Title: Germs on a journey: what pathogens can tell us about population movements and human evolution

Authors: Charlotte J. Houldcroft1,2*, Jean-Baptiste Ramond3, Riaan F. Rifkin3 and Simon J. Underdown4,5.

Affiliations:

1. Division of Biological Anthropology, Department of Archaeology and Anthropology, University of Cambridge, Pembroke Street, Cambridge, CB2 3QG, UK.
2. McDonald Institute of Archaeological Research, University of Cambridge, UK.
3. Centre for Microbial Ecology and Genomics, Genomic Research Institute, Department of Genetics, University of Pretoria, Hatfield, South Africa
4. Human Origins and Palaeoenvironmental Research Group (HOPE), Department of Anthropology & Geography, Oxford Brookes University, Oxford, UK
5. Leverhulme Centre for Human Evolutionary Studies, Henry Wellcome Building, Fitzwilliam Street, Cambridge, UK.

*Corresponding author: ch504@cam.ac.uk

ORCIDs:

CJ Houldcroft orcid.org/0000-0002-1833-5285

JB Ramond orcid.org/0000-0003-4790-6232

RF Rifkind orcid.org/0000-0003-1791-3706

SJ Underdown orcid.org/0000-0001-6056-2353

Biographies of authors:

Charlotte J. Houldcroft is an unestablished lecturer in the Division of Biological Anthropology, University of Cambridge, UK. She received her Ph.D. in molecular biology from the Wellcome Trust Sanger Institute, UK. Her first postdoctoral position was at the Institute of Child Health, University College London, UK, where she researched the role of genomic variation in pathology caused by the three biggest infectious killers of paediatric transplant recipients: Epstein–Barr virus, cytomegalovirus and adenovirus. She also has an interest in which infectious diseases afflicted ancient hominins.

Jean-Baptiste Ramond holds a PhD in microbiology and microbial ecology from the University of Rouen, France. He focuses his research on desert soil microbial ecology and is now using this expertise to study host and pathogen evolution using ancient DNA sequences retrieved from archeologically-relevant South African soils.

Riaan F. Rifkin has a PhD in archaeology from the Evolutionary Studies Institute, University of the Witwatersrand, and has employed diverse experimental and ethnographic research methods to explore the role of symbolic material culture in promoting and maintaining complex social relations amongst early modern southern African *Homo sapiens* societies. His current research focuses on the evolutionary history of human diseases, based on the incidence of arthropod vectors and pathogen DNA in prehistoric sub-Saharan African anthropogenic sediments.

Simon J. Underdown has a PhD in Biological Anthropology from the University of Cambridge and is senior lecturer in Biological Anthropology at Oxford Brookes University. His research centres around human evolution and especially the co-evolution of humans and disease and how aDNA can be used to understand not just the human evolutionary process but also how it has helped to shape what being human means.

Abstract (195/200 words)

The biology of human migration can be observed from our co-evolutionary relationship with infectious diseases. While many pathogens are brief, unpleasant visitors to our bodies, others have the ability to become life-long human passengers. The story of a pathogen's genetic code may therefore provide insight into the history of its human host. The evolution and distribution of disease in Africa is of particular interest, because of the deep history of human evolution in Africa, the presence of a variety of non-human primates, and tropical reservoirs of emerging infectious diseases.

Here, we explore which pathogens leave traces in the archaeological record, and whether there are realistic prospects that these pathogens can be recovered from sub-Saharan African archaeological contexts. We then present three stories of germs on a journey. The first is the story of HIV's spread on the back of colonialism and the railway networks over the last 150 years. The second involves the spread of *Schistosoma mansoni*, a parasite which shares its history with the trans-Atlantic slave trade and the origins of fresh-water fishing. Finally, we discuss the tantalising hints of hominin migration and interaction found in the genome of human herpes simplex virus 2.

Words: 6408

Introduction

There is significant evidence that human populations currently suffer, and may have suffered for millions of years, from infectious diseases shared with or closely related to the infectious diseases of wild primates (Wolfe et al. 2007). These conclusions are drawn largely by studying the phylogenetic relationships of extant pathogens, both those that are exclusively human pathogens, and those we share with other primates. This has led researchers to conclude that many infectious diseases have been co-evolving with humans and our ancestors for millennia (Houldcroft & Underdown 2016). Reviews by (Wolfe et al. 2007; Trueba & Dunthorn 2012; Harkins & Stone 2015; Houldcroft & Underdown 2016) highlight the gaps in our understanding of the origins of diseases, and especially their relationship with human evolution, behaviour and migration in Africa. Besides being the cradle of behaviourally modern *Homo sapiens* (Mourre et al. 2010; Henshilwood et al. 2011; Henshilwood et al. 2009; D'Errico et al. 2012) sub-Saharan Africa also brings together exceptionally rich biodiversity with pathogen abundance (Just et al. 2014). Prehistoric sub-Saharan African populations who inhabited the region over the past 150 000 years are therefore believed to characterise the human ancient disease landscape.

The first modern human dispersals occurred within Africa during MIS 5 (Marine Isotope Stage 5) some 135 000 to 75 000 years ago (ka). The increasing aridity experienced during MIS 5 likely played a role in the expansion of human populations in central and eastern Africa, ultimately triggering the dispersal of humans out of Africa after c. 65 ka. The development of modernity in early human populations has been linked to various phases of technological and behavioural innovation. While the triggers for these sporadic pulses of technological innovation are not obvious, the incidence of innovations appears to be linked to instances of abrupt climate change (Ziegler et al. 2013). When rainforests expanded during MIS 5, hunters of grassland species moved north and south, taking bifacial technology to North Africa (the Aterian), and South Africa (the Still Bay) (Wadley 2007). Thus, and by the beginning of MIS 5, two behaviourally fully modern human populations were isolated at the opposite ends of Africa (Rito et al. 2013). One thrived on the southern coastal plain in South Africa after 145 ka (Marean et al. 2007; Wadley 2007; Henshilwood et al. 2009; Henshilwood et al. 2011) and the other prospered in the Maghreb, North Africa after 140 ka (Osborne et al. 2008; Barton et al. 2009; Castaneda et al. 2009; Garcea 2012). It is from these isolated populations that the earliest archaeological indications of 'fully modern' and symbolic human behaviour derive. While the regional distributions of projectile point

styles may indicate the existence of complex social networks, the first cultural traditions emerge just before 100 ka, as shown by the engraved ochres from Blombos Cave (Henshilwood et al. 2009), Klein Kliphuis Rock Shelter (Mackay & Welz 2008), Pinnacle Point Cave 13B (Watts 2010) and Klasies River Cave (D'Errico et al. 2012). From 92 ka to 72 ka, evidence for personal ornamentation, in the form of perforated marine shell beads, appears for the first time during the Still Bay and the Aterian (Bouzouggar et al. 2007; Bar-Yosef Mayer et al. 2009; Zilhão et al. 2010).

While much is known about the evolution of human technological competence and symbolic capacity, the influence that diseases had on the biological and social evolution of our species is an essential and often overlooked aspect of our developmental history. Disease certainly appears to have played a significant role in the evolution and geographic distribution of ancient behaviourally-modern Africans. In fact, the scarcity of evidence for human occupation in the tropical regions of central and West Africa (Webb Jr 2005) has been attributed to disease, specifically malaria (Spriggs 2008). Malaria infection occurs when female mosquitoes inject saliva containing plasmodial sporozoites (parasites) into the host during feeding (Tolle 2009). Of the roughly 250 *Plasmodium* species, *P. vivax*, *P. malariae*, *P. falciparum* and *P. ovale* are highly anthropophilic (Ollomo et al. 2009). *P. falciparum* is closely related to *P. reichenowi* and possibly originated from parasites specific to chimpanzees (Rich et al. 2009) and bonobos (Krief et al. 2010) some 3 million years ago (mya), although, data from faecal sampling suggests that gorillas were the likely host species for *P falciparum*, before a cross-species transmission event to humans or our ancestors (Liu et al. 2010). *P malariae* diverged from a parasite of chimps, or both chimps and hominins, around 3.5 mya (Rutledge et al. 2017). The presence of malaria in sub-Saharan Africa therefore predates the emergence of anatomically modern humans 200 ka (White et al. 2003; McDougall et al. 2005), and mitochondrial mtDNA analyses confirm that early forms of *P. falciparum* were present by at least 100 ka (Silva et al. 2011). The parasite subsequently spread from Africa to the Near East and Asia between 90 ka and 80 ka, and to Europe after 40 ka (Tanabe et al. 2010). These ages are consistent with current hypotheses concerning the spread of *Homo sapiens* (Armitage et al. 2011). *Plasmodium vivax*, which today is principally a pathogen of Asia and Latin America, evolved in chimpanzees or gorillas in central Africa. *P vivax* then radiated across the world, most likely as the result of human migration (Liu et al. 2010).

For humans, the avoidance of ecological niches conducive to mosquitoes presents an obvious means to prevent the risk of malaria infection, and also other mosquito-borne diseases such as dengue fever,

West Nile virus, Chikungunya and yellow fever. In Namibia and Botswana, Singer (Singer 1960) observed that San hunter-gatherers who lived more than 25 km from water sources were not susceptible to mosquito-borne diseases. Besides ambient temperature and precipitation (Tonnang et al. 2010), proximity to water plays an important role in the prevalence of malaria. Singer (1960) therefore suggested that Kalahari San migratory routes were deliberately structured to avoid waterlogged areas during summer. There is no conclusive evidence for the incidence of the HbS, b-thalassemia or G6PD traits amongst the Kalahari San (Jenkins et al. 1968; Tishkoff et al. 2001; Kwiatkowski et al. 2005), suggesting that malaria did not exert selective pressure on these groups. Similarly, prehistoric humans may simply have circumvented areas prone to seasonal malaria transmission (including the tropical region of Central Africa), and in this regard the absence of the HbS sickle-cell, b-thalassemia and the G6PD traits amongst the Kalahari San is significant. In addition, whereas the Duffy negative blood group locus is widespread amongst sub-Saharan populations, both the FY*A and FY*B antigens are rare amongst the San of the Kalahari Desert (Howes et al. 2011). This seems confirm the notion that ecological niche avoidance restricted the susceptibility of humans to mosquito-borne diseases (Singer 1960; Dugassa et al. 2009; Wadley 2012).

The relevance of ancient African diseases for modern human society

Resembling our co-evolutionary history with malaria parasites, gaining information about the incidence of disease in prehistoric Africa is important as tropical pathogens and parasites had, and still exert, a significant impact on the evolution of our species. Our rise to being the predominant species on earth is the result of complex interactions between biological and cultural processes, and during the initial stages of our cognitive, technological and cultural evolutionary history, all these processes occurred in sub-Saharan Africa. Current epidemiologic transition models tend to associate the emergence of most human diseases with changes in living conditions associated with agricultural innovation and higher population densities during the Neolithic Period, c. 12 ka (Omran 1971). As a result, the search for the origins of diseases has focussed primarily on domestic animals and environments outside Africa. But, many of these tropical infections are likely to have played a role in the human evolutionary process for much lengthier periods of time (Barrett et al. 1998).

Of the approximately 2100 species of microorganisms that interact directly with humans (Wardeh et al. 2015), 1415 species are known to be pathogenic, including 217 viruses and prions, 538 bacteria, 307

fungi, 66 protozoa and 287 helminths (Taylor et al. 2001; Woolhouse & Gowtage-Sequeria 2005). Approximately 65% of these are zoonotic (Lloyd-Smith et al. 2009) and 177 (8.4%) cause emerging infectious diseases (Dutour 2013). Of these, at least 20 have certain to probable African origin, including hepatitis B, measles, cholera, dengue fever, *P. falciparum* malaria and leishmaniasis, plague and smallpox (Wolfe et al. 2007; Houldcroft & Underdown 2016). The potential impact of disease on prehistoric humans is illustrated by the fact that ~60% of modern hunter-gatherers succumb to disease before reaching reproductive age (c. 15 years) (Gurven & Kaplan 2007).

But how are we affected by disease today? And what can we learn from the study of ancient human disease pathogens? The current global disease burden is dominated by both ancestral (Wolfe et al. 2007; Houldcroft & Underdown 2016) and novel emerging or infectious diseases (Langwig et al. 2015; Plummer et al. 2016). Pathogens result in nearly 11 million human deaths per annum and are responsible for 51% of years of life lost globally (Dunn et al. 2010). Research concerning ancient pathogens can contribute significantly to our understanding of infectious disease evolution in a number of ways. These include improving our understanding of when (and how) virulence evolves in pathogens, when certain pathogens (or parasites) became human pathogens and even whom to prioritise during vaccination campaigns. Studying ancient pathogen DNA (apDNA) is therefore not just of interest to archaeologists, but is also of relevance to public health researchers and molecular biologists.

Because our temporal frame of reference is restricted, and since changes in disease aetiology (including virulence and communicability) frequently occur over longer time periods (Achtman 2016), we do not fully comprehend the processes implicated in disease evolution and emergence. Comparative genomics can be used to reconstruct short-term evolutionary histories of pathogen clades whose diversity converges towards a 'most recent common ancestor' (MRCA) that existed decades or even millennia ago (Der Sarkissian et al. 2015). Genetic changes can be observed in the genomes of bacteria, viruses and parasites and occur through single nucleotide mutations, insertions or deletions or genomic rearrangements. Since mutations play an important role in pathogen evolution and virulence, information derived from apDNA sequences have incredible epidemiological potential. Prehistoric pathogen research can therefore contribute to our understanding of infectious disease evolution by providing chronologically-secure (dated) sequence data to integrate into phylogenetic reconstructions. For a number of reasons, studies of extant pathogen genomes and estimations of the age of a pathogen based on genetic data represent only minimal estimates of the age of a taxon (Achtman 2016). But by

anchoring pathogen emergence dates and mutation rates, apDNA research can provide crucial calibration points to estimate the timing of divergence events (DeWitte et al. 2016).

For example, the sequencing of the 1918-1919 Spanish influenza (H1N1) virus genome yielded new insight into virus biology and pathogenesis (Taubenberger et al. 2012). It is believed that H1N1 emerged in 1893 and that, by 1918, the virus had already accumulated ~375 mutations, at a rate of approximately 15 mutations per year! One of these involved the acquisition of mutations derived from the H5N1 avian virus, and the result was the 1918 Spanish influenza pandemic. Subsequent human-to-human transmission barriers were crossed by the novel zoonotic influenza virus, finally triggering the 1918 pandemic. This illustrates the devastating consequences of influenza virus cross-species transmission (Reperant et al. 2012). The ensuing pandemic viruses of 1957, 1968, and 2009 all descended from the original 1918 virus. The reconstruction of the 1918 virus facilitated the rapid assessment of the potential virulence of the 2009 H1N1 pandemic virus (Medina et al. 2010). The 1918 Spanish influenza virus-specific B cell clones could still be recovered from elderly survivors 90 years after their exposure to the virus but before their exposure to the 2009 pandemic virus (Taubenberger et al. 2012). This realisation provided a scientific rationale for targeting the initial 2009 H1N1 pandemic vaccine to those who needed it most, namely younger persons who had not been exposed to the cross-protective 1918 virus or to its seasonally prevalent descendants. Thus, early in the 2009 pandemic, limited vaccine supplies that might have been misdirected to the traditional (elderly) risk group was administered to younger persons, who benefitted most.

Studying the evolution of infectious diseases, through modern and apDNA, has implications for chronic disease too. There are a number of cancers that are partly or wholly attributable to oncogenic viruses and bacteria (e.g. Kaposi's sarcoma herpesvirus (KSHV) and *Helicobacter pylori*), and in sub-Saharan Africa, up to 33% of all cancers are attributable to infections (Plummer et al. 2016). The long-term tracing of genetic adaptations and rates of evolutionary change are therefore highly informative in understanding how a pathogen becomes virulent or transmissible, providing insights into how we can effectively manage future epidemics (Boire et al. 2014; Andam et al. 2016). In addition to a long list of known vectors and pathogens responsible for epidemic and pandemic influenza, cholera, Ebola, plague, Rift Valley fever, Yellow fever, babesiosis and tuberculosis, the influence of increasingly warmer global temperatures on the re-emergence and prevalence of novel bacterial and viral pathogens is cause for great concern (Wu et al. 2016). This realisation validates the potential of information derived from palaeopathogenic research on sub-Saharan African archaeological contexts.

Ancient African pathogens: Is DNA recovery possible?

There are numerous reasons to study infectious diseases in human prehistory, but this relies on evidence of the disease persisting in the archaeological record. The likelihood of detecting ancient diseases poses significant complications as most diseases are invisible in the archaeological record. Individuals who succumb to death shortly after disease onset will not display skeletal indications of infection, while those that did survive long after the onset of symptoms might, in some instances, have developed skeletal indications of disease (Brothwell 2012). Even those that do affect human skeletal morphology (e.g. *Yersinia pestis*, *Mycobacterium tuberculosis*, *M. leprae*, *Treponema pallidum*, *Brucella melitensis*, *Plasmodium falciparum*, *Trypanosoma cruzi*) are often misdiagnosed. Taphonomic alterations also mimic disease conditions that can induce interpretation errors (pseudo-pathologies), even for experienced palaeo-pathologists (Dutour 2008). Thus, and on account of this 'osteological paradox' (Wood et al. 1992), disease incidence is often unnoticed or misinterpreted, and this can lead to unverified statements that some diseases were either rare or non-existent in prehistory. Unless detected with innovative archaeometric techniques such as X-ray synchrotron micro-tomography (Odes et al. 2016; Randolph-Quinney et al. 2016) or molecular (DNA) analyses, evidence of ancient disease incidence is basically imperceptible.

The application of state-of-the-art molecular analytical techniques to archaeological remains has transformed hominin evolutionary research. Examples of developments in the field of ancient DNA (aDNA) includes the recovery of aDNA from equid remains dated to ~700 ka (Orlando et al. 2013), the sequencing of the oldest hominin nuclear DNA from Sima de los Huesos (Spain) dated to 430 ka (Meyer et al. 2016) and the oldest-known *H. sapiens* genome which was extracted from a human femur recovered from the banks of the Irtysh River in Siberia, dated to 45 ka (Fu et al. 2014). These techniques have also been applied to the emerging field of apDNA and have contributed significantly to our understanding of historical epidemiological etiology (Schuenemann 2013; Devault et al. 2014; Bos et al. 2015; Harkins & Stone 2015; Rasmussen et al. 2015)

The detection of pro-viral sequences (human T-cell lymphotropic virus type I (HTLV-I)) integrated in the genomes of a 1500 year-old Andean mummy (Li et al. 1999), and sequences from the human endogenous retrovirus K (HERV-K) in Neanderthal and Denisovan genomes exemplifies the utility of molecular analytical techniques (Agoni et al. 2012). Ultimately, such 'markers' can be used to study the migration and co-evolution of (prehistoric) humans and their pathogens. For example, Agoni and colleagues (2012) detected that one HERK-V provirus sequence was shared by Neanderthals and

244 Denisovans, providing confirmation that they shared a common ancestor (Reich et al. 2010). Other
245 relevant examples include the reconstruction of the bacterial genome responsible for the Black Death in
246 the Middle Ages (*Y. pestis*) (Bos et al. 2011) and the near-complete genome of the medieval leprosy
247 agent *Mycobacterium leprae* (Schuenemann 2013). These studies, however, rely on the destructive
248 sampling of preserved human skin (Li et al. 1999) , dental pulp (Bos et al. 2011) or bone (Agoni et al.
249 2012) for aDNA extraction and sequencing. Because ancient human remains are rare and therefore
250 valuable, it is necessary to consider the analyses of alternative sources of human aDNA and apDNA. In
251 this regard, ancient human (and animal) coprolites and also anthropogenic sediments (soils derived from
252 caves and rock-shelters inhabited by prehistoric humans) present an unexplored and potentially highly
253 viable alternative.

254 Extracting aDNA from ancient sediments is dependent on various post-depositional processes.
255 Specifically, extremely cold and highly arid conditions are most suited to the preservation of aDNA. In
256 fact, all the oldest known examples of sedimentary aDNA have been recovered from permafrost
257 environments (Thomsen & Willerslev 2015). The discovery and re-animation of two 30,000-year-old
258 viruses (*Pithovirus sibericum* and *Mollivirus sibericum*) from Siberian permafrost (Legendre et al. 2015)
259 not only highlights the preservative capacity of frozen environments, but also illustrates the imaginable
260 severity of the impact that an increasingly warmer world might have on pathogen prevalence. Similarly,
261 the study by Bellemain and colleagues (Bellemain et al. 2013) on the palaeodiversity of fungi in arctic
262 permafrost has detected multiple sequences related to known plant and insect pathogens.

263 This obviously does not sound encouraging for African aDNA studies, particularly when attempting to
264 track down ancient apDNA and correlate these instances with human evolutionary processes. Ancient
265 biomolecules have however been recovered from warm tropical environments. Currently, the oldest
266 known palaeo-protein sequences are dated to 3.8 mya and originate from ostrich eggshell fragments
267 excavated in Tanzania (Demarchi et al. 2016). In addition, the genome of a 4,500 man has recently been
268 sequenced from a cave in Ethiopia (Llorente et al. 2015). Similarly, the sequencing of tropical aDNA from
269 ~1000 year-old extinct tortoise shells has led to the near-complete reconstruction of ancient tortoise
270 mitochondrial genomes (Kehlmaier et al. 2017). This specific study provides proof that, under specific
271 micro-environmental conditions (e.g., anoxic deposits and under relatively stable temperatures) can
272 preserve aDNA for long periods in tropical environments. This is promising for African aDNA research as
273 the caves where our ancestors lived also present conditions suitable for aDNA conservation. For
274 example, the preservation of organic (non-fossilised) human remains comprising the oldest modern

human burial (dated to >70 ka) from Border Cave (d'Errico & Backwell 2016), and the preservation of 40,000 year-old antelope (*Damaliscus niro*) horn (Thackeray & Brink 2004) at Wonderwerk Cave in South Africa, suggest that sheltered and extremely arid sedimentary conditions could contribute to the preservation of aDNA in ancient African archaeological contexts.

Many bacterial, fungal and parasitic pathogens have been isolated from archaeological contexts (Mitchell 2013); some examples are summarised in Table 1. But while the DNA of bacteria and fungi and the remains of parasitic eggs are likely to be detected in ancient African cave sediments, viral DNA is not as likely to be preserved. Unlike the double-stranded DNA structures of bacteria, viral genetic information is encoded in a variety of structures, including double- or single-stranded DNA or RNA genomes. Viral aDNA is more likely to preserve than viral aRNA because DNA degrades more slowly over time than RNA, except when integrated in their host's genome (Li et al., 1999; Agoni et al., 2012) as for example HHV-6 (Arbuckle et al. 2010). Double-stranded viral DNA can furthermore be sequenced along with host aDNA in a single reaction, without additional reverse transcription or steps such as preparation of a single-stranded library (Houldcroft et al. 2017). Ancient single-stranded or RNA genome viruses in archaeological samples may occur when preservation conditions are exceptional (Guy 2014), for example in caves which have a cool and constant temperature, or where soft tissue has been preserved, such as the stomach contents of the Tyrolean Iceman (Maixner et al. 2016). While the hepatitis B virus, which causes hepatocellular carcinoma, has been recovered from human mummified remains, and may be much older than previously estimated (Littlejohn et al. 2016), the likelihood of recovering ancient viral RNA is largely predicted as there is currently little data to support this theory. Scientists have however speculated that hepatitis C virus, despite its single-stranded RNA genome, may preserve in archaeological remains as it has been detected in the tooth pulp of living humans with HCV, and HCV RNA may therefore be preserved in teeth after death (Siravenha et al. 2016).

Extracting aDNA and apDNA from African contexts

In Africa, the extraction of aDNA from archaeological samples has proven challenging, largely because of an extreme shortage of appropriate aDNA extraction facilities. Internationally, there are more than 65 laboratories dedicated specifically to aDNA research (https://palaeogenomics.wordpress.com/ancient-dna-labs/). Yet, excluding Antarctica, Africa is the only continent on which only a single aDNA laboratory (at the Egyptian Museum in Cairo) exists. Extracting and sequencing aDNA necessitates access to highly specialised research facilities and involves strict analytical protocols and complex bioinformatic tools (Willerslev & Cooper 2005; Der Sarkissian et al. 2015; Llamas et al. 2017). This is largely a result of the

fact that, owing to post-mortem decay, aDNA molecules are typically short (less than 100 base-pairs in length), physically damaged (with specific fragmentation and base modification patterns) (Willerslev & Cooper 2005) and, most problematically, that ancient samples are susceptible to contamination by modern DNA derived from environmental sources such as modern sedimentary aDNA, airborne bacterial and fungal spores, microorganisms derived from the human scientists handling the samples and also contaminated sampling equipment and laboratory spaces (Hofreiter et al. 2001; Willerslev & Cooper 2005; Llamas et al. 2017).

Box: Why are aDNA laboratories so special?

Before entering any aDNA-dedicated facility, stringent sampling protocols must be followed to avoid the contamination of the ancient samples with modern DNA or nucleases (Willerslev & Cooper 2005)*. The former contaminant would out-compete aDNA molecules during the sequencing reaction and the latter would either digest or destroy aDNA molecules. Appropriate sampling procedures are detailed in* (Llamas et al. 2017)*. Briefly, researchers performing aDNA sampling activities should wear full body DNA-free protective gear including face-masks, gloves, biologically-impervious body-suits and shoe-covers. All the tools used to collect each individual sample should be treated with specific chemicals that degrade exogenous DNA on their surfaces, such as the use of a 3% bleach-solution or commercial products such as DNA-Away. Importantly, no modern DNA should ever be allowed to enter the aDNA laboratory. This implies that aDNA specific laboratories must a) be physically isolated from any laboratories working with modern DNA, b) be cleaned and decontaminated after each analytical session and treated with UV-lights and c) that access must be restricted only to researchers that have been trained in aDNA analyses and workflow protocols* (Llamas et al. 2017)*. Willerslev and Cooper* (Willerslev & Cooper 2005) *also provide aDNA laboratory operational protocols to follow to evaluate the validity of any aDNA-based research results.*

Given the fact that southern and also eastern Africa forms the focus of human evolutionary research, one would expect the continent to play a key role in the discovery and analyses of aDNA and apDNA. This is even more emphasised by the fact the first ever aDNA sequences studied in the early 1980s originated from the Quagga (*Equus quagga*), an extinct southern African zebra subspecies (Higuchi et al. 1984). While the current shortage of aDNA facilities in Africa leads to international collaboration, it still necessitates the acquisition of substantial funding which limits the usage and development of local expertise. While collaborations undoubtedly increase the quality of research, the lack of aDNA facility in southern Africa impedes the development of local research expertise and knowledge in a highly active

and innovative scientific field that produces high impact and sometimes revolutionary research (Reich et al. 2010; Prüfer et al. 2014; Llorente et al. 2015). Until archaeologists and scientists manage to resolve the problems concerning access to local African aDNA facilities, fragmented aDNA and poor aDNA preservation, understanding African pathogen evolution will depend on studying extant pathogen genomes and using phylogenetics to work backwards in time.

Pathogens and migration

Pathogens (e.g. *Helicobacter pylori* and *Mycobacterium tuberculosis*) and also conspecific human parasites (e.g. human lice) have been used to track human population movements and have provided invaluable information regarding human migrations out of sub-Saharan Africa (Moodley et al. 2012; Comas et al. 2013) and into the New World in particular (Raoult et al. 2008). But exactly which species were brought from Africa to the rest of the world after *H. sapiens* left the continent at c. 100 ka, and again at c. 65 ka, remains unclear. While many of the major modern human diseases that originated in Africa (Wolfe et al. 2007; Houldcroft & Underdown 2016)exerted a profound influence on human evolutionary history, most are still implicated in the deaths of millions of people annually.

Importantly, different pathogens tell stories on different timescales. Large bacterial and parasitic genomes can tell very old stories as these tend to contain substantial amounts of molecular and genomic information, and many might have deep coalescent times. Smaller viral genomes, especially viruses which mutate rapidly (e.g. RNA viruses), 'turn over' their viral genomes so fast that ancient variation has all been lost or replaced before we can observe it (Biek et al. 2015). Such instances however present an opportunity to study more recent pathogen population histories on shorter present-day time scales.

Three stories of migration and disease from Africa

1. <u>HIV, colonialism and male migration</u>

The zoonotic origins of HIV have been reviewed thoroughly elsewhere (Hemelaar 2012). What is of interest is the story of changing human behaviour and migration patterns that allowed HIV to spread and become a global pandemic. Pandemic HIV is caused by viruses from group M, which jumped from wild chimpanzees in Cameroon to humans some 150 years ago, although primate (especially the great apes, i.e. gorillas, chimpanzees and bonobos) to human transmission of SIVs must have been occurring for hundreds and possibly thousands of years before this. This likely resulted in infections which were

poorly adapted to their new human host and therefore unable to spread efficiently between humans. In Cameroon the virus did eventually adapt to spread from human-to-human, and then began to be transmitted southwards across Africa, beginning its journey along the route of the Sangha river to Kinshasa in the Democratic Republic of the Congo (DRC). The transmission of HIV along this route was likely driven by German colonialism in Cameroon which promoted increased movement of goods such as rubber (and also HIV-positive individuals) into the DRC by river. This period of colonial rule is a plausible explanation for the dating and location of the most recent common ancestor (MRCA) of HIV group M strains derived from Kinshasa in around 1920 (Faria et al. 2014).

Kinshasa became the epicentre for the HIV pandemic and, within 20 years, HIV had spread to Brazzaville, Lubumbashi and Mbuji-Mayi. Increasing population mobility due to urbanisation and new transportation methods, such as railways, further facilitated the spread of HIV. In fact, historical data on population movements by rail and river supports data derived from modelling the spread of HIV genetically. HIV reached cities receiving high volumes of rail and river migration from Kinshasa earlier (by an estimated 15-20 years) than cities which received lower volumes of river traffic from Kinshasa. A disease which had once been transmitted only sporadically, first from primates to bush-meat hunters in tropical Africa, has now become a fully human-adapted disease free to infect residents in densely populated areas with increasing mobility, allowing it to spread across Africa (Faria et al. 2014).

International mobility was instrumental in the development of HIV from an epidemic to a pandemic. Human migration was at the heart of the early global spread of HIV, and this is reflected in the genetic structure of HIV strains collected from current and historic HIV cases. Professionals from Haiti who travelled back and forth to post-colonial Congo carried a specific lineage of HIV group M (subtype B) with them; subtype B HIV was detectable in Haiti by 1964. From Haiti, subtype B was able to spread to the USA (Faria et al. 2014) around 1970, 10 years before the first US cases were recognised (Worobey et al. 2016). While many aspects of human behaviour are also integral to the spread of HIV (such as bush-meat hunting, sex work, and unsafe medical practices that led to extensive needle reuse), migration over smaller and larger geographic distances, facilitated by mass transportation and trade, enabled HIV group M to spread widely and rapidly.

2. New ecological niches: flukes, fishing and farming in the African Pleistocene

S mansoni is a trematode blood fluke found across sub-Saharan Africa, the Caribbean and parts of South America, infecting 250 million people worldwide and killing a small proportion of infected individuals every year by causing chronic inflammation of the spleen and/or liver. *S mansoni* requires an intermediate freshwater snail host to complete its life cycle, which means that populations who engage in behaviour which leads them to spend periods of time wading in fresh water are at risk of infection. There are a number of different behaviours which bring humans in to contact with *S mansoni*, including wading through irrigated fields (Hibbs et al. 2011), which lead to human *S mansoni* infections in ancient Nubia, and also fishing at the edge of fresh water lakes and rivers (Crellen et al. 2016), a food exploitation technique which dates to between 74-111KYA (Yellen et al. 1995; Brooks et al. 1995) for freshwater resources and even earlier for marine resources (152–176KYA) (Marean et al. 2007).

The origins of S mansoni as a human infection are intimately tied up with human migration and changing human behaviour. *S mansoni*'s closest relative is *S rodhaini*, a rodent trematode, and it is likely that the last common ancestor of these two species was a rodent parasite. This ancestor was able to switch hosts and infect humans following changing human behaviour, leading to the speciation of the trematode ancestor into *mansoni* and *rodhaini* around 125kya. The most genetically diverse isolates of *S mansoni* come from lakes Victoria and Albert in Uganda, suggesting that it was in east Africa – where some of the earliest anatomically modern human remains are found (McDougall et al. 2005) – that *S mansoni* was able to switch hosts, likely after humans began to exploit fresh water lakes and rivers through fishing and dwelling on the edges of lakes to hunt the fauna who came there to drink

The spread of farming in Africa lead to movements of people and the spread of technology, and this in turn helped to spread *S mansoni* across sub-Saharan Africa. Isolates of *S mansoni* from east and west Africa were a single population (likely endemic in east Africa, where the original host-switch in AMH occurred) until 7KYA. This could reflect a movement of infected people as part of the expansion of farming and pastoralism (Marshall & Hildebrand 2002), and the population expansions around 6KYA of the Yoruba and Luhya (Crellen et al. 2016).

Genetic data also charts the path of *S mansoni* from Africa to its other foci. The genetic diversity of *S mansoni* in the Caribbean reflects spread of this parasite during the trans-Atlantic slave trade. *S mansoni* found in Guadeloupe diverged from *S mansoni* from Senegal and Cameroon between ~1100-1750CE. This coincides with the beginning of the French colonial slave trade to the French Caribbean, from 1669-1864. The mass enslavement, forced migration and then forced labour in the French Caribbean of at

425 least 20,000 West Africans therefore seems the most plausible explanation for the spread of *S mansoni*
426 to the Caribbean (Crellen et al. 2016).

427

428 3. Chimpanzees, hominins and herpes

429 Herpes simplex virus 2 (HSV2) is a sexually transmitted human DNA virus that causes genital lesions and,
430 rarely, encephalitis (Tang et al. 2003) and is associated with increased risk of HIV acquisition (Freeman et
431 al. 2006). After primary infection, the virus adopts a life cycle of latency punctuated by periods of lytic
432 replication when new hosts can be infected through genital contact. The virus is related to the human
433 oral pathogen herpes simplex 1 (HSV1). Both are alphaherpesviruses, which are found in many primates
434 (Wertheim et al. 2014).

435 HSV2 was originally thought to have co-speciated with humans when our lineage diverged from
436 chimpanzees and bonobos, but recent comparisons of the HSV1, HSV2 and chHV1 genomes suggests
437 that HSV2 is in fact more closely related to chimpanzee herpesvirus than human herpesvirus (Wertheim
438 et al. 2014). A similar pattern can be seen between genetically distinct clades of *Pediculus humanus* (the
439 human head and body louse), with clade divergence times which pre-date the emergence of the *Homo*
440 *sapiens* in Africa and unusual geographic distributions (Reed et al. 2004). Taken together, these data are
441 highly suggestive of close bodily interaction between hominin species (Ashfaq et al. 2015).

442 The analysis of human and chimp simplex viruses found that HSV2 diverged from ChHV1 between 1.4
443 and 3 MYA, and an intermediate hominin is inferred to have served as a host for proto-HSV2 before it
444 introgressed into the ancestors of modern humans. Humans are susceptible to infection by primate
445 simplex viruses from bite injuries, suggesting that hunting of chimpanzees by hominins could have been
446 one transmission route. However, it is unclear how recently HSV2 introgressed into the modern human
447 population (and whether that transmission was sexual or the result of bite injuries during hunting), but
448 given HSV2's global distribution, it seems likely that this virus infected the human lineage before the
449 migration of AMH out of Africa. There is evidence of interbreeding and genetic introgression between
450 anatomically modern humans and unknown hominins in Africa around 35kya (Hammer et al. 2011;
451 Hsieh et al. 2016), too late to be the HSV2-transmitting hominin; but together, these different lines of
452 evidence illustrate that interbreeding between different groups of hominins, not all known from fossils,
453 was occurring across the globe in the Pleistocene. It also predicts certain patterns of migration, as

different hominins interacted as climate conditions changed, changing resource distribution and generating the potential for inter-species conflict.

Unfortunately, ancient DNA evidence will not resolve the identity of the hominin(s) who transmitted proto-HSV2 to the ancestors of modern humans. There is an 'event horizon' for ancient DNA preservation, predicted by factors such as the age of the fossil, the heat and humidity it is exposed to, and the potential for degradation by microbes (Allentoft et al. 2012). This means that recovering authentic aDNA or apDNA from an African fossil more than 1 million years old is highly unlikely.

The issue is further complicated by the high human-to-virus DNA ratios experienced when trying to sequence herpesviruses from living humans, meaning only a tiny proportion of the total DNA within a sample would come from HSV2 (Houldcroft et al. 2017) if samples from archaeological specimens were available; explicit enrichment of HSV2 DNA by PCR or target capture would be required (Houldcroft & Breuer 2015; Depledge et al. 2011; Ebert et al. 2013).

However, this is not the end of the story: the examples of HIV and *S mansoni* demonstrate the power of genomic analysis to reveal aspects of human and pathogen co-evolution. Computational modelling is increasingly being applied within archaeology to reconstruct past events (eg (Crema et al. 2016; Bortolini et al. 2016)), uniting many data sources. For example, modelling has been used to predict the movement of anatomically modern humans out of Africa based on climate data and patterns of extant human genetic diversity, without relying on fossil or other archaeological data (Eriksson et al. 2012). Computational modelling and knowledge of areas of Africa with particular important for the understanding of human evolution would allow for more HSV2 genomes to be collected from humans and ChHV in an evolutionarily informed manner and the potential transmission route reconstructed (Underdown et al. 2017). This would allow researchers to focus on areas of Africa where particularly ancient HSV2 lineages are predicted to be found; sequencing of a bonobo herpes simplex genome would also aid in reconstructions of the history of ChHV1 and HSV2.

Conclusions

It is evident that ancient biomolecular research can contribute to existing genome databases which may have public health benefits by providing tools for developing therapeutics, particularly if virulent forms of ancient diseases re-emerge. This is important as history has taught us that disease is by far the most effective eradicator of our species. Past pandemics are much more than just ancient history. They are important drivers of human genetic diversity and natural selection (Pittman et al. 2016). At the time of

484 writing this Review, a report entitled 'Killer bird flu has spread across Europe - are humans next?'
485 appeared in New Scientist (https://www.newscientist.com/article/2113725-killer-bird-flu-has-spread-
486 across-europe-are-humans-next/). While rather sensationalist, the H5N8 virus, lurking in domestic and
487 wild avian populations since 2014, has rapidly spread along avian migration routes into India, the Middle
488 East and Europe and it certainly does hold the potential to develop into a global influenza pandemic. The
489 potentially severe economic and social repercussions of disease epidemics are further demonstrated by
490 both historical (e.g. plaque, smallpox, influenza etc.) and current (i.e. Zika, Ebola, SARS etc.) examples.
491 But the biological origin of a many prehistoric, historical and even contemporary pathogens remains
492 mysterious. The emphasis should therefore also be on the development of sub-Saharan capabilities to
493 detect, predict, prevent and control potential infectious disease epidemics rather than waiting for
494 known diseases to threaten global human health. This is particularly important given the current global
495 interconnectedness, which can put people at risk of diseases that emerge in distant locales.

496 The recent retrieval of the first ancient African genome from Mota Cave in Ethiopia (Llorente et al. 2015)
497 dated to c. 4,500 years suggests that the prospect of retrieving both human and apDNA from sub-
498 Saharan African contexts is becoming progressively more promising. Temperate and Arctic regions have
499 yielded more aDNA sequences than tropical regions, partially because conditions are more favourable to
500 the preservation of aDNA, but also because they have been researched more intensively (Slatkin &
501 Racimo 2016). Because of the paucity of aDNA sequences from Africa, any novel pathogen genomes will
502 provide novel revelations concerning human-pathogen co-evolutionary processes. As this review has
503 shown, evidence from even modern African pathogen genomes can shine a light on changes in human
504 behaviour and migration. The unique combination of an unrivalled archaeological record and a thriving
505 and highly skilled academic community places southern African archaeologists, geneticists and medical
506 scientists in a prime position to explore past pathogenic influences and to contribute to the
507 improvement of human quality of life and longevity.

508 **Funding**

509 CJH was funded by the University of Cambridge. JBR was funded by the University of Pretoria and the
510 South African National Research Foundation (ERTTG 2016 grant UID 105197). RFR was funded by the
511 National Geographic Society/Waitt Fecundation Scientific Exploration Grant (Nr. W420-15) and the
512 University of Pretoria. SJU was funded by Oxford Brookes University.

513

514 **References**

515 Achtman, M., 2016. How old are bacterial pathogens? *Proceedings of the Royal Society of London B:*
516 *Biological Sciences*, 283(1836).

517 Adler, C.J. et al., 2013. Sequencing ancient calcified dental plaque shows changes in oral microbiota with
518 dietary shifts of the Neolithic and Industrial revolutions. *Nature genetics*, 45(4), p.450–5, 455e1.
519 Available at: http://dx.doi.org/10.1038/ng.2536 [Accessed May 28, 2014].

520 Agoni, L. et al., 2012. Neandertal and Denisovan retroviruses. *Current Biology*, 22(11).

521 Allentoft, M.E. et al., 2012. The half-life of DNA in bone: measuring decay kinetics in 158 dated fossils.
522 *Proceedings of the Royal Society of London B: Biological Sciences*, 279(1748).

523 Andam, C.P. et al., 2016. Microbial Genomics of Ancient Plagues and Outbreaks. *Trends in Microbiology*,
524 24(12), pp.978–990.

525 Arbuckle, J.H. et al., 2010. The latent human herpesvirus-6A genome specifically integrates in telomeres
526 of human chromosomes in vivo and in vitro. *Proceedings of the National Academy of Sciences of*
527 *the United States of America*, 107(12), pp.5563–8. Available at:
528 http://www.ncbi.nlm.nih.gov/pubmed/20212114 [Accessed January 17, 2017].

529 Armitage, S.J. et al., 2011. The Southern Route "Out of Africa": Evidence for an Early Expansion of
530 Modern Humans into Arabia. *Science*, 453(6016), pp.453–156.

531 Ashfaq, M. et al., 2015. High diversity and rapid diversification in the head louse, Pediculus humanus
532 (Pediculidae: Phthiraptera). *Scientific Reports*, 5, p.14188. Available at:
533 http://www.nature.com/articles/srep14188 [Accessed January 26, 2017].

534 Bar-Yosef Mayer, D.E., Vandermeersch, B. & Bar-Yosef, O., 2009. Shells and ochre in Middle Paleolithic
535 Qafzeh Cave, Israel: indications for modern behavior. *Journal of Human Evolution*, 56(3), pp.307–
536 314.

537 Barrett, R. et al., 1998. EMERGING AND RE-EMERGING INFECTIOUS DISEASES: The Third Epidemiologic
538 Transition. *Annual Review of Anthropology*, 27(1), pp.247–271. Available at:
539 http://www.annualreviews.org/doi/abs/10.1146/annurev.anthro.27.1.247 [Accessed August 5,
540 2014].

Barton, R.N.E. et al., 2009. OSL dating of the Aterian levels at Dar es-Soltan I (Rabat, Morocco) and implications for the dispersal of modern Homo sapiens. *Quaternary Science Reviews*, 28(19–20), pp.1914–1931.

Bellemain, E. et al., 2013. Fungal palaeodiversity revealed using high-throughput metabarcoding of ancient DNA from arctic permafrost. *Environmental Microbiology*, 15(4), pp.1176–1189. Available at: http://doi.wiley.com/10.1111/1462-2920.12020 [Accessed January 26, 2017].

Biek, R. et al., 2015. Measurably evolving pathogens in the genomic era. *Trends in Ecology & Evolution*, 30(6), pp.306–313.

Boire, N.A. et al., 2014. Lessons Learned from Historic Plague Epidemics: The Relevance of an Ancient Disease in Modern Times. *Journal of Ancient Diseases & Preventive Remedies*, 2(2). Available at: http://www.esciencecentral.org/journals/lessons-learned-from-historic-plague-epidemics-the-relevance-of-an-ancient-disease-in-modern-times-2329-8731.1000114.php?aid=30293 [Accessed January 25, 2017].

Bortolini, E. et al., 2016. The Tell-Tale Genome. *bioRxiv*. Available at: http://biorxiv.org/content/early/2016/12/02/091074 [Accessed January 26, 2017].

Bos, K.I. et al., 2011. A draft genome of Yersinia pestis from victims of the Black Death. *Nature*, 478(7370), pp.506–10. Available at: http://www.nature.com/doifinder/10.1038/nature10549 [Accessed May 25, 2014].

Bos, K.I. et al., 2015. Parallel detection of ancient pathogens via array-based DNA capture. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, 370(1660), p.20130375. Available at: http://rstb.royalsocietypublishing.org/content/370/1660/20130375.long.

Boutellis, A., Abi-Rached, L. & Raoult, D., 2014. The origin and distribution of human lice in the world. *Infection, genetics and evolution : journal of molecular epidemiology and evolutionary genetics in infectious diseases*, 23, pp.209–17. Available at: http://www.sciencedirect.com/science/article/pii/S1567134814000203 [Accessed July 8, 2014].

Bouzouggar, A. et al., 2007. 82,000-year-old shell beads from North Africa and implications for the origins of modern human behavior. *Proceedings of the National Academy of Sciences of the United States of America*, 104(24), pp.9964–9969. Available at:

http://www.ncbi.nlm.nih.gov/pubmed/17548808.

Brooks, A. et al., 1995. Dating and context of three middle stone age sites with bone points in the Upper Semliki Valley, Zaire. *Science*, 268(5210).

Brothwell, D., 2012. Tumors: Problems of Differential Diagnosis in Paleopathology. In *A Companion to Paleopathology*. pp. 420–433.

Castaneda, I.S. et al., 2009. Wet phases in the Sahara/Sahel region and human migration patterns in North Africa. *Proceedings of the National Academy of Sciences of the United States of America*, 106(48), pp.20159–20163.

Comas, I. et al., 2013. Out-of-Africa migration and Neolithic coexpansion of Mycobacterium tuberculosis with modern humans. *Nature Genetics*, 45(10), pp.1176–1182. Available at: http://www.nature.com/doifinder/10.1038/ng.2744%5Cnhttp://dx.doi.org/10.1038/ng.2744 [Accessed May 24, 2014].

Compton, A.A. et al., 2013. Host gene evolution traces the evolutionary history of ancient primate lentiviruses. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, 368(1626), p.20120496. Available at: http://www.ncbi.nlm.nih.gov/pubmed/23938749 [Accessed September 1, 2016].

Crellen, T. et al., 2016. Whole genome resequencing of the human parasite Schistosoma mansoni reveals population history and effects of selection. *Scientific Reports*, 6, p.20954. Available at: http://www.nature.com/articles/srep20954 [Accessed October 31, 2016].

Crema, E.R., Kandler, A. & Shennan, S., 2016. Revealing patterns of cultural transmission from frequency data: equilibrium and non-equilibrium assumptions. *Scientific reports*, 6, p.39122. Available at: http://www.ncbi.nlm.nih.gov/pubmed/27974814 [Accessed January 26, 2017].

d'Errico, F. & Backwell, L., 2016. Earliest evidence of personal ornaments associated with burial: The Conus shells from Border Cave. *Journal of Human Evolution*, 93, pp.91–108.

D'Errico, F., Garc??a Moreno, R. & Rifkin, R.F., 2012. Technological, elemental and colorimetric analysis of an engraved ochre fragment from the Middle Stone Age levels of Klasies River Cave 1, South Africa. *Journal of Archaeological Science*, 39(4), pp.942–952.

596 Demarchi, B. et al., 2016. Protein sequences bound to mineral surfaces persist into deep time. *eLife*, 5,
597 pp.4724–4733. Available at: http://www.ncbi.nlm.nih.gov/pubmed/27668515 [Accessed January
598 26, 2017].

599 Depledge, D.P. et al., 2011. Specific Capture and Whole-Genome Sequencing of Viruses from Clinical
600 Samples R. Jhaveri, ed. *PLoS ONE*, 6(11), p.e27805. Available at:
601 http://dx.plos.org/10.1371/journal.pone.0027805 [Accessed November 7, 2016].

602 Devault, A.M. et al., 2014. Ancient pathogen DNA in archaeological samples detected with a Microbial
603 Detection Array. *Scientific reports*, 4, p.4245. Available at:
604 http://www.nature.com/articles/srep04245 [Accessed November 1, 2016].

605 DeWitte, S.N. et al., 2016. Disease epidemics: lessons for resilience in an increasingly connected world.
606 *Journal of public health (Oxford, England)*, p.fdw044. Available at:
607 http://www.ncbi.nlm.nih.gov/pubmed/27226010 [Accessed December 14, 2016].

608 Dugassa, S. et al., 2009. Field investigation on the repellent activity of some aromatic plants by
609 traditional means against Anopheles arabiensis and An. pharoensis (Diptera: Culicidae) around
610 Koka, central Ethiopia. *Acta Tropica*, 112(1), pp.38–42.

611 Dunn, R.R. et al., 2010. Global drivers of human pathogen richness and prevalence. *Proceedings.*
612 *Biological sciences / The Royal Society*, 277(1694), pp.2587–95.

613 Dutour, O., 2008. Archaeology of human pathogens: Palaeopathological appraisal of
614 palaeoepidemiology. In *Paleomicrobiology: Past Human Infections*. pp. 125–144.

615 Dutour, O., 2013. Paleoparasitology and paleopathology. Synergies for reconstructing the past of human
616 infectious diseases and their pathocenosis. *International Journal of Paleopathology*, 3, pp.145–149.

617 Ebert, K. et al., 2013. Mode of virus rescue determines the acquisition of VHS mutations in VP22-
618 negative herpes simplex virus 1. *Journal of virology*, 87(18), pp.10389–93. Available at:
619 http://www.ncbi.nlm.nih.gov/pubmed/23864617 [Accessed December 13, 2016].

620 Eriksson, A. et al., 2012. Late Pleistocene climate change and the global expansion of anatomically
621 modern humans. *Proceedings of the National Academy of Sciences of the United States of America*,
622 109(40), pp.16089–94. Available at: http://www.ncbi.nlm.nih.gov/pubmed/22988099 [Accessed
623 December 13, 2016].

624 Faria, N.R. et al., 2014. The early spread and epidemic ignition of HIV-1 in human populations. *Science*,
625 346(6205).

626 Freeman, E.E. et al., 2006. Herpes simplex virus 2 infection increases HIV acquisition in men and women:
627 systematic review and meta-analysis of longitudinal studies. *AIDS (London, England)*, 20(1), pp.73–
628 83. Available at: http://www.ncbi.nlm.nih.gov/pubmed/16327322 [Accessed January 26, 2017].

629 Fu, Q. et al., 2014. Genome sequence of a 45,000-year-old modern human from western Siberia. *Nature*,
630 514(7523), pp.445–9. Available at: http://dx.doi.org/10.1038/nature13810.

631 Garcea, E.A.A., 2012. Successes and failures of human dispersals from North Africa. *Quaternary*
632 *International*, 270, pp.119–128.

633 Gurven, M. & Kaplan, H., 2007. Longevity Among Hunter- Gatherers: A Cross-Cultural Examination.
634 *Population and Development Review*, 33(2), pp.321–365.

635 Guy, P.L., 2014. Prospects for analyzing ancient RNA in preserved materials. *Wiley Interdisciplinary*
636 *Reviews: RNA*, 5(1), pp.87–94. Available at: http://doi.wiley.com/10.1002/wrna.1199 [Accessed
637 January 17, 2017].

638 Hammer, M.F. et al., 2011. Genetic evidence for archaic admixture in Africa. *Proceedings of the National*
639 *Academy of Sciences of the United States of America*, 108(37), pp.15123–8. Available at:
640 http://www.ncbi.nlm.nih.gov/pubmed/21896735 [Accessed January 26, 2017].

641 Harkins, K.M. & Stone, A.C., 2015. Ancient pathogen genomics: insights into timing and adaptation.
642 *Journal of Human Evolution*, 79, pp.137–149.

643 Hemelaar, J., 2012. The origin and diversity of the HIV-1 pandemic. *Trends in molecular medicine*, 18(3),
644 pp.182–92. Available at: http://www.sciencedirect.com/science/article/pii/S1471491411002103
645 [Accessed July 13, 2014].

646 Henshilwood, C.S. et al., 2011. A 100,000-Year-Old Ochre-Processing Workshop at Blombos Cave, South
647 Africa. *Science*, 334(6053), pp.219–222.

648 Henshilwood, C.S., d'Errico, F. & Watts, I., 2009. Engraved ochres from the Middle Stone Age levels at
649 Blombos Cave, South Africa. *Journal of Human Evolution*, 57(1), pp.27–47.

650 Hibbs, A.C. et al., 2011. Irrigation and infection: The immunoepidemiology of schistosomiasis in ancient

Nubia. *American Journal of Physical Anthropology*, 145(2), pp.290–298. Available at: http://doi.wiley.com/10.1002/ajpa.21493 [Accessed October 31, 2016].

Higuchi, R. et al., 1984. DNA sequences from the quagga, an extinct member of the horse family. *Nature*, 312(5991), pp.282–284.

Hofreiter, M. et al., 2001. Ancient DNA. *Nature reviews. Genetics*, 2(5), pp.353–9. Available at: http://www.ncbi.nlm.nih.gov/pubmed/15568989.

Houldcroft, C. & Underdown, S., 2016. Neanderthal genomics suggests a pleistocene time frame for the first epidemiologic transition. *American Journal of Physical Anthropology*.

Houldcroft, C.J., Beale, M.A. & Breuer, J., 2017. Clinical and biological insights from viral genome sequencing. *Nature Reviews Microbiology*. Available at: http://www.nature.com/doifinder/10.1038/nrmicro.2016.182 [Accessed January 25, 2017].

Houldcroft, C.J. & Breuer, J., 2015. Tales from the crypt and coral reef: the successes and challenges of identifying new herpesviruses using metagenomics. *Frontiers in Microbiology*, 6. Available at: http://journal.frontiersin.org/Article/10.3389/fmicb.2015.00188/abstract [Accessed November 15, 2016].

Howes, R.E. et al., 2011. The global distribution of the Duffy blood group. *Nature communications*, 2, p.266. Available at: http://dx.doi.org/10.1038/ncomms1265.

Hsieh, P. et al., 2016. Model-based analyses of whole-genome data reveal a complex evolutionary history involving archaic introgression in Central African Pygmies. *Genome research*, 26(3), pp.291–300. Available at: http://www.ncbi.nlm.nih.gov/pubmed/26888264 [Accessed January 26, 2017].

Jenkins, T. et al., 1968. Some hereditary red-cell traits in Kalahari Bushmen and Bantu: hemoglobins, glucose-6-phosphate dehydrogenase deficiency, and blood groups. *American journal of human genetics*, 20(4), pp.299–309. Available at: http://www.ncbi.nlm.nih.gov/pubmed/4969414 [Accessed January 25, 2017].

Just, M., Norton, J. & Traud, A., 2014. Global biogeographic regions in a human-dominated world: the case of human diseases. *Ecosphere*, 5(November), pp.1–21. Available at: http://www.esajournals.org/doi/abs/10.1890/ES14-00201.1.

Kehlmaier, C. et al., 2017. Tropical ancient DNA reveals relationships of the extinct Bahamian giant tortoise Chelonoidis alburyorum. *Proceedings of the Royal Society of London B: Biological Sciences*, 284(1846).

Krief, S. et al., 2010. On the Diversity of Malaria Parasites in African Apes and the Origin of Plasmodium falciparum from Bonobos L. D. Sibley, ed. *PLoS Pathogens*, 6(2), p.e1000765. Available at: http://dx.plos.org/10.1371/journal.ppat.1000765 [Accessed January 25, 2017].

Kwiatkowski, D.P. et al., 2005. How malaria has affected the human genome and what human genetics can teach us about malaria. *American journal of human genetics*, 77(2), pp.171–92. Available at: http://www.ncbi.nlm.nih.gov/pubmed/16001361 [Accessed January 25, 2017].

Langwig, K.E. et al., 2015. Context-dependent conservation responses to emerging wildlife diseases. *Frontiers in Ecology and the Environment*, 13(4), pp.195–202.

Legendre, M. et al., 2015. In-depth study of Mollivirus sibericum, a new 30,000-y-old giant virus infecting Acanthamoeba. *PNAS*, 112(38), pp.5327–5335. Available at: http://www.pnas.org/content/112/38/E5327.short.

Li, H.-C. et al., 1999. The presence of ancient human T-cell lymphotropic virus type I provirusDNA in an Andean mummy. *Nature Medicine*, 5(12), pp.1428–1432. Available at: http://www.nature.com/doifinder/10.1038/71006 [Accessed January 25, 2017].

Littlejohn, M., Locarnini, S. & Yuen, L., 2016. Origins and Evolution of Hepatitis B Virus and Hepatitis D Virus. *Cold Spring Harbor perspectives in medicine*, 6(1), p.a021360. Available at: http://www.ncbi.nlm.nih.gov/pubmed/26729756 [Accessed November 17, 2016].

Liu, W. et al., 2010. Origin of the human malaria parasite Plasmodium falciparum in gorillas. *Nature*, 467(7314), pp.420–425. Available at: http://dx.doi.org/10.1038/nature09442 [Accessed August 3, 2014].

Llamas, B. et al., 2017. From the field to the laboratory: Controlling DNA contamination in human ancient DNA research in the high-throughput sequencing era. *STAR: Science & Technology of Archaeological Research*, 3(1), pp.1–14. Available at: https://www.tandfonline.com/doi/full/10.1080/20548923.2016.1258824 [Accessed January 25, 2017].

706 Llorente, M.G. et al., 2015. Ancient Ethiopian genome reveals extensive Eurasian admixture throughout
707 the African continent. *Science*, 350(6262), pp.820–2. Available at:
708 http://science.sciencemag.org/content/350/6262/820.abstract.

709 Lloyd-Smith, J.O. et al., 2009. Epidemic Dynamics at the Human-Animal Interface. *Science*, 326(5958),
710 pp.1362–1367. Available at:
711 http://science.sciencemag.org/content/326/5958/1362.abstract%5Cnhttp://www.sciencemag.org/
712 content/326/5958/1362.abstract.

713 Mackay, A. & Welz, A., 2008. Engraved ochre from a Middle Stone Age context at Klein Kliphuis in the
714 Western Cape of South Africa. *Journal of Archaeological Science*, 35(6), pp.1521–1532.

715 Maixner, F. et al., 2016. The 5300-year-old Helicobacter pylori genome of the Iceman. *Science*,
716 351(6269), pp.162–165. Available at: http://www.sciencemag.org/content/351/6269/162.abstract.

717 Marean, C.W.C.W. et al., 2007. Early human use of marine resources and pigment in South Africa during
718 the Middle Pleistocene. *Nature*, 449(7164), pp.905–908. Available at:
719 http://www.nature.com/doifinder/10.1038/nature06204 [Accessed October 31, 2016].

720 Marshall, F. & Hildebrand, E., 2002. Cattle Before Crops: The Beginnings of Food Production in Africa.
721 *Journal of World Prehistory*, 16(2), pp.99–143. Available at:
722 http://link.springer.com/10.1023/A:1019954903395 [Accessed January 26, 2017].

723 McDougall, I., Brown, F.H. & Fleagle, J.G., 2005. Stratigraphic placement and age of modern humans
724 from Kibish, Ethiopia. *Nature*, 433(7027), pp.733–736. Available at:
725 http://www.nature.com/doifinder/10.1038/nature03258 [Accessed January 25, 2017].

726 Medina, R.A. et al., 2010. Pandemic 2009 H1N1 vaccine protects against 1918 Spanish influenza virus.
727 *Nature Communications*, 1(3), pp.1–6. Available at:
728 http://www.nature.com/doifinder/10.1038/ncomms1026 [Accessed January 25, 2017].

729 Meyer, M. et al., 2016. Nuclear DNA sequences from the Middle Pleistocene Sima de los Huesos
730 hominins. *Nature*, 531(7595), pp.504–507. Available at:
731 http://www.nature.com/doifinder/10.1038/nature17405.

732 Mitchell, P.D., 2013. The origins of human parasites: Exploring the evidence for endoparasitism
733 throughout human evolution. *International Journal of Paleopathology*, 3(3), pp.191–198. Available

at: http://www.sciencedirect.com/science/article/pii/S187998171300082X [Accessed May 29, 2014].

Moodley, Y. et al., 2012. Age of the association between Helicobacter pylori and man. H. Ochman, ed. *PLoS Pathogens*, 8(5), p.e1002693. Available at: http://dx.plos.org/10.1371/journal.ppat.1002693 [Accessed May 25, 2014].

Mourre, V., Villa, P. & Henshilwood, C.S., 2010. Early Use of Pressure Flaking on Lithic Artifacts at Blombos Cave, South Africa. *Science*, 330(6004).

Odes, E.J. et al., 2016. Earliest hominin cancer: 1.7-million-year-old osteosarcoma from Swartkrans Cave, South Africa. *South African Journal of Science*, 112(7/8), p.5 Pages.

Ollomo, B. et al., 2009. A new malaria agent in African hominids. *PLoS Pathogens*, 5(5).

Omran, A., 1971. The epidemiologic transition: a theory of the epidemiology of populatuon change. *The Milbank Memorial Fund Quarterly*, 49(1), pp.509–538.

Orlando, L. et al., 2013. Recalibrating Equus evolution using the genome sequence of an early Middle Pleistocene horse. *Nature*, 499(7456), pp.74–78. Available at: http://www.ncbi.nlm.nih.gov/pubmed/23803765.

Osborne, A.H. et al., 2008. A humid corridor across the Sahara for the migration of early modern humans out of Africa 120,000 years ago. *Proceedings of the National Academy of Sciences of the United States of America*, 105(43), pp.16444–7. Available at: http://www.ncbi.nlm.nih.gov/pubmed/18936490 [Accessed January 25, 2017].

Pittman, K.J. et al., 2016. The Legacy of Past Pandemics: Common Human Mutations That Protect against Infectious Disease W. E. Goldman, ed. *PLOS Pathogens*, 12(7), p.e1005680. Available at: http://dx.plos.org/10.1371/journal.ppat.1005680 [Accessed January 25, 2017].

Plummer, M. et al., 2016. Global burden of cancers attributable to infections in 2012: a synthetic analysis. *The Lancet Global Health*, 4(9), pp.e609–e616.

Prüfer, K. et al., 2014. The complete genome sequence of a Neanderthal from the Altai Mountains. *Nature*, 505(7481), pp.43–9. Available at: http://dx.doi.org/10.1038/nature12886 [Accessed May 27, 2014].

Randolph-Quinney, P.S. et al., 2016. Osteogenic tumour in Australopithecus sediba: Earliest hominin evidence for neoplastic disease. *South African Journal of Science*, 112(7/8), p.7 Pages.

Raoult, D. et al., 2008. Molecular identification of lice from pre-Columbian mummies. *The Journal of infectious diseases*, 197(4), pp.535–43. Available at: http://www.ncbi.nlm.nih.gov/pubmed/18254682.

Rasmussen, S. et al., 2015. Early Divergent Strains of Yersinia pestis in Eurasia 5,000 Years Ago. *Cell*, 163(3). Available at: http://linkinghub.elsevier.com/retrieve/pii/S0092867415013227 [Accessed October 31, 2016].

Reed, D.D.L. et al., 2004. Genetic analysis of lice supports direct contact between modern and archaic humans. Nick Barton, ed. *PLoS biology*, 2(11), p.e340. Available at: http://dx.plos.org/10.1371/journal.pbio.0020340 [Accessed July 23, 2014].

Reich, D. et al., 2010. Genetic history of an archaic hominin group from Denisova Cave in Siberia. *Nature*, 468(7327), pp.1053–1060.

Reperant, L.A., Kuiken, T. & Osterhaus, A.D.M.E., 2012. Adaptive pathways of zoonotic influenza viruses: From exposure to establishment in humans. *Vaccine*, 30(30), pp.4419–4434.

Rich, S.M. et al., 2009. The origin of malignant malaria. *Proceedings of the National Academy of Sciences*, 106(35), pp.14902–14907. Available at: http://www.ncbi.nlm.nih.gov/pubmed/19666593%5Cnhttp://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC2720412.

Rito, T. et al., 2013. The first modern human dispersals across Africa. *PLoS ONE*, 8(11).

Rutledge, G.G. et al., 2017. Plasmodium malariae and P. ovale genomes provide insights into malaria parasite evolution. *Nature*. Available at: http://www.nature.com/doifinder/10.1038/nature21038 [Accessed January 26, 2017].

Der Sarkissian, C. et al., 2015. Ancient genomics. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, 370(1660), p.20130387. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=4275894&tool=pmcentrez&rendertype=abstract.

Schuenemann, V.J., 2013. Genome-Wide Comparison of medieval and modern Mycobacterium leprae. *Science*, 341(July), pp.179–184.

Silva, J.C. et al., 2011. Genome sequences reveal divergence times of malaria parasite lineages. *Parasitology*, 138(13), pp.1737–49. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3081533&tool=pmcentrez&rendertype=abstract.

Singer, R., 1960. Some Biological Aspects of the Bushman. *Zeitschrift für morphologie und Anthropologie*, 1(6).

Siravenha, L.G.L.Q. et al., 2016. Detection of HCV Persistent Infections in the Dental Pulp: A Novel Approach for the Detection of Past and Ancient Infections E. M. Shankar, ed. *PLOS ONE*, 11(10), p.e0165272. Available at: http://dx.plos.org/10.1371/journal.pone.0165272 [Accessed November 1, 2016].

Slatkin, M. & Racimo, F., 2016. Ancient DNA and human history. *Proceedings of the National Academy of Sciences*, 2016(23), pp.1–8.

Spriggs, M., 2008. Ethnographic parallels and the denial of history. *World Archaeology*, 40(4), pp.538–552.

Tanabe, K. et al., 2010. Plasmodium falciparum Accompanied the Human Expansion out of Africa. *Current Biology*, 20(14), pp.1283–1289.

Tang, J.W. et al., 2003. Brain stem encephalitis caused by primary herpes simplex 2 infection in a young woman. *Journal of neurology, neurosurgery, and psychiatry*, 74(9), pp.1323–5. Available at: http://www.ncbi.nlm.nih.gov/pubmed/12933947 [Accessed January 26, 2017].

Taubenberger, J.K. et al., 2012. Reconstruction of the 1918 influenza virus: Unexpected rewards from the past. *mBio*, 3(5).

Taylor, L.H., Latham, S.M. & Woolhouse, M.E., 2001. Risk factors for human disease emergence. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, 356(1411), pp.983–9. Available at: http://www.ncbi.nlm.nih.gov/pubmed/11516376%5Cnhttp://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC1088493.

Thackeray, J. & Brink, J., 2004. Damaliscus niro horns from Wonderwerk Cave and other Pleistocene sites: morphological and chronological considerations.

Thomsen, P.F. & Willerslev, E., 2015. Environmental DNA – An emerging tool in conservation for monitoring past and present biodiversity. *Biological Conservation*, 183, pp.4–18.

Tishkoff, S.A. et al., 2001. Haplotype diversity and linkage disequilibrium at human G6PD: recent origin of alleles that confer malarial resistance. *Science (New York, N.Y.)*, 293(5529), pp.455–62. Available at: http://www.sciencemag.org/content/293/5529/455.full [Accessed August 25, 2014].

Tolle, M.A., 2009. Mosquito-borne Diseases. *Current Problems in Pediatric and Adolescent Health Care*, 39(4), pp.97–140.

Tonnang, H.E.Z., Kangalawe, R.Y.M. & Yanda, P.Z., 2010. Predicting and mapping malaria under climate change scenarios: the potential redistribution of malaria vectors in Africa. *Malaria journal*, 9, p.111. Available at: http://www.ncbi.nlm.nih.gov/pubmed/20416059 [Accessed January 25, 2017].

Trueba, G. & Dunthorn, M., 2012. Many neglected tropical diseases may have originated in the Paleolithic or before: new insights from genetics. S. Brooker, ed. *PLoS neglected tropical diseases*, 6(3), p.e1393. Available at: http://dx.plos.org/10.1371/journal.pntd.0001393 [Accessed May 25, 2014].

Underdown, S., Kumar, K. & Houldcroft, C., 2017. Network analysis of the hominin origin of Herpes Simplex virus 2 from fossil data. *bioRxiv*.

Wadley, L., 2007. Announcing a Still Bay industry at Sibudu Cave, South Africa. *Journal of Human Evolution*, 52(6), pp.681–689.

Wadley, L., 2012. Two “moments in time” during Middle Stone Age occupations of Sibudu, South Africa. Available at: http://www.sahumanities.org.za [Accessed January 25, 2017].

Wardeh, M. et al., 2015. Database of host-pathogen and related species interactions, and their global distribution. *Scientific Data*, 2, p.150049. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=4570150&tool=pmcentrez&rendertype=abstract.

Watts, I., 2010. The pigments from Pinnacle Point Cave 13B, Western Cape, South Africa. *Journal of*

Human Evolution, 59(3–4), pp.392–411.

Webb Jr, J.L., 2005. Malaria and the Peopling of Early Tropical Africa Malaria and the Peopling of Early Tropical Africa*. *Source Journal of World History*, 16111185(3), pp.269–291. Available at: http://www.jstor.org/stable/20079330 [Accessed January 25, 2017].

Wertheim, J.O. et al., 2014. Evolutionary origins of human herpes simplex viruses 1 and 2. *Molecular Biology and Evolution*, 31(9), pp.2356–2364. Available at: http://mbe.oxfordjournals.org/cgi/doi/10.1093/molbev/msu185 [Accessed September 1, 2016].

White, T.D. et al., 2003. Pleistocene Homo sapiens from Middle Awash, Ethiopia. *Nature*, 423(6941), pp.742–747.

Willerslev, E. & Cooper, A., 2005. Review Paper. Ancient DNA. *Proceedings of the Royal Society of London B: Biological Sciences*, 272(1558).

Wolfe, N.D., Dunavan, C.P. & Diamond, J., 2007. Origins of major human infectious diseases. *Nature*, 447(7142), pp.279–83. Available at: http://dx.doi.org/10.1038/nature05775 [Accessed May 27, 2014].

Wood, J., Milner, G. & Harpending, H., 1992. The osteological paradox: problems of inferring prehistoric health from skeletal samples. *Current Anthropology*, 33(4), pp.343–370. Available at: http://www.jstor.org/stable/10.2307/2743861 [Accessed August 5, 2014].

Woolhouse, M.E.J. & Gowtage-Sequeria, S., 2005. Host Range and Emerging and Reemerging Pathogens. *Emerging Infectious Diseases*, 11(12), pp.1842–1847. Available at: http://wwwnc.cdc.gov/eid/article/11/12/05-0997_article.htm [Accessed January 25, 2017].

Worobey, M. et al., 2016. 1970s and “Patient 0” HIV-1 genomes illuminate early HIV/AIDS history in North America. *Nature*, 539(7627), pp.98–101. Available at: http://www.nature.com/doifinder/10.1038/nature19827 [Accessed October 31, 2016].

Wu, X. et al., 2016. Impact of climate change on human infectious diseases: Empirical evidence and human adaptation. *Environment International*, 86, pp.14–23.

Yellen, J. et al., 1995. A middle stone age worked bone industry from Katanda, Upper Semliki Valley, Zaire. *Science*, 268(5210).

870 Ziegler, M. et al., 2013. Development of Middle Stone Age innovation linked to rapid climate change.
871 *Nature Communications*, 4(1905), pp.1–9.

872 Zilhão, J. et al., 2010. Symbolic use of marine shells and mineral pigments by Iberian Neandertals.
873 *Proceedings of the National Academy of Sciences of the United States of America*, 107(3), pp.1023–
874 8. Available at: http://www.pnas.org/cgi/content/long/107/3/1023.

875 **Table 1**

876

Pathogen	**Modern nucleic acids**	**Ancient nucleic acids**	**Fossils**
RNA viruses	Primate lentiviruses HTLV (Compton et al. 2013); HCV, tooth pulp (Siravenha et al. 2016)	Plant pathogens (Barley Stripe Mosaic Virus) [Smith, Clapham, 2014] and tomato mosaic tobamovirus RNA in ancient glacial ice. Polar Biol [Castello et al, 1999]	
DNA viruses	Papillomaviruses, herpesviruses, adenoviruses, polyomaviruses [Houldcroft, 2017]	Coprolite {Appelt}, Smallpox (variola virus), permafrost (Biagini et al., 2012)	
Bacteria	Coalescent analysis of modern genetic diversity, trematode blood flukes (Crellen et al. 2016)	MTB, Y. pestis (Bos et al. 2011; Rasmussen et al. 2015); oral microbiome (Adler et al. 2013); cholera from 1849 preserved intestine {Devault, 2014}; syphilis (T. pallidum) [Montiel, 2012]	
Parasites	Body and hair lice (Boutellis et al. 2014)	Helminth egg aDNA [Loreille 2001]; malaria aDNA from blood slides [Gelabert, 2016] and Roman-era teeth [Marciniak, 2016]	Eggs (Mitchell 2013)