Yersinia pestis is transmitted from fleas to rodents when the bacterium develops an extensive biofilm in the foregut of a flea, starving it into a feeding frenzy, or, alternatively, during a brief period directly after feeding on a bacteremic host. These two transmission modes are in a trade-off regulated by the amount of biofilm produced by the bacterium. Here by investigating 446 global isolated Y. pestis genomes, including 78 newly sequenced isolates sampled over 40 years from a plague focus in China, we provide evidence for strong selection pressures on the RNA polymerase ω-subunit encoding gene rpoZ. We demonstrate that rpoZ variants have an increased rate of biofilm production in vitro, and that they evolve in the ecosystem during colder and drier periods. Our results support the notion that the bacterium is constantly adapting—through extended phenotype changes in the fleas—in response to climate-driven changes in the niche.
Plague is an endemic disease in many wildlife rodents across Asia’s montane steppes, semi-arid deserts and mountain ecosystems. In our current understanding, Yersinia pestis, the causative agent of plague, persists in these natural foci through an unbroken chain of transmission, where fleas become infected when sucking blood from plague-infected rodents and transfer the disease to other rodents during subsequent blood meals. Flea-borne transmission of plague has long been known to occur by two main modes. The first is biofilm-dependent transmission, where Y. pestis forms a partial or complete biofilm in the proventriculus (foregut) of the flea over a period of days to weeks that hampers, or stops new blood meals from reaching this organ. The flea, deprived of food and fluids, begins to feed more frantically, regurgitating infected blood in the process. The second transmission mode is known as early-phase transmission. While the existence of the transmission route is well-documented, the mechanism of transmission is as of yet unknown. Early-phase transmission in fleas fed on mice blood has been shown to transmit in similar efficiency in biofilm-deficient Y. pestis strains as in biofilm-producing strains. Both modes of transmission are extensively reviewed by Hinnebusch et al.5.

The two transmission modes are in a trade-off with each other. Increased levels of biofilm formation lead to a better ability of the bacterium to maintain itself in the foregut of the flea, and thus improve its chances to persist in the flea gut for long enough to successfully transmit through blockage-induced transmission. However, these increased levels of biofilm formation decreases the efficiency of early-phase transmission. This trade-off between blockage-induced transmission and early-phase transmission occurs in the domain of normal-to-high levels of biofilm production. In the domain of normal-to-reduced levels of biofilm production, the relationship between blockage-induced transmission and early-phase transmission becomes more complex, depending on the host blood source. For mice, normal levels of biofilm formation, or even the complete absence of biofilm formation do not seem to negatively impact early-phase transmission, while they do negatively impact blockage-induced transmission. In rats and guinea pigs, low or absent levels of biofilm formation also negatively impact early-phase transmission, as some degree of biofilm formation appears to be involved in an interaction between host blood and Y. pestis that drastically boosts the efficiency of early-phase transmission. In addition, at reduced levels of biofilm production, more fleas will be at a stage of partial, rather than complete blockage of the proventriculus. As partially blocked fleas can still hydrate themselves, yet also spread plague, reduced levels of biofilm production increase the bacterium’s ability to survive in hibernating fleas, and extends the lifespan of infected fleas—both factors improving the persistence of plague in the ecosystem. Outside of the flea vector, biofilm formation may play a role in rodent hosts, possibly as an early defense against the innate immune system, and could play an as-of-yet unknown role for plague in the hypothetical soil compartment.

The rate of biofilm production is a genetically determined trait of Y. pestis, and alters the feeding behavior of the flea by blocking its proventriculus. Among bacteria, the ability to modulate the fleas’ feeding behavior appears to be a feature specific to Y. pestis. As such, this transmission trade-off is an example of variation in an extended phenotype trait (as originally introduced by Dawkins) of the bacterium expressed in the flea (its vector). According to the extended phenotype perspective, the genotype of the bacterium influences the phenotype and/or the behavior of the host (flea) species through the type of biofilm formation affecting the flea’s “feeding” behavior.

To study whether evolutionary trade-offs like the one above occur in Y. pestis in a natural setting, we investigated the temporal dynamics of Y. pestis within a wildlife plague reservoir of ground squirrels (Spermophilus undulatus) and their fleas (Citelophillus tesquorum). The Guertu plague reservoir is located in the Tien Shan mountains in north–west China. It exhibits strong seasonality and large interannual variation in temperature and precipitation, large fluctuations in the population density of ground squirrels and their fleas, as well as in the prevalence of Y. pestis (Supplementary Fig. 1, Supplementary Data 1). We hypothesized that these large fluctuations in the environment might exert sufficient selection pressure on the bacterium for such signals to be visible in a longitudinal time series of bacterial genome.

To trace the associations between evolution of Y. pestis and the ecological factors, we collected Y. pestis isolates and climate information over 40 years from a single natural plague focus. Comparing the Y. pestis genomes, we find a particularly strong selection signal in the RNA polymerase ω-subunit encoding gene rpoZ. In vitro, these rpoZ variants display an overdeveloped biofilm-producing capability, which suggests that these variants could have an altered extended phenotype in terms of flea transmission. As such, this transmission trade-off might be visible in a longitudinal time series of bacterial genome.

Results

Low genetic diversity present in natural isolates of plague.

Between 1967 and 2006, 78 Y. pestis isolates were collected from the Guertu natural plague reservoir (Fig. 1a), and their genomes were sequenced in 2013 (Supplementary Data 2). Of these isolates, 71 had a highly consistent genome size, but seven isolates had lost the putative pathogenicity island GI03 or the putative pathogenicity island pgm locus, and one lost its pMT plasmid, which we marked as the accessory genome (Supplementary Fig. 2). In the shared core genome of 4.3 million base pairs, we found a total of 54 single-nucleotide polymorphisms (Supplementary Data 3) and 76 insertions or deletions (Supplementary Data 4). A phylogenetic tree of the isolates constructed based on the 54 SNPs using a Bayesian relaxed-clock model in BEAST2 shows that between 1967 and 2006 multiple lineages of Y. pestis existed in parallel within the ecosystem, without any introductions of new strains from outside (Fig. 1b).

Positive selection in a biofilm-associated gene.

In order to assess whether there were signs of positive selection pressure within these 78 plague isolates, we screened for clusters of variations that occurred in such small regions of the circular chromosome that they had a low probability of occurring under a neutral substitution model, in which variations are assumed to be randomly distributed across the genome (Table 1). The most significant cluster (P < 0.00001 by permutation testing, see the Methods section and the code repository for more details) was a cluster of eight independent variations, including three SNPs and five indels, in the small RNA polymerase ω-subunit rpoZ gene (276 bp). As most strains experienced less than five passages before being kept as freeze-dried powder, we strongly suspect that the observed clustering of variations in the rpoZ gene is not an artifact of laboratory passages of the bacterium, but is caused by selection pressure from natural plague ecosystem. The rpoZ gene is known to alter colony morphology in both Streptomyces kasugaensis and
Mycobacterium smegmatis17,18, and deletion of the rpoZ gene lowered biofilm formation in M. smegmatis18. Such a selection pressure on the rpoZ gene in the Guertu plague ecosystem is of interest with regard to the trade-off between biofilm-induced blockage and early-phase transmission of Y. pestis.

rpoZ variants are found worldwide. Notably, the high frequency of polymorphisms in the rpoZ gene is not confined to the Guertu strains. By comparing 368 public available genome sequences of Y. pestis strains from across the globe (Supplementary Data 5), we found three additional variations in the rpoZ gene (one SNP and two small clusters) across the globe.
two indels) distributed across 35 Y. pestis strains, most of which occurred in the 0.0E phylogroup in the Former Soviet Union (FSU) countries (Supplementary Fig. 3, Supplementary Data 6). Interestingly, two FSU strains carried the exact same variation as one of the Guertu rpoZ variants (a single-nucleotide deletion at 52672 at position in CO92 genome, Supplementary Fig. 4). This may indicate a signal of convergent evolution (including selection for independent de novo mutations or selection for standing variation) and further supports that the rpoZ gene is under strong selection.

In the Guertu ecosystem, none of the rpoZ variations were found in more than one sample, and no rpoZ variations were found post 1987. In part, the uneven sampling distribution plays a role here (lower panel, Fig. 1b), with just 29 samples collected after 1987. Given the relatively low odds of finding rpoZ variants in general (8/78 samples in Guertu, 35/368 samples across sequenced Y. pestis strains), not finding any rpoZ variants in the 29 samples collected post 1987 are low, but not highly unlikely \[
\left(1 - \frac{8}{78}\right)^{29} = 0.043\]. Furthermore, those low odds depend on the assumption that rpoZ variants are equally likely to be selected for in every year, which we show in the climate analysis below not to be the case.

Biofilm production rates vary among natural isolates. Comparing the eight rpoZ variants and two rpoZ references (strains that have an rpoZ sequence identical to the CO92 reference strain) shows that the former had a significantly higher level of biofilm formation after 24 h, as measured by crystal violet staining (ANOVA test, \(F=161.48, P<0.001\)) (Fig. 1c). Furthermore, the ten tested isolates showed a large amount of variation in biofilm production rate between them, despite their small genetic differences in just 25 loci across the whole genome (Supplementary Table 1). Many of these 25 loci appear to be involved in processes known to affect biofilm formation (Supplementary Table 1).

rpoZ variants correlate with colder and drier climates. A general caveat when comparing biofilm-formation rates in vitro is that effect sizes can be different from those in vivo\(^9\). However, we do find additional support for a phenotypic effect of the rpoZ variants in the distinct ecological circumstances during which these variants were selected for. As none of the rpoZ variations was observed in more than one sample, each rpoZ variation can be described as being selected for somewhere in the period between the phylogenetically estimated date when the rpoZ variant sample branched off from the main tree and the sampling date of the variant (Fig. 1b). We statistically compare the average local climate conditions during the period in which the rpoZ variants developed with the same period calculated for the rpoZ references. We extend the comparison period for all samples by ±1.3 years to account for uncertainty in the sampling date (samples were collected between June and August), to allow for a year of trophic cascade effects prior to the branching off to play a potential effect, and to round the duration up to full months—see the Methods section and the code repository for more details. We find that the phylogenetic branches that contained the rpoZ variants existed during periods in which the average monthly temperature and average monthly precipitation levels were lower than that were typical for the rpoZ references (Fig. 2a). The statistical significance of a correlation between climate and rpoZ variants was tested for the colder and drier climate hypothesis and seven other possible climate hypothesis (colder, warmer, wetter, or drier climate, or the four possible combinations thereof). Statistical significance was estimated through permutation testing by counting how frequently the climate of a random set of eight rpoZ references (sampled without return out of the 70 rpoZ references) was more extreme than that of the rpoZ variants (more extreme here means that if we are testing for colder and drier climate, the random sample would have to be both colder and drier). We found the climate in the Guertu ecosystem to be significantly more often colder and drier (\(P\)-value < 0.0092, permutation test) during the time periods in which the phylogenetic branches existed in which rpoZ variations occurred, than was the case for the rpoZ references. The other seven climate hypotheses were borderline significant or nonsignificant, with the two most significant of these (at \(P\)-values of \(P<0.0396\) and \(P<0.0304\), permutation tests) indicating that rpoZ variants arose during climate periods that were drier, and warmer and drier, respectively. While neither of these hypotheses would remain significant after correcting for multiple testing, they may indicate that a lack of precipitation is the more important factor in the rise of rpoZ variants, and that the role of temperature is less certain.
Applying a Bonferroni correction to correct for multiple testing of the different climate hypotheses would be overly stringent here, as our eight hypotheses are correlated with each other. We therefore corrected for multiple testing by using another layer of permutation testing, in which we repeatedly marked eight randomly selected samples (without return) as our “samples of interest” (like the eight rpoZ variants in our main result), ran the earlier described permutation test for the eight climate hypotheses again. We then scored how often we would find significant correlations between these randomly generated “samples of interest” and one of the eight climate hypothesis that had a lower P-value than the P-value < 0.0092, we found when using the actual rpoZ variants as our sample of interest (Fig. 2b). The frequency with which we find that these randomly generated “samples of interest” result in P-values < 0.0092 is then our multiple test-corrected P-value. After correcting, we estimated the P-value for the climate signal of colder and drier years correlated to the rise of rpoZ variants to be P < 0.0423 (estimated by permutation testing). For more details on the permutation testing, see the Methods section and the code repository.

Discussion
How exactly colder and drier years affect the different trophic levels in the Guertu plague ecosystem is hampered by the limited amount of surveillance data available (Supplementary Fig. 1, Supplementary Data 1). From other studied plague ecosystems, located in the Prebalkhash desert of Kazakhstan, in Inner Mongolia in China and in Montana and Colorado in the USA20–22, we know that the strength of the trophic cascade relationships between climate and plague can differ between ecosystems, being evident in some, and absent in other plague ecosystems. For Colorado, no relationships with climate were found23, and for the Mongolian gerbil in Inner Mongolia, the cascade effects were complicated and the overall effect of increases in temperature and precipitation on plague prevalence not reported. For Kazakhstan and Mongolia, we see positive effects of increased precipitation, as well as positive effects of increased temperatures (up to a limit) for plague prevalence20,21,23. The studies in Kazakhstan further elucidate the relationship between climate and plague prevalence, linking it to an increase in both the flea burden per rodent20 and the overall rodent density21, which in turn facilitates the spread of plague24. Extrapolating from these findings, the relationship we report here between a colder and drier average climate in Guertu and the evolution of rpoZ variants could indeed relate to depressed rodent and flea abundances during the period in which rpoZ the selected variants persisted in the ecosystem, up to the date of their sampling.

Using a combination of sequencing, surveillance, and ecological information, we present a rare look into the evolutionary dynamics of Y. pestis within a natural plague reservoir. We find continuous and substantial genetic variation, similar to that reported in the urban reservoir of plague below Mahajanga, Madagascar25,26. The observed continuous occurrence of genetic variation in both Guertu and Mahajanga is in sharp contrast to the unresolved puzzle of the mostly clonal expansion of Y. pestis through millions of medieval Europeans during the Black Death pandemic27,28.

In the Guertu natural plague focus, eight independently evolved rpoZ variants appear to have been selected during colder and drier periods in the ecosystem. The limited ecosystem surveillance data and overlap thereof with the time frame during which the rpoZ variants likely evolved, make it difficult to test for causal links between the different trophic layers. However, based on numerous studies on the flea-borne transmission of plague, reviewed in Hinnebusch et al.3, a causal link between biofilm formation capability and the trade-off between biofilm-induced and early-phase flea-borne transmission of Y. pestis seems plausible. In this respect, the extended phenotype triggered by the full blockage of the foregut of the fleas is one of frenzied feeding and regurgitation of the bacteria, and is estimated to increase the probability of transmission of the disease to 25–50% per flea per bite, compared with the low 0–10% probability of early-phase transmission per flea during the 3-day time window of early-phase transmission29. The probability of plague transmission is ultimately dependent on multiple factors, such as flea species, temperature, and blood source30,31.

Theoretical explorations of the conditions under which blockage-induced transmission is favored over early-phase transmission are still in their early stages, but do suggest that both transmission strategies can persist simultaneously in a fluctuating environment, and become more or less dominant, due to genetic selection based on the conditions that favor their spread32. A possible hypothesis, in need of further (experimental) investigation, is therefore that the selection for increased biofilm production through variations in the rpoZ gene is caused by a trophic cascade, where cold and dry climate negatively affects rodent and flea densities, which in turn favors biofilm-induced plague transmission over early-phase transmission. Although there is strong experimental evidence for the importance of biofilm production in the flea-borne transmission of plague, we cannot exclude that the selection of the rpoZ variants in the Guertu ecosystem may be related to other factors, such as the poorly understood ability of the bacterium to survive in the soil11,12, or the role that the expression of biofilm-related proteins plays in the rodent host9,10.

The ability of Y. pestis to regulate its preferred path of flea-borne transmission through an extended phenotype linked to biofilm production is a remarkable adaptation to its local environment—essentially through an environmentally/climatically driven response. The consequence of this adaptation stretches beyond the local flea species when plague spillovers reach ectoparasites associated with humans, domestic or peri-domestic animals, and alters their vector efficiency, for better or worse.

Methods
Characteristics of the surveillance. Plague-surveillance work in the Guertu region began in 1964. Routine surveillance, including the type and distribution of host/vector, host animal density, flea index, and bacteria isolation was performed each year, and the first Y. pestis strain in Guertu was isolated in 1967. The sentinel site in Guertu was established in 1983. Plague-surveillance staff lived and worked at the sentinel site from May to October each year, when Spermidesmus undulatus and Marmota baibacina, the hosts of Y. pestis, emerge from hibernation. Once the sentinel site was established, seropositivity rate was also measured. Host density, flea index, seropositivity rate, and bacteria isolation were determined by standard protocol described in the National Scheme of Plague Surveillance released by the National Health and Family Planning Commission of the People’s Republic of China.

Plague-surveillance data are kept by the Xinjiang CDC, and some of it has been published in previous reports33–35. This study used data from 1982 to 2010 (Supplementary Fig. 1, Supplementary Data 1). Sentinel Y. pestis strains were stored and provided to us by the Chinese Medical Bacterium Center of Management and Preservation in Xining, China (Supplementary Data 2).

Host density. Quadrat sampling was used to determine ground squirrel density of Guertu plague focus since 1970s until 201739. Each quadrat had an area of 10,000 m2. Five to ten quadrats were randomly selected from the Guertu region each year, with a distance of greater than two kilometers between each quadrant. For each quadrat, the number of ground squirrels was counted 1 h after sunrise by telescope. The count lasted for 2 h, and was repeated for 2 consecutive days. The highest count recorded within the 2 days was used to calculate host density by dividing the number of ground squirrels by the quadrant area. Host density for the entire Guertu region was calculated as the average density between quadrats. The counts were performed twice each year (in May and July, respectively), and the average value of the two measurements denoted the annual host density.
**Capture of host animals.** The host animals were captured in region that excluded the quadrats using for determining the host density. The mouse traps were set near the burrows of host animals, and in every half an hour it was checked. The captured animal was executed in the field, and the body was sealed in a hermetic bag to be carried back to laboratory for counting fleas, collecting serum, and organ samples for further investigation.

**Flea index.** The bodies of captured animals in the hermetic bags were fumigated by using diethyl ether. For each captured animal, the number of fleas was counted and recorded independently. The flea index of the year was calculated as the total number of collected fleas divided by the total number of captured animals in the year.

**Bacteria isolation.** For each captured animal, samples of lymph nodes, blood, heart, liver, spleen, kidney, and lungs were divided into two portions. One portion was used to inoculate onto selective agar plates (BIN agar, a brain heart infusion agar base with the added selective agents of irgasan and cholate salts)40. The other portion was used to inoculate onto the selective plates. The cultures were observed for 5 days, and those which increased were considered to be positive. The selected colonies of *Y. pestis* were observed. Suspected colonies of *Y. pestis* were transferred onto LB plates to ensure a pure culture.

**Sequencing and assembly.** Paired-end libraries with an insert size of 500 bp were constructed for each Guertu isolate, and whole-genome sequencing was performed by Illumina HiSeq 2000 (Illumina Inc. USA) following the manufacturer’s instructions. The speed was about 90 bp, and >500 Mb raw data were generated for each genome. The paired-end sequencing reads were assembled de novo using SOAPdenovo62 (Supplementary Data 2).

**Variation detection.** SNPs were identified by aligning each assembled Guertu genome against a reference genome, CO92 (accession number: NC_003143.1) using MUMmer software.58 SNPs located in repetitive regions, or supported by less than ten high-quality paired-end reads (quality score > 20), were removed.44 In total, 53112 SNPs (28 non-synonymous substitutions and 20 non-synonymous substitutions) were identified in 78 Guertu genomes (Supplementary Data 3). We selected one Guertu genome, 42003, as a reference genome and repeated the SNP calling process, but no new SNPs were identified. The ancestral state for all SNPs was obtained by the program Tracer in BEAST2. The final effective sample sizes of all inferred parameters were above 200.
Statistics of the ecological analysis. Among all 446 Y. pestis genomes analyzed in this study, the sequence of rpoZ gene is fully identical in 403 genomes. Here we defined the major allelic type of the rpoZ as “rpoZ reference” (identical to the CO92 reference strains). The other 43 genomes that carried minor allele types of rpoZ were defined as “rpoZ variants”. As described in the article, we used permutation testing (i.e., resampling without returns, also known as re-randomization testing) to gain insight into the question whether the climate in Guettu was different (that is colder, warmer, wetter, or drier climate, or the four possible combination thereof) during the estimated time periods since the phylogenetic branches that contained the rpoZ variants split off from the main tree, compared with the climate in the time periods preceding the rpoZ references. We extended the actual climate period to compare between samples to be somewhat longer than the branch length indicated by the phylogenetic tree. We did so because the exact sampling dates are not known for the samples (we know sampling occurred in June–July–August), and we did not want to accidentally exclude months that might have been relevant for the selection for rpoZ variants. So, all samples were treated as having been sampled on the 1st of September, and thus had three additional months added to the duration indicated by the branch length to cover the sampling period. We also rounded up the number of duration of the periods to include the whole month (e.g., if the branch length indicated to consider the climate back to the 25th of June 1979, we would include the whole of June). Finally, we added a full year to the duration period considered, to allow for trophic cascade effects, where the climate could exert its selection pressure through affecting the conditions for rodent and or flea populations in a way that would take time to express themselves in changes in rodent and flea densities. On average, thus the duration of the climate considered for each sample was the branch length indicated by the phylogenetic tree (on average 4.9 years) + 3 months + 0.5 month + 1 year, which when added together is on average 6 years, 2 months, and 10 days, counting back from the 1st of September.

The date of diverge for the internal nodes (and thus of the branch lengths between the main tree and the sampling date of the rpoZ variants and the rpoZ references) in trees generated by BEAST comes with a substantial amount of uncertainty. To capture some of this uncertainty in our analysis, we used permutation testing of phylogenetic trees generated by the MCMC process of BEAST, under the same settings as were used in a later iteration of the phylogenetic analysis that used BEAST2 (i.e., statistically the same output). This log gave us 8000 potential phylogenetic trees to work (the log file held 10,000 trees collected during the MCMC, each 6000 iterations apart). We dropped the first 20%. In each of these trees, we calculated the average temperature and precipitation (on a monthly resolution) for the eight rpoZ variants combined and for eight randomly selected (without resampling) rpoZ reference samples combined, and recorded how the randomly selected set compared with the eight rpoZ variant samples in terms of precipitation, temperature, or combinations thereof. We iterated 300 times over all 8000 trees, resulting in a permutation test with a total of 2.4 million iterations. In precipitation, temperature, or combinations thereof. We iterated 300 times over all wetter, warmer, and drier, etc), compared with the corrected for multiple testing with another layer of permutation testing, where we “flaing” the plague: adaptations of Yersinia pestis to its insect vector that lead to transmission. Ann. Rev. Microbiol. 71, 215–232 (2017).

References
1. Stenseth, N. C. et al. Plague: past, present, and future. PLoS Med. 5, e3 (2008).
2. Brantley, R., Stenseth, N. C., Wallau, L. & Xu, L. In Yersinia pestis: RETrospective and Perspective (eds Yang, R. & Anisimov, A.) 1–26 (Springer, 2016).
3. Brubaker, R. R. Mutation rate to nonpigmentation in Pasteurella pestis. J. Bacteriol. 198, 1404–1406 (1969).
4. Vetter, S. M. et al. Biometrical evaluation is not required for early-phase transmission of Yersinia pestis. Microbiology 156, 2216–2225 (2010).
5. Hinnebusch, B. J., Jarrett, C. O. & Bland, D. M. “Fealing” the plague: adaptations of Yersinia pestis to its insect vector that lead to transmission. Annu. Rev. Microbiol. 71, 215–232 (2017).
6. Erickson, D. L., Jarrett, C. O., Callison, J. A., Fischer, E. R. & Hinnebusch, B. J. Loss of a biofilm-inhibiting glycosyldihydrolase during the emergence of Yersinia pestis. J. Bacteriol. 190, 8163–8170 (2018).
7. Bland, D. M., Jarrett, C. O., Bosio, C. F. & Hinnebusch, B. J. Infectious blood source alters early forget infection and regurgitative transmission of Yersinia pestis by rodent fleas. PLoS Pathog. 14, e1006859 (2018).
8. Bazanova, L. F., Nikitin, A. Y., Popkov, A. F. & Maeski, M. P. Seasonal dynamics of epizootic process of the plague agent Yersinia pestis transmission to the long-tailed suslik Citellus undulatus by the flea Cletillopses tesselatus in Tuva. Entomological Rev. 87, 685–691 (2007).
9. Vadyvaloo, V., Jarrett, C., Sturdevant, D. E., Sebban, F. & Hinnebusch, B. J. Transist through the flea vector induces a pretransmission innate immunity resistance phenotype in Yersinia pestis. PLoS Pathog. 6, e1000783 (2010).
10. Yoong, P., Cywes-Bentley, C. & Pier, G. B. Poly-N-acetylglucosamine expression by wild-type Yersinia pestis is maximal at mammalian, not flea, temperatures. MBio 3, e00217–00212 (2012).
11. Eisen, R. J. et al. Persistence of Yersinia pestis in soil under natural conditions. Environ. Microbiol. 14, 94–103 (2012).
12. Benavides-Montano, J. A. & Vadyvaloo, V. Yersinia pestis resists predation by acanthomastia castellania and exhibits prolonged intracellular survival. Appl. Environ. Microbiol. 83, e00593–e00517 (2017).
13. Darby, C. Uniquely insidious: Yersinia pestis biofilms. Trends Microbiol. 16, 158–164 (2008).
14. Dawkins, R. The Extended Phenotype: The Gene as the Unit of Selection (Freeman, 1982).
15. Bouckaert, R. et al. BEAST 2: a software platform for Bayesian evolutionary analysis. PLoS Comput. Biol. 10, e1003537 (2014).
16. Zhou, T., Enyart, P. J. & Wilke, C. O. Detecting clusters of mutations. PLoS ONE 3, e3765 (2008).
17. Kojima, I. et al. The rpoZ gene, encoding the RNA polymerase omega subunit, is required for antibiotic production and morphological differentiation in Streptomyces kasugaensis. J. Bacteriol. 184, 6417–6423 (2002).
18. Mathew, R., Mukherjee, R., Balachander, R. & Chatterji, D. Deletion of the rpoZ gene, encoding the omega subunit of RNA polymerase, results in pleiotropic surface-related phenotypes in Mycobacterium smegmatis. Microbiology 152, 1741–1750 (2006).
19. Darby, C., Ananth, S. L., Tan, L. & Hinnebusch, B. J. Identification of genlh, a Yersinia pestis gene required for flea blockage, by using a Caenorhabditis elegans biofilm system. Infect. Immun. 73, 7236–7242 (2005).
20. Stenseth, N. C. et al. Plague dynamics are driven by climate variation. Proc. Natl Acad. Sci. USA 103, 13110–13115 (2006).
21. Kaurud, K. L. et al. Climatically driven synchony of gerbil populations allows large-scale plague outbreaks. Proc. Biol. Sci. 274, 1963–1969 (2007).
22. Xu, L. et al. The trophic responses of two different rodent-vector-plague systems to climate change. Proc. Biol. Sci. 282, 20141846 (2015).
23. Collinge, S. et al. Testing the generality of a trophic-cascade model for plague. EcoHealth 2, 102–112 (2005).
24. Reijniers, J. et al. A curve of thresholds governs plague epizootics in Central Asia. Ecol. Lett. 15, 554–560 (2012).
25. Vogler, A. J. et al. A decade of plague in Mahajanga, Madagascar: insights into the global maritime spread of pandemic plague. mbio 4, e00623–00612 (2013).
30. Zhang, Y. et al. Transmission efficiency of the plague pathogen (Y. pestis) by the flea, Xenopsylla cheopis, to mice and great gerbils. Parasit. Vectors 8, 256 (2015).
31. Williams, S. K. et al. Effects of low-temperature flea maintenance on the transmission of Yersinia pestis by Oropsylla montana. Vector Borne Zoonotic Dis. 13, 468–478 (2013).
32. Gandon, S., Heitzmann, L. & Sebbane, F. To block or not to block: the adaptive manipulation of plague transmission. Evol. Lett. 3, 152–161 (2019).
33. Xu, F., Liu, C., Tian, W., Chen, J. & Zhao, F. 1995 plague surveillance report of Guerut, Xinjiang, China. Endem. Dis. Bull. 11, 102–103 (1996).
34. Maituohuti, A. et al. 1996 plague surveillance report of Guerut, Xinjiang, China. Endem. Dis. Bull. 12, 67–68 (1997).
35. Tang, J., Xu, B., Lin, J., Lei, G. & Chen, J. 1994 plague surveillance report of Guerut, Xinjiang, China. Endem. Dis. Bull. 12, 91–92 (1997).
36. Xu, F., Maituohuti, A., Sun, S., Lin, J. & Chen, J. 1997 plague surveillance report of Guerut, Xinjiang, China. Endem. Dis. Bull. 13, 62–64 (1998).
37. Maituohuti, A. et al. 1998 plague surveillance report of Guerut, Xinjiang, China. Endem. Dis. Bull. 15, 58–59 (2000).
38. Wang, X. et al. 2008 plague surveillance report of Guerut, Xinjiang, China. Endem. Dis. Bull. 24, 46–47 (2009).
39. Ji, S. Plague. People’s Medical Publishing House, 1988.
40. Ber, R. et al. Development of an improved selective agar medium for isolation of Yersinia pestis. Appl. Environ. Microbiol. 69, 5787–5792 (2003).
41. Garcia, E. et al. The genome sequence of Yersinia pestis bacteriophage phiA1122 reveals an intimate history with the coliphage T3 and T7 genomes. J. Bacteriol. 185, 5248–5262 (2003).
42. Luo, R. et al. SOAPdenovo2: an empirically improved memory-efficient short-read de novo assembler. Gigascience 1, 18 (2012).
43. Delcher, A. L., Salzberg, S. L. & Phillippy, A. M. Using MUMmer to identify similar regions in large scale sets. Curr. Protoc. Bioinform. Chapter 10, Unit 10, 13, https://doi.org/10.1002/0471250953.bi1003s00 (2003).
44. Cui, Y. et al. Historical variations in mutation rate in an epidemic pathogen, Yersinia pestis. Proc. Natl Acad. Sci. USA 110, 577–582 (2013).
45. Cui, Y. et al. Genetic variations of live attenuated plague vaccine strains (Yersinia pestis EV76 lineage) during laboratory passages in different countries. Infect. Genet. Evol. 26, 172–179 (2014).
46. Harris, R. S. Improved pairwise alignment of genomic DNA: The Pennsylvania State University (2007).
47. Li, H. & Durbin, R. Fast and accurate long-read alignment with burrows-wheeler transform. Bioinformatics 26, 589–595 (2010).
48. Hinnebusch, B. J., Perry, R. D. & Schwan, T. G. Role of the hemin storage (phiA1122 reveals an intimate history with the coliphage T3 and T7 genomes. J. Bacteriol. 185, 5248–5262 (2003).
49. Tong, Z. Z. et al. Genetic variations in the pgm locus among natural isolates of Yersinia pestis. J. Gen. Appl. Microbiol. 51, 11–19 (2005).
50. Parkhill, J. et al. Genome sequence of Yersinia pestis, the causative agent of plague. Nature 413, 523–527 (2001).
51. Vogler, A. J. et al. Mutations, mutation rates, and evolution at the hypervariable VNTR loci of Yersinia pestis. Mutat. Res. 616, 145–158 (2007).
52. Guindon, S. et al. New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. Syst. Biol. 59, 307–321 (2010).
53. Letunic, I. & Bork, P. Interactive tree of life (iTOL) v3: an online tool for the display and annotation of phylogenetic and other trees. Nucleic Acids Res. 44, W242–W245 (2016).
54. Li, Y. et al. Features of variable number of tandem repeats in Yersinia pestis and the development of a hierarchical genotyping scheme. PLoS ONE 8, e66567 (2013).

Acknowledgements
We appreciate all three generations of Xinjiang CDC employees for their >50 years of dedicated plague surveillance work. We thank Joe Hinnebusch, Amine Namouchi, Jukka Corander, and John Michael Koomer for helpful discussions. This work is supported by the National Natural Science Foundation of China (No. 31430006), National Key Research & Development Program of China (2016YFC1200100), funding from the State Key Laboratory of Pathogen and Biosecurity (No. SKLPBS1405), Beijing Advanced Innovation Program for Land Surface Science; Beijing Natural Science Foundation (JJ18025); Young Elite Scientist Sponsorship Program by CAST (YESS) (2018QYRC001); the European Research Council (ERC) under the FP7- IDEASS50 ERC Program (Grant 324249), the Research Council of Norway (FRIMEDBIO project 288551), the University of Oslo, MLS grant #152950, and the Centre for Ecological and Evolutionary Synthesis (CEES) funded through the Research Council of Norway.

Author contributions
R.Y., N.C.S., K.S.J. and Y.C. designed the study and coordinated the project. Specifically, R.Y. designed the study and the collection of the data; N.C.S. developed the theoretical framework for interpreting the data; H.C., X.D. and Y.Z. contributed strains and surveillance information for analysis; Y.C., B.V.S., C.G., S.H., K.S.J., H.T., Y.W., C.Y., J.Y. and X.Y. analyzed the data; Z.D., H.F., W.L., Z.Q., M.W. and Q.Z. performed the experiments; W.R.E., Y.S., and B.X. provided insightful comments, B.V.S., Y.C., K.S.J., N.C.S. and R.Y. wrote the paper. All authors approved the final version of the paper.

Competing interests
The authors declare no competing interests.

Additional information
Supplementary information is available for this paper at https://doi.org/10.1038/s41467-019-14099-w.

Correspondence and requests for materials should be addressed to K.S.J., Y.Z., N.C.S. and R.Y.

Peer review information Nature Communications thanks Ana Bento, Isabella Cattadori and Kenneth Gage for their contribution to the peer review of this work. Peer reviewer reports are available.

Reprints and permission information is available at http://www.nature.com/reprints

Publisher’s note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.