Yersinia pestis and the Plague of Justinian 541–543 AD: a genomic analysis

David M Wagner*, Jennifer Klunk*, Michaela Harbeck, Alison Devault, Nicholas Waglechner, Jason W Sahl, Jacob Enk, Dawn N Birdsell, Melanie Kuch, Candice Lumibao, Debi Poinar, Talima Pearson, Mathieu Fourment, Brian Golding, Julia M Riehm, David J D Earn, Sharon DeWitte, Jean-Marie Rouillard, Gisela Grupe, Ingrid Wiechmann, James B Bliska, Paul S Keim, Holger C Scholz, Edward C Holmes, Hendrik Poinar

Summary

Background *Yersinia pestis* has caused at least three human plague pandemics. The second (Black Death, 14–17th centuries) and third (19–20th centuries) have been genetically characterised, but there is only a limited understanding of the first pandemic, the Plague of Justinian (6–8th centuries). To address this gap, we sequenced and analysed draft genomes of *Y pestis* obtained from two individuals who died in the first pandemic.

Methods Teeth were removed from two individuals (known as A120 and A76) from the early medieval Aschheim-Bajuwarenring cemetery (Aschheim, Bavaria, Germany). We isolated DNA from the teeth using a modified phenol-chloroform method. We screened DNA extracts for the presence of the *Y pestis*-specific *pla* gene on the pPCP1 plasmid using primers and standards from an established assay, enriched the DNA, and then sequenced it. We reconstructed draft genomes of the infectious *Y pestis* strains, compared them with a database of genomes from 131 *Y pestis* strains from the second and third pandemics, and constructed a maximum likelihood phylogenetic tree.

Findings Radiocarbon dating of both individuals (A120 to 533 AD [plus or minus 98 years]; A76 to 504 AD [plus or minus 61 years]) places them in the timeframe of the first pandemic. Our phylogeny contains a novel branch (100% bootstrap at all relevant nodes) leading to the two Justinian samples. This branch has no known contemporary representatives, and thus is either extinct or unsampled in wild rodent reservoirs. The Justinian branch is interleaved between two extant groups, 0.ANT1 and 0.ANT2 , and is distant from strains associated with the second and third pandemics.

Interpretation We conclude that the *Y pestis* lineages that caused the Plague of Justinian and the Black Death 800 years later were independent emergences from rodents into human beings. These results show that rodent species worldwide represent important reservoirs for the repeated emergence of diverse lineages of *Y pestis* into human populations.

Funding McMaster University, Northern Arizona University, Social Sciences and Humanities Research Council of Canada, Canada Research Chairs Program, US Department of Homeland Security, US National Institutes of Health, Australian National Health and Medical Research Council.

Introduction

Between 541 and 543 AD, the Plague of Justinian, traditionally regarded as the first of three human plague pandemics, spread from either central Asia or Africa across the Mediterranean basin into Europe, killing an estimated 100 million people according to the contemporary scholar Procopius (although this history is disputed[1]), contributing to the end of the Roman empire, and marking the transition from the classical to the Medieval period.[2–4] Subsequent outbreaks of this disease occurred in 8–12 year cycles for two centuries after the initial epidemic[5,6] with estimated mortality of 15–40%.[6] Although suspected to have been caused by *Yersinia pestis*, uncertainties in historical epidemiology have led some researchers to propose other pathogens as the causative agent of the Plague of Justinian, such as an influenza virus.[7] However, the symptoms described by Procopius are very similar to those later reported during the second pandemic of the 14–17th centuries (ie, the Black Death),[4] from which *Y pestis* has been genomically characterised,[8] implicating this bacterial species as the infectious agent of the first pandemic.

Although the presence of *Y pestis* in victims of the Plague of Justinian has been confirmed by the sequencing of *Y pestis*-specific genomic regions from skeletal material of individuals buried in Sens, France,[9] and Bavaria, Germany,[10,11] a detailed genomic characterisation of the strains that caused the first pandemic and their relation to subsequent plague pandemics is absent. In particular, it is unclear whether the Justinian strains were direct ancestors of the Black Death strains and later plague outbreaks, or form a novel lineage of *Y pestis*. Similarly, it is unclear whether the second and third pandemics were the result of continued re-emergences of one pandemic strain, or were novel emergences from diverse rodent reservoirs. To address these questions we report and analyse draft genomes of *Y pestis* obtained from individuals who died in the first pandemic.

Methods

DNA extraction, screening, preparation, and indexing

Teeth were removed from two individuals (known as A120 and A76) from the Aschheim-Bajuwarenring

SC, USA; Department of Chemical Engineering, University of Michigan, Ann Arbor, MI, USA (J-M Rouillard PhD); Mycroarray, Ann Arbor, MI, USA (J-M Rouillard); Department Biology I, Biodiversity Research/Anthropology (Prof G Grupe), and Department of Veterinary Sciences, Institute of Palaeoanatomy, Domestication Research and the History of Veterinary Medicine (Ingrid Wiechmann PhD), Ludwig-Maximilian University of Munich, Martinsried, Germany; and Department of Molecular Genetics and Microbiology and Center for Infectious Diseases, Stony Brook University, Stony Brook, NY, USA (Prof J B Bliska PhD)

Correspondence to:
Dr Hendrik Poinar, McMaster Ancient DNA Centre, Department of Anthropology, McMaster University, Hamilton, ON L8S 4L8, Canada
poinarh@mcmaster.ca
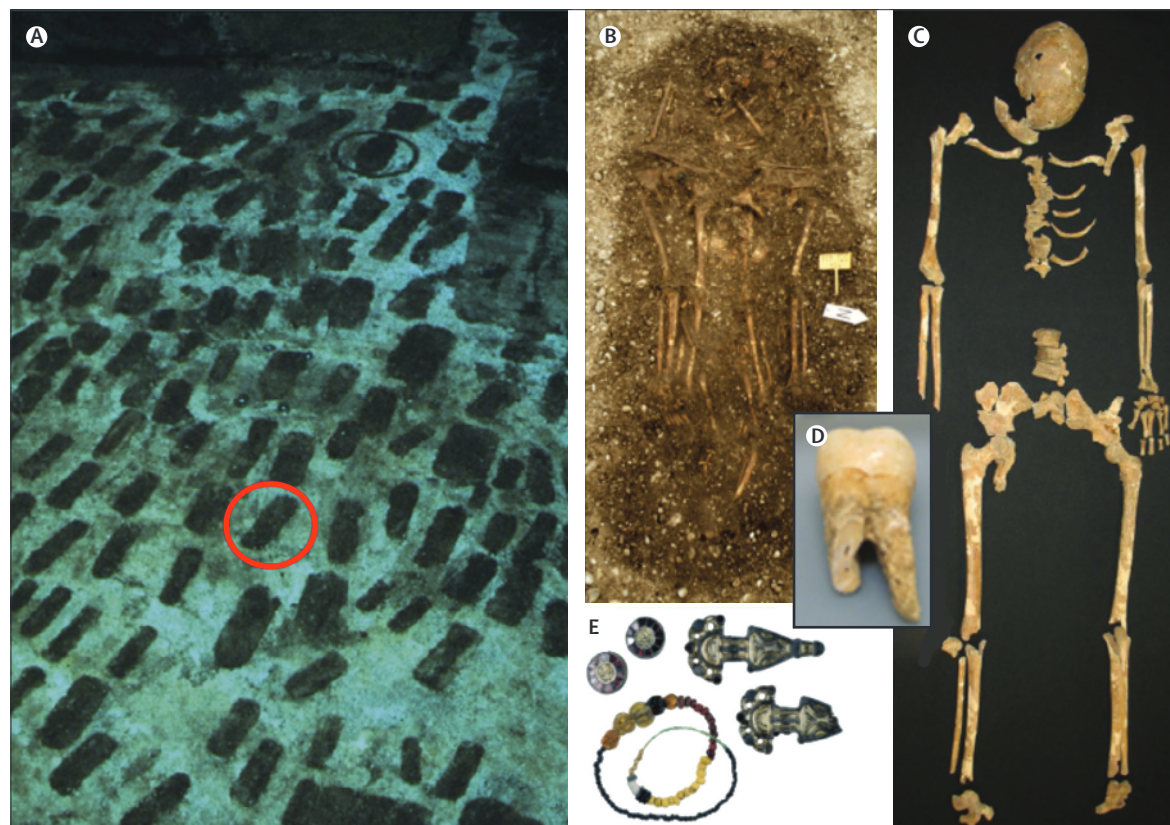
Figure 1: Source of materials
(A) Aerial photograph showing the location of the grave of individual A120 in the Aschheim-Bajuwarenring cemetery, Bavaria, Germany. (B) Shared grave of A120 and two other individuals. (C) Skeleton of A120. (D) Sampled tooth from A120. (E) Goods[23] obtained from the grave of A120 that were used to estimate the age of this burial to 525–550 AD. Parts (A) and (B) reproduced with permission of Hans Peter Volpert. Part (E) reproduced with permission of Doris Gutsmiedl-Schümann.

cemetery, an early medieval graveyard in Aschheim, Bavaria, Germany, which contained 438 individuals buried in single and a few multiple graves.[12] All laboratory procedures were done at the McMaster Ancient DNA Centre (McMaster University, Hamilton, ON, Canada).

We extracted DNA using a modified phenol-chloroform method as previously described.[13] We screened DNA extracts for the presence of the *Y pestis*-specific *pla* gene on the pPCP1 plasmid using primers and standards from an established assay[14] with minor modifications of the cycling conditions (appendix 1 p 2). We followed a previously published protocol[15] with recommended modifications[16] for library preparation of all our DNA extracts (appendix 1 pp 2–3).

DNA enrichment
We designed a 1 million Agilent SureSelect DNA capture array (Agilent, Santa Clara, USA) using the program PanArray (version 1.0)[17] as per the manufacturer's instructions. We aimed to target and capture sequences from the core *Y pestis* chromosome (strain CO92,[18] NC-003143), three primary plasmids (pPCP1, pMT1, and pCD1), and genes from 155 other *Y pestis* strains not present

in the core chromosomal genome (all targets are listed in appendix 1 pp 18–21). Although this array contained probes representing all unique genomic components from 156 *Y pestis* strains it was not capable of capturing unique genomic sequences that might have been present in the Justinian lineage. 2 million 60-mer probes were tiled every five bases and distributed in genomically contiguous sections across two 1 million arrays. In-solution enrichment was done according to the manufacturer's protocol (Mycroarray, Ann Arbor, MI, USA). We designed three in-solution bait sets for the complete human mitochondrial genome, the complete *Y pestis* CO92 genome, and all 2298 previously confirmed polymorphic sites[19] plus all novel single-nucleotide polymorphisms (SNPs) discovered in our novel strains. We did 46 serial targeted enrichments from 23 uracil DNA glycosylase (UDG)-treated DNA libraries derived from four teeth from individual A120 and one tooth from A76.

We sequenced the enriched libraries on the Illumina HiSeq 1500 platform (Illumina, San Diego, CA, USA). Adapter sequences 3′ of the inserts were trimmed with cutadapt (version 1.0).[20] We then merged the reads using FLASH (version 1.2.4).[21] Finally, we mapped the trimmed

merged reads using BWA (version 0.7.5;[22] appendix 1 p 5). All read files have been deposited at the National Centre for Biotechnology Information Sequence Read Archive (accession number SRP033879).

To ensure a minimum of five-times coverage of all novel SNPs discovered in the Justinian strains in addition to known SNPs within *Y pestis*,[19] we designed 7083 80-mer in-solution baits and serially enriched and sequenced 14 A120 and A76 UDG-treated DNA libraries. We identified SNPs specific to the Justinian samples using Geneious (version 6.1.2) at five-times read coverage and 90% variant frequency and Freebayes (version 0.9.9)[23] at five-times read coverage and 90% variant frequency with the parameters: use mapping quality, exclude unobserved genotypes, alternate allele frequency, ploidy of one, minimum mapping quality of 20, minimum base quality of 20, exclude multiple nucleotide variants, no complex events, no indels, and disabled population priors. SNPs that were called by only one of the two programs were visually examined for authenticity and were only included in subsequent phylogenetic analysis when total variant calls were at least 90% (final SNPs are in appendix 2).

A maximum likelihood phylogenetic tree was estimated for 133 strains of *Y pestis* with an additional strain of *Y pseudotuberculosis* (IP32953) used as an outgroup (appendix 1 p 17). We did analysis with the PhyML program (version 3)[24] using the generalised time-reversible GTR+γ_4 model of nucleotide substitution with substitution rates (GTR) and among-site rate variation (γ_4) estimated from the empirical data. To infer the rate and timescale of *Y pestis* evolution we used a Bayesian Markov chain Monte Carlo approach (BEAST program, version 1.7.5).[25]

Role of the funding source

The funders of the study had no role in study design, data collection, data analysis, data interpretation, or writing of the report. All authors had full access to all the data in the study. DMW, ECH, and HP made the final decision to submit for publication.

Results

Radiocarbon dating of both individuals (A120 to 533 AD [plus or minus 98 years]; A76 to 504 AD [plus or minus 61 years]), and the dating of associated grave goods (figure 1[12]), places them in the timeframe of the first pandemic.[11] To assess the authenticity of the DNA isolated from the remains, we used in-solution enrichment targeting the human mitochondrial genome and sequenced samples to 218-times (A120) and 1600-times (A76) coverage. The consensus sequences obtained were distinct haplotypes matching closest to J1c3 (A76) and H8c (A120; appendix 1 pp 22–23), both of
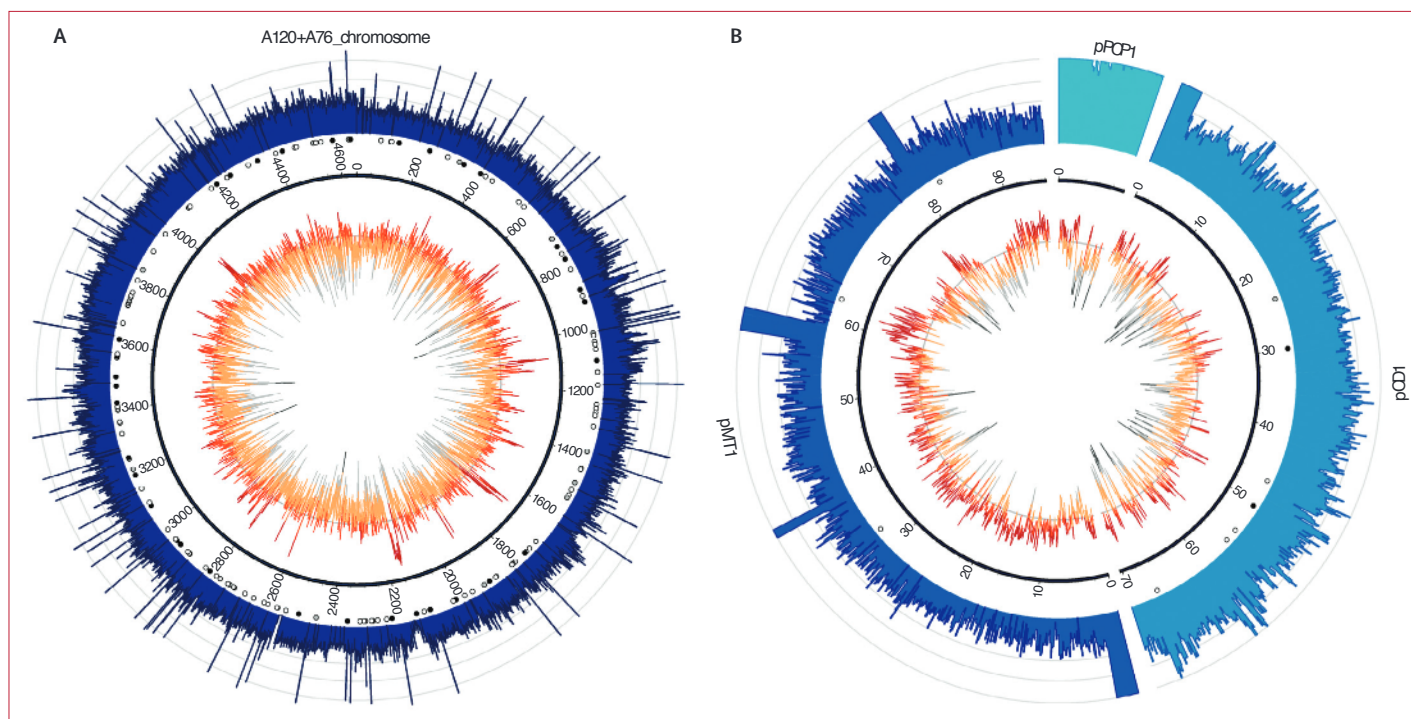
Figure 2: **Coverage plots (log scale, outer ring) and GC percentage content (orange inner ring) for the draft genome of Justinian sample A120+A76**
(A) Mapped to the chromosome of the reference strain CO92. (B) Mapped to the three plasmids (pPCP1, pCD1, and pMT1) of the reference strain CO92. The chromosome and plasmid plots are not shown to same scale. Single-nucleotide polymorphism colours: black=non-synonymous, grey=synonymous, white=non-coding region. Coverage axes (outer ring) for both chromosome and plasmids at 1 (ten-times), 1·47 (30-times), and 2 (100-times) GC content: axis at 0·5. Colour gradation from greater than 0·55 (red) to less than 0·3 (black). Coverage and GC content calculated in: 1000 bp windows from chromosome, 100 bp windows for plasmids. Figures produced using Circos.[27]
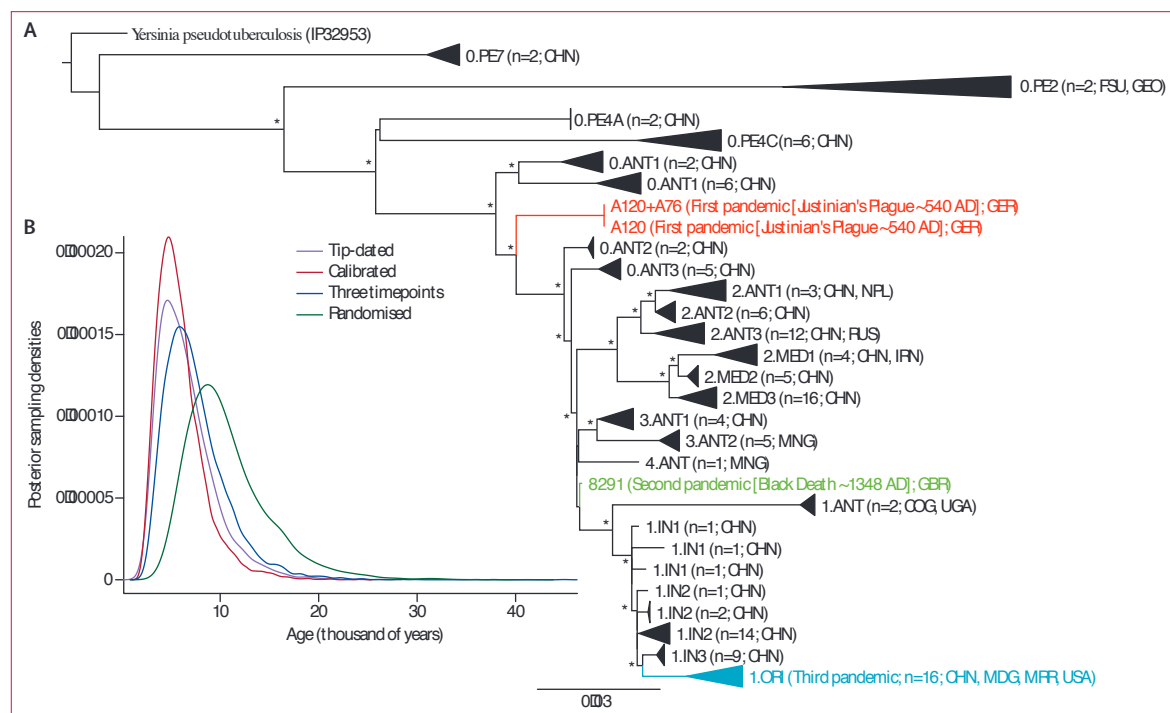
Figure 3: (A) Maximum likelihood tree and (B) posterior distributions of times to most recent common ancestry for different BEAST runs
(A) Maximum likelihood tree of 2268 single-nucleotide polymorphisms (SNPs) from 133 strains of *Yersinia pestis*. The tree is rooted using a genome sequence[28] from the soil-dwelling *Y pseudotuberculosis,* the probable ancestor of *Y pestis*.[19,29] The branch leading to *Y pseudotuberculosis* is artificially short because no SNPs were identified in this species; it was therefore excluded from the molecular clock and root-to-tip regression analyses. The branches associated with first, second, and third pandemics are coloured red, green, and blue, respectively, and, when possible, branches have been collapsed to improve clarity. Phylogenetic groups (eg, 0.PE7) are designated as described previously,[19] with the first number indicating the major branch (0–4) along which that group is found; the number of strains and the countries where the strains were isolated are indicated for each group. (B) The different BEAST[25] runs were calibrated using: (1) all *Y pestis* tip dates in the dataset ("Tip-dated"), (2) dates of the Plague of Justinian and Black Death as nodal "fossils" with all modern strains assumed to be isochronous (year 2000; "Calibrated"), (3) tip dates of the three ancient strains incorporated with all modern strains assumed to be isochronous (year 2000; "Three timepoints"), and (4) tip dates randomised by swapping the sampling date of the Justinian and Black Death samples ("Randomised"). The overlapping distributions, particularly with the randomised data, show the absence of temporal signal in the data. CHN=China. FSU=Former Soviet Union. GEO=Georgia. GER=Germany. NPL=Nepal. RUS=Russia. IRN=Iran. MNG=Mongolia. GBR=Great Britain. COG=Republic of Congo. UGA=Uganda. MDG=Madagascar. MRR=Myanmar.

which are identified at high frequencies in contemporary European and Middle Eastern populations.[26] Damage profiles (appendix 1 pp 12–15) of libraries not treated with UDG were consistent with ancient damaged DNA and support the authenticity of our samples.[14]

DNA enrichment and sequencing produced 511 842 485 bp of sequence. After trimming and merging, raw mapped reads with identical 5′ and 3′ end coordinates and strand origin (direction) were collapsed to establish unique counts. This procedure resulted in 683 919 reads of an average length of 51·8 bp that mapped to the CO92 reference, with an average of 7·6-times coverage for 91·5% of the chromosome, 99·9% for pCD1, 96·7% for pMT1, and 100·0% for pPCP1 (figure 2;[27] appendix 1 p 22). Because the overall quality of the sequences and the total coverage obtained for strain A76 was much lower than that for strain A120, we analysed data from A76 and A120 together as well as A120 individually.

We used two approaches to infer the placement of Justinian strains (A120 and A120+A76) in the *Y pestis* phylogeny. First, we queried the allele state in A120 and

A120+A76 at 2298 SNPs identified in a large-scale phylogenetic analysis of *Y pestis*.[19] Of these SNPs, 1509 positions from A120 and 1675 positions from A120+A76 were covered at more than four-times depth and were present in at least 90% of the sequences obtained (appendix 1 pp 25–27), yielding identical and unambiguous placement of the Justinian samples along branch N04-N05 in the phylogeny of Cui and colleagues.[19] Second, we inferred a maximum likelihood phylogenetic tree (figure 3) using 2268 polymorphic positions in 130 contemporary *Y pestis* strains,[19] the draft genome from the second pandemic strain,[8] and the two Justinian samples (appendix 3). One (representative) *Y pseudotuberculosis* genome[28] was used as an outgroup. Our phylogeny is much the same as that obtained previously[19] but contains a novel branch (100% bootstrap at all relevant nodes) leading to the two Justinian samples. Notably, this branch has no known contemporary representatives, and so is either extinct or unsampled in wild rodent reservoirs. The Justinian branch is interleaved between two extant groups, 0.ANT1

and 0.ANT2 (figure 3), and distant from strains associated with the second and third pandemics. These extant groups are identified in long-term plague foci in the Xinjiang region of China[29] where they are ecologically established within rodent and flea populations.[30] These findings are compatible with the theory that the first pandemic emerged from rodents in long-term plague foci in China, not Africa as was originally proposed by Procopius,[3] and spread east to Europe along established trade routes (eg, the Silk Road; figure 4).

To further assess whether the *Y pestis* strains associated with the Justinian plague represented a novel emergence event with no later sampled descendants we compared the maximum likelihood tree with a second phylogeny in which the Justinian strains were constrained to be the direct ancestor of those associated with the second and third pandemics. The likelihood of the maximum likelihood tree (figure 3) was significantly better than the competing phylogeny (p<0·001, Shimodaira-Hasegawa test[31]), strongly supporting the idea that the Plague of Justinian was an independent emergence that did not give rise to later human pandemics. Thus, the Black Death is the result of a separate emergence of *Y pestis*, some 800 years later, from rodents into human beings that led to pandemic plague.

The contrasting phylogenetic positions of the Plague of Justinian and the Black Death strains are noteworthy (figure 3). The Black Death strain sits near the base of branch 1, strongly suggesting that the second pandemic gave rise to all later *Y pestis* strains along branch 1, including those directly responsible for the third

pandemic (1.ORI group; figure 3). These descendants include the 1.ANT group, which is currently ecologically established in rodent populations in central Africa; the 1.IN1, 1.IN2, and 1.IN3 groups, currently identified in rodent populations in western, central, and southern China, respectively;[19,29] and the 1.ORI group, which arose in southeastern China and spread from Hong Kong worldwide during the third pandemic, leading to the establishment of new rodent foci in Africa, Asia, North America, and South America that persist to this day.[29] A viable explanation for these phylogeographic patterns is that *Y pestis* spread from Asia to Europe during the second pandemic and then from Europe or northern Africa to central Africa establishing the 1.ANT group. Under this scenario, *Y pestis* also spread back to China from Europe (or northern Africa) along the same established trade routes giving rise to the 1.IN populations, which later gave rise to the 1.ORI group that caused the third pandemic (figure 4). Historical records are consistent with this scenario because the 1.IN3 group is identified in the Yunnan province of China, where the third pandemic originated.[32] Additionally, some 1.IN3 strains exhibit the derived orientalis phenotype,[19] which is conserved in all 1.ORI strains but not identified in any other groups.[19,29]

As in other *Y pestis* phylogenies,[19] our maximum likelihood tree is notable for a substantial variation in branch length that has no correlation with sampling time (eg, the long branch leading to the 0.PE2 group; figure 3). In particular, the lineage of the Justinian strains is roughly the same distance from the root of the tree as the strain from the Black Death, despite being about 800 years older.
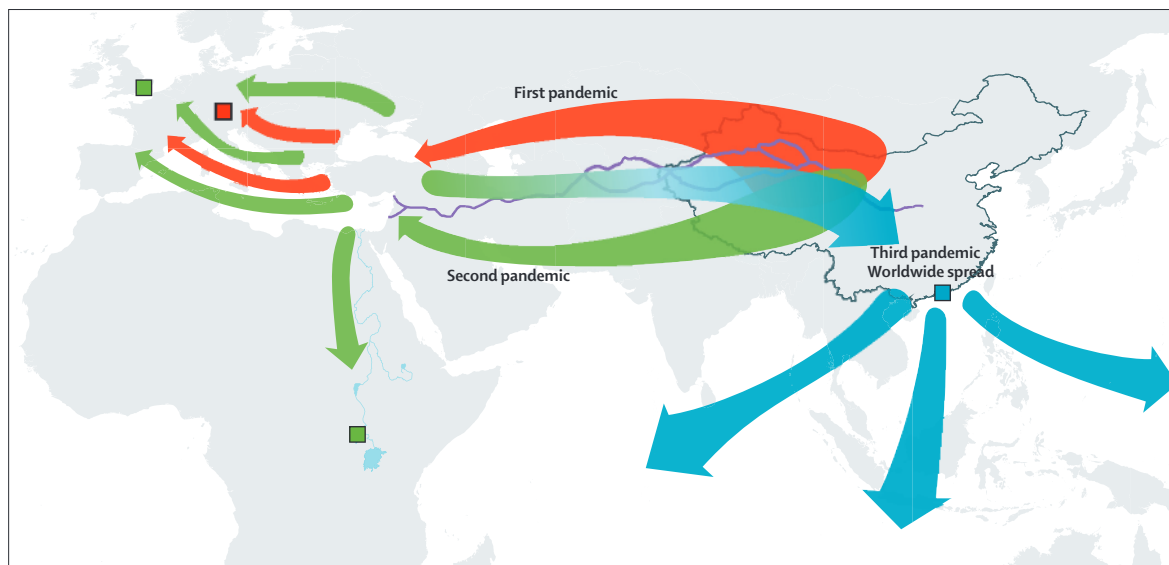
Figure 4: **Hypothetical scenario for the geographic spread of *Yersinia pestis***
The lineage associated with the first pandemic might have arisen in Asia (China outlined) and then travelled along established trade routes, such as the Silk Road (purple lines), to Europe. The lineage associated with the second pandemic might also have spread to Europe from Asia along trade routes and from Europe via trade routes along the Nile river to central Africa. It might also have travelled back to China to establish the 1.IN populations, which later gave rise to the 1.ORI group involved in the third pandemic. The orange square shows the location of the Justinian samples, the green square in the UK shows the location of the East Smithfield Black Death samples, the green square in central Africa shows the location of the extant 1.ANT group, and the blue square in China shows Hong Kong, from which the third pandemic was disseminated globally.

Panel: Research in context

Systematic review
We searched PubMed for articles published up to Jan 9, 2014, using the search terms "Yersinia pestis", "first pandemic", "ancient DNA", and "whole genome sequencing". This search yielded no reports. We repeated the search using the search terms "plague", "first pandemic", "ancient DNA", and "whole genome sequencing". Again, this search yielded no reports. Whole genome sequencing of *Y pestis* from samples of the first plague pandemic does not exist because of the rarity of these samples and the technical difficulties associated with generating whole genome sequences from degraded ancient DNA.

Interpretation
Our results show that the lineage of *Y pestis* that caused the first plague pandemic is distinct from the lineage that caused the second pandemic (including the Black Death) 800 years later. These findings show how rodent-based plague foci worldwide represent important reservoirs for re-emergence of *Y pestis* in human populations.

In our Bayesian Markov chain Monte Carlo analysis, use of various methods to account for rate variation did not produce a robust estimate for either the rate of nucleotide substitution or times to common ancestry (figure 3; appendix 1 pp 6–7). An absence of temporal structure, which complicates all molecular clock dating, was confirmed through regressions of date (ie, year) of sampling against root-to-tip genetic distance on the maximum likelihood phylogeny for all strains (correlation coefficient 0·25), contemporary strains only (0·27), and the Black Death strain and its descendants (0·20). We therefore conclude, as have others,[19] that the evolution of *Y pestis* is characterised by extensive rate variation, probably attributable to a combination of fluctuations in the number of bacterial replication cycles (generations) per unit time, such as between endemic and epidemic cycles, or changes in selection pressures through time, including localised adaptive evolution, or combinations of these factors. Whatever the cause of this rate variation, these data suggest that previous molecular-clock-derived estimates of the timescale of *Y pestis* evolution,[8,19] including the date of its divergence from *Y pseudotuberculosis*, might be erroneous. Such extensive rate variation contrasts with studies of other bacteria such as *Mycobacterium leprae* in which samples collected over about 1000 years show substantial temporal structure.[33] Importantly, without a firm timescale for *Y pestis* evolution we cannot exclude that earlier dated epidemics (Plague of Athens of 430 BC) or pandemics (Antonine Plague 165–180 AD) of unknown cause are potentially derived from repeated, novel emergences of *Y pestis* into human populations as was the case with later pandemics.

Estimated mortality from the Justinian Plague was up to 40%,[3] which might mean that the strains of *Y pestis* associated with this pandemic were of intrinsically high virulence. Of the 66 unique chromosomal and plasmid SNPs identified in A120 and A76 combined, 23 were non-synonymous, two of which are identified in well-characterised virulence factors—*ail* and *yopJ* (appendix 1 pp 27–29). Ail is an outer membrane protein that promotes bacterial resistance to complement-mediated killing and adhesion to host cells. A novel Arg→Gly aminoacid substitution occurs in an extracellular region (loop 3) of the Ail protein from the Justinian strains, which could potentially increase resistance to complement-mediated killing or adherence, or both.[34,35] YopJ is an effector protein with acetyltransferase activity that is secreted into host cells. Compared with CO92, the Lys→Glu substitution in YopJ of the Justinian (and Black Death) strains increases the antihost activity of this effector.[36] Additionally, using our new pan-array we were able to identify a roughly 15 kb genomic island in the Justinian strains and also in the Black Death strain that we previously sequenced (appendix 1 p 16).[8] This genomic island, referred to as DFR 4 (different region 4),[37] was lost in some *Y pestis* strains, such as CO92. DFR 4 contains several genes with potential roles in virulence, including *ccm2A*, which encodes a multidrug ABC transporter ATPase. DFR 4 is currently identified in *Y pseudotuberculosis* and some *Y pestis* strains (in 0.PE4 and other groups on branch 2). This island is an example of the decay common to the highly plastic *Y pestis* genome.[28] What part DFR 4 and novel non-synonymous SNPs might have played in virulence is unknown. However, additional sequences from ancient plague pandemics will allow us to assess the virulence capacity of these mutations. In the present report we have focused exclusively on possible genomic differences within our ancient *Y pestis* strains that might have accounted for the high mortality associated with the first plague pandemic. However, host and environmental factors also probably had a role in determining mortality.

Discussion

Our genomic analysis shows that the strains of *Y pestis* involved in the Plague of Justinian form a novel branch on the *Y pestis* phylogeny. This branch has no known contemporary representatives, and so is either extinct or unsampled in wild rodent reservoirs. The Justinian branch is interleaved between two extant groups, 0.ANT1 and 0.ANT2, and is distant from strains associated with the second and third pandemics. Our phylogeny showed a substantial variation in branch length that had no correlation with sampling time. We identified 23 non-synonymous SNPs in the Justinian strains, two of which occurred in well-characterised virulence factors—*ail* and *yopJ*. The Justinian strains also include the genomic island known as DFR 4, which contains several genes with potential roles in virulence. These genomic components, combined with host and environmental factors, might have contributed to the high (up to 40%) mortality associated with the Plague of Justinian.

Our study provides a new perspective on one of the most devastating disease outbreaks in human history and the dynamics of plague emergences in general (panel). We have shown that the strain associated with the Plague of Justinian is distinct from those associated with later human plague pandemics. Thus, we infer that *Y pestis* has emerged from rodent reservoirs at several timepoints in history to cause pandemics in human beings. In the case of the lineage associated with the first pandemic, this emergence seems to be an evolutionary dead-end, because no extant descendants of this lineage have been identified so far. By contrast, the lineage associated with the second pandemic was more successful in that it gave rise to new rodent foci, one of which was later responsible for the third pandemic, in turn establishing new endemic rodent foci worldwide. Thus, each successive pandemic is associated with increasingly successful spatial dissemination of *Y pestis*; the first seems to have died out, the second spread back and forth between Asia, Europe, and Africa, whereas the third became established worldwide (figure 4).

Why the *Y pestis* lineage associated with the Plague of Justinian eventually died out is unclear. That it probably caused human epidemics for several centuries[5,6] suggests that it was well adapted to human transmission. As a consequence, several viable explanations for its extinction might be a scarcity of susceptible hosts (people, or rodents, or both), insufficient numbers of susceptible hosts in a background of widespread population immunity, or mutations that arose and spread in the human genome that conferred resistance to this particular plague strain. The success of the lineage (or lineages) associated with the two most recent pandemics is probably attributable partly to human mobilisation,[38] since increasing trade between countries and continents is known to have moved people and rodents with *Y pestis* infection around the world.[32]

The Plague of Justinian and, indeed, the emergence of all three plague pandemics, might be tightly linked to climatic instability; all were preceded by periods of exceptional rainfall[39–41] and ended during periods of climatic stability (around 700–1000 AD in the case of the Plague of Justinian).[39] Irrespective of the effect of climate, the epidemiological pattern that we propose suggests that several *Y pestis* lineages, which are currently ecologically established in rodent foci worldwide, remain capable of emerging and igniting epidemics of plague in human beings, as they have repeatedly in the past.

Contributors
DMW, HCS, MH, ECH, PSK, J-MR, and HP designed the study. JK, MH, DNB, AD, JE, MK, CL, IW, J-MR, and GG, gathered the data. JK, JWS, MF, AD, BG, NW, JE, MK, JBB, HCS, SDW, TP, DMW, ECH, and HP analysed the data. DP, DJDE, ECH, DMW, and HP wrote the initial draft and all authors contributed to subsequent drafts. DMW, ECH, and HP funded the work.

Conflicts of interest
J-MR has financial interest in Mycroarray (Ann Arbor, MI, USA) who provided in-solution capture baits used in this work. All other authors declare that they have no conflicts of interest.

References
1   Zietz BP, Dunkelberg H. The history of the plague and the research on the causative agent *Yersinia pestis*. *Int J Hyg Env Health* 2004; **207:** 165–78.
2   Russell JC. That earlier plague. *Demography* 1968; **5:** 174–84.
3   Little LK, ed. Plague and the end of antiquity: the pandemic of 541–750. Cambridge: Cambridge University Press, 2007.
4   Duncan CJ, Scott S. What caused the Black Death? *Postgrad Med J* 2005; **81:** 315–20.
5   Sussman GD. Was the Black Death in India and China? *Bull History Med* 2011; **85:** 319–55.
6   Perry RD, Fetherston JD. *Yersinia pestis*—etiologic agent of plague. *Clin Microbiol Rev* 1997; **10:** 35–66.
7   Altschuler EL, Kariuki YM. Was the Justinian Plague caused by the 1918 flu virus? *Med Hypotheses* 2009; **72:** 234.
8   Bos KI, Schuenemann VJ, Golding GB, et al. A draft genome of *Yersinia pestis* from victims of the Black Death. *Nature* 2011; **478:** 506–10.
9   Drancourt M, Signoli M, Dang LV, et al. *Yersinia pestis* orientalis in remains of ancient plague patients. *Emerg Infect Dis* 2007; **13:** 332–33.
10   Wiechmann I, Grupe G. Detection of *Yersinia pestis* DNA in two early medieval skeletal finds from Aschheim (Upper Bavaria, 6th century AD). *Am J Phys Anthropol* 2005; **126:** 48–55.
11   Harbeck M, Seifert L, Hänsch S, et al. *Yersinia pestis* DNA from skeletal remains from the 6th century AD reveals insights into Justinianic Plague. *PLoS Pathog* 2013; **9:** e1003349.
12   Gutsmiedl-Schümann D. Das frühmittelalterliche Gräberfeld Aschheim-Bajuwarenring. Kallmünz/Oberpfalz: Lassleben, 2010.
13   Schwarz C, Debruyne R, Kuch M, et al. New insights from old bones: DNA preservation and degradation in permafrost preserved mammoth remains. *Nucleic Acids Res* 2009; **37:** 3215–29.
14   Schuenemann VJ, Bos K, DeWitte S, et al. Targeted enrichment of ancient pathogens yielding the pPCP1 plasmid of *Yersinia pestis* from victims of the Black Death. *Proc Natl Acad Sci USA* 2011; **108:** E746–52.
15   Meyer M, Kircher M. Illumina sequencing library preparation for highly multiplexed target capture and sequencing. *Cold Spring Harbor Protocol* 2010; **2010:** pdb prot5448.
16   Kircher M, Sawyer S, Meyer M. Double indexing overcomes inaccuracies in multiplex sequencing on the Illumina platform. *Nucleic Acids Res* 2012; **40:** e3.
17   Phillippy AM, Deng X, Zhang W, Salzberg SL. Efficient oligonucleotide probe selection for pan-genomic tiling arrays. *BMC Bioinformatics* 2009; **10:** 293.
18   Parkhill J, Wren BW, Thomson NR, et al. Genome sequence of *Yersinia pestis*, the causative agent of plague. *Nature* 2001; **413:** 523–27.
19   Cui Y, Yu C, Yan Y, et al. Historical variations in mutation rate in an epidemic pathogen, *Yersinia pestis*. *Proc Natl Acad Sci USA* 2013; **110:** 577–82.
20   Martin M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet J* 2011; **17:** 10–12.
21   Magoc T, Salzberg SL. FLASH: fast length adjustment of short reads to improve genome assemblies. *Bioinformatics* 2011; **27:** 2957–63.
22   Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* 2009; **25:** 1754–60.
23   Garrison E, Marth G. Haplotype-based variant detection from short-read sequencing. 2012. http://arxiv.org/abs/1207.3907 (accessed July 1, 2013).

24   Guindon S, Dufayard JF, Lefort V, Anisimova M, Hordijk W, Gascuel O. New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Syst Biol* 2010; **59:** 307–21.

25   Drummond AJ, Suchard MA, Xie D, Rambaut A. Bayesian phylogenetics with BEAUti and the BEAST 1.7. *Mol Biol Evol* 2012; **29:** 1969–73.

26   Pala M, Olivieri A, Achilli A, et al. Mitochondrial DNA signals of late glacial recolonization of Europe from near eastern refugia. *Am J Hum Genet* 2012; **90:** 915–24.

27   Krzywinski M, Schein J, Birol I, et al. Circos: an information aesthetic for comparative genomics. *Genome Res* 2009; **19:** 1639–45.

28   Chain PS, Carniel E, Larimer FW, et al. Insights into the evolution of *Yersinia pestis* through whole-genome comparison with *Yersinia pseudotuberculosis*. *Proc Natl Acad Sci USA* 2004; **101:** 13826–31.

29   Morelli G, Song Y, Mazzoni CJ, et al. *Yersinia pestis* genome sequencing identifies patterns of global phylogenetic diversity. *Nat Genet* 2010; **42:** 1140–43.

30   Department of Diseases Control Ministry of Health. The atlas of plague and its environment in the People's Republic of China. Beijing: Science Press, 2004.

31   Swofford D. PAUP*. Phylogenetic analysis using parsimony (and other methods), 4th edn. Sunderland, MA: Sinauer Associates, 2003.

32   Pollitzer R. Plague studies. 1. A summary of the history and survey of the present distribution of the disease. *Bull World Health Org* 1951; **4:** 475–533.

33   Schuenemann VJ, Singh P, Medum TA, et al. Genome-wide comparison of medieval and modern *Mycobacterium leprae*. *Science* 2013; **341:** 179–83.

34   Yamashita S, Lukacik P, Barnard TJ, et al. Structural insights into Ail-mediated adhesion in *Yersinia pestis*. *Structure* 2011; **19:** 1672–82.

35   Miller VL, Beer KB, Heusipp G, Young BM, Wachtel MR. Identification of regions of Ail required for the invasion and serum resistance phenotypes. *Mol Micro* 2001; **41:** 1053–62.

36   Zheng Y, Lilo S, Brodsky IE, et al. A Yersinia effector with enhanced inhibitory activity on the NF-kappaB pathway activates the NLRP3/ASC/caspase-1 inflammasome in macrophages. *PLoS Pathog* 2011; **7:** e1002026.

37   Radnedge L, Agron PG, Worsham PL, Andersen GL. Genome plasticity in *Yersinia pestis*. *Microbiology* **148:** 1687–98.

38   Keim PS, Wagner DM. Humans and evolutionary and ecological forces shaped the phylogeography of recently emerged diseases. *Nat Rev Microbiol* 2009; **7:** 813–21.

39   Buntgen U, Tegel W, Nicolussi K, et al. 2500 years of European climate variability and human susceptibility. *Science* 2011; **331:** 578–82.

40   Gage KL, Burkot TR, Eisen RJ, Hayes EB. Climate and vectorborne diseases. *Am J Prevent Med* 2008; **35:** 436–50.

41   Stenseth NC, Samia NI, Viljugrein H, et al. Plague dynamics are driven by climate variation. *Proc Natl Acad Sci USA* 2006; **103:** 13110–15.