



# No evidence for persistent natural plague reservoirs in historical and modern Europe

Nils Chr. Stenseth<sup>a,b,1,2</sup>, Yuxin Tao<sup>c,1</sup>, Chutian Zhang<sup>d,e,1</sup>, Barbara Bramanti<sup>a,f</sup>, Ulf Büntgen<sup>g,h,i,j</sup>, Xianbin Cong<sup>k</sup>, Yujun Cui<sup>l</sup>, Hu Zhou<sup>m</sup>, Lorna A. Dawson<sup>n</sup>, Sacha J. Mooney<sup>o</sup>, Dong Li<sup>c</sup>, Henry G. Fell<sup>p</sup>, Samuel Cohn<sup>q</sup>, Florent Sebbane<sup>r</sup>, Phillip Slavin<sup>s</sup>, Wannian Liang<sup>d,e</sup>, Howell Tong<sup>g,t</sup>, Ruifu Yang<sup>l,2</sup>, and Lei Xu<sup>d,e,2</sup>

Contributed by Nils Stenseth; received June 8, 2022; accepted October 7, 2022; reviewed by Barbel Finkenstadt, Paul Slack, and David M. Wagner

Caused by *Yersinia pestis*, plague ravaged the world through three known pandemics: the First or the Justinianic (6th–8th century); the Second (beginning with the Black Death during c.1338–1353 and lasting until the 19th century); and the Third (which became global in 1894). It is debatable whether *Y. pestis* persisted in European wildlife reservoirs or was repeatedly introduced from outside Europe (as covered by European Union and the British Isles). Here, we analyze environmental data (soil characteristics and climate) from active Chinese plague reservoirs to assess whether such environmental conditions in Europe had ever supported “natural plague reservoirs”. We have used new statistical methods which are validated through predicting the presence of modern plague reservoirs in the western United States. We find no support for persistent natural plague reservoirs in either historical or modern Europe. Two factors make Europe unfavorable for long-term plague reservoirs: 1) Soil texture and biochemistry and 2) low rodent diversity. By comparing rodent communities in Europe with those in China and the United States, we conclude that a lack of suitable host species might be the main reason for the absence of plague reservoirs in Europe today. These findings support the hypothesis that long-term plague reservoirs did not exist in Europe and therefore question the importance of wildlife rodent species as the primary plague hosts in Europe.

*Yersinia pestis* | natural plague reservoirs | Europe | environmental conditions | rodent diversity

The plague’s agent, *Yersinia pestis*, is primarily found in wildlife mammals but occasionally spills over to human populations. Its wildlife host species typically consist of a range of burrow-dwelling rodents (1), comprising of 279 species. Human infections usually occur after rodent–host populations remain above certain population thresholds for a few years before collapsing (2). Between high prevalence of the bacterium in the wild and subsequent plague outbreaks in humans, the bacteria can survive in what is typically described as a “natural plague reservoir,” i.e., a place where suitable rodent hosts, their flea vectors, and the pathogen (either within or outside their hosts and vectors) can exist indefinitely.

*Yersinia pestis* has caused at least three extensive human plague pandemics (3–7): the First Plague Pandemic starting with the Justinianic Plague from 541–2 to ca. 750 CE; the Second Plague Pandemic begun with the Black Death of c.1338–1353 and was followed by numerous outbreaks until the mid-19th century CE; and the Third Plague Pandemic (8) that started in 1772 as a local epidemic in Yunnan but became a global pandemic only in 1894 and continues today in various parts of the world. Plague has been reported historically to be present on all continents, except Antarctica (9). It remains contested, however, whether plague reservoirs existed in Europe in the past (2, 10–12). It has been argued that medium-term plague reservoirs may have existed in Europe, such as the one hypothesized for South-Central Germany in the later 14th century and another in Central Europe (possibly the Alps) in the late 15th–early 17th centuries, before becoming extinct (see Table 1) (13–16). However, it is not clear whether the plague bacterium settled within Europe during the late Middle Ages or was repeatedly reintroduced (through human transport) from beyond its borders (possibly from the same region) (11).

To advance the debate in understanding whether plague reservoirs could have existed given the ecology of Europe in its premodern past, we have statistically evaluated the environmental conditions characterizing active Chinese rodent-based reservoirs today. For this analysis, we have used data on current soil composition, meteorology, and elevation from active plague reservoirs in China (see *Materials and Methods*) and have applied the conditional minimum average variance estimation (MAVE) (20) method to reduce the dimension of the covariates and establish the model (see *Materials and Methods*).

We then validated our China-fitted model by checking whether it could predict the currently active plague reservoirs in the United States. We then employed the China-fitted model (validated by the US data) to predict the presence and geographic distribution of

## Significance

Plague killed millions of people during the three pandemics in the past two millennia. Despite much research, it remains unclear whether persistent natural plague reservoirs existed in Europe. To examine this question, we have developed a statistical model based on high-resolution and long-term environmental data. From it, we have found no evidence for persistent natural plague reservoirs in historical or contemporary Europe. This suggests that the plague bacterium was repeatedly introduced to Europe, although it might have survived in local medium-term reservoirs. Finally, we question the importance of wildlife rodents as the main hosts in Europe. These findings have wide-ranging significance for the study of human plague through history and provide new tools for resolving century-long enigmas posed by plague.

Author contributions: N.C.S., R.Y., and L.X. designed research; N.C.S., Y.T., C.Z., B.B., U.B., R.Y., and L.X. performed research; H.T. contributed new reagents/analytic tools; Y.T. and C.Z. analyzed data; and N.C.S., Y.T., C.Z., B.B., U.B., X.C., Y.C., H.Z., L.D., S.J.M., D.L., H.G.F., S.C., F.S., P.S., W.L., H.T., R.Y., and L.X. wrote the paper.

Reviewers: B.F., University of Warwick; P.S., Lincoln College, University of Oxford; and D.M.W., Northern Arizona University.

Competing interest statement: The authors declare a competing interest, the authors have additional information to disclose, D.W. and N.C.S. are coauthors on a 2020 review article.

Copyright © 2022 the Author(s). Published by PNAS. This open access article is distributed under Creative Commons Attribution-NonCommercial-NoDerivatives License 4.0 (CC BY-NC-ND).

<sup>1</sup>N.C.S., Y.T., and C.Z. contributed equally to this work.

<sup>2</sup>To whom correspondence may be addressed. Email: n.c.stenseth@mn.uio.no, ruifuyang@gmail.com, or xu\_lei@tsinghua.edu.cn.

This article contains supporting information online at <https://www.pnas.org/lookup/suppl/doi:10.1073/pnas.2209816119/-DCSupplemental>.

Published December 12, 2022.

**Table 1. Previously proposed localities for putative plague reservoirs in Western–Central Europe in the past**

Pandemic	Period	Place	Sources
First	Late 6th – mid-8th century	Iberian Peninsula	(17)
Second	c.1349–c.1400	South–Central Germany	(13, 14)
Second	c.1348–c.1640	Southern Alps* & Pyrenees	(15)
Second	c.1460–c.1640	Central Europe	(13)
Second	16th–early 19th century	The Balkans	(18)
Second	c.1348–c.1500	Temperate climate zone of Europe, including East Anglia (UK)	(16)
Third	c.1906–1918	East Suffolk (UK)**	(16,19)

\*The Maritime and French Alps.

\*\*The several cases reconstructed for the period 1906–1918 have been clarified to have been due to repeated introductions of plague on regional waterways (8,19).

plague reservoirs in Europe (as covered by the European Union plus the British Isles). For Europe, we further considered two annually resolved and absolutely dated summer temperatures and hydroclimate reconstructions to assess whether climatic conditions during the past 2,000 y allowed plague reservoirs to form and persist (see *Materials and Methods*).

## Results and Discussion

In the China-fitted model, all covariates exhibit some degrees of nonlinearity. In general, higher soil concentrations of cadmium (Cd), copper (Cu), iron (Fe), magnesium (Mg), sodium (Na), antimony (Sb), and uranium (U), together with an overall high soil pH, tended to be associated with higher chances of plague reservoir formation, while the opposite was true with concentrations of calcium (Ca), cerium (Ce), molybdenum (Mo), and yttrium (Y). Furthermore, elevation was associated with higher chances of plague reservoir emergence, while the opposite was true for annual mean temperature (AMT) and annual precipitation (AP) (see *SI Appendix, Table S16* and *Fig. S26* and further discussion in *SI Appendix*).

Previous studies have found relationships between specific soil attributes and plague reservoirs (21). Some of these are consistent with our results based on the Chinese conditions, while others are not. For example, it has been found that plague reservoirs in China tend to lie in areas where soils enriched with calcium or iron are prevalent (22). Consistently, Pauling et al. (23) discovered a strong positive correlation between the quantity of soil calcium carbonate and the occurrence of plague in black-tailed prairie dogs (*Cynomys ludovicianus*) in the Western United States. The importance of calcium could be associated with its role in the formation of the cuticle of insects, including fleas. Another study in the Western United States (24), based on an ecological niche model, also supports the thesis that long-term persistence of plague reservoirs depends on soil conditions, such as lower soil concentrations of Fe and Na. Moreover, a field survey in Algeria (25) and a modeling study linking plague cases and outbreaks to specific soil characteristics in the United States demonstrated that *Y. pestis* could survive outside its host in saline soil (26) (e.g., with concentrations of NaCl higher than 40 g/L), implying that the salt tolerance of *Y. pestis* might play a role in maintaining plague

reservoirs, possibly by rendering the bacterium into an L form. In addition, Mezentsev et al. (27) conducted animal experiments (using three species of gerbil, *Rhombomys opimus*, *Meriones meridianus*, and *M. tamariscinus*) providing evidence of the effect of different metals or metal combinations on the course of *Y. pestis* infection. These results support the hypothesis that the metal composition and availability (determined by soil texture, moisture content, and pH) in the soil environment, where it can be ingested by rodents, may be an important parameter influencing plague development in wild (sylvatic) rodents. Metal concentrations in soils might change the concentrations of the corresponding ions in both animals and humans through bioaccumulation along the food chain (28–30). Moreover, plague-competent host rodents were found in similar numbers in areas with and without plague reservoirs, suggesting that appropriate host species and suitable soil characteristics are necessary for the maintenance of plague reservoirs.

In addition to metal elements in soil, several previous studies suggested that soil pH supports the development of a plague reservoir, as it can impact *Y. pestis* and possibly flea larvae development (23, 24, 31). These results are consistent with the fact that soil pH greatly affects the availability and mobility of various metals in soil, including several key elements associated with plague, such as Ca and Fe. The importance of these elements may partially explain why soil pH appears in the China-fitted model. Finally, few studies have addressed the role of other soil metals (including Cd, Mg, Sb, U, Ce, Mo, and Y), which may be essential to the formation of plague reservoirs. Additional on-site observations and experiments are necessary to determine the effects of other metals on *Y. pestis* in vivo. To gain further insights, maps overlaying the spatial distribution of current plague reservoirs with selected soil attributes are provided in *SI Appendix*.

Plague reservoirs are common at high elevation, such as in the highlands of the Tian-Shan, Pamir, and Altai ranges (2,300–4,000 m above sea level) (32, 33), the Qinghai–Tibet plateau (3,000–5,000 m above sea level) (34, 35), the plateaus in the Great Lakes highland region of East Africa (1,200–2,000 m above sea level) (36), and the Central and Northern Highlands of Madagascar (higher than 800 m above sea level) (37). Elevation may not directly influence plague transmission, but it can create favorable conditions for plague hosts and their vectors to coexist (38). Likewise, temperature and precipitation may indirectly affect rodent and flea population dynamics and behavior (39–41). The relationship between host species and plague is complex (42), and the presence of competent host species does not guarantee the presence of plague (43), which we are currently investigating. However, on a continental scale, host distributions play an important role in the distribution of plague (44).

We tested a reduced model by removing the covariate nHost (the number of species of mammalian plague hosts) from the final model. It resulted in a decrease in the percentage of explained variation from 91.7 to 83.6%, demonstrating the high explanatory power of nHost (see further discussion in *SI Appendix*). The number of plague-host species is thus one of the key predictors of the presence of plague reservoirs. A lower number of plague-host species may imply the lack of plague-competent rodent-host species (*SI Appendix, Table S4*). Mahmoudi, et al. (1) recently collected taxonomic details on plague reservoir species and mapped their distribution across the world, finding that these species are predominantly of the order of *Rodentia*. The species richness of *Rodentia*, and presumably of their associated flea species, in China and the United States is significantly higher than those in Europe ( $P < 0.001$ ) (derived from *SI Appendix, Fig. S8*). This suggests that

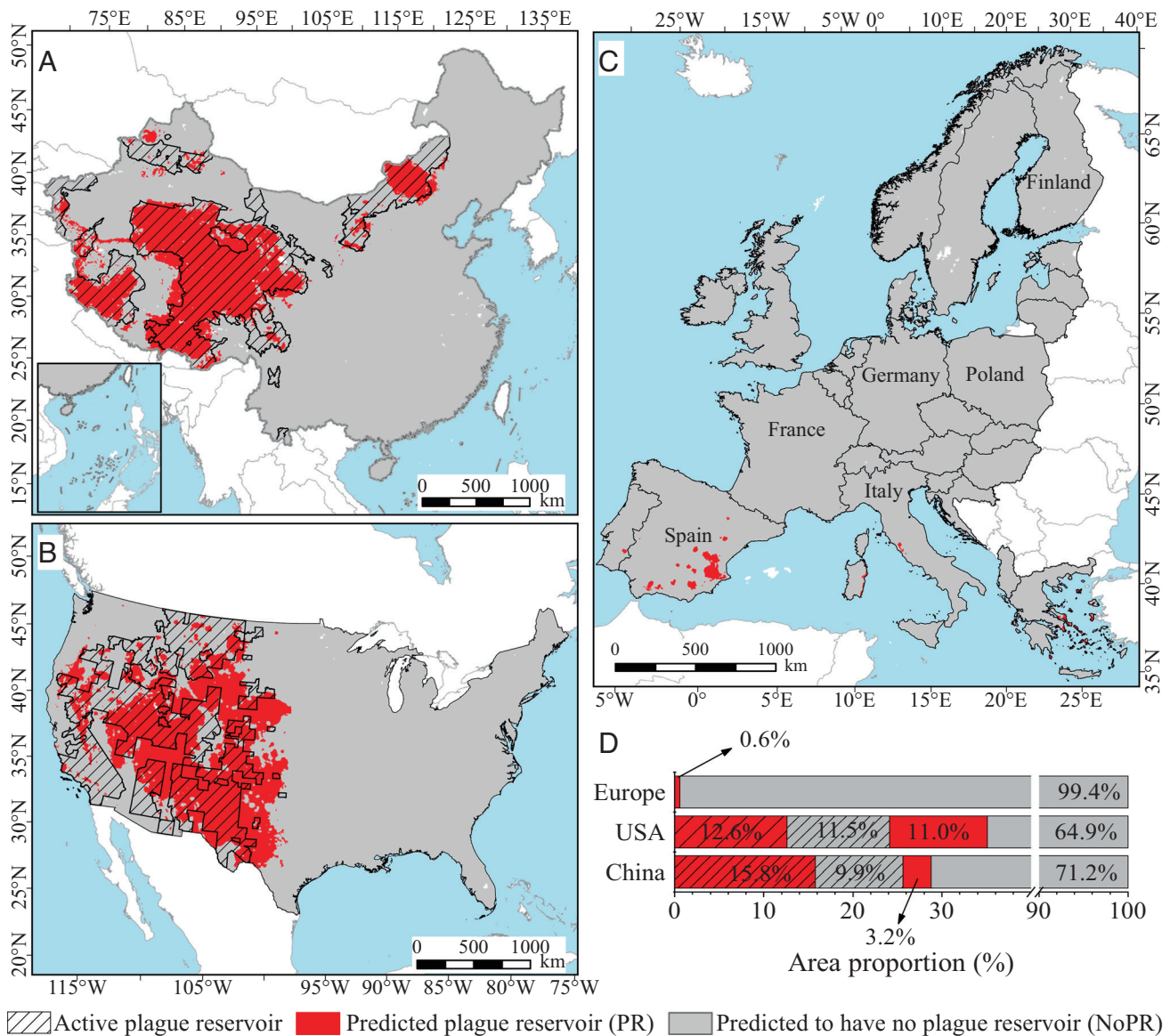
European rodent and ectoparasite communities may not be conducive to plague reservoirs to develop in Europe. In addition, the mean values of nHost (per area) are significantly different ( $P < 0.001$ , *SI Appendix, Table S3*) between China (mean value: 15), the United States (mean value: 20), and Europe (mean value: 12). Despite the robustness of our results, further studies in the field are needed to determine the mechanisms underlying the patterns documented above.

We then constructed confidence intervals and tested for each estimated probability of a plague reservoir in China, the United States, and Europe (Fig. 1). Our predictions also relied on tests of significance that compared the locations of real plague reservoirs (see *Materials and Methods*). The level of accuracy for China and the United States was 87.0% and 77.5%, respectively. The model predicted negligible plague reservoirs for Europe (see Fig. 1). The model is validated, with any variations in predictability contingent

on differing climate niches of plague between China and the United States (44) (see *Materials and Methods*). According to the same analysis, only 0.6% of Europe was predicted to have conditions commensurate with plague reservoirs; these potential locations appear in parts of Spain, Portugal, Southern France, Western/Central Italy, and Eastern Greece.

We then combined the model-based metrics of the current climate envelope of plague risk (i.e., AMT, AP, isothermality (ISO), temperature seasonality, and precipitation seasonality) with past climate proxy records to assess whether more suitable conditions for plague reservoirs in Europe might have occurred over the past 2,000 y (see *Materials and Methods*).

Fig. 2 illustrates these percentages over time for various regions of Europe. They remain low, and, in fact, lower than what is predicted for the presence of plague reservoirs in Europe today (0.6%). Nonetheless, a few peaks appear around the early 4th,



**Fig. 1.** Spatial distribution of predicted plague reservoirs (in red) and areas predicted to have no plague reservoirs (in gray) in China (A), the United States (B), and Europe (C), and results of significance tests of estimated probabilities of current plague reservoirs compared with real surveillance data (D). The data from China, the United States, and Europe were used as the training set, the validation set, and the prediction set, respectively. In China, 74.4% of the areas are nonplague reservoirs (71.2% are correctly predicted and 3.2% incorrectly predicted) and 25.7% of the areas are plague reservoirs (15.8% are correctly predicted and 9.9% incorrectly predicted). (Note: The total of 100.1% is due to rounding errors.) In the United States, 24.1% of the areas are plague reservoirs (12.6% are correctly predicted and 11.5% incorrectly predicted) and 75.9% of the areas are nonplague reservoirs (64.9% were correctly predicted and 11.0% incorrectly predicted). Note that no data is available from Taiwan or the Balkans (excluding Croatia, Albania and Greece).



11th, and 16th centuries. Yet, none of these correspond with the regions or time periods for which plague reservoirs have been hypothesized for Europe (see Table 1).

Fig. 3 depicts the regions hypothesized as having had plague reservoirs (see Table 1), corresponding to the Iberian Peninsula, late 6th – mid-8th century (Fig. 3A); South-Central Germany, c.1349–c.1400 (Fig. 3B); the Southern Alps, from c.1348 through the 16th century (Fig. 3C); Central Europe from c.1460 to c.1640 and the temperate climate zone of Europe, c.1348–c.1500 (Fig. 3D); the Balkans from 16th through 19th centuries (Fig. 3E), and East Suffolk, UK, c.1906–1918 (Fig. 3F). The figure graphically illustrates the disparity between those earlier studies and the prediction results of our analyses. Precipitation was a stronger factor than temperature in sustaining plague reservoirs (SI Appendix). The evidence shown in Fig. 3 (together with the annual prediction results over the past 2,000 y) fails to support any of the present claims for historical long-term plague reservoirs in Europe.

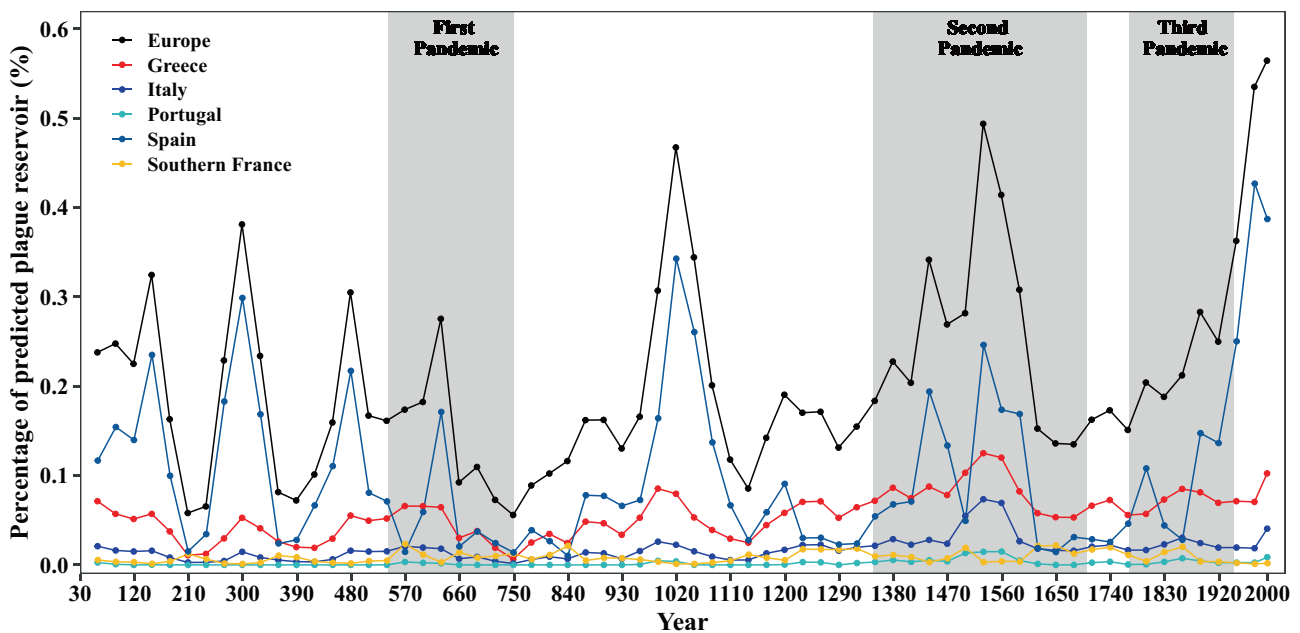
## Conclusions

Our analyses strongly suggest that local environmental factors in Western and Central Europe, including the chemical composition of the soil, altitude, and climates, did not provide favorable conditions for persistent long-term plague reservoirs maintained by wild rodents and their ectoparasites (11, 12, 45, 46). Indeed, no local European plague outbreaks have arisen over the last century (8, 47), while such outbreaks continued to occur across Asia, Africa, and the Americas (48). Our analyses, however, do not eliminate the possible existence of putative short- to medium-term reservoirs in the Mediterranean regions of Spain, Greece, Italy, and France, as well as in Central Europe in the 14th–17th centuries, and the Balkans in the 16th–19th centuries—as have been previously proposed (13–15,18). To appreciate how such putative European reservoirs might have existed for decades or even centuries, historians, archaeologists, soil scientists, and biologists need

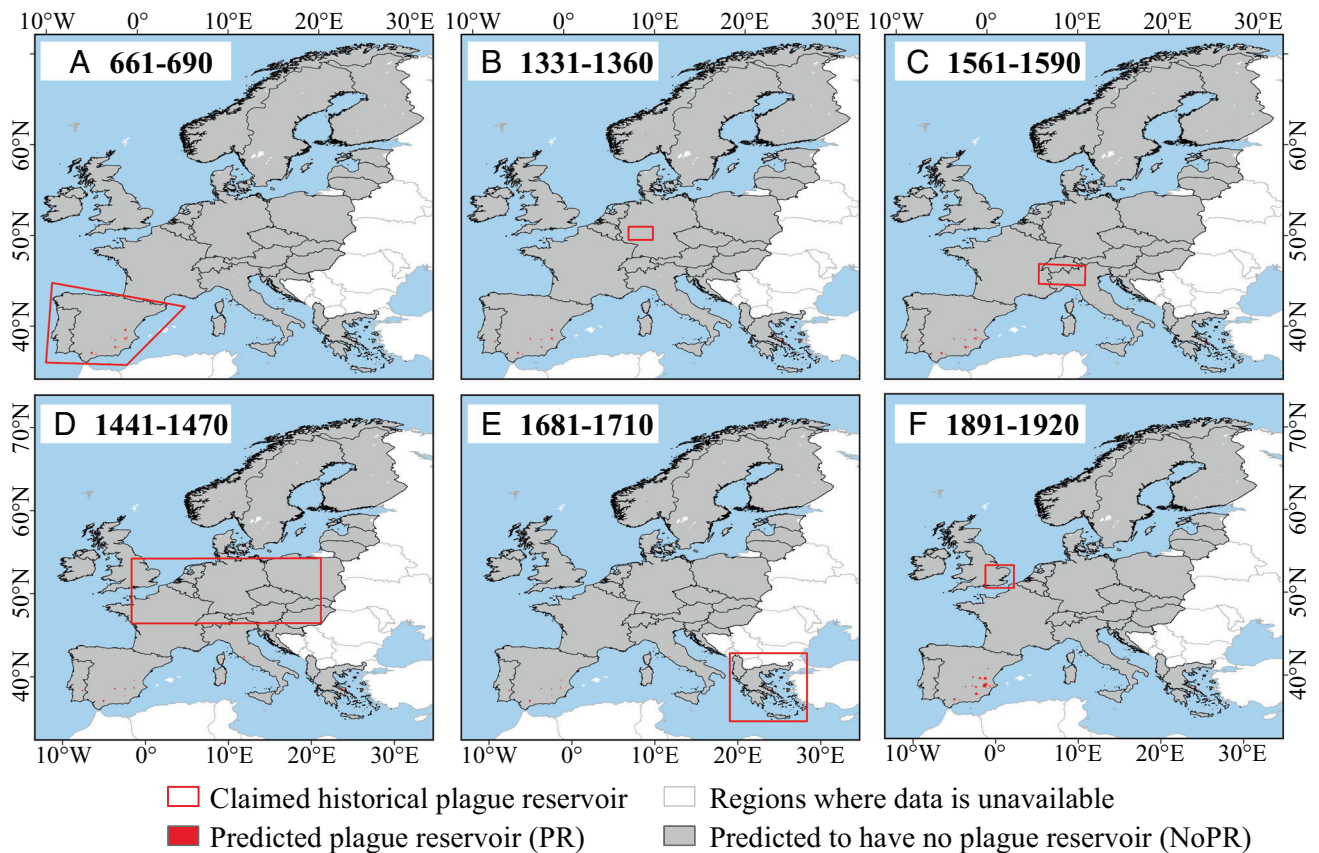
to collaborate to investigate other factors that sustained them (49). This will require exploring the complex relationships between European rodents, their fleas, and their ecologies, along with humans and their ectoparasites, as possible plague hosts and vectors after *Y. pestis* arrived in European ports from natural wildlife reservoirs in the East. Their existence may explain some sub-branches of the extant *Y. pestis* phylogeny. We argue that such medium-term historic plague reservoirs in Europe were geographically limited and not persistent, as compared with the long-lasting extensive reservoirs in large regions of China and the United States.

Such multidisciplinary collaboration would open another new field in plague studies—not the least regarding why human plague was so prevalent in Europe during the medieval and early modern period. We now know that plague initially arrived into Europe from Central Asia in the late 1340s (50). On the one hand, this initial introduction does not mean plague was necessarily reintroduced multiple times into Europe. Indeed, there is some phylogenetic evidence suggesting that it may have dwelled in native European wild animal reservoir or reservoirs, capable of maintaining the bacterium for a long period (13, 14). On the other hand, and as previously suggested (2, 8, 11, 45, 47), human movements may have provided the primary conduits for the reintroduction and persistence of plague in Europe. Human movements may also account for the current *Y. pestis* phylogeny (11, 51–53), if we consider that plague was reimported more than once (via commercial routes) from faraway reservoir(s) (11) and spread in Europe by human chains of transmission due to contacts with infected persons or pets, their ectoparasites, or their belongings, as suggested during the Third Plague Pandemic (8, 54).

Importantly, both medium-term native rodent reservoir(s) and reintroductions are not mutually inclusive scenarios. We should, indeed, not exclude the possibility that an animal reservoir may have existed: native European reservoir or reservoirs capable of maintaining the bacterium for a long period thus remain a



**Fig. 2.** Area percentage of predicted plague reservoirs of Europe over the past 2,000 y, together with the percentage of predicted plague reservoirs in each country (i.e., France, Greece, Italy, Portugal, and Spain) with different countries indicated by different colors. (Note that the percentage values are averaged values for every 30 y.) The three historical plague pandemic periods—the First, Second, and Third Plague Pandemics—are marked in gray.



**Fig. 3.** Marked in red are the regions in Europe proposed to have possessed historical plague reservoirs over the three pandemics and are compared with our predictions of environmental conditions suitable for the formation and persistence of plague reservoirs. (A) The earliest authenticated plague pandemic (the Justinianic Plague, late 6th – mid-8th century). Historians and scientists have claimed that plague reservoirs in Europe during the Second Plague Pandemic include (B) South–Central Germany (c.1349–c.1400), (C) Southern Alps (c.1348–c.1640), (D) Central Europe (c.1460–c.1640), (E) the Balkans (16th–early 19th century), and (F) England (c.1348–c.1500 and c.1906–1918) (see Table 1). Note that the years labeled in this figure are just 30-y periods which overlap with the beginning or activation of the putative reservoirs, rather than the entire periods during which the same reservoirs are hypothesised to have existed in Table 1.

possibility. This possibility may explain some molecular evidences suggesting that plague was not necessarily reintroduced multiple times into Europe. This controversy remains to be settled.

Multidisciplinary research, involving soil scientists, archaeologists, historians, and biologists, is now needed to answer the questions and controversies we have posed above and the broader implications of our findings. It is clear though from our analyses that the persistence of environmental conditions concerning soil and climate in Europe meant that persistent, long-term natural plague reservoirs over the past 2,000 y would have been even less likely in Europe historically than they are today, when it seems clear that no plague reservoir exists in that continent.

## Materials and Methods

**Plague Reservoir Data.** Across China’s mainland, 319 counties host plague reservoirs and have been classified into four categories according to the risk of plague outbreaks (*SI Appendix, section 1.1*). We devised a statistical model for China’s mainland because it has the following characteristics: (a) long-term continuous records of plague since 1772 (41), (b) stable plague reservoirs (22), and (c) an excellent monitoring system (55). We overlaid a fine grid mesh of approximately  $10 \times 10 \text{ km}^2$  (i.e., 5 arc minutes) with the boundaries of the 319 counties to construct a binary dependent variable, “*bActive*”. Currently active reservoirs are classified as belonging to categories I and II (*Fig. S2*). In particular, *bActive* is assigned the value 1, if the geometric center of one grid is within the county of an active plague reservoir; otherwise, it is 0. We also tested other methods of assignment; for instance, we assigned *bActive* to 1, if the geometric center of one grid is within a county with reservoirs in categories I, II, or III; otherwise, it is 0. That method did not affect the results of this study. Meanwhile, we have

extracted the plague reservoirs known in the United States (*Fig. S3*) by digitizing and merging two maps: one showing positive rodent samples in California (56); the other representing plague-positive counties in Western United States (57). Note that the map shown in *Fig. S1* was just used to show the general spatial distribution of plague reservoirs around the world, rather than to construct the independent variable *bActive*.

**Environmental Data.** To assess environmental impacts as comprehensively as possible, we began with a large dataset of 44 environmental variables representing soil properties, meteorological data, elevation, and spatial distribution of plague hosts, as listed below. To be consistent with the above plague reservoir data, all the environmental data were converted to a raster layer at a spatial resolution of approximately  $10 \times 10 \text{ km}^2$ . Consequently, the sample sizes were 139,064, 118,365, and 87,127 for China, the United States, and Europe, respectively.

**Soil geochemical data.** Three different datasets, all collected by authoritative organizations (*SI Appendix, Table S1*), were used to extract the soil geochemical data of 33 elements, for example Ca, Fe, Mg, and Na. Because of data availability in the soil geochemical dataset, we restricted our study to areas of China, the United States, and Europe (*SI Appendix, Fig. S2*). The soil geochemical data for China were obtained by digitizing the maps of background values of soil elements at 1:25,000,000 scale (58). Then, the digitized shapefile was converted to raster data based on the maximum area method. For the United States (59) and Europe (60), the original sampled data were downloaded and interpolated through the inverse distance weighted method in ArcGIS 10.3 (see *SI Appendix, sections 1.2.1 and 1.2.3* for more detailed descriptions of the processing of soil geochemical datasets). Note that we have such comprehensive soil geochemical data only for China’s mainland (i.e., no data for Taiwan), the continental United States (i.e., no data for Alaska, Hawaii, or Puerto Rico), and some European countries (i.e., no

data for Eastern Europe). To the best of our knowledge and up to the time of our study, no such data were available for the rest of the world. Regarding our use of the terms China, the United States, and Europe in this study, it should be noted that we are not referring to their geographical entireties.

**Additional soil attributes.** Other soil attributes, such as cation exchange capacity (CEC), clay content (clay, 0–2  $\mu\text{m}$ ), pH value, and soil organic carbon content (SOC), may also affect the development and persistence of plague reservoirs. To be consistent with the sampling depth (*SI Appendix, Table S1*) of the above geochemical data, the data for CEC, clay, pH, and SOC from that depth were extracted from SoilGrids (61), a widely used global gridded soil dataset. Data with a spatial resolution of 0.1 arc degree (approximately 10 km) were selected and downloaded. They were then aligned by the “nearest neighbor” method to spatially match the plague reservoir data. See *SI Appendix, Table S3* and *SI Appendix* for a more detailed description of the data for CEC, clay, pH, and SOC.

**Bioclimatic variables.** Climatic conditions have long been regarded as one of the most important factors in ecological studies. We extracted data for five variables from WorldClim 2.0 (62) with a spatial resolution of 5 arc minutes: AP, AMT, ISO, temperature seasonality, and precipitation seasonality. Only three raw variables in WorldClim 2.0 are used to extract a total of 19 variables: average minimum temperature, average maximum temperature, and total precipitation for each month. Thus, they are highly correlated with each other. The five variables we selected are the ones which best represent annual trends and seasonality for temperature and precipitation and best describe the principal components of all the WorldClim variables. Adding more WorldClim variables will not improve the predictability of the model within the framework of the sufficient dimension reduction method, such as MAVE, which assumes that the directions of the reduced central subspace are orthogonal to each other (see *SI Appendix, Table S3* and *SI Appendix* for further detail).

**Elevation.** Global Multi-resolution Terrain Elevation Data (GMTED2010) (63)—a suitable data source for information on the elevation for working at the continental scale—was obtained at a spatial resolution of 30 arc second (approximately 1 km) and resampled by the bilinear method to spatially match the plague reservoir data. (A more detailed description of elevation can be found in *SI Appendix*.)

**Spatial distribution of plague host.** A list of plague hosts was kindly provided by Mahmoudi et al. (1) The list was compared with species in the IUCN Red List of Threatened Species (64), resulting in a total of 339 plague hosts. Based on the spatial distribution data of these identified hosts, the number of species of mammalian hosts (denoted as nHost) at every fine grid node was counted; the fine grids were the same as those of the plague reservoir data. (A description of the nHost data obtained can be found in *SI Appendix*.)

**Modeling Approach.** A sufficient dimension reduction method, MAVE (20) (conditional MAVE), is used to establish the quantitative relationships between the spatial distribution of active plague reservoirs and potential (and yet to be identified) environmental conditions. A regression-type model for dimension reduction can be written as:

$$y = g(B_0^T X) + \epsilon,$$

where  $g$  is an unknown smooth (typically nonlinear) link function,  $B_0 = (\beta_1, \dots, \beta_D)$  is a  $p \times D$  orthogonal matrix ( $B_0^T B_0 = I_{D \times D}$ ) with  $D < p$  and  $E(\epsilon | X) = 0$  almost surely. When it holds, the projection of the  $p$ -dimensional covariates  $X$  onto the  $D$ -dimensional subspace  $B_0^T X$  captures all the information that is provided by  $X$  on  $y$ . The  $D$ -dimensional subspace  $B_0^T X$ , which is also called the effective dimension reduction (EDR) space, is what we are going to find, i.e., to estimate the matrix  $B_0$ .

MAVE is an adaptive estimation, based on semiparametric models, of the EDR space (20,65). Compared with other dimension reduction methods, MAVE has the advantage of enabling a faster consistency rate without undersmoothing the nonparametric link function estimator. MAVE is applicable to a wide range of models, with fewer restrictions on the distribution of covariates. Since it enables a faster rate of consistency for the parameter estimators, it is possible to estimate the dimension of the space consistently, with fewer observations. As it is a regression-type model, it captures the most relevant information contained in the covariates  $X$  for the prediction/explanation of the response variable  $y$ . We use the R package “MAVE” and “earth” for our exercise of MAVE method.

Because  $y$  is a binary variable, MAVE will return an estimate of probability value  $P(y = 1)$  directly (66). In the literature, it is recommended that transforming covariates  $X$  closer to normality is optimal (67). The crossvalidation method is used to find the best choice of dimension of EDR space. Thereafter, we derive the estimated probability of occurrence of a plague reservoir, denoted by  $\hat{y}_i$ . By assuming that the error term  $\epsilon$  in the model follows a normal distribution, the estimated  $\hat{y}_i$  follows a normal distribution, with standard deviation of the fitted residuals as its estimated standard deviation. Then the confidence interval for  $\hat{y}_i$  can be constructed and a significance test conducted (see *SI Appendix* for more detail). All analyses were carried out in the R environment (version 4.0.3).

**Selection of Explanatory Variables and Dimension of EDR Space.** The MAVE method was first applied to all the 44 covariates to conduct dimension reduction. Crossvalidation approaches are commonly used to penalize the complexity of the model and determine the dimension of EDR space (20,65). Using the crossvalidation method, we have found that the best choice of dimension of EDR space is 4. Using the regression-type model of dimension reduction with dimension 4, we conducted predictions for China, the United States, and Europe. (The prediction works well for China, but not equally well for the United States. The poorer performance for the latter is probably because some of the 44 covariates do not supply any information on the occurrence of plague reservoirs, thereby introducing noise to the modeling.) We then streamlined the number of covariates by reference to the direction estimates of MAVE and reran MAVE. As a result, 29 covariates with relatively larger absolute values of coefficients of EDR direction estimates were selected (*SI Appendix*).

We conducted dimension reduction for the 29 selected covariates. Using the crossvalidation method, EDR was achieved at dimension 5 (*SI Appendix, Table S7*). Note that  $d$  with the smallest  $CV(d)$  value is the estimated dimension, with  $d =$  dimension. As we use standardized covariates, we can interpret the  $CV(d)$  value as a percentage of unexplained variation.

In addition, we also explored the possible influence on the model prediction by incorporating new variables, such as the length of the growing season. However, the result shows that there is little benefit in terms of improving the prediction capability of the model by adding new variables (*SI Appendix, section 3.5*).

**Predictions of Plague Reservoirs and Test of Conditional Effects.** A robust and accurate model of China's plague reservoirs was established via MAVE, illustrating the relationship between the occurrence of reservoirs and explanatory variables. To validate this model, it was then used to obtain the probability of plague reservoir occurrence for the United States. In other words, the China data and the USA data were used as the training set and the validation set, respectively. We calculated the AUC value using predicted probabilities and the real surveillance data to evaluate the performance of prediction. (The AUC value signifies the area under the receiver operating characteristic curve, which is used to evaluate the relationship between sensitivity and specificity of a binary classifier with ranges between 0 and 1, with 1 being the best.) The AUC value is 0.932 for China and 0.851 for the United States, suggesting that the model performs very well. Significance tests were carried out (at the 95%, 90%, and 80% levels of significance) for each site, using the standard deviation of the fitted residuals ( $\hat{\sigma} = 0.30$ ). This evidence suggests that the Chinese model can be used for the European data. Consequently, we have observed remarkably low estimated probabilities of plague reservoirs occurring across Europe (see Fig. 1 and more in *SI Appendix*.)

The qualitative conditional effect of each explanatory variable was further investigated by fixing all elements but one and varying the single free element. The fixed elements were each set at their average values for the known active Chinese reservoirs. For the free element, we varied that continuously across the obtained range in China. Plots were generated, indicating the positive, negative, or nonlinear relationship between each covariate and the probability of occurrence of plague reservoir. The result was further verified and adjusted by partial variable selection, linear combinations of variables, and logarithmic transformations for compositional data (*SI Appendix*).

**Predictions for the Past 2,000 Y.** With the availability of climate reconstruction data, we can assess whether there might have been more conducive conditions for plague reservoirs to form and persist over the past 2,000 y. Using the Annual Eurasian JJA Temperature (68) and Annual Central European JJA scPDSI anomalies (69), we built a linear relationship between the two climate reconstruction variables and the five bioclimatic variables in the model to



adjust the value of bioclimatic variables with equivalent amount of climatic difference between the past and the present. A 30-y moving average of each climatic variable has been used to capture the climatic change for a period instead of a single year. Three regression models were established to cope with different climatic changes for Northern Europe, Central Europe, and Southern Europe. After the adjustment, we standardized the bioclimatic variables, using the same parameters as earlier (i.e., mean, standard deviation, parameter for Box-Cox transformation, etc.) in validating the use of the model. Prediction and significance tests were then conducted, following the previous procedure for the 30-y periods. We also employed 20-y and 40-y moving averages and have discussed the results (*SI Appendix*).

Historians have claimed several hypothesized plague reservoirs in Europe during the three pandemics, in the Mediterranean region, Central Germany, Southern Alps, England, and the Balkans. Predictions for Europe over the past 2,000 y are compared with these historically claimed plague reservoirs according to their periods of occurrence. This paper offers new insights into this.

**Data, Materials, and Software Availability.** The R code used for statistical analyses is listed in Code S1, and all the relevant data are listed in [Datasets S1–S7](#).

**ACKNOWLEDGMENTS.** This work is supported by the National Natural Science Foundation of China (Grant 72091514) and Vanke School of Public Health, Tsinghua University, to L.X. and C.Z., the University of Oslo and the Research Council of Norway to the Centre for Ecological and Evolutionary Synthesis (CEES) to N.C.S., the National Natural Science Foundation of China (Grant

71973077) to Y.T. and D.L., the European Research Council (ERC) MedPlag (Grant AdG #324249) to B.B., and the SustES project (Grant CZ.02.1.01/0.0/0.0/16\_019/0000797) and European Research Council (ERC) MONOSTAR (Grant AdG #882727) to U.B. Atle Mysterud is thanked for providing comments on an earlier version of the paper. Barbara Kiser is thanked for language editing of the manuscript.

Author affiliations: <sup>a</sup>Department of Biosciences, Centre for Ecological and Evolutionary Synthesis, University of Oslo, Oslo 0316, Norway; <sup>b</sup>Department of Earth System Science, Ministry of Education Key Laboratory for Earth System Modeling, Tsinghua University, Beijing 100084, China; <sup>c</sup>Department of Industrial Engineering, Center for Statistical Science, Tsinghua University, Beijing 100084, China; <sup>d</sup>Vanke School of Public Health, Tsinghua University, Beijing 100084, China; <sup>e</sup>Institute of Healthy China, Tsinghua University, Beijing 100084, China; <sup>f</sup>Department of Environmental and Prevention Sciences, University of Ferrara, Ferrara 44121, Italy; <sup>g</sup>Department of Geography, University of Cambridge, Cambridge CB2 3EN, UK; <sup>h</sup>Global Change Research Institute (CzechGlobe), Czech Academy of Sciences, Brno 603 00, Czech Republic; <sup>i</sup>Department of Geography, Faculty of Science, Masaryk University, Brno 611 37, Czech Republic; <sup>j</sup>Swiss Federal Research Institute (WSL), Birmensdorf 8903, Switzerland; <sup>k</sup>The First Institute of Endemic Disease Control and Prevention of Jilin Province, Baicheng 137000, Jilin Province, China; <sup>l</sup>State Key Laboratory of Pathogen and Biosecurity, Beijing Institute of Microbiology and Epidemiology, Beijing 100071, China; <sup>m</sup>Department of Soil and Water Sciences, China Agricultural University, Beijing 100193, China; <sup>n</sup>Forensic Soil Science Group, James Hutton Institute, Aberdeen AB15 8QH, UK; <sup>o</sup>School of Biosciences, University of Nottingham, Loughborough, Leics LE12 5RD, UK; <sup>p</sup>School of Geography, University of Nottingham, University Park, Nottingham NG7 2RD, UK; <sup>q</sup>Department of History, University of Glasgow, Glasgow G12 8QQ, UK; <sup>r</sup>Univ. Lille, CNRS, Inserm, CHU Lille, Institut Pasteur Lille, U1019 - UMR 9017 - CILL - Center for Infection and Immunity of Lille, Lille F-59000, France; <sup>s</sup>Division of History and Politics, University of Stirling, Stirling FK9 4LJ, UK; and <sup>t</sup>London School of Economics, London, WC2A 2AE, UK

1. A. Mahmoudi *et al.*, Plague reservoir species throughout the world. *Integr. Zool.* **16**, 820–833 (2021).
2. B. V. Schmid *et al.*, Climate-driven introduction of the Black Death and successive plague reintroductions into Europe. *Proc. Natl. Acad. Sci. U.S.A.* **112**, 3020–3025 (2015).
3. N. C. Stenseth *et al.*, Plague: Past, present, and future. *PLoS Med.* **5**, e3 (2008).
4. S. Haenschel *et al.*, Distinct clones of *Yersinia pestis* caused the Black Death. *PLoS Pathog.* **6**, e1001134 (2010).
5. K. I. Bos *et al.*, A draft genome of *Yersinia pestis* from victims of the Black Death. *Nature* **478**, 506–510 (2011).
6. M. Harbeck *et al.*, *Yersinia pestis* DNA from skeletal remains from the 6th century AD reveals insights into Justinianic plague. *PLoS Pathog.* **9**, e1003349 (2013).
7. D. M. Wagner *et al.*, *Yersinia pestis* and the plague of Justinian 541–543 AD: A genomic analysis. *Lancet Infect. Dis.* **14**, 319–326 (2014).
8. B. Bramanti, K. R. Dean, L. Walloe, N. C. Stenseth, The third plague pandemic in Europe. *Proc. R. Soc. B Biol. Sci.* **286**, 8 (2019).
9. L. Xu *et al.*, Historical and genomic data reveal the influencing factors on global transmission velocity of plague during the Third Pandemic. *Proc. Natl. Acad. Sci. U.S.A.* **116**, 11833–11838 (2019).
10. K. Linné Kausrud *et al.*, Climatically driven synchrony of gerbil populations allows large-scale plague outbreaks. *Proc. R. Soc. B Biol. Sci.* **274**, 1963–1969 (2007).
11. B. Bramanti, Y. R. Wu, R. F. Yang, Y. J. Cui, N. C. Stenseth, Assessing the origins of the European plagues following the Black Death: A synthesis of genomic, historical, and ecological information. *Proc. Natl. Acad. Sci. U.S.A.* **118**, e2101940118 (2021).
12. R. Barbieri, Origin, transmission, and evolution of plague over 400 y in Europe. *Proc. Natl. Acad. Sci. U.S.A.* **118**, e2114241118 (2021).
13. P. Slavín, Reply: Out of the West – and neither East, nor North, nor South. *Past & Present* **256**, 325–360 (2022).
14. P. Slavín, Out of the west: Formation of a permanent plague reservoir in South-Central Germany (1349–1356) and its implications. *Past & Present* **252**, 3–51 (2021).
15. A. G. Carmichael, "Plague persistence in western Europe: A hypothesis" in *Pandemic Disease in the Medieval World: Rethinking the Black Death*, M. H. Green, C. Symes, Eds. (ARC, Amsterdam University Press, 2014), pp. 157–192, 10.1515/9781942401018-009.
16. K. Pribyl, *Farming, Famine and Plague: The Impact of Climate in Late Medieval England* (Springer International Publishing, 2017), 10.1007/978-3-319-55953-7.
17. K. Harper, *The Fate of Rome: Climate, Disease, and the End of an Empire* (Princeton University Press, 2017), 10.1515/9781400888917.
18. D. Panzac, La peste dans l'Empire ottoman (1700–1850). *Collection Turcica* (1985).
19. J. Black, D. Black, Plague in east Suffolk 1906–1918. *J. R. Soc. Med.* **93**, 540–543 (2000).
20. Y. C. Xia, H. Tong, W. K. Li, L. X. Zhu, An adaptive estimation of dimension reduction space. *J. R. Stat. Soc. Ser. B Stat. Methodol.* **64**, 363–410 (2002).
21. M. Kosoy, D. Biggins, Plague and trace metals in natural systems. *Int. J. Environ. Res. Public Health* **19**, 9979 (2022).
22. Y. P. Liu, J. A. Tan, E. L. Shen, *The Atlas of Plague and Its Environment in the People's Republic of China* (Science Press, Beijing, 2000).
23. C. D. Pauling, D. L. Finke, D. M. Anderson, Interrelationship of soil moisture and temperature to sylvatic plague cycle among prairie dogs in the Western United States. *Integr. Zool.* **16**, 852–867 (2021).
24. C. J. Carlson, S. N. Bevins, B. V. Schmid, Plague risk in the western United States over seven decades of environmental change. *Glob. Change Biol.* **28**, 1–17 (2021).
25. M. A. Malek *et al.*, *Yersinia pestis* halotolerance illuminates plague reservoirs. *Sci Rep.* **7**, 40022 (2017).
26. R. Barbieri, G. Texier, C. Keller, M. Drancourt, Soil salinity and aridity specify plague foci in the United States of America. *Sci. Rep.* **10**, 6186 (2020).
27. V. M. Mezentsev, E. V. Rotshil'd, G. A. Medzykhovskii, A. K. Grazhdanov, The effect of trace elements on the infectious process in plague in an experiment. *Microbiol. Epidemiol. Immunobiol.* **41–45**, (2000).
28. V. R. Angelova, R. V. Ivanova, J. M. Todorov, K. I. Ivanov, Lead, cadmium, zinc, and copper bioavailability in the soil-plant-animal system in a polluted area. *Sci. World J.* **10**, 318203 (2010).
29. X. Y. Shen, Y. K. Chi, K. N. Xiong, The effect of heavy metal contamination on humans and animals in the vicinity of a zinc smelting facility. *PLoS One* **14**, e0207423 (2019).
30. X. P. Ge *et al.*, Evaluation of pasture allowance of manganese for ruminants. *Environ. Sci. Pollut. Res.* **28**, 56906–56914 (2021).
31. B. H. John *et al.*, "Selected soil properties for prediction of plague vectors and reservoirs in Mavumo area, Lushoto District, Tanzania" in *Second RUFORUM Biennial Meeting* (Entebbe, Uganda, 2010), pp. 1199–1202.
32. G. A. Eroshenko *et al.*, *Yersinia pestis* strains of ancient phylogenetic branch 0.ANT are widely spread in the high-mountain plague foci of Kyrgyzstan. *PLoS One* **12**, e0187230 (2017).
33. G. A. Eroshenko *et al.*, Natural mega-focus of *Yersinia pestis* main subspecies, antique biovar, phylogenetic line 4. ANT in Gorny Altai. *Probl. Particularly Dangerous Infect.* **2**, 49–56 (2018).
34. X. Q. Xu *et al.*, Genetic diversity and spatial-temporal distribution of *Yersinia pestis* in Qinghai Plateau, China. *Plos Neglect. Trop. Dis.* **12**, e0006579 (2018).
35. S. L. Wang, F. J. Hou, Burrow characteristics and ecological significance of Marmota himalayana in the northeastern Qinghai-Tibetan Plateau. *Ecol. Evol.* **11**, 9100–9109 (2021).
36. A. M. Winters *et al.*, Spatial risk models for human plague in the West Nile region of Uganda. *Am. J. Trop. Med. Hyg.* **80**, 1014–1022 (2009).
37. V. Andrianivoarimanana *et al.*, Understanding the persistence of plague foci in Madagascar. *Plos Neglect. Trop. Dis.* **7**, e2382 (2013).
38. A. Laudisoit *et al.*, Plague and the human flea, Tanzania. *Emerg. Infect. Dis.* **13**, 687–693 (2007).
39. S. Neerincx *et al.*, Predicting potential risk areas of human plague for the Western Usambara Mountains, Lushoto District, Tanzania. *Am. J. Trop. Med. Hyg.* **82**, 492–500 (2010).
40. N. C. Stenseth *et al.*, Plague dynamics are driven by climate variation. *Proc. Natl. Acad. Sci. U.S.A.* **103**, 13110–13115 (2006).
41. L. Xu *et al.*, Nonlinear effect of climate on plague during the third pandemic in China. *Proc. Natl. Acad. Sci. U.S.A.* **108**, 10214–10219 (2011).
42. D. J. Salkeld, M. Salathé, P. Stapp, J. H. Jones, Plague outbreaks in prairie dog populations explained by percolation thresholds of alternate host abundance. *Proc. Natl. Acad. Sci. U.S.A.* **107**, 14247 (2010).
43. S. P. Maher, C. Ellis, K. L. Gage, R. E. Ensore, A. T. Peterson, Range-wide determinants of Plague distribution in North America. *Am. J. Trop. Med. Hyg.* **83**, 736–742 (2010).
44. H. G. Fell *et al.*, Biotic factors limit the invasion of plague's pathogen (*Yersinia pestis*) in novel geographical settings. *Glob. Ecol. Biogeogr.* **1–13** (2021).
45. R. Dean Katharine *et al.*, Human ectoparasites and the spread of plague in Europe during the Second Pandemic. *Proc. Natl. Acad. Sci. U.S.A.* **115**, 1304–1309 (2018).
46. R. Barbieri, M. Drancourt, D. Raoult, Plague, camels, and lice. *Proc. Natl. Acad. Sci. U.S.A.* **116**, 7620–7621 (2019).
47. N. C. Stenseth, K. R. Dean, B. Bramanti, The end of plague in Europe. *Centauros* **1**, 23–24 (2022).
48. C. A. Nelson *et al.*, Antimicrobial treatment of human plague: A systematic review of the literature on individual cases, 1937–2019. *Clin. Infect. Dis.* **70**, S3–S10 (2020).
49. N. Varlik, "New science and old sources: Why the Ottoman experience of plague matters" in *Pandemic Disease in the Medieval World: Rethinking the Black Death*, H. G. Monica S. Carol, Eds. (ARC, Amsterdam University Press, 2015), pp. 193–228, 10.1515/9781942401018-010.
50. M. A. Spyrou *et al.*, The source of the Black Death in fourteenth-century central Eurasia. *Nature* **606**, 718–724 (2022).

51. B. Bramanti, N. C. Stenseth, L. Walløe, X. Lei, "Plague: A disease which changed the path of human civilization" in *Yersinia pestis: Retrospective and Perspective*, R. Yang, A. Anisimov, Eds. (Springer, Netherlands, Dordrecht, 2016), pp. 1–26.
52. A. Namouchi, Integrative approach using *Yersinia pestis* genomes to revisit the historical landscape of plague during the medieval period. *Proc. Natl. Acad. Sci. U.S.A.*, **115**, E11790–E11797 (2018).
53. M. Guellil *et al.*, A genomic and historical synthesis of plague in 18th century Eurasia. **117**, 28328–28335 (2020).
54. K. R. Dean, F. Krauer, B. V. Schmid, Epidemiology of a bubonic plague outbreak in Glasgow, Scotland in 1900. *R. Soc. Open Sci.*, **6**, 181695 (2019).
55. X. B. Cong, C. Ju, *Human Plague in China* (People's Medical Publishing House, Beijing, 2018).
56. A. C. Holt, D. J. Salkeld, C. L. Fritz, J. R. Tucker, P. Gong, Spatial analysis of plague in California: Niche modeling predictions of the current distribution and potential response to climate change. *Int. J. Health Geogr.* **8**, 38 (2009).
57. S. N. Bevins, J. A. Baroch, D. L. Nolte, M. Zhang, H. X. He, *Yersinia pestis*: Examining wildlife plague surveillance in China and the USA. *Integr. Zool.* **7**, 99–109 (2012).
58. National Environmental Protection Agency of the People's Republic of China, *The Atlas of Soil Environmental Background Value in People's Republic of China* (China Environmental Science Press, Beijing, 1994).
59. D. B. Smith, S. Federico, L. G. Woodruff, W. F. Cannon, K. J. Ellefsen, "Geochemical and mineralogical maps, with interpretation, for soils of the conterminous United States: U.S" in *Geological Survey Scientific Investigations Report 2017–5118*, (Science Publishing Network, Denver, 2019).
60. R. Salminen, *Geochemical atlas of Europe: Background information, methodology and maps, Geochemical Atlas of Europe* (Geological Survey of Finland, Espoo, 2005).
61. T. Hengl *et al.*, SoilGrids250m: Global gridded soil information based on machine learning. *PLoS One* **12**, e0169748 (2017).
62. S. E. Fick, R. J. Hijmans, WorldClim 2: New 1-km spatial resolution climate surfaces for global land areas. *Int. J. Climatol.* **37**, 4302–4315 (2017).
63. J. J. Danielson, D. B. Gesch, Global multi-resolution terrain elevation data 2010 (GMTED2010). *U.S. Geological Survey Open-File Report 2011–1073* (2011).
64. IUCN, The IUCN Red List of Threatened Species. <https://www.iucnredlist.org>.
65. B. Li, *Sufficient dimension reduction: Methods and applications with R* (CRC Press, Boca Raton, 2018), 10.1201/9781315119427.
66. A. Antoniadis, S. Lambert-Lacroix, F. Leblanc, Effective dimension reduction methods for tumor classification using gene expression data. *Bioinformatics* **19**, 563–570 (2003).
67. R. D. Cook, H. Lee, Dimension reduction in binary response regression. *J. Am. Stat. Assoc.* **94**, 1187–1200 (1999).
68. U. Büntgen *et al.*, Prominent role of volcanism in Common Era climate variability and human history. *Dendrochronologia* **64**, 125757 (2020).
69. U. Büntgen *et al.*, Recent European drought extremes beyond Common Era background variability. *Nat. Geosci.* **14**, 190–196 (2021).