



Looking Backward To Move Forward: the Utility of Sequencing Historical Bacterial Genomes

Rebecca J. Bennett,^a Kate S. Baker^a

^aDepartment of Functional and Comparative Genomics, Institute of Integrative Biology, University of Liverpool, Liverpool, United Kingdom

ABSTRACT Many pathogens that caused devastating disease throughout human history, such as *Yersinia pestis*, *Mycobacterium tuberculosis*, and *Mycobacterium leprae*, remain problematic today. Historical bacterial genomes represent a unique source of genetic information and advancements in sequencing technologies have allowed unprecedented insights from this previously understudied resource. This minireview brings together example studies which have utilized ancient DNA, individual historical isolates (both extant and dead) and collections of historical isolates. The studies span human history and highlight the contribution that sequencing and analysis of historical bacterial genomes have made to a wide variety of fields. From providing retrospective diagnosis, to uncovering epidemiological pathways and characterizing genetic diversity, there is clear evidence for the utility of historical isolate studies in understanding disease today. Studies utilizing historical isolate collections, such as those from the National Collection of Type Cultures, the American Type Culture Collection, and the Institut Pasteur, offer enhanced insight since they typically span a wide time period encompassing important historical events and are useful for the investigating the phylodynamics of pathogens. Furthermore, historical sequencing studies are particularly useful for looking into the evolution of antimicrobial resistance, a major public health concern. In summary, although there are limitations to working with historical bacterial isolates, especially when utilizing ancient DNA, continued improvement in molecular and sequencing technologies and the resourcefulness of investigators mean this area of study will continue to expand and contribute to the understanding of pathogens.

KEYWORDS Murray Collection, ancient DNA, historical isolates, whole-genome sequencing

LOOKING BACKWARDS TO MOVE FORWARD: THE UTILITY OF SEQUENCING HISTORICAL BACTERIAL GENOMES

The first genome sequence of a bacterium (*Haemophilus influenzae*) was completed in 1995 (1) and now, just over 20 years later, the required technology and computation have advanced such that Public Health England (PHE) have implemented whole-genome sequencing for routine microbiological surveillance (2). In this public health context, rapid outbreak detection and intervention targeting are achieved through genome sequencing and analysis of our most modern bacterial isolates. However, this powerful technology can also be applied backward in time, and by analyzing the genomes of historical bacteria we gain relevant insight on some of the world's most devastating diseases.

Historical bacterial genome material is available from a wide range of sources, from long-buried bones and teeth to extant isolates of bacteria isolated long ago which are held in pathogen repositories. Depending on when these bacteria existed on the timescale of human and medical history, their genomic information can shed light on long-term trends in pathogen evolution and epidemiology, as well as evaluate the

Citation Bennett RJ, Baker KS. 2019. Looking backward to move forward: the utility of sequencing historical bacterial genomes. *J Clin Microbiol* 57:e00100-19. <https://doi.org/10.1128/JCM.00100-19>.

Editor Colleen Suzanne Kraft, Emory University

Copyright © 2019 American Society for Microbiology. All Rights Reserved.

Address correspondence to Kate S. Baker, kate.baker@liverpool.ac.uk.

Accepted manuscript posted online 15 May 2019

Published 26 July 2019

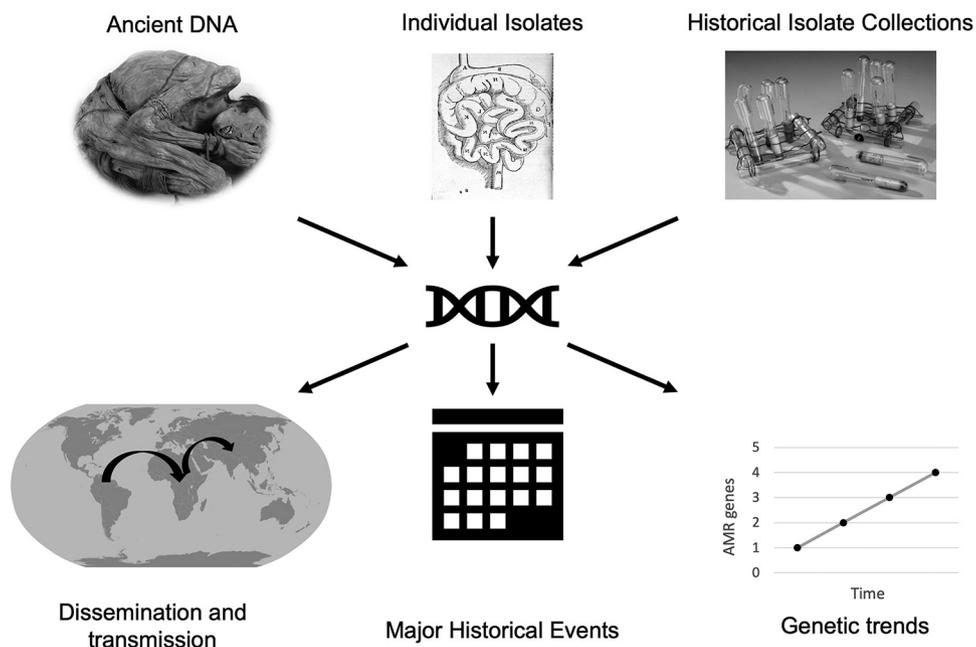


FIG 1 Graphical overview of the types of historical isolates, illustrating some of the insights that can be gained through sequencing of these isolates. Images: Mundinus, Anatomia Mundini; Peruvian mummified male (ca. 1200 to 1400); set of six test tubes with stand, used by Joseph Lister. In the top row, the left and middle images are from the Wellcome Collection (<https://wellcomecollection.org/works/n974azec> and <https://wellcomecollection.org/works/mppkfwb3>), and the right image is from the Science Museum, London, United Kingdom (<https://wellcomecollection.org/works/z5spdfcm>). These images are distributed pursuant to the Creative Commons Attribution 4.0 international license (<https://creativecommons.org/licenses/by/4.0/>).

impact of major historical events, such as migrations and medical advances, including the application of antimicrobials.

Here, in this noncomprehensive minireview, we give an overview of some of the insights which have been gained through the analysis of historical bacterial genome sequences sampled across human history (Fig. 1 and 2 and Table 1). For the purposes of this review, we define “historical” as having initially existed before the widespread clinical use of modern antimicrobial compounds (since the 1940s). We highlight the insights gained across many fields of biology, including evolution, epidemiology, and genetic diversity, and discuss the limitations and potential future directions of this approach.

BONES AND BACTERIA

In recent decades the increased study of ancient DNA (aDNA) has been revolutionizing our view of the genetic history and evolution of many species of plants and animals, as well as microorganisms. For example, aDNA has been used to retrospec-

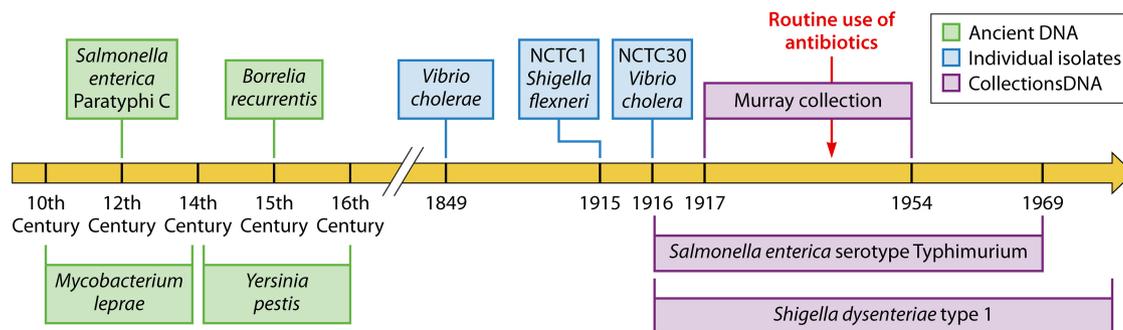


FIG 2 Timeline of examples used to highlight the utility of sequencing historical bacterial genomes.

TABLE 1 Comprehensive list of studies utilized in this minireview to highlight the utility of historical genomes

Sample type	Disease	Pathogen	Reference(s)
aDNA	Leprosy	<i>Mycobacterium leprae</i>	9
	Tuberculosis	<i>Mycobacterium tuberculosis</i>	11, 13
	Plague	<i>Yersinia pestis</i>	3, 15, 17–19
	Louse-borne relapsing fever	<i>Borrelia recurrentis</i>	20
	Salmonellosis	<i>Salmonella enterica</i> Paratyphi C	23, 24
Individual isolates	Cholera	<i>Vibrio cholerae</i>	26, 27
	Dysentery	<i>Shigella flexneri</i>	31
Historical collections	Listeriosis	<i>Listeria monocytogenes</i>	35–38
	Dysentery	<i>Shigella dysenteriae</i>	40
	Salmonellosis	<i>Salmonella enterica</i> serotype Typhimurium	41
	Opportunistic infections	<i>Klebsiella pneumoniae</i>	43

tively diagnose the cause of ancient epidemics. Today, diagnosis is approached through detection of microorganisms in infected tissues, which are understandably not available for the majority of ancient samples. Certain materials, however, such as bones and teeth, can provide a place of lasting refuge for historical bacteria. In a seminal study, Drancourt et al. demonstrated the efficacy of these sample types for studying ancient microorganisms. DNA extractions were completed on 12 teeth extracted from skeletal remains from the 16th to the 18th centuries (3). To confirm whether the extracted DNA contained *Yersinia pestis*, the causative agent of plague, PCR targeting two *Y. pestis* genes—*rpoB* and *pla*—was conducted and confirmed the presence of this disease in France at the end of the 16th century (3). This study was the first nucleic acid-based confirmation of ancient septicemia in material other than bone lesions, highlighting the utility of dental pulp as suitable historical material for studying historical disease outbreaks.

Although aDNA is a unique and rich source of genetic information, it poses some challenges, specifically, degradation and contamination. Typically, DNA from ancient sources is degraded from exposure, having suffered damage such as the fragmentation and modification of DNA bases, particularly the deamination of cytosine residues (4, 5). Contamination of aDNA samples can occur from a variety of sources. Even before skeletal remains are excavated, aDNA is exposed to the surrounding soil leading to aDNA from microbes of interest (i.e., those that may have caused infection/death) being diluted with environmental DNA, typically from soil microorganisms (4). Furthermore, during excavation, every archaeologist and researcher involved is another potential source of contamination. Such contamination at the point of collection would mean DNA purification would result in modern DNA being mixed with aDNA (4).

Recent advances in sampling strategies, molecular techniques and sequencing technologies have helped to reduce the limitations of studying aDNA. The methods utilized to molecularly extract aDNA from various sources have been improved greatly through increasing yield and decreasing destruction or degradation of samples. Techniques such as EDTA predigestion of powdered bone and teeth have been shown to increase the proportion of endogenous DNA severalfold (6). Not only has molecular extraction greatly improved yield, but sequencing technologies have also been essential in the recent advances into aDNA studies. For example, short-read next-generation sequencing (NGS) technologies require only short fragment lengths and can provide large amounts of sequence data from low concentrations of aDNA. Metagenomic sequencing in particular is helpful for aDNA studies since it allows analysis of the composition of a DNA extract to determine the relative abundance of endogenous DNA to contaminant DNA (5). Thus, since the advent of NGS, more genome studies of aDNA have been utilized effectively to study the causes of serious illnesses, and some examples are highlighted below. Furthermore, the use of bioinformatic techniques to distinguish between aDNA and environmental DNA has been invaluable in studies utilizing these historical genomes. DNA fragments isolated from ancient samples typically have a characteristic pattern of DNA damage which differs from present-day

organisms (7). In particular, methods such as metagenomic deep sequencing and cognate bioinformatic analyses (such as the MEGAN Alignment Tool) have revolutionized the process of reconstructing genomes from aDNA.

Leprosy. Leprosy is a debilitating, chronic infection affecting the peripheral nerves and skin. Renowned as a worldwide health problem in the 11th to the 16th centuries, it remains a modern-day problem as a World Health Organization (WHO)-listed neglected tropical disease (8).

Analysis of aDNA from the causative bacterium, *Mycobacterium leprae*, has been used to study the dissemination and decline of leprosy. In one study, aDNA was recovered using bead capture-based approaches on bone and teeth extracts from skeletons in cemeteries in Sweden, Denmark, and the United Kingdom from between the 10th and 14th centuries. Five samples contained enough *M. leprae* aDNA (>80% coverage, >5-fold depth) to progress to genome-wide comparison and were used to probe the origins of leprosy through comparisons with modern-day isolates (9). This comparison revealed remarkable levels of conservation, with only 755 single nucleotide polymorphisms (SNPs) and 57 InDels (<7 bp) found among the ancient European *M. leprae* genomes. In addition, this comparison revealed that there was no loss of known virulence genes over time (9), suggesting that the decline of leprosy in the 16th century was not due to pathogen evolution, but other factors, such as improvements in sanitation and social conditions. Temporal phylogenetic inference showed that most of the 16 ancient European *M. leprae* clustered into four distinct branches (9). Two new phylogeographic conclusions could be drawn from the reconstructed *M. leprae* phylogeny. First, three ancient strains were found to cluster closely among modern *M. leprae* genomes only isolated from Iran and Turkey, suggesting a link between ancient European *M. leprae* and present-day Middle-Eastern strains (9). Second, there was striking relatedness between 2 ancient *M. leprae* strains and 52 other modern strains from the United States, consistent with a European origin for leprosy in the Americas.

Thus, the use of aDNA from *M. leprae* provided support for existing historical theories and gave new insight on the spread of this ancient pathogen and suggested that its historical decline was likely not attributable to pathogen evolution.

Tuberculosis. One of the greatest public health challenges over the course of human history is caused by another mycobacterium, *Mycobacterium tuberculosis*, which is thought to have originated more than 150 million years ago (10). Tuberculosis is a contagious, infectious disease characterized by its persistence in the body and today afflicts 2 billion people worldwide (10). Despite its global importance, uncertainty remains as to how the interaction between *M. tuberculosis* and humans began and how this organism was globally disseminated, and studies of ancient strains of *M. tuberculosis* have been used to gain insight into the origin story of this prominent public health pathogen.

The earliest models of *M. tuberculosis* origins suggested that the human-niche-adapted pathogen evolved from zoonotic transfer of *Mycobacterium bovis* (11). However, genomic studies supported *M. bovis* and other animal mycobacteria having arisen from human strains and suggested that following an early origin in Africa, human *M. tuberculosis* disseminated from the old to the new world with human migration (12). However, this “out of Africa” theory was irreconcilable with the fossil record of *M. tuberculosis*-afflicted skeletons in the Americas before European contact (13).

This key question of the earliest disseminations of *M. tuberculosis* was addressed in a study by Bos et al. through the study of 68 New World skeletal samples that were screened representing pre- and post- (European) contact sites. Of these 68, only 3 mycobacterial genomes were recovered which showed convincing preservation of tuberculosis DNA and were isolated from 11th-century (pre-Columbian) Peruvian human skeletons (13). Phylogenetic reconstructions revealed that the pre-Columbian *M. tuberculosis* genomes did not cluster close to other human lineages and were more closely related to animal lineages with *M. pinnipedii*, which primarily infects seals (13). In addition to this phylogenetic clustering, the genomic architecture revealed a region of variable deletion pattern common to all animal lineages (13). The three ancient

genomes were found to share five unique SNPs (i.e., not found in modern strains) which are nonsynonymous; this suggests that these strains derive from a common progenitor. The fact that these ancient genomes shared a recent common ancestor with strains restricted to marine mammal species suggested that these animals were a novel reservoir and an historical route of entry for *M. tuberculosis* into human populations in both the Old World and the New World. All three genomes share an ancestor that predated the radiocarbon age of the skeletal material by more than 100 years, and two SNPs showed potential signatures of adaptation. These observations could support a single zoonotic transfer from pinnipeds to humans between AD 700 and AD 1000. Subsequent host adaptation and dissemination are a compelling prospect for future work. If confirmed, this would constitute the first example of a zoonotic transfer followed by readaptation to the human host in the *M. tuberculosis* complex.

This study provided some possible explanations for the introduction of *M. tuberculosis* into human populations. First, this may have occurred via seals, followed by bacterial adaptation to the human host and then subsequent transmission throughout the Americas. These theories can only be confirmed through further genomic study and comparisons. Hence, in this case, the study of aDNA from *M. tuberculosis* suggested a novel and unexpected potential ancient zoonotic reservoir for the disease that may be relevant for other pathogens.

Plague. The causative bacterium of the invasive infection known as the plague is *Yersinia pestis*. Historical data suggest that three human plague pandemics have occurred: the Justinian plague of the 6th century; the 2nd pandemic, which included the “Black Death” of the 14th to 17th centuries; and the ongoing modern plague, which started in the 19th century (14).

Despite their critical role in shaping human history, the origins of plague pandemics and the mechanisms underpinning the high mortality rates are relatively unknown, and aDNA studies have been used to unpick these crucial questions. The “Black Death” represents one of the most well-known historical pandemics and was characterized by its high mortality rate and rapid dissemination (15). Thought to have arrived in Britain from the Crimea in 1347, the Black Death spread rapidly to reach London in 1348 (16). In 1986, a large cemetery located outside the Royal Mint in East Smithfield was excavated, and two mass burial trenches and a burial pit densely packed with skeletons were discovered (16). This cemetery is one of the emergency burial grounds created to cope with the devastating effects of the Black Death and is currently the largest excavated Black Death cemetery in England (16). Thus, the skeletal remains within represent a unique opportunity for genetic investigations into *Y. pestis*. One investigation utilized multiple DNA extracts from teeth found in the cemetery to evaluate capture-based methods and reconstruct a complete ancient *Y. pestis* genome (30-fold coverage) (15). Single nucleotide differences between this ancient genome and the reference only consisted of 97 chromosomal positions, which highlights the tight genetic conservation of this organism over the past 600 years. To place the ancient genome in a phylogenetic context, the genome was compared to those from aggregate base call data for 17 publicly available *Y. pestis* genomes and the ancestral *Y. pseudotuberculosis* (15). Phylogenetic trees placed the genome close to the ancestral node of all extant human-pathogenic *Y. pestis* strains. Using a Bayesian approach, temporal estimates were made which indicated that all *Y. pestis* shared a common ancestor sometime between 668 and 729 years ago. Having the common ancestor for all *Y. pestis* strains existing between 600 and 700 years ago implies that this medieval plague was the main historical event that introduced human populations to the common ancestor of all known pathogenic strains of *Y. pestis* (15).

Following the Black Death, during which nearly half of the European population was killed, the second pandemic persisted as a series of smaller, apparently disjointed epidemics, and key questions are how the plague persisted in Europe during this time and why it suddenly declined thereafter. The analysis of climatic and historical epidemiological data supports the theory of repeated introductions of *Y. pestis* to harbors in

Europe through maritime trade from climate-driven, large-scale outbreaks in Asia (with a delay of around 15 years) (17). However, studies of aDNA from *Y. pestis* outbreaks subsequent to the Black Death (though still in the 2nd pandemic) support the notion that plague persistence in Europe was likely due to a now-extinct European (or juxtaposed) reservoir. Several studies into the phylogenetic analysis of aDNA genomes from throughout the duration of the 2nd pandemic seem to point to a single common ancestor and suggest that this lineage is now extinct (18, 19). For example, Spyrou et al. (19) screened DNA extracts from teeth of individuals found in mass graves in Spain, Russia, and Germany for the presence of *Y. pestis* and ultimately reconstructed 53 *Y. pestis* genomes. The evolutionary history of these ancient genomes was inferred through phylogenetic analysis alongside >130 modern and 7 other historical *Y. pestis* strains. All of these reconstructed genomes clustered closely onto a single branch which has been previously identified in historical genomes from the Black Death, confirming them as ancient. The clustering of genomes from different geographic locations and their having no detectable differences with other Black Death strains from London support the notion that the Black Death and subsequent outbreaks in Europe were caused by a single lineage.

By combining geographical, temporal, and phylogenetic data, these studies have facilitated high-level debate on a key question of global disease ecology that has continued relevance for many pathogens today. Within these studies, aDNA has provided valuable insight into the key features of a prominent historic pathogen and the legacy it left behind.

Louse-borne relapsing fever. Sequencing of aDNA can be used to investigate the relative virulence of historical and modern strains, which can in turn explain the long-term emergence and evolution of a pathogen. A study examining this evolution of virulence was performed on an ancient microbial genome recovered by metagenomic sequencing of dental pulp from a 15th-century skeleton in Oslo (20). The microbial genome recovered from bioinformatic analysis of the metagenome was *Borrelia recurrentis*, which causes louse-borne relapsing fever (LBRF). Throughout the course of European history, LBRF was a prominent problem, and the disease still causes sporadic outbreaks in Eastern Africa (21). The relapsing nature of the disease is driven by the alteration of microbial surface proteins to facilitate host immune evasion potential, and the number of relapse episodes is limited by the microbe's ability to generate new serotypes, which is determined by the key virulence determinants, i.e., the plasmid-borne antigenic phase variation genes (20).

Phylogenetics and comparative genomics among the recovered medieval European isolate, modern African *B. recurrentis*, and phylogenetically distinct tick-borne relapsing fevers revealed important insights into the evolution of LBRFs. Specifically, absent or potentially degraded numbers of antigenic-phase-variation genes relative to modern strains potentially indicate a less-virulent phenotype where the medieval strain may have generated fewer serotypes and thus bouts of fever, which is consistent with historical records (20). Furthermore, the study hints at the importance of reductive evolution in facilitating *B. recurrentis*' epidemic potential and its increased virulence compared to other relapsing fevers. Reductive evolution has been noted with other louse-to-human transmission specialists, such as *Bartonella quintana* and *Rickettsia prowazekii*, and signatures of *B. recurrentis* reductive evolution were characterized by accelerated rates of genome degradation, thought to be driven by adaptation to host-restricted vectors and functional trade-offs (22). Hence, with only a single ancient LBRF genome, signatures of changing virulence phenotypes and reductive genome evolution that led to the development of our modern-day pathogens could be seen.

Salmonellosis. Infections with *Salmonella enterica* subspecies *enterica* serovar Paratyphi C cause a wide range of clinical syndromes from enteric fever (bacteremia) to urinary tract infections. In the modern day, the infection is largely restricted to Asia and Africa, though it is unknown whether European populations were historically affected

(23). However, Zhou et al. recovered a *Salmonella enterica* Paratyphi C genome sequence from a Norwegian 13th-century skeleton, confirming the presence of this bacterium within Europe (23). Recovery of this aDNA suggested that the individual died from enteric fever caused by *S. enterica* Paratyphi C, providing evidence that this bacterium caused invasive salmonellosis in Europe over 800 years ago, thus showing the potential of aDNA to establish the presence of pathogens historically where records are not available.

Disease caused by *S. enterica* Paratyphi C continued throughout history, and this pathogen was also identified as the cause of an important epidemic in the New World through an aDNA study. In the 16th century, a large epidemic (cocoliztli) responsible for millions of deaths in modern-day Mexico occurred and, for the past 500 years, the cause of this epidemic has been unknown. Through the use of a novel and comprehensive metagenomic screening approach (the MEGAN Alignment Tool), *S. enterica* Paratyphi C was identified in nonenriched DNA sequence data from dental pulp, collected from individuals excavated from the (European) contact era, but not from individuals in precontact era burial sites at Teposcolula-Yucundaa (24). Without the sequencing and subsequent analysis of aDNA, it would not have been possible to establish the bacterial cause of this devastating epidemic.

Thus, retrospective epidemiological investigations utilizing aDNA demonstrated that infections with *S. enterica* Paratyphi C occurred in Europe in the precontact era, and this is one of the pathogens responsible for the collapse of native populations following its introduction to the New World.

INDIVIDUAL ISOLATES: THE LIVING AND THE DEAD

In addition to aDNA studies on pathogens from the 10th to the 17th centuries, studies on more recent historical isolates from the 19th and 20th centuries have been performed. Genomes of individual isolates have been recovered from these later centuries and, due to their more recent origin, some of the limitations of working with aDNA are overcome. Despite only being individual genomes, sequences from older bacteria still provide invaluable historical context for studying the evolution of the modern-day pathogens. Moreover, when the historical bacteria are extant (living), phenotypic testing can be performed to confirm hypotheses developed from studying the genomic information.

Cholera. In the 19th century, the Gram-negative bacterium *Vibrio cholerae* was a prominent pathogen which has caused seven cholera pandemics that swept across Europe, Asia, North America, and India, with the seventh pandemic still causing disease today (25). The cholera epidemic of 1832 to 1833 in London resulted in between 4,000 and 7,000 deaths alone. The cause of these outbreaks and the genomic strains responsible are not well understood, but a study by Devault et al. (26) utilized a second-pandemic strain of *V. cholerae* from the 1849 outbreak in Philadelphia to provide insight on the organism which caused these 19th-century pandemics. High-throughput sequencing was used to reconstruct the genome of a (nonviable) *V. cholerae* from the preserved intestine of a cholera victim. Comparative genomics showed that this genome belonged to the classical rather than the El Tor biotype (the latter of which is responsible for the seventh, current pandemic) of serogroup O1 (26). The reconstructed genome shared 95 to 97% similarity with the classical O395 genome, differing by only ~200 SNPs (26). The little variation between them indicates that modern *V. cholerae* evolution has been subjected to substantial selective constraint since the mid-19th century, similar to that exhibited by other pathogens such as *Y. pestis*, demonstrated by its long-term genome conservation.

The other type of individual historical isolate that can be recovered is an extant (living) isolate, which enables phenotypic testing. The sixth cholera pandemic occurred between 1899 and 1923, covering the period during which World War I (WW1) took place. Despite the widespread dissemination of cholera globally, it has been observed that very few soldiers in the British Expeditionary Forces contracted cholera between 1914 and 1918. In 1917 a *V. cholerae* strain was isolated from a British soldier in Egypt.

The strain is held at the National Collection of Type Cultures (NCTC) held by Public Health England under accession NCTC30. This 102-year-old strain, which is thought to be the oldest publicly available live *V. cholerae* strain in existence, was recently revived and sequenced (27). A pangenome was constructed using a collection of >190 *V. cholerae* sequences, and the resulting core genome alignment was used to generate maximum-likelihood phylogeny. This phylogenetic reconstruction showed that NCTC30 is more closely related to *Vibrio cholerae* sequences than to other members of the *Vibrio* genus, although NCTC30 was distinct from most *V. cholerae* strains included in the study. The genome was explored to determine whether NCTC30 had the etiological genetic determinants to cause “choleraic diarrhoea” which was observed in the patient. The strain was lacking CTX_φ, the bacteriophage that encodes the cholera toxin (27), so the researchers looked for secondary virulence factors, including heat-stable accessory virulence genes, to determine whether there was an alternative factor causing the pathogenic phenotype. This analysis revealed a genomic island encoding a putative type 3 secretion system (T3SS), similar to the T3SS found in the genome of the *Vibrio parahaemolyticus* strain (27). It was thought that the T3SS may be responsible for the clinical symptoms observed; however, coinfection by another bacterium could not be ruled out.

Upon culturing, it was determined that NCTC30 lacked the typical flagella, and an explanatory frameshift mutation was identified in the *flrC* 3' region which is thought to have prevented the expression of flagellum biosynthesis genes (27). Previous phenotypic studies on this strain reported that NCTC30 was resistant to penicillin (28), and Dorman et al. (27) used ResFinder to identify a β -lactamase gene potentially responsible for this phenotype. NCTC30 predates the discovery of penicillin, so its presence was not attributable to the selective pressure of clinical antibiotic usage, consistent with the ancient origins of antimicrobial resistance (AMR) genes as being important for survival in the environment (29).

These studies into historical *V. cholerae* show what a powerful tool historical isolates can be in not only shedding light on transmission routes and dissemination pathways but also providing insight into the evolution of *V. cholerae* as a nonpandemic and pandemic pathogen.

Dysentery. Dysentery is the now-outdated syndromic term for disease characterized by inflammation of the intestine, blood in the stool, and necrosis of the colonic mucosa. Bacillary dysentery is typically caused by *Shigella* species, and epidemics of the disease have occurred throughout human history and continue to this day, particularly in young children in lower-to-middle-income nations (30). Genome sequencing of an extant historical isolate of bacillary dysentery, NCTC1, and its comparisons to modern-day strains showed exquisitely targeted accumulation of genes over time.

NCTC1 was recovered in 1915 from a British Forces soldier who died from dysentery on the Western Front during WW1 (31). NCTC1 belonged to the *Shigella* group *S. flexneri*, the type responsible for the largest *Shigella* disease burden today. The isolate represents the oldest extant *Shigella flexneri* isolate in existence and, since it could be cultured, enough DNA was obtained to completely sequence the genome using long-read sequencing technology (31). Gene content analyses showed that NCTC1, although isolated prior to the discovery of penicillin, contained the genes required for conferring resistance against both penicillin and the more modern antimicrobial compound erythromycin, phenotypic hypotheses confirmed through phenotypic testing (31). Phylogenetic analysis of NCTC1 showed that it belonged to the still-prevalent 2a sublineage of *Shigella flexneri*, and gene accumulation in this sublineage over the 100 years since WW1 showed the gain of genomic islands associated primarily with AMR, virulence, and immune evasion.

In this case, sequencing of a historical isolate provided an invaluable benchmark for the study of the still epidemiologically important *S. flexneri* 2a sublineage, demonstrating, through comparisons with contemporary isolates, the accumulation of genes conferring pathogenic functions, including AMR.

THE COLLECTIONS

Given the insights that can be gained from only a few aDNA genomes and single isolates, it is easy to conceptualize how much more information can be gained from studying collections of historical isolates. These collections can provide a more thorough historical perspective of the evolution and genetic architecture of pathogens. Though typically limited to more recent centuries (Fig. 1), collections of historical bacteria may still span key historical and geographical events, such as wars, human migrations, and key advances in technology and medicine, including the introduction of antimicrobials.

There are several large national pathogen repositories containing bacterial strains, such as the National Collection of Type Cultures (NCTC), the American Type Culture Collection (ATCC), and the collections of the Institut Pasteur. The NCTC is held by PHE and houses approximately 5,100 type and reference bacterial strains, many of which are invaluable for medical and scientific research and are stipulated in internationally recognized standardized methods as control strains. The U.S. equivalent collection, the ATCC, is a similar global biological materials and standards organization that hosts over 18,000 bacterial strains used in a variety of fields. Another large reference collection is the Collection de l'Institut Pasteur (CIP), which first began collecting bacterial strains in 1892. Currently housing >12,000 bacterial strains, with new ones being added regularly, the CIP represents the largest collection at the Institut Pasteur.

In addition to these growing repositories, one of the most studied historical bacterial collections is the Murray Collection. The Murray Collection comprises several hundred *Enterobacteriaceae* strains collected in the preantibiotic era (1917 to 1954) from a wide range of geographic areas (32). This now-public collection, held by the NCTC, provides a unique genomic resource to provide insight into an era before the routine use of antibiotics and so is particularly useful for studies investigating the evolution and accumulation of AMR. AMR is a global public health crisis, with the World Health Organization warning of a "postantibiotic era" in the future where common infections could kill (33).

In fact, the utility of studying the Murray Collection for providing insight into the emergence of pathogen AMR has already been demonstrated. For example, the Murray Collection was instrumental in investigating the relationship between plasmids and the dissemination of AMR after the clinical introduction of antimicrobials. A seminal study on the collection utilized 692 strains from the collection to show that conjugative plasmid types that carry AMR were just as commonly found in bacteria of the "preantibiotic" era compared to modern strains (34). This finding revealed that emergence of plasmid-borne AMR is likely attributable to acquisition of AMR determinants on existing plasmids, rather than the dissemination of novel plasmid types already carrying AMR.

Listeriosis. The common foodborne disease listeriosis is caused by the bacterium *Listeria monocytogenes*. A historical collection of >6,000 *Listeria* strains isolated between 1921 and 1987 was amassed by H. P. Seeliger. This Special *Listeria* Culture Collection (SLCC) was recently resuscitated, digitized, and translated (from German to English) to facilitate the accessibility of the collection, with the metadata being publicly released (35).

A subsequent study used whole-genome sequencing to determine the evolutionary relationships among 20 outbreak-associated clinical isolates of *Listeria monocytogenes* (36). For placement of these 20 genomes into the larger context of *L. monocytogenes* genetic diversity, other isolates were included in the study; of particular interest were 9 SLCC isolates from 1924 to 1983. The use of these isolates allowed an *L. monocytogenes* phylogeny to be constructed that clustered the 20 genomes into 10 distinct phylogenetic groups (36). From this study, an evolutionary framework was generated to better understand how *L. monocytogenes* outbreak-associated serotypes 1/2a and 1/2b were related to one another and to other clones. Here, the historical isolates provided a much-needed historical background and perspective that contributed to our understanding of the evolution of these serotypes over time. A better understand-

ing of the genetic basis for the traits that enable higher rates of transmission and virulence is valuable for the control and prevention of human health pathogens. The use of the SLCC isolates in this study was important in providing a larger context for comparisons of the newly sequenced isolates.

Other genomic investigations using isolates from this collection are also advancing insights into listeriosis, including the generation of a draft genome of a 94-year-old isolate, which may provide historical context for future studies (37), and the use of the isolates in increasing the genetic diversity available to characterize the population structure of *L. monocytogenes* (38).

Dysentery. There have been multiple recorded dysentery epidemics associated with warfare, including during Napoleon's retreat from Moscow and the attacks on the Gallipoli peninsula during WW1, from which ~120,000 victims, most of which had bacillary dysentery, were evacuated in 1915 (39). These epidemics were caused by the bacterium *Shigella dysenteriae* type 1 (Sd1), a bacterium that causes severe disease through elaborating the cytotoxic Shiga toxin (39). Within the second half of the 20th century, large outbreaks of Sd1 dysentery were still occurring, with death tolls reaching 20,000 in a 1969-1973 outbreak in Central America (40).

The geographic dissemination, evolution and origins of Sd1 were recently elucidated by large-scale genomic analyses incorporating historical isolates. Specifically, a study by Njamkepo et al. (40) conducted whole-genome sequence analysis on 331 Sd1 isolates from around the world, sampled between 1915 and 2011. These isolates were selected from over 35 collections, including 10 WW1 isolates from the historic Murray Collection (40). Bayesian phylogenetic inference was used to define different phylogenetic lineages of the pathogen and utilized to estimate nucleotide substitution rates and divergence times. This suggested that the most recent common ancestor for Sd1 first emerged ca. 1747 (40). In addition to demonstrating that Sd1 existed since at least the 18th century, the temporal phylogenetic analysis (when combined with geographical metadata) suggested that the global dissemination of Sd1 predated WW1 (40).

Looking over the wide time period bridging the clinical introduction of antimicrobials also facilitated investigations into the evolution of AMR. One of the key features in the genome evolution of Sd1 was the acquisition of AMR genes. The first antibiotic-resistant Sd1 isolates were isolated from Asia and America in the 1960s and rapidly became dominant, with <1% of isolates remaining susceptible from 1991 to 2011 (40). These genomes also revealed how AMR genes were being disseminated, with the first AMR genes in Sd1 being identified on small plasmids encoding resistance against streptomycin and sulfonamides. Following this, through the mid-1960s to the 1980s, resistance was acquired via larger plasmids (40).

In performing these phylogenetic and AMR analyses, the collections provided a thorough historical perspective of a significant pathogen, elucidated the origins of outbreaks, and provided insight into the accumulation of AMR in this globally important pathogen.

Salmonellosis. *Salmonella* isolates can be found in multiple collections, and drawing information from multiple sources can provide important insights into the emergence and evolution of AMR in prominent public health pathogens. A large-scale whole-genome sequencing study analyzed 288 *S. enterica* serotype Typhimurium isolates collected between 1911 and 1969, again spanning the preantibiotic era (41). These isolates originated from a variety of different collections, including the French National Reference Centre for *Escherichia coli*, *Shigella*, and *Salmonella*; the CIP; and the WHO Collaborating Centre for Research on *Salmonella*. The study hoped to provide insights into the mechanisms of resistance to ampicillin, an antibiotic which came into routine clinical use in 1961 (41). Again, combining phylogenetic and genetic and phenotypic AMR studies, the authors found the multiphyletic presence of diverse ampicillin resistance genes and contexts across *S. enterica* Typhimurium. Eleven of the isolates tested were resistant to ampicillin; seven of these were also resistant to other antibiotics, including tetracycline and trimethoprim (41). From analysis of the whole-genome

sequences, the β -lactamase *bla*_{TEM-1B} gene was found on two different plasmids (IncX1 and IncF) in three isolates collected between 1959 and 1960 (41). Thus, this study again revealed that the presence of an AMR phenotype (in this case ampicillin resistance) predated its routine clinical use. In contrast to the findings from individual isolates of dysentery and cholera above, however, analysis of the collection of isolates here was able to show that a diversity of vectors was involved in the dispersal of ampicillin resistance. Specifically, the early emergence of *S. enterica* Typhimurium was not due to the expansion of a single clonal population with a particular resistance plasmid but was attributable to multiple independent acquisitions of *bla*_{TEM}-carrying plasmids (41). Thus, this study demonstrated that early ampicillin resistance emerged multiple times in this important pathogen, a key finding that is likely relevant for other organisms and vital in combatting the global threat of AMR.

***Klebsiella pneumoniae*.** The Murray Collection bacteria largely belong to the family *Enterobacteriaceae* and include 37 *Klebsiella* isolates. This genus of Gram-negative bacteria is of global health importance due to the recent rise of invasive and multidrug-resistant strains, which constitute a significant clinical threat (42). *Klebsiella pneumoniae* is a common commensal gut bacterium and can also cause disease, such as pneumonia, wound infections, and meningitis, as an opportunistic pathogen.

Wand et al. studied *K. pneumoniae* isolates ($n = 37$) from the collection by combining the available genome sequence data with phenotypic experimentation (43). In this way, phenotypic studies of these historical isolates were complemented by the genome information. For example, ca. 30% of the preantibiotic-era *K. pneumoniae* isolates were resistant to the β -lactam antibiotic penicillin, and the genomic analyses showed the presence of diverse *bla*_{SHV} β -lactamase genes among these isolates. Further phenotypic testing revealed variable levels of susceptibility among Murray Collection *Klebsiella* and more modern strains to skin antiseptics and triclosan. Examination of the whole-genome sequences shows that lower MICs of chlorhexidine for the Murray Collection isolates could be linked to 9- to 18-bp insertions in the *cepA* gene, which is known to be associated with disinfectant resistance (43). Finally, the sequences in the Murray Collection also provided insights into virulence. Phylogenetic analysis of the Murray *Klebsiella* strains showed marked diversity, including a large clade of isolates from the same sequence type (ST82). However, a single isolate belonged to a sublineage now associated with virulent clinical isolates; the high-virulence clade CC23 showed that although this lineage was present in the preantibiotic era, it was not epidemiologically dominant (43). Thus, the importance of the use of historical collections of bacteria (rather than individual isolates) is highlighted here; had only a single isolate been studied, the presence of this modern-day virulent clade might have been missed, since CC23 was historically underrepresented.

Thus, sequencing of collections provided a more representative historical perspective relative to the insights gained by individual isolate studies, facilitating studies of the evolution of important traits such as AMR and virulence in pathogens that are highly relevant for public health today.

DISCUSSION

Within this noncomprehensive minireview, the utility of sequencing historical bacterial isolates has been explored using examples from a range of different time periods and sample types.

There is very little written record of the spread of some large ancient outbreaks, and so genomic analysis can be instrumental in understanding the evolutionary and epidemiological pathways of ancient pathogens. With the examples provided above for *Y. pestis* and leprosy, it is evident that aDNA can be used to explain the dissemination and decline of important pathogens. Moreover, in some circumstances, recovery of aDNA from victims also provided retrospective diagnosis, such as with the Mexican outbreak of *Salmonella*. Without aDNA, this would have been impossible. aDNA clearly represents a unique and exciting resource from which to gain a historical perspective on ancient outbreaks and pathogens. However, aDNA study has some limitations. There

is still a concern with contamination, which can occur at different points, as well as low recovery yields and degradation of the genetic material, meaning reconstructed genomes may not be complete or may not be fully representative.

Individual isolates tend to be from later periods of history, typically the 19th and 20th centuries, and so are less susceptible to degradation and potentially present in greater quantities, particularly in the case of extant isolates. In addition to achieving greater genome reconstruction (through the higher DNA yields possible), extant strains can also be utilized for phenotypic testing. Phenotypic testing is extremely useful for investigations into the evolution of characteristics critical to addressing the public health challenges of today, such as AMR and virulence. Future studies of historical isolates will continue to elucidate these issues, particularly as new potential sources of historical genetic material come to light, such as preserved infected tissue as used in the study of *V. cholerae*.

Investigations utilizing individual isolates, however, may not give a full representation of pathogen evolution due to the focus on a single isolate. It is possible that there may be vital elements that would be over- or underestimated in their importance. The historical insight gained by individual isolates can be augmented by the use of collections, which are available in multiple countries and often contain thousands of bacterial strains spanning large periods of time. Investigations utilizing these collections give a more representative view of the evolutionary dynamics and epidemiology of historical pathogens. This was shown for both the *Klebsiella* and the *Salmonella* examples described above, where the long-term presence of a now-important sublineage and diverse contexts for an important AMR phenotype were determined, where they might have been missed using fewer or individual isolates from those collections.

Since many historical collections span the antibiotic era, they have been particularly useful for demonstrating the evolutionary dynamics of bacteria and genes of interest to AMR. Collections such as the preantibiotic-era Murray Collection represent a unique insight into a time before antibiotic use and can be used to track the accumulation of AMR determinants and virulence factors via comparison with modern isolates. With the future of the efficacy of antibiotics in the balance, it is now more important than ever to understand the means by which AMR developed since the preantibiotic era and to investigate other ways in which pathogenic bacteria have evolved over this critical time period.

ACKNOWLEDGMENTS

R.J.B. is funded by a Biotechnology and Biological Sciences Research Council Doctoral Training Partnership studentship (BB/M011186/1). K.S.B. is funded by a Wellcome Trust Career Development Fellowship (106690/A/14/Z).

We are also grateful to three anonymous reviewers for their constructive feedback on the manuscript.

REFERENCES

1. Fleischmann R, Adams M, White O, Clayton R, Kirkness E, Kerlavage A, Bult C, Tomb J, Dougherty B, Merrick JM, et al. 1995. Whole-genome random sequencing and assembly of *Haemophilus influenzae* Rd. *Science* 269:496–512. <https://doi.org/10.1126/science.7542800>.
2. Grant K, Jenkins C, Arnold C, Green J, Zambon M. 2018. Implementing pathogen genomics: a case study. Public Health England, London, United Kingdom.
3. Drancourt M, Aboudharam G, Signoli M, Dutour O, Raoult D. 1998. Detection of 400-year-old *Yersinia pestis* DNA in human dental pulp: an approach to the diagnosis of ancient septicemia. *Proc Natl Acad Sci U S A* 95:12637–12640. <https://doi.org/10.1073/pnas.95.21.12637>.
4. Gorge O, Bennett EA, Massilani D, Daligault J, Pruvost M, Geigl EM, Grange T. 2016. Analysis of ancient DNA in microbial ecology. *Methods Mol Biol* 1399:289–315. https://doi.org/10.1007/978-1-4939-3369-3_17.
5. Knapp M, Hofreiter M. 2010. Next generation sequencing of ancient DNA: requirements, strategies, and perspectives. *Genes (Basel)* 1:227–243. <https://doi.org/10.3390/genes1020227>.
6. Damgaard PB, Margaryan A, Schroeder H, Orlando L, Willerslev E, Al-lentoft ME. 2015. Improving access to endogenous DNA in ancient bones and teeth. *Sci Rep* 5:11184. <https://doi.org/10.1038/srep11184>.
7. Skoglund P, Northoff BH, Shunkov MV, Derevianko AP, Pääbo S, Krause J, Jakobsson M. 2014. Separating endogenous ancient DNA from modern day contamination in a Siberian Neandertal. *Proc Natl Acad Sci U S A* 111:2229–2234. <https://doi.org/10.1073/pnas.1318934111>.
8. White C, Franco-Paredes C. 2015. Leprosy in the 21st century. *Clin Microbiol Rev* 28:80–94. <https://doi.org/10.1128/CMR.00079-13>.
9. Schuenemann VJ, Singh P, Mendum TA, Krause-Kyora B, Jäger G, Bos KI, Herbig A, Economou C, Benjak A, Busso P, Nebel A, Boldsen JL, Kjellström A, Wu H, Stewart GR, Taylor GM, Bauer P, Lee O-C, Wu HHT, Minnikin DE, Besra GS, Tucker K, Roffey S, Sow SO, Cole ST, Nieselt K, Krause J. 2013. Genome-wide comparison of medieval and modern *Mycobacterium leprae*. *Science* 341:179–183. <https://doi.org/10.1126/science.1238286>.
10. Barberis I, Bragazzi NL, Galluzzo L, Martini M. 2017. The history of tuberculosis: from the first historical records to the isolation of Koch's bacillus. *J Prev Med Hyg* 58:E9–E12.
11. Stead WW, Eisenach KD, Cave MD, Beggs ML, Templeton GL, Thoen CO,

- Bates JH. 1995. When did *Mycobacterium tuberculosis* infection first occur in the New World? An important question with public health implications. *Am J Respir Crit Care Med* 151:1267–1268. <https://doi.org/10.1164/ajrccm.151.4.7697265>.
12. Brosch R, Gordon SV, Marmiesse M, Brodin P, Buchrieser C, Eiglmeier K, Garnier T, Gutierrez C, Hewinson G, Kremer K, Parsons LM, Pym AS, Samper S, van Soolingen D, Cole ST. 2002. A new evolutionary scenario for the *Mycobacterium tuberculosis* complex. *Proc Natl Acad Sci U S A* 99:3684–3689. <https://doi.org/10.1073/pnas.052548299>.
 13. Bos KI, Harkins KM, Herbig A, Coscolla M, Weber N, Comas I, Forrest SA, Bryant JM, Harris SR, Schuenemann VJ, Campbell TJ, Majander K, Willbur AK, Guichon RA, Wolfe Steadman DL, Cook DC, Niemann S, Behr MA, Zumarraga M, Bastida R, Huson D, Nieselt K, Young D, Parkhill J, Buikstra JE, Gagneux S, Stone AC, Krause J. 2014. Pre-Columbian mycobacterial genomes reveal seals as a source of New World human tuberculosis. *Nature* 514:494–497. <https://doi.org/10.1038/nature13591>.
 14. Seifert L, Wiechmann I, Harbeck M, Thomas A, Grupe G, Projahn M, Scholz HC, Riehm JM. 2016. Genotyping *Yersinia pestis* in historical plague: evidence for long-term persistence of *Y. pestis* in Europe from the 14th to the 17th century. *PLoS One* 11:e0145194. <https://doi.org/10.1371/journal.pone.0145194>.
 15. Bos KI, Schuenemann VJ, Golding GB, Burbano HA, Wagelchner N, Coombes BK, McPhee JB, DeWitte SN, Meyer M, Schmedes S, Wood J, Earn DJ, Herring DA, Bauer P, Poinar HN, Krause J. 2011. A draft genome of *Yersinia pestis* from victims of the Black Death. *Nature* 478:506–510. <https://doi.org/10.1038/nature10549>.
 16. Antoine D. 2008. The archaeology of “plague.” *Med Hist Suppl* (27): 101–114.
 17. Schmid BV, Büntgen U, Easterday WR, Ginzler C, Walløe L, Bramanti B, Stenseth NC. 2015. Climate-driven introduction of the Black Death and successive plague reintroductions into Europe. *Proc Natl Acad Sci U S A* 112:3020–3025. <https://doi.org/10.1073/pnas.1412887112>.
 18. Bos KI, Herbig A, Sahl J, Wagelchner N, Fourment M, Forrest SA, Klunk J, Schuenemann VJ, Poinar D, Kuch M, Golding GB, Dutour O, Keim P, Wagner DM, Holmes EC, Krause J, Poinar HN. 2016. Eighteenth century *Yersinia pestis* genomes reveal the long-term persistence of an historical plague focus. *Elife* 5:e12994. <https://doi.org/10.7554/eLife.12994>.
 19. Spyrou MA, Tukhbatova RI, Feldman M, Drath J, Kacki S, Beltrán de Heredia J, Arnold S, Sitdikov AQ, Castex D, Wahl J, Gazimzyanov IR, Nurgaliev DK, Herbig A, Bos KI, Krause J. 2016. Historical *Yersinia pestis* genomes reveal the European black death as the source of ancient and modern plague pandemics. *Cell Host Microbe* 19:874–881. <https://doi.org/10.1016/j.chom.2016.05.012>.
 20. Guellil M, Kersten O, Namouchi A, Bauer EL, Derrick M, Jensen AO, Stenseth NC, Bramanti B. 2018. Genomic blueprint of a relapsing fever pathogen in 15th century Scandinavia. *Proc Natl Acad Sci U S A* 115: 10422–10427. <https://doi.org/10.1073/pnas.1807266115>.
 21. Antinori S, Mediannikov O, Corbellino M, Grande R, Parravicini C, Bestetti G, Longhi E, Ricaboni D, Ehououd CB, Fenollar F, Raoult D, Rimoldi SG. 2016. Louse-borne relapsing fever (*Borrelia recurrentis*) in a Somali refugee arriving in Italy: a re-emerging infection in Europe? *PLoS Negl Trop Dis* 10:e0004522. <https://doi.org/10.1371/journal.pntd.0004522>.
 22. Lescot M, Audic S, Robert C, Nguyen TT, Blanc G, Cutler SJ, Winkler P, Couloux A, Claverie J-M, Raoult D, Drancourt M. 2008. The genome of *Borrelia recurrentis*, the agent of deadly louse-borne relapsing fever, is a degraded subset of tick-borne *Borrelia duttonii*. *PLoS Genet* 4:e1000185. <https://doi.org/10.1371/journal.pgen.1000185>.
 23. Zhou Z, Lundstrom I, Tran-Dien A, Duchene S, Alikhan NF, Sergeant MJ, Langridge G, Fotakis AK, Nair S, Stenoien HK, Hamre SS, Casjens S, Christophersen A, Quince C, Thomson NR, Weill FX, Ho SYW, Gilbert MTP, Achtman M. 2018. Pan-genome analysis of ancient and modern *Salmonella enterica* demonstrates genomic stability of the invasive Para C lineage for millennia. *Curr Biol* 28:2420–2428.e10. <https://doi.org/10.1016/j.cub.2018.05.058>.
 24. Vågene ÅJ, Herbig A, Campana MG, Robles García NM, Warinner C, Sabin S, Spyrou MA, Andrades Valtueña A, Huson D, Tuross N, Bos KI, Krause J. 2018. *Salmonella enterica* genomes from victims of a major sixteenth-century epidemic in Mexico. *Nat Ecol Evol* 2:520–528. <https://doi.org/10.1038/s41559-017-0446-6>.
 25. Dziejman M, Balon E, Boyd D, Fraser CM, Heidelberg JF, Mekalanos J. 2002. Comparative genomic analysis of *Vibrio cholerae*: genes that correlate with cholera endemic and pandemic disease. *Proc Natl Acad Sci U S A* 99:1556–1561. <https://doi.org/10.1073/pnas.042667999>.
 26. Devault A, Brian Golding G, Wagelchner N, Enk J, Kuch M, Tien JH, Shi M, Fisman D, Dhody A, Forrest S, Bos K, Earn D, Holmes E, Poinar H. 2014. Second-pandemic strain of *Vibrio cholerae* from the Philadelphia cholera outbreak of 1849. *New Engl J Med* 370:334–340.
 27. Dorman M, Kane L, Dorman D, Turnbull JD, Cormie C, Fazal MA, Goulding DA, Russell JE, Alexander S, Thomson NR. 2019. The history, genome, and biology of NCTC 30: a non-pandemic *Vibrio cholerae* isolate from World War I. *Proc R Soc B* 286:20182025. <https://doi.org/10.1098/rspb.2018.2025>.
 28. Davis GH, Park RW. 1962. A taxonomic study of certain bacteria currently classified as *Vibrio* species. *J Gen Microbiol* 27:101–119. <https://doi.org/10.1099/00221287-27-1-101>.
 29. Barlow M, Hall BGJ. 2002. Phylogenetic analysis shows that the OXA β -lactamase genes have been on plasmids for millions of years. *J Mol Evol* 55:314–321. <https://doi.org/10.1007/s00239-002-2328-y>.
 30. Rohmer L, Jacobs MA, Brittnacher MJ, Fong C, Hayden HS, Hocquet D, Weiss EJ, Radey M, Germani Y, Talukder KA, Hager AJ, Kemner JM, Sims-Day EH, Matamouros S, Hager KR, Miller SI. 2014. Genomic analysis of the emergence of 20th century epidemic dysentery. *BMC Genomics* 15:355. <https://doi.org/10.1186/1471-2164-15-355>.
 31. Baker KS, Mather AE, McGregor H, Coupland P, Langridge GC, Day M, Deheer-Graham A, Parkhill J, Russell JE, Thomson NR. 2014. The extant World War I dysentery bacillus NCTC1: a genomic analysis. *Lancet* 384: 1691–1697. [https://doi.org/10.1016/S0140-6736\(14\)61789-X](https://doi.org/10.1016/S0140-6736(14)61789-X).
 32. Baker KS, Burnett E, McGregor H, Deheer-Graham A, Boinett C, Langridge GC, Willan AM, Cain AK, Thomson NR, Russell JE, Parkhill J. 2015. The Murray collection of pre-antibiotic era *Enterobacteriaceae*: a unique research resource. *Genome Med* 7:97. <https://doi.org/10.1186/s13073-015-0222-7>.
 33. Reardon S. 2014. Antibiotic resistance sweeping developing world: bacteria are increasingly dodging extermination as drug availability outpaces regulation. *Nature* 509:141–143. <https://doi.org/10.1038/509141a>.
 34. Datta N, Hughes VM. 1983. Plasmids of the same Inc groups in enterobacteria before and after the medical use of antibiotics. *Nature* 306: 616–617. <https://doi.org/10.1038/306616a0>.
 35. Haase JK, Murphy RA, Choudhury KR, Achtman M. 2011. Revival of Seeliger's historical “Special Listeria Culture Collection.” *Environ Microbiol* 13:3163–3171. <https://doi.org/10.1111/j.1462-2920.2011.02610.x>.
 36. Bergholz TM, den Bakker HC, Katz LS, Silk BJ, Jackson KA, Kucerova Z, Joseph LA, Turnsek M, Gladney LM, Halpin JL, Xavier K, Gossack J, Ward TJ, Frace M, Tarr CL. 2016. Determination of evolutionary relationships of outbreak-associated *Listeria monocytogenes* strains of serotypes 1/2a and 1/2b by whole-genome sequencing. *Appl Environ Microbiol* 82: 928–938. <https://doi.org/10.1128/AEM.02440-15>.
 37. Hyden P, Pietzka A, Allerberger F, Springer B, Sensen C, Ruppitsch W. 2016. Draft genome sequence of a 94-year-old *Listeria monocytogenes* isolate, SLCC208. *Genome Announc* 4:e01572-15.
 38. Hyden P, Pietzka A, Lennkh A, Murer A, Springer B, Blaschitz M, Indra A, Huhulescu S, Allerberger F, Ruppitsch W, Sensen CW. 2016. Whole-genome sequence-based serogrouping of *Listeria monocytogenes* isolates. *J Biotechnol* 235:181–186. <https://doi.org/10.1016/j.jbiotec.2016.06.005>.
 39. Manson-Bahr PH. 1942. Dysentery and diarrhoea in wartime. *Br Med J* 2:346–348. <https://doi.org/10.1136/bmj.2.4263.346>.
 40. Njamkepo E, Fawal N, Tran-Dien A, Hawkey J, Strockbine N, Jenkins C, Talukder KA, Bercion R, Kuleshov K, Kolinská R, Russell JE, Kaftyreva L, Accou-Demartin M, Karas A, Vandenberg O, Mather AE, et al. 2016. Global phylogeography and evolutionary history of *Shigella dysenteriae* type 1. *Nat Microbiol* 1:16027. <https://doi.org/10.1038/nmicrobiol.2016.27>.
 41. Tran-Dien A, Le Hello S, Bouchier C, Weill F-X. 2018. Early transmissible ampicillin resistance in zoonotic *Salmonella enterica* serotype Typhimurium in the late 1950s: a retrospective, whole-genome sequencing study. *Lancet Infect Dis* 18:207–214. [https://doi.org/10.1016/S1473-3099\(17\)30705-3](https://doi.org/10.1016/S1473-3099(17)30705-3).
 42. Temkin E, Fallach N, Almagor J, Gladstone BP, Tacconelli E, Carmeli Y. 2018. Estimating the number of infections caused by antibiotic-resistant *Escherichia coli* and *Klebsiella pneumoniae* in 2014: a modelling study. *Lancet Global Health* 6:e969–e979. [https://doi.org/10.1016/S2214-109X\(18\)30278-X](https://doi.org/10.1016/S2214-109X(18)30278-X).
 43. Wand ME, Baker KS, Benthall G, McGregor H, McCowen JW, Deheer-Graham A, Sutton JM. 2015. Characterization of pre-antibiotic era *Klebsiella pneumoniae* isolates with respect to antibiotic/disinfectant susceptibility and virulence in *Galleria mellonella*. *Antimicrob Agents Chemother* 59:3966–3972. <https://doi.org/10.1128/AAC.05009-14>.