, waste deposits, c750-870 CE, 16 samples
  Samples received from Rebecca Nicholson, Oxford Archaeology
- UKISP-Med: Medieval graveyard Ipswich, grave deposits, c14th c CE, 50 samples
  Samples received from Rebecca Nicholson, Oxford Archaeology

Silchester

- RS: Roman latrine deposit, dried latrine sample, c1st c CE, 1 sample
  Sample received from Mark Robinson, Environmental Archaeology, University of Oxford

Switzerland

Zurich

- ZPO: Parking Opéra, Neolithic stilt settlement deposits, 3630-3070 BCE, 15 samples
  Samples received from Niels Bleicher, Hochbauamt der Stadt Zürich

Germany

Lübeck

- HLGV: Gründungsviertel der Hansestadt Lübeck, latrine deposits, c 1200-1650 CE, 33 samples
  Samples received from Dirk Rieger, Stadtarchäologie Lübeck

Herbertingen

- ST: Burial of celtic age noble woman "Swabian Troy", grave sample, c590 BCE, 1 sample
  Sample received from Joachim Wahl, Landesamt für Denkmalpflege Baden-Württemberg and University of Tübingen

Bad Mergentheim

- SDBM: Neolithic (corded ware culture) burials Bad Mergentheim, grave samples, c 2950 BCE, 7 samples
  Samples received from Joachim Wahl, Landesamt für Denkmalpflege Baden-Württemberg and University of Tübingen

Ellwangen-Jagst

- DEEJ: Ellwangen-Jagst, Medieval grave samples, c 1400-1600 CE, 121 samples (of which 29 analysed at submission)
  Sample received from Joachim Wahl, Landesamt für Denkmalpflege Baden-Würtenberg (Konstanz) and University of Tübingen

Region Upper Swabia/Lake Constance

- SDNeo: Collection of Neolithic stilt settlement deposits from 7 sites, c3920-2900 BCE , 19 samples
  Received from Helmut Schlichterle, Landesdenkmalamt Baden-Würtenberg (Memmingen)

Czech Republic

Brno

- CZVB: Brno-Vídeňská, medieval graves, c 14thc CE, 14 samples
  Samples received from Petra Urbanová, Masaryk University, Brno

Břeclav-Pohansko (CZBP)

- CZBP: Slavic (Great Moravian) elite stronghold of Břeclav-Pohansko, grave samples, c850-950 CE, 102 samples (of which 61 analysed on submission)
  Samples received from Renáta Přichystalová, Masaryk University, Brno

Italy
Herculaneum

- ITRH: Sample from the sewage system of Herculaneum, dried sediment sample, c1st c CE, 1 sample
  Sample received from Mark Robinson, Environmental Archaeology, University of Oxford

Denmark
Copenhagen

- VDKKH: Viking settlement near Copenhagen, latrine samples, c1020 CE, 9 samples
  Samples received from Peter Nejsum, University of Copenhagen

Viborg

- VDKVS: Viborg Søndersø temporary Viking settlement, c1030-1040 CE, 4 samples
  Samples received from Annie Moltson, Natur og Kultur, Copenhagen

Langeland

- VDKLB: Langeland Bogøvej Viking burials, grave samples, c10th c CE, 29 samples
  Samples received from Pia Bennike, University of Copenhagen

Norway
Gokstad nedre Sandefjord

- VNOGNS, Viking settlement near Gokstad, settlement/latrine samples, c10th c CE, 28 samples
  Samples received from Jan Bill and Christian Løchsen Rødsrud, Kulturhistorisk Museum, University of Oslo

Latvia
Riga

Various sites, samples received from Gunita Zarina, Latvian University in Riga

Non-European samples
Lesotho

- Cave samples from Lesotho, sediment samples, dated c 17.000-70.000 BCE, 30 samples
  Samples received from Charles Arthur and Peter Mitchell (School of Archaeology, University of Oxford)