

ARTICLE

# Germs, genomes, and global history in the time of COVID-19<sup>†</sup>

Kyle Harper

Department of Classics and Letters, University of Oklahoma, 650 Parrington Oval, CARN 110, Norman, OK 73019-4042, USA

Corresponding author. E-mail: [kyleharper@ou.edu](mailto:kyleharper@ou.edu)

## Abstract

The COVID-19 pandemic presents opportunities and challenges for historians working on the global history of health and disease. This article argues that the history of disease will benefit from interdisciplinary work that brings together historians, microbiologists, and archaeologists. Genomes are historical archives, in two complementary ways. Palaeogenomics provide direct access to genomes of the past, while phylogenetics furnish historical insight from evolutionary relationships. Both palaeogenomics and phylogenetics have already contributed enormously to the history of disease, helping us understand how human ecological transformation drives the evolution of our microparasites.

**Keywords:** history of coronavirus; history of measles virus; palaeogenomics; phylogenetics

## Historical pandemics in a time of COVID-19

Even in normal times, to be a Roman historian is to be practised in responding gently to opinions about ‘parallels’ between past and present. Being a student of Roman pandemics in the age of COVID-19, though, has offered a new platform to disappoint history buffs everywhere by sharing the home truth that we are, for better or worse, not the Roman empire. The Roman empire was indeed struck by a succession of important pandemics, including the Antonine Plague (c. 165–80 CE), the Plague of Cyprian (c. 250–65 CE), and the Justinianic Plague (c. 541–44 CE, the opening salvo in the first plague pandemic, which lasted until the 740s CE). But the differences between the ancient world and our own are overwhelming. The Roman empire was a loosely integrated congeries of subsistence-level societies where most people, most of the time, died of infectious disease; medical controls were non-existent or immaterial for disease outcomes; mortality tolls from epidemics were vastly higher; and the sociopolitical consequences were highly variable. There may be generalizable patterns to be extrapolated from the history of pandemics, but the truth is that we are a long way from having a comparative sociology of epidemic disease.

Richer comparisons are probably possible for the last two centuries or so, starting with the cholera epidemics that threatened to undermine the fragile control over crisis mortality that had recently been achieved in western Europe, and continuing with the influenza pandemics of the twentieth century. The further into the past we travel, the greater the sociological differences and empirical uncertainties become. Yet the conclusion to be drawn is hardly that pre-modernists

---

<sup>†</sup>I would like to thank Peter Sarris, Heidi Tworek, and Tina Warinner for reading a draft of this article and offering valuable suggestions.

should sit on the sidelines. The COVID-19 pandemic underscores the importance of bringing together history, ecology, and biology on deep timescales. Infectious disease is an inherently biological phenomenon. It would be impossible to understand the COVID-19 pandemic in full without a basis in the virology of SARS-CoV-2. The evolution of this virus, with its particular biological properties, is unambiguously the proximate cause of the current pandemic. At the same time, the global ecology that we have created, a product of deep human history, is the context for the evolution, the transmission, and ultimately the impact of this medical crisis. Pandemics happen at the intersection of evolution, ecology, and history.

Global historians have long worked to help us understand the interplay of human history and disease ecology. William McNeill's *Plagues and Peoples* of 1976 remains the model, and the architecture of McNeill's narrative has proven remarkably durable.<sup>1</sup> The human dispersal out of Africa and its tropical environments, the domestication of animals in the Neolithic Revolution, the 'confluence of the civilized disease pools' of Eurasia in the Iron Age (which we might now date earlier than McNeill, with new kinds of evidence), the role of central Asian nomads in the origins of the Black Death, the catastrophic health dimensions of the Columbian Exchange (already traced by Crosby and others when McNeill wrote): these have remained indispensable themes in the global history of disease, precisely because they connect history with disease ecology in a biologically meaningful way.<sup>2</sup> The history of infectious disease inherently lends itself to global history.<sup>3</sup> It is impossible to write this kind of history without foregrounding the interplay of biogeography, dispersal, and exchange. Microbes have no respect for the conventional periods into which historians divide the past, not to mention political, cultural, and linguistic borders.

If the COVID-19 pandemic stirs interest in the global history of infectious disease, we are fortunate to work at a time when new kinds of evidence for such a history are rapidly coming into the picture. In the same year that *Plagues and Peoples* appeared, a Belgian microbiologist named Walter Fiers sequenced the genome of a small virus.<sup>4</sup> It was a scientific milestone, the first complete genome of any organism to be sequenced. But, for decades, genome sequencing was a plodding and expensive undertaking. Not until 1995 was the first complete bacterial genome sequenced. The Human Genome Project took thirteen years, roughly US\$1 billion, and a massive collaboration to map the human genome. Over the last decade or so, however, technologies known as high-throughput sequencing (often called 'next generation' sequencing in its first iteration) have fundamentally changed the equation.<sup>5</sup> A genome can be sequenced in less than a day, and the expense has plummeted. The amount of data is staggering, and it doubles every eighteen months.<sup>6</sup>

Genomic data are inherently historical; genes are passed from one generation to the next (with variation) for humans as for our microscopic pathogens. Genomes are a kind of historical archive, an organic record of evolutionary history. The genomes being sequenced in ever greater numbers represent a potential trove of historical insight. Broadly, microbial genomes can inform the history of infectious disease through two complementary but distinct approaches: palaeogenomics and phylogenetics. Palaeogenomics is the study of historical (or prehistorical) genetic material.

<sup>1</sup>William McNeill, *Plagues and Peoples* (Harmondsworth: Penguin, 1976).

<sup>2</sup>Alfred Crosby, *The Columbian Exchange: Biological and Cultural Consequences of 1492* (Westport: Greenwood Press, 1973).

<sup>3</sup>E.g. Monica Green, 'The Globalisations of Disease', in *Human Dispersal and Species Movement: From Prehistory to the Present*, ed. Nicole Boivin, Rémy Crassard, and Michael Petraglia (Cambridge: Cambridge University Press, 2017), 494–520.

<sup>4</sup>Specifically *Escherichia* virus MS2, a virus that infects the bacterium *E. coli*: see Walter Fiers *et al.*, 'Complete Nucleotide Sequence of Bacteriophage MS2 RNA: Primary and Secondary Structure of the Replicase Gene', *Nature* 260 (1976): 500–7.

<sup>5</sup>Jason Reuter, Damek Spacek, and Michael Snyder, 'High-Throughput Sequencing Technologies', *Molecular Cell* 58, no. 4 (2015): 586–97.

<sup>6</sup>Measured by the number of bases in GenBank, <https://www.ncbi.nlm.nih.gov/genbank/statistics/>.

In the last decade, high-throughput sequencing has enabled revolutionary methodological advances in the retrieval and validation of ancient DNA, generally recovered from archaeological contexts (though also from museum specimens).<sup>7</sup> Ancient DNA has been evocatively called ‘an evolutionary photo album’: rather than having to infer things about an ancestor from its descendants, you can look directly at the ancestor.<sup>8</sup> Phylogenetics is the study of evolutionary relationships between groups of organisms: for instance, between species or between populations (such as different strains of a virus). As genome sequencing has become technologically ubiquitous, genome-wide molecular data have become the basis for reconstructing phylogenetic trees.<sup>9</sup> Phylogenies can be constructed from contemporary samples; they do not require ancient or archaeological DNA, although palaeogenomic samples can greatly enrich phylogenies. Phylogenies – as we will explore below – can be rich sources of historical information.

The symmetry between McNeill’s book and the first viral genome is a starting point to ask how the insights generated by genomic data might inform and enrich the kind of project represented by *Plagues and Peoples*. The genomic evidence is particularly useful for tracing historical connections that are invisible or obscured by the parochial and constrictive nature of traditional documents. The phylogeny of the coronavirus family itself is an example of how genomic data open up completely new possibilities for historians. There is no ‘history of coronaviruses’, and this family of viruses is completely absent in every canonical history of disease. But the phylogeny of the coronavirus family makes it possible, at least in a limited fashion, to trace the connections between human transformation of planetary ecologies and the emergence of viruses like SARS-CoV-2. Recent work on the phylogeny of the measles virus provides a point of comparison discussed below. Measles is also a respiratory virus, which emerged in response to the evolutionary opportunities created by human settlements. Its early history is impossible to trace from traditional documents alone but, with the help of phylogenetics, we can at least narrow down the context of its emergence.

Ancient history does not offer ‘parallels’ or ‘lessons’. But the deep history of disease can give us perspective and insight, especially when that history is global in scope and conversant with new genomic data. The history of infectious disease is a prime opportunity for consilience, the ‘jumping together’ of knowledge from different domains and disciplines.<sup>10</sup> Historians bring to the table domain knowledge of human demographic, economic, and social history that is indispensable for understanding the ecological context of biological evolution. They can trace how technology, trade, war, and urbanization (to take just some examples) have shaped, and been shaped by, the history of disease. And they bring a sharp sensibility for contingency, conjuncture, and the burden of the past that is perfectly at home in evolutionary biology.<sup>11</sup> This kind of history will require our field to value collaboration and team science and to be comfortable with co-authorship, not to mention risk-taking. Historians should not surrender this project to others, nor hope to do it except by serious engagement with the natural sciences.

<sup>7</sup>Johannes Krause and Svante Pääbo, ‘Genetic Time Travel’, *Genetics* 203 (2016): 9–12; Courtney Hofman and Christina Warinner, ‘Ancient DNA 101: An Introductory Guide in the Era of High-Throughput Sequencing’, *SAA Archaeological Record* 19 no. 1 (2019): 18–25.

<sup>8</sup>Oliver Smith and Thomas Gilbert, ‘Ancient RNA’, in *Paleogenomics: Genome-Scale Analysis of Ancient DNA*, ed. C. Lindqvist and Om Rajora (Cham: Springer, 2019), 54.

<sup>9</sup>Ziheng Yang and Bruce Rannala, ‘Molecular Phylogenetics: Principles and Practice’, *Nature Reviews: Genetics* 13 (2012): 303–14.

<sup>10</sup>E. O. Wilson, *Consilience* (New York: Alfred Knopf, 1998); for a conspectus of the possibilities for historians, see Michael McCormick, ‘History’s Changing Climate: Climate Science, Genomics, and the Emerging Consilient Approach to Interdisciplinary History’, *Journal of Interdisciplinary History* 42, no. 2 (2011): 251–73.

<sup>11</sup>See the ‘Introduction’ to ‘AHR Roundtable: History Meets Biology’, *American Historical Review* 119, no. 5 (2014): 1492–99.

## History with genes

Coronaviruses are human pathogens without a history, in the traditional sense of history. There is no entry for coronaviruses in standards like the *Cambridge World History of Human Disease*, and human coronaviruses had only been recently discovered when McNeill wrote *Plagues and Peoples*. They are a family of RNA viruses so called because protein spikes protruding from the envelope of the virion give the impression of a 'crown' (Latin *corona*). Coronaviruses infect a wide range of birds and mammals. Several species cause serious disease in livestock. Bats are particularly important wild reservoirs of coronaviruses.<sup>12</sup> There are seven coronaviruses known to infect humans.<sup>13</sup> Four of these (NL63, 229E, OC43, HKU1) cause mild respiratory illness. Coronaviruses are basically ubiquitous members of the human disease pool, one of the major causes of the 'common cold'.<sup>14</sup> Most of us have probably suffered a coronavirus infection. But three species of coronavirus have only recently crossed to humans: MERS (Middle East Respiratory Syndrome), SARS-CoV-1 (the cause of the 2003 pandemic), and SARS-CoV-2, the agent of the COVID-19 pandemic. These three species all cause severe disease. It seems possible that SARS-CoV-2 could join the first four as a permanent constituent of the human disease pool.

Before considering what we know about the history of coronaviruses, it bears putting in deeper context just how unusual the human disease pool is. Relative to other primates, humans have a huge number of pathogens.<sup>15</sup> And a disproportionate number of these are viruses and bacteria (rather than worms and protozoa, like malaria). In the words of Charles Nunn and Sonia Altizer, primates 'reportedly harbor only a tiny fraction of the diversity of parasites infecting contemporary human populations. Furthermore, whereas the greatest diversity of parasites reported from wild primates is captured by helminths and protozoa ... the majority of modern-day human pathogens are bacteria, viruses, and fungi, many of which cause acute infections and are often associated with contact-based transmission.'<sup>16</sup> And a startling number of these directly transmitted pathogens are *specialists* in the infection of human beings, rather than generalists that circulate among a number of host species.

The number of respiratory pathogens endemic in human populations is strange, and it is a product of the interplay between evolution, ecology, and history in deep time. Compare the ecology of *Homo sapiens* to its closest relative, the chimpanzee. Chimps live in group sizes of 150 or less, have a global population of under half a million, are well adapted to a reliable base of forest foods, have a ranging lifestyle, and are not surrounded by an array of domesticated and commensal animals.<sup>17</sup> Humans live in groups of millions, have a population of over 7.5 billion, wantonly harvest the biosphere for food, have a sedentary lifestyle, and are surrounded by domesticates and commensals (such as house flies, mice, etc.). From the perspective of a parasite, which is to say from the blind perspective of evolution, we are a far more attractive host. One of the hardest evolutionary challenges for any microparasite is transmitting from one host to the next.<sup>18</sup> Respiratory pathogens have to make it from one respiratory tract to the next. Chimpanzees have few viruses,

<sup>12</sup>V. M. Corman, 'Evidence for an Ancestral Association of Human Coronavirus 229E with Bats', *Journal of Virology* 89 (2015): 11858–70.

<sup>13</sup>Coronaviridae Study Group of the International Committee on Taxonomy of Viruses, 'The Species Severe Acute Respiratory Syndrome-Related Coronavirus: Classifying 2019-nCoV and Naming It SARS-CoV-2', *Nature: Microbiology* 5 (2020): 536–44.

<sup>14</sup>Terho Heikkinen and Asko Järvinen, 'The Common Cold', *Lancet* 361, no. 9351 (2003): 51–9.

<sup>15</sup>Although there is no satisfactory tally. See Maya Wardeh, Claire Risley, Marie Kirsty McIntyre, Christian Setzkorn, and Matthew Baylis, 'Database of Host–Pathogen and Related Species Interactions, and Their Global Distribution', *Scientific Data* 2 (2015), <https://doi.org/10.1038/sdata.2015.49>.

<sup>16</sup>Charles Nunn and Sonia Altizer, *Infectious Diseases in Primates: Behavior, Ecology and Evolution* (Oxford: Oxford University Press, 2006), 255.

<sup>17</sup>See, in general, *ibid.*; Jessica Brinkworth and Ekaterina Pechenkina, eds., *Primates, Pathogens, and Evolution* (New York: Springer, 2013).

<sup>18</sup>Paul Ewald, 'Transmission Modes and the Evolution of Virulence', *Human Nature* 2 (1991): 1–30.

and most of those are either vector-borne, extreme generalists, or picked up incidentally from humans.<sup>19</sup> Human settlements, by contrast, are an irresistible network of densely packed lungs. Chimps have few respiratory viruses; we have many. We are the sneezing ape.

It is virtually impossible to recover the history of specific respiratory pathogens from written sources. In historical nosologies, especially before germ theory, symptoms (like fever) were often considered a disease, not a physiological phenomenon caused by a specific infectious agent.<sup>20</sup> Retrospective diagnosis from historical texts is methodologically fraught, even in the case of diseases with relatively distinct pathologies, such as bubonic plague, smallpox, or leprosy.<sup>21</sup> For diseases that present with a relatively generic course of illness, as do most respiratory or gastroenteric diseases, identification from historical sources alone is almost hopeless. Hence, the history of disease has tended to focus on glamorous villains with distinct personalities, to the neglect of the mundane workaday constituents of the human disease pool. Call it the ‘great germ’ theory of history.

Palaeogenomics and phylogenetics are starting to enrich the history of infectious disease and to fill in some of these blanks. The impact of palaeogenomics on the history of disease has already been profound.<sup>22</sup> Thus far, discovery has been weighted toward bacteria and protozoa, which are bigger, have bigger genomes, and preserve better, so are more likely to be recovered from archaeological contexts. The study of the plague bacterium, *Yersinia pestis*, has always been in the vanguard of palaeogenomics, and the revelations have come non-stop. To name just a few instances, palaeogenomic evidence has put to rest long-standing debates about the causative agent of the Justinianic Plague (541–44 CE, the initial outbreak of the first pandemic) and the Black Death, securely identifying the bacterium *Yersinia pestis* as the microbial cause of both pandemics.<sup>23</sup> Palaeogenomic evidence has also, unexpectedly, documented the presence of plague in Stone Age and Bronze Age Europe, opening new questions about the connection between and movement of populations across the steppe in prehistory, and suggesting that McNeill’s ‘confluence’ of disease pools started earlier than the written record alone lets us see.<sup>24</sup> It has allowed history at the molecular scale, tracing the acquisition of critical virulence factors that made the plague so efficient and so deadly.<sup>25</sup> Palaeogenomic evidence helps to answer some of the most important and otherwise intractable questions about the first pandemic, such as whether it penetrated far to the

<sup>19</sup>See the Global Primate Parasite Database, <https://parasites.nunn-lab.org>.

<sup>20</sup>On fever, see Dale Smith, ‘Medical Science, Medical Practice, and the Emerging Concept of Typhus in Mid-Eighteenth-Century Britain’, *Medical History* 25, supplement no. 1 (1981): 121–34; or, more expansively, Christopher Hamlin, *More Than Hot: A Short History of Fever* (Baltimore: Johns Hopkins University Press, 2015).

<sup>21</sup>For a highly critical assessment of the possibilities of retrospective diagnosis, see Jon Arrizabalaga, John Henderson, and R. K. French, *The Great Pox: The French Disease in Renaissance Europe* (New Haven: Yale University Press, 2014). More generally, see Piers Mitchell, ‘Retrospective Diagnosis and the Use of Historical Texts for Investigating Disease in the Past’, *International Journal of Paleopathology* 1, no. 2 (2011): 81–8; and recently A. J. Lerner, ‘Retrospective Diagnosis: Pitfalls and Purposes’, *Journal of Medical Biography* 27, no. 3 (2019): 127–8.

<sup>22</sup>Stephanie Marciniak and Hendrik Poinar, ‘Ancient Pathogens through Human History: A Palaeogenomic Perspective’, in *Paleogenomics*, ed. Lindqvist and Rajora, 115–38.

<sup>23</sup>On the Black Death, see Kirsten I. Bos *et al.*, ‘A Draft Genome of *Yersinia pestis* from Victims of the Black Death’, *Nature* 478 (2011): 506–10. On Justinianic Plague, see Ingrid Wiechmann and Gisela Grupe, ‘Detection of *Yersinia pestis* DNA in Two Early Medieval Skeletal Finds from Aschheim (Upper Bavaria, 6th Century A.D.)’, *American Journal of Physical Anthropology* 126 (2005): 48–55; Michael Feldman *et al.*, ‘A High-Coverage *Yersinia pestis* Genome from a Sixth-Century Justinianic Plague Victim’, *Molecular Biology and Evolution* 33 (2016): 2911–23; David M. Wagner *et al.*, ‘*Yersinia pestis* and the Plague of Justinian 541–543 AD: A Genomic Analysis’, *Lancet Infectious Diseases* 14 (2014): 319–26. Some earlier claims to have sequenced ancient plague are problematic.

<sup>24</sup>Simon Rasmussen *et al.*, ‘Early Divergent Strains of *Yersinia pestis* in Eurasia 5,000 Years Ago’, *Cell* 163 (2015): 571–82; Aida Andrades Valtueña *et al.*, ‘The Stone Age Plague and Its Persistence in Eurasia’, *Current Biology* 27 (2017): 3683–91; Nicolás Rascovan *et al.*, ‘Emergence and Spread of Basal Lineages of *Yersinia pestis* during the Neolithic Decline’, *Cell* 176 (2019): 295–305.

<sup>25</sup>Maria A. Spyrou *et al.*, ‘Analysis of 3800-Year-Old *Yersinia pestis* Genomes Suggests Bronze Age Origin for Bubonic Plague’, *Nature Communications* 9 (2018): 2234.



west and deep into the countryside (emphatically, yes).<sup>26</sup> It speaks to critical questions about the geographic origins, itinerary, and subsequent persistence of the plague germ in western Eurasia: we now have reason to believe that the strain of the germ leading to the Black Death was spreading for decades before the eruption of the human pandemic and that the pathogen probably became established in animal reservoirs in western Eurasia for centuries.<sup>27</sup> The case of plague is *sui generis*, but it showcases the revolutionary potential of palaeogenomics.

Palaeogenomic evidence has also shown that some unsuspected and less famous pathogens may have played an important role in human history. For instance, victims from a mass grave in sixteenth-century southern Mexico yielded evidence for one of the bacterial agents of paratyphoid fever (*Salmonella enterica* Paratyphi C).<sup>28</sup> The discovery suggests that this enteric fever contributed to the devastating epidemics of the 1540s and needs to be reckoned among the deadly introductions of the conquistadors. Palaeogenomic evidence for viruses is still much rarer, but there is reason for optimism. Human parvovirus B19 – a single-stranded DNA virus – has been reported from a Neolithic sample, and Hepatitis B virus from the Bronze Age.<sup>29</sup> Smallpox, caused by the double-stranded DNA virus *Variola major*, has been recovered from a seventeenth-century skeleton.<sup>30</sup> Most of the work on viruses to date has focused on DNA viruses. RNA viruses tend to have smaller genomes that degrade much faster, and the laboratory methods used to retrieve ancient DNA do not recover RNA.<sup>31</sup> The oldest sample of a human RNA virus sequenced so far is the measles virus, recently extracted from formalin-fixed lung tissue preserved in a Berlin collection from 1912.<sup>32</sup> For the most part, RNA viruses are a blind spot in current palaeogenomic studies.

There are, of course, other technical limitations of palaeogenomic approaches. So far most of the pathogen DNA recovered from archaeological samples has been retrieved from the pulp cavity, where hard enamel creates an environment conducive to preservation. But this means that, to be preserved at all, the pathogen must have been circulating abundantly in the blood at the time of death, which is often not the case for important infectious diseases, especially the acute diseases that are of interest to historians. Palaeoproteomics, the study of all proteins associated with an ancient organism, may open new doors to ancient biomolecules beyond DNA and preservation environments outside the tooth; proteins preserve better than DNA and might ultimately be the best palaeomolecular avenue for uncovering the history of RNA viruses, but there are still daunting technical challenges at present.<sup>33</sup>

In the meantime, phylogenetics can enrich and expand what is possible in the history of disease. A phylogeny is a representation of evolutionary relationships and is thus essentially historical. In the genomic age, phylogenies are based on patterns of genetic relatedness, and the proliferation of

<sup>26</sup>Marcel Keller *et al.*, 'Ancient *Yersinia pestis* Genomes from across Western Europe Reveal Early Diversification during the First Pandemic (541–750)', *Proceedings of the National Academy of Sciences* 116, no. 25 (2019): 12363–72.

<sup>27</sup>Maria A. Spyrou *et al.*, 'Phylogeography of the Second Plague Pandemic Revealed through Analysis of Historical *Yersinia pestis* Genomes', *Nature Communications* 10 (2019): 1–13, <https://doi.org/10.1038/s41467-019-12154-0>. See, for example, historians integrating these new insights: Monica Green, 'Taking "Pandemic" Seriously: Making the Black Death Global', *Medieval Globe* 1 (2014): 27–61.

<sup>28</sup>Åshild Vågane *et al.*, '*Salmonella enterica* Genomes from Victims of a Major Sixteenth-Century Epidemic in Mexico', *Nature Ecology and Evolution* 2, no. 3 (2018): 520–8.

<sup>29</sup>Barbara Mühlemann *et al.*, 'Ancient Human Parvovirus B19 in Eurasia Reveals Its Long-Term Association with Humans', *Proceedings of the National Academy of Sciences* 115 (2018): 7557–62; Barbara Mühlemann *et al.*, 'Ancient Hepatitis B Viruses from the Bronze Age to the Medieval Period', *Nature* 577 (2018): 418–23.

<sup>30</sup>Ana T. Duggan *et al.*, '17th Century *Variola* Virus Reveals the Recent History of Smallpox', *Current Biology* 26 (2016): 3407–12.

<sup>31</sup>Smith and Gilbert, 'Ancient RNA'.

<sup>32</sup>Ariane Düx *et al.*, 'Measles Virus and Rinderpest Virus Divergence Dated to the Sixth Century BCE', *Science* 368 (2020): 1367–70.

<sup>33</sup>Christina Warinner *et al.*, 'Pathogens and Host Immunity in the Ancient Human Oral Cavity', *Nature: Genetics* 46 (2014): 336–44. I am grateful to Tina Warinner for illuminating conversation on these points.

genomic data has enormously expanded the application of phylogenetics. This is especially true for microorganisms. Since molecular phylogenetics is so much more efficient than older laboratory methods of characterizing viruses, bacteria, and protozoa, faster and cheaper genomic sequencing has revolutionized our understanding of the diversity, ubiquity, and evolutionary history of microorganisms. This is true of disease-causing microbes but also of symbiotic, commensal, and environmental microorganisms.

Phylogenetics can be a powerful tool, and while the applications for history are still novel and fluid, we can characterize at least three interrelated ways in which evolutionary relationships have so far been valuable for understanding the history of human disease. First, phylogenies can reveal the ancestral hosts of microorganisms that have evolved to become pathogens of humans. Second, they can place evolutionary relationships in space and thus be integrated with the history of human migrations, contacts, and exchanges. Third, they can be situated in time, allowing evolutionary events such as disease emergence to be placed in historical context. A brief word on each of these is in order.

Phylogenies illuminate the evolutionary background of our pathogens by identifying the animals that serve or served as hosts of closely related microbes. Excepting a few parasites that have been with us from hominin origins, human diseases have zoonotic origins. In some cases, pathogens that cause disease in humans remain truly zoonotic diseases; they are ill-adapted to human hosts and cannot transmit between humans, so humans are only transient, dead-end hosts.<sup>34</sup> The rabies virus is an example of an important disease that remains a zoonotic disease, with every case representing a cross-species infection. At the other end of the spectrum is a disease like measles, which is well adapted to humans and has no known animal reservoirs. But such human diseases had animal origins, and phylogenies can clarify the animal hosts of closely related or ancestral pathogens. Smallpox, for example, is closely related to a gerbil virus and a camel virus.<sup>35</sup> *Falciparum malaria* is most closely related to a parasite of gorillas and, it has been argued, adapted to humans only in the Neolithic.<sup>36</sup> A number of human diseases evolved from microbes that are naturally hosted by bats and rodents, reflecting the simple abundance of these orders, as well as perhaps their large group sizes.<sup>37</sup>

Human and animal health are closely intertwined, and our history has dramatically changed the configuration of these relationships, from the Neolithic revolution to the rise of industrial farming. Domestication was a major turning point. The very strain of paratyphoid fever that was found in sixteenth-century Mexico probably evolved in the Old World Neolithic following the domestication of pigs. But phylogenetics has complicated the story in many cases. For instance, human and bovine tuberculosis are closely related, but it turns out that bovine tuberculosis evolved from the human form, and not vice versa. Moreover, it is probably better to see domestic livestock as important evolutionary bridges, rather than ultimate sources, of new human diseases. The phylogeny of the measles virus (see below) suggests that it was a parasite of rodents that adapted to goats, sheep, and cattle before becoming a human specialist. Sometimes, the

<sup>34</sup>Nathan Wolfe, Claire Panosian Dunavan, and Jared Diamond, 'Origins of Major Human Infectious Diseases', *Nature Reviews* 447 (2007): 279–83.

<sup>35</sup>Igor V. Babkin and Irina N. Babkina, 'A Retrospective Study of the Orthopoxvirus Molecular Evolution', *Infection, Genetics and Evolution* 12 (2012): 1597–1604; Gianguglielmo Zehender *et al.*, 'Bayesian Reconstruction of the Evolutionary History and Cross-Species Transition of *Variola* Virus and Orthopoxviruses', *Journal of Medical Virology* 90 (2018): 1134–41.

<sup>36</sup>Dorothy Loy *et al.*, 'Out of Africa: Origins and Evolution of the Human Malaria Parasites *Plasmodium falciparum* and *Plasmodium vivax*', *International Journal for Parasitology* 47 (2017): 87–97.

<sup>37</sup>See, e.g., Bastiaan Meerburg, Grant R. Singleton, and Aize Kijlstra, 'Rodent-Borne Diseases and Their Risks for Public Health', *Critical Reviews in Microbiology* 35, no. 3 (2009): 221–70; Angela D. Luis *et al.*, 'A Comparison of Bats and Rodents as Reservoirs of Zoonotic Viruses: Are Bats Special?', *Proceedings of the Royal Society B* 280 (2013), <https://doi.org/10.1098/rspb.2012.2753>; Barbara Han, 'Rodent Reservoirs of Future Zoonotic Diseases', *Proceedings of the National Academy of Sciences* 112, no. 22 (2015): 7039–44; David Hayman, 'Bats as Viral Reservoirs', *Annual Reviews in Virology* 3 (2016): 77–99.

relationship between human pathogens and wild animal reservoirs has become more complicated. The typhus bacterium, for example, is found with some regularity in North American flying squirrels, a fact whose historical import remains unresolved. Was typhus a New World germ introduced to Europe after Columbus, or did it establish itself in New World rodents after European contact? Are there other mammal reservoirs (and, if so, why have these not caused documented spillover events)?<sup>38</sup>

Second, phylogenies can provide spatial coordinates to evolutionary history. Geographic structure in the patterns of genetic diversity can be especially revealing for historians, since it often mirrors or reveals structures of human movement and connectivity.<sup>39</sup> This approach has proven especially fertile in studying the microbial dimensions of the Columbian Exchange. Phylogenies resolve questions about the Old World origins of the major malaria parasites, for instance. The genomic evidence also illuminates the migration of diseases that otherwise have almost no history, such as the virus that causes Hepatitis C, which has been revealed as unintended cargo on the early modern slave trade across the Atlantic.<sup>40</sup> Such applications are not limited to the Columbian Exchange. In a recent study, Monica Green and Lori Jones have shown how much might be learned from phylogenetics about the dispersal of disease throughout the Indian Ocean world. The geography of tuberculosis lineages, for instance, tells a story in which the disease was a presence in South Asia long before European colonialism, then partly overwritten by the introduction of more virulent strains by merchants and imperialists: ‘European global expansion . . . and the more virulent TB strains that accompanied it, seems to have added a further health burden to this already epidemiologically challenged region of the world.’<sup>41</sup> The genomic evidence might thus inspire us to search the traditional evidence with more fine-tuned questions: even if tuberculosis was already present in a population, are there indications of new, deadlier strains associated with human contacts and movements?

Third, phylogenies can be situated in time. Sometimes palaeogenomic samples allow what Johannes Krause and Svante Pääbo have called ‘genetic time travel’, anchoring specific evolutionary transitions to more or less precise, archaeologically defined strata of the past.<sup>42</sup> But, even without ancient genomes, it is possible to measure the evolutionary distance between two related organisms using molecular clock methods. Molecular clocks estimate the time that it has taken for genetic differences to accumulate. Consider a crude analogy. Imagine that you are copying Shakespeare’s *Hamlet* by hand once a day (the generation time) and every day you accidentally copy exactly one word in error (the substitution rate), uncorrected in subsequent copies: ‘To be or snot to be.’ After ten days, there are ten differences between the latest version and the original. If you only have a much later copy of unknown age, and then counted one hundred differences between it and the original, you could treat the differences as an estimate of time: that copy was made one hundred days after the original. Generation times and substitution rates vary across species. They can also change over time, according to the evolutionary context, as do population sizes.<sup>43</sup> Errors may become saturated or get weeded out by purifying selection. And it is important to underscore that molecular clocks only *estimate* the time from which two sampled groups shared a common ancestor using computational methods. Such estimates may lack the precision that

<sup>38</sup>E. Angelakis, ‘The History of Epidemic Typhus’, *Microbiology Spectrum* 4, no. 4 (2016): 1–9.

<sup>39</sup>For applications by historians, see Green, ‘Taking “Pandemic” Seriously’; Philip Slavin, ‘Death by the Lake: Mortality Crisis in Early Fourteenth-Century Central Asia’, *Journal of Interdisciplinary History* 50, no. 1 (2019): 59–90.

<sup>40</sup>Peter Markov *et al.*, ‘Colonial History and Contemporary Transmission Shape the Genetic Diversity of Hepatitis C Virus Genotype 2 in Amsterdam’, *Journal of Virology* 86, no. 14 (2012): 7677–87.

<sup>41</sup>Monica Green and Lori Jones, ‘The Evolution and Spread of Major Human Diseases in the Indian Ocean World’, in *Disease Dispersion and Impact in the Indian Ocean World*, ed. Gwyn Campbell and Eva-Maria Knoll (Cham: Springer, 2020), 35.

<sup>42</sup>Krause and Pääbo, ‘Genetic Time Travel’.

<sup>43</sup>Jade Vincent Membrebe, Marc A. Suchard, Andrew Rambaut, Guy Baele, and Philippe Lemey, ‘Bayesian Inference of Evolutionary Histories under Time-Dependent Substitution Rates’, *Molecular Biology and Evolution* 36 (2019): 1793–1803.



historians are comfortable with, but they can be highly revealing nonetheless, especially for understanding the age of human pathogens.

The value and complexity of molecular clock analysis can be illustrated by recent work on the virus that causes measles, a human disease whose evolutionary story throws light on the history of respiratory pathogens. Measles is a major infectious disease that has been largely controlled (though the anti-vaccination movement has triggered outbreaks in Europe and the US); prior to the vaccination era, measles infection was a rite of passage for children.<sup>44</sup> The course of illness starts with a prodromal fever lasting two to three days, accompanied by a cough, coryza, and conjunctivitis. A red, maculopapular rash spreads from the head to the entire body, lasting around eight days. Measles mortality can depend on nutritional status and comorbidity, but absent good medical treatment it was (and still is) a serious disease. Measles is spread by the respiratory route and is extremely contagious – one of the most contagious diseases known. But survivors acquire strong and lasting immunity to subsequent infection.

Measles is an organism that should not exist, in the sense that it would have gone quickly extinct had it specialized in the infection of almost any host other than humans. Being a directly transmitted, highly contagious virus that causes acute illness and confers lifelong immunity but lacks animal reservoirs is not usually a sound evolutionary strategy. Measles exists because we exist, or, more properly, measles avoided hasty extinction because it happened to adapt to a host that had learned to concentrate populations in enormous numbers. The epidemiology of measles has been particularly well studied, and the virus goes locally extinct wherever the population is below about 250,000, which is the critical threshold for the virus to establish permanently in a population.<sup>45</sup>

Measles is a morbillivirus whose closest relative is the Rinderpest virus, a devastating pathogen of cattle that humans deliberately drove to extinction; in turn, these two viruses are closely related to the agent of *peste des petits ruminants*, a disease that infects goats and sheep as well as some wild animals.<sup>46</sup> Broader genomic surveillance of wild animals has identified a range of morbilliviruses, many of which infect rodents.<sup>47</sup> The branch including *peste des petits ruminants*, Rinderpest, and measles diverged from an ancestral rodent virus. Molecular clock analysis shows that the measles virus evolved relatively recently. The oldest estimate for the divergence of Rinderpest and measles had been the ninth century CE.<sup>48</sup> But a new analysis likely improves on this estimate, both by acquiring older samples (including the aforementioned lung tissue from 1912) and by using more complex evolutionary models that try to account for changes in the rate of genetic substitution.<sup>49</sup> These refinements push back the divergence of measles virus and Rinderpest virus to the first millennium BCE, with the mean estimate at 528 BCE (which is the mean within a range of probabilities, not an exact date). What is more remarkable is that this divergence aligns with the period in history – the last centuries BCE – when human societies first built cities above the critical population threshold of the measles virus. Even though there are inevitable margins of error in understanding ancient city sizes, there is a degree of consensus that in the last centuries BCE human settlements of several hundred thousand persons arose.<sup>50</sup> Cities such as ancient Rome and

<sup>44</sup>R. E. Davaro, 'Measles Virus', in *Viruses and the Lung: Infections and Non-Infectious Viral-Linked Lung Disorders*, ed. Armando E. Fraire, Bruce A. Woda, Raymond M. Welsh, and Richard L. Kradin (Berlin: Springer-Verlag, 2014), 71–8.

<sup>45</sup>Andrew Cliff, Peter Haggett, and Matthew Smallman-Raynor, *Measles: An Historical Geography of a Major Human Viral Disease: From Global Expansion to Local Retreat, 1840–1990* (Oxford: Blackwell, 1993).

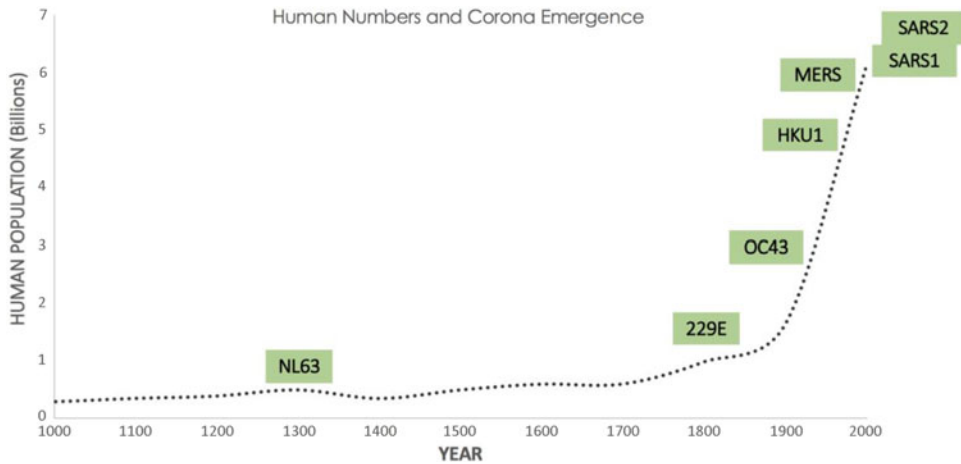
<sup>46</sup>Düx *et al.*, 'Measles Virus and Rinderpest Virus'.

<sup>47</sup>Wisseem Ghawar *et al.*, 'Insight into the Global Evolution of Rodentia Associated Morbilli-Related Paramyxoviruses', *Scientific Reports* 7 (2017): 1–12.

<sup>48</sup>Yuki Furuse, Akira Suzuki, and Hitoshi Oshitani, 'Origin of Measles Virus: Divergence from Rinderpest Virus between the 11th and 12th Centuries', *Virology Journal* 7 (2010): 52.

<sup>49</sup>Düx *et al.*, 'Measles Virus and Rinderpest Virus'.

<sup>50</sup>Hiroko Inoue *et al.*, 'Urban Scale Shifts since the Bronze Age: Upsweeps, Collapses, and Semiperipheral Development', *Social Science History* 39 (2015): 175–200.



**Figure 1.** Human population numbers and emergence of coronavirus species.

Sources: HYDE 3.1, <https://themasites.pbl.nl/tridion/en/themasites/hyde/index.html>; Forni *et al.*, 'Molecular Evolution of Human Coronavirus Genomes'.

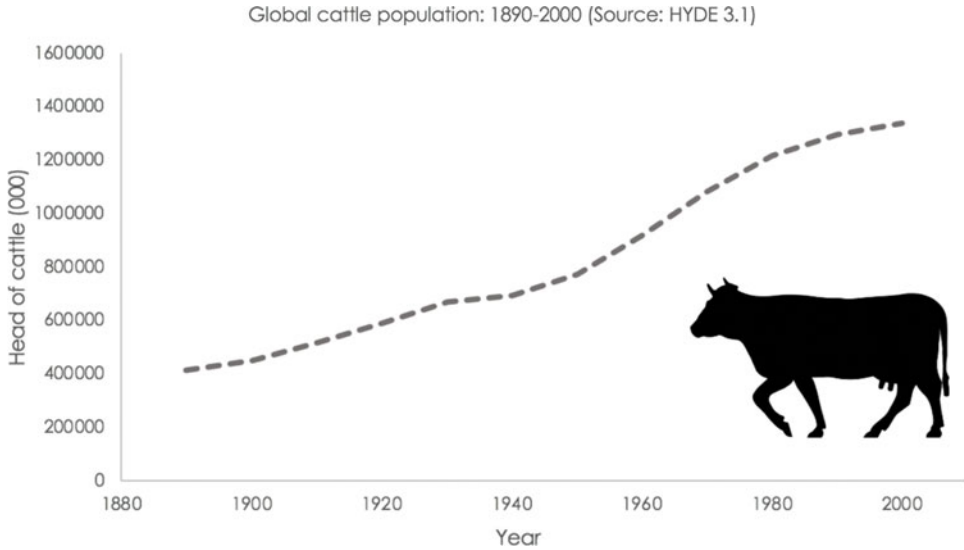
Chang'an were among the first to benefit from the political technologies (such as regimes of food distribution) that allowed human settlement on this scale. While it would be overinterpreting the results to infer that the virus became established in precisely these cities, the alignment is striking and underscores the close relationship between human urbanism and this important respiratory pathogen. These cities were an unprecedented ecosystem of interconnected human lungs, and measles was the evolutionary by-product of these new ecosystems.

At present, molecular clock methods are the only way to probe the history of human coronaviruses, from the recent appearance of SARS-CoV-2, which first infected humans in late 2019, to human coronavirus NL63, which seems to be the oldest of the seven species that can infect humans (see figure 1).<sup>51</sup> The phylogeny permits us to make several observations about the history of the family. Bats, and to some extent rodents, seem to be important natural hosts of coronaviruses. Several coronaviruses have adapted to humans via intermediate hosts, especially domesticated animals (such as camels and cattle). Human population growth has accelerated the emergence of coronaviruses, and it has also reduced the probability of extinction for any coronaviruses that crossed to humans; while speculative, it is plausible that pre-modern human populations were exposed to coronaviruses but lacked the population density necessary for the virus to maintain a continuous chain of transmission. The entire array of human-infecting coronaviruses, therefore, must be seen as the product of history, ecology, and evolution.

There are fifteen or so well-attested planetary waves of acute respiratory disease from 1500 to the present.<sup>52</sup> These global pandemics have generally been diagnosed as influenza, and reasonably so. But, given the impact of the COVID-19 pandemic, it is worth noting that the diagnoses of influenza (before the 1918 pandemic) are based on inference from patterns of transmission and generic symptoms. This evidence is not incompatible with the emergence of other respiratory pathogens, such as a human coronavirus. Recovery of genomic or proteomic evidence from archaeological or museum samples could provide confirmation or disconfirmation, and also furnish a chance for closer evolutionary and social comparisons with COVID-19. Notably, molecular clock dating puts the emergence of human coronavirus OC43, closely related to a bovine

<sup>51</sup>Diego Forni, Rachele Cagliani, Mario Clerici, and Manuela Sironi, 'Molecular Evolution of Human Coronavirus Genomes', *Trends in Microbiology* 25, no. 1 (2017): 35–48.

<sup>52</sup>David Morens and Jeffery Taubenberger, 'Historical Thoughts on Influenza Viral Ecosystems, or Behold a Pale Horse, Dead Dogs, Failing Fowl, and Sick Swine', *Influenza and Other Respiratory Viruses* 4, no. 6 (2010): 327–37.



**Figure 2.** Global cattle population, 1890–2000.

Source: HYDE 3.1, <https://themasites.pbl.nl/tridion/en/themasites/hyde/index.html>.

coronavirus, right around the year 1890, and the possibility has been mooted that the supposed influenza of 1889–92 (an event which helped bring the very word ‘pandemic’ into widespread use) was really the emergence of this now ubiquitous human respiratory pathogen.<sup>53</sup> Again, this kind of question is amenable to resolution, as the molecular key undoubtedly lurks in specimens somewhere.

Regardless, the evolution of this human coronavirus from a bovine ancestor underscores the close connections between human and animal health and the role of our transformation of the biosphere in driving disease emergence. The human lust for beef meant that by 1900 there were nearly half a billion head of cattle on planet earth (figure 2).<sup>54</sup> The rise of industrial farming in the nineteenth century and the links fostered by global capitalism created new disease ecologies; truly global panzootics in this period (of everything from equine influenza to Rinderpest and bovine pleuropneumonia: the phenomenon remains understudied) represent the unintended consequence of humanity’s preference for a handful of birds and mammals.<sup>55</sup> Industrial farms are

<sup>53</sup>On OC43, see Leen Vijgen *et al.*, ‘Complete Genomic Sequence of Human Coronavirus OC43: Molecular Clock Analysis Suggests a Relatively Recent Zoonotic Coronavirus Transmission Event’, *Journal of Virology* 79, no. 3 (2005): 1595–1604. On the ‘influenza’ of 1889–92, see K. David Patterson, *Pandemic Influenza, 1700–1900: A Study in Historical Epidemiology* (Totowa, NJ: Rowman & Littlefield: 1986), 49–82; and Gerald Pyle, *The Diffusion of Influenza: Patterns and Paradigms* (Totowa, NJ: Rowman & Littlefield, 1986), 31–3. The word pandemic came into widespread use in distinction to epidemic in the later nineteenth century, driven by the spread of cholera and influenza. As David Morens, Gregory Folkers, and Anthony Fauci, ‘What Is a Pandemic?’, *Journal of Infectious Diseases* 200, no. 7 (2009): 1018–21, argued, the violent spread of the 1889–92 influenza pandemic – closely covered by the popular press – brought the term ‘into general use’.

<sup>54</sup>Data from HYDE 3.1, <https://themasites.pbl.nl/tridion/en/themasites/hyde/landusedata/livestock/index-2.html>. For the environmental history of beef production in the United States, see Joshua Specht, *Red Meat Republic: A Hoof-to-Table History of How Beef Changed America* (Princeton: Princeton University Press, 2020). For an earlier period, a rich exploration of modern agriculture and animal health is Karl Appuhn, ‘Ecologies of Beef: Eighteenth-Century Epizootics and the Environmental History of Early Modern Europe’, *Environmental History* 15 (2010): 268–87.

<sup>55</sup>On Rinderpest, see e.g. Amanda McVety, *The Rinderpest Campaigns: A Virus, Its Vaccines, and Global Development in the Twentieth Century* (Cambridge: Cambridge University Press, 2018). More broadly, see Alan Olmstead and Paul Rhode, *Arresting Contagion: Science, Policy, and Conflicts over Animal Disease Control* (Cambridge, MA: Harvard University Press, 2015).

massive evolutionary stewpots that have been, and are, dangerous incubators of microbial evolution – if perhaps less sensational and exotic to Euro-American eyes than ‘bushmeat’ and ‘wet markets’.

### The big picture in the time of COVID-19

The interplay of human social development and infectious disease is one of the main threads of history, and the relationship is always two-way. Humans alter the ecological context of disease evolution, and disease evolution collides with human history.<sup>56</sup> The proliferation of genomic data and advances in the recovery of ancient biomolecules allow us to imagine new archives that can bring into greater focus and specificity the relationship between ecology and biology in human history. Such an interdisciplinary history provides a firmer basis to answer the kind of comparative questions that we are inevitably called upon to deliver amid the anxiety of the present pandemic.

A biological framework also helps to explain why the comparative study of responses to epidemics is so challenging. Even in my own research on the impact of pandemics in the Roman empire, I found that the consequences of sudden mortality were extremely variable and always conditioned by a range of human and natural factors.<sup>57</sup> But the importance of human structures does not diminish the importance of biological shocks, and ‘shock’ is a deliberate choice of word. Biological systems operate on non-linear and unpredictable timescales; often, the work of history is to identify the nature of these conjunctures between human and natural systems. The sudden, repeated interruption of human history by seemingly exogenous biological shocks, highly contingent in their occurrence and impact, is itself the deeper pattern.

An interdisciplinary history of infectious disease also puts the novelty of our current environmental predicament in deeper perspective by underscoring the microbial dimension of what has been called the Great Acceleration.<sup>58</sup> Human population growth and dominance of the biosphere, achieved over the last several thousand years, and put into hyperdrive during the last three generations, has accelerated the emergence of new human pathogens.<sup>59</sup> Paradoxically, the control of infectious disease has helped to precipitate the dizzying expansion of human numbers by driving down mortality rates.<sup>60</sup> I have described this process as the imposition of extended immunity onto the environment itself – a combination of social, chemical, and biomedical interventions that aim to make the world habitable for humans, even as we incentivize microbes to adapt to us as hosts. New respiratory viruses fit into this story in a peculiar way, because they are hard to prevent without vaccines and hard to treat, and thus can evade our extended immunity in distinct ways. This perspective puts the human coronaviruses in context and helps us to understand why a pathogen with the biological properties of SARS-CoV-2 was bound to be able to exploit our imperfect control of infectious disease.

Finally, a history of infectious disease can underscore that it is not only the ‘great germs’ that play an important role in the human story. We might envision the emergence of SARS-CoV-2 not

<sup>56</sup>Ron Barrett and George J. Armelagos, *An Unnatural History of Emerging Infections* (Oxford: Oxford University Press, 2013).

<sup>57</sup>Kyle Harper, *The Fate of Rome: Climate, Disease, and the End of an Empire* (Princeton: Princeton University Press, 2017).

<sup>58</sup>John McNeill and Peter Engelke, *The Great Acceleration: An Environmental History of the Anthropocene since 1945* (Cambridge, MA: Belknap Press of Harvard University Press, 2016).

<sup>59</sup>Michael R. Gillings and Ian T. Paulsen, ‘Microbiology of the Anthropocene’, *Anthropocene* 5 (2014): 1–8, is one of the only attempts to consider the Anthropocene explicitly through the lens of its relationship with microbial biodiversity.

<sup>60</sup>Roger Schofield, David Reher, and Alain Bideau, eds., *The Decline of Mortality in Europe* (Oxford: Clarendon Press, 1991); James C. Riley, *Rising Life Expectancy: A Global History* (Cambridge: Cambridge University Press, 2001); Simon Szreter, *Health and Wealth: Studies in History and Policy* (Rochester: University of Rochester Press, 2005); David Cutler, Angus Deaton, and Adriana Lleras-Muney, ‘The Determinants of Mortality’, *Journal of Economic Perspectives* 20 (2006): 97–120.

as the dramatic arrival of a powerful new microbe that will have a discrete impact, but rather as an important addition to the rich portfolio of human microparasites. In the *longue durée*, it is the ecology of infectious disease that matters. Perhaps vaccination will bring the COVID-19 pandemic to heel, but, even so, we may have to learn to live and sometimes die with the latest parasite that has acquired the ability to replicate its genes in the global environment that we have created.

Kyle Harper is Professor of Classics and Letters at the University of Oklahoma. He is a scholar of the Roman empire and late antiquity working on economic and environmental history with a focus on health and disease. He is the author of *Slavery in the Late Roman World, AD 275–425* (2011), *From Shame to Sin: The Christian Transformation of Sexual Morality in Late Antiquity* (2013), and *The Fate of Rome: Climate Disease, and the End of an Empire* (2017).