Evolution of immune genes is associated with the Black Death

Jennifer Klunk, Tauras P Vilgalys, Christian E Demeure, Xiaoheng Cheng, Mari Shiratori, Julien Madej, Rémi Beau, Derek Elli, Maria I Patino, Rebecca Redfern, et al.

Black Death shaped the evolution of immune genes

Jennifer Klunk[1,2]*, Tauras P. Vilgalys[3]*, Christian E. Demeure[4], Xiaoheng Cheng [5], Mari Shiratori[3], Julien Madej[4], Rémi Beau[4], Derek Elli[6], Rebecca Redfern[7], Sharon N. DeWitte[8], Julia A. Gamble[9], Jesper L. Boldsen[10], Ann Carmichael[11], Nükhet Varlik[12], Katherine Eaton[1], Jean-Christophe Grenier[13], G. Brian Golding[1], Alison Devault[2], Jean-Marie Rouillard[2,14], Vania Yotova[15], Renata Sindeaux[15], Anne Dumaine[3], Jessica F Brinkworth[16,17], Dominique Missiakas[6], Guy A. Rouleau[18], Matthias Steinrücken[5,19], Javier Pizarro-Cerdá[4], Hendrik N. Poinar[1,20,21#], Luis B. Barreiro[3,19,22,23#]

*These authors contributed equally to this work, and are presented in alphabetical order
[#]These authors jointly supervised this work

Correspondence should be addressed to: lbarreiro@uchicago.edu; poinarh@mcmaster.ca

Author Affiliations
[1]*McMaster Ancient DNA Centre, Departments of Anthropology Biology and Biochemistry, McMaster University, Hamilton, Ontario, Canada L8S4L9*
[2]*Daicel Arbor Biosciences, Ann Arbor, MI, USA*
[3]*Section of Genetic Medicine, Department of Medicine, University of Chicago, Chicago, IL, USA*
[4]*Yersinia Research Unit, Institut Pasteur, Paris, France*
[5]*Department of Ecology and Evolution, University of Chicago, Chicago, IL, USA*
[6]*Department of Microbiology, Ricketts Laboratory, University of Chicago, Lemont, IL, USA*
[7]*Centre for Human Bioarchaeology, Museum of London, London, UK, EC2Y 5HN*
[8]*Department of Anthropology, University of South Carolina, Columbia, SC, USA*
[9]*Department of Anthropology, University of Manitoba, Winnipeg, Manitoba, R3T2N2*
[10]*Department of Forensic Medicine, Unit of Anthropology (ADBOU), University of Southern Denmark, Odense S, 5260, Denmark*
[11]*History Department, Indiana University, Bloomington, IN, USA*
[12]*Department of History, Rutgers University-Newark, NJ, USA*
[13]*Montreal Heart Institute, Faculty of Medicine, Université de Montréal, Montréal, Quebec, Canada, H1T 1C7*
[14]*Department of Chemical Engineering, University of Michigan Ann Arbor, Ann Arbor, MI, USA*
[15]*Centre Hospitalier Universitaire Sainte-Justine, Montréal, Quebec, Canada, H3T 1C5*
[16]*Department of Anthropology, University of Illinois Urbana-Champaign, Urbana, IL, USA*
[17]*Carl R Woese Institute for Genomic Biology, University of Illinois at Urbana-Champaign, Urbana, IL, USA*
[18]*Montreal Neurological Institute-Hospital, McGill University, Montréal, Quebec, Canada, H3A 2B4*
[19]*Department of Human Genetics, University of Chicago, Chicago, IL, USA*
[20]*Michael G. DeGroote Institute of Infectious Disease Research, McMaster University, Hamilton, Ontario, Canada L8S4L9*
[21]*Canadian Institute for Advanced Research, Toronto, Canada*
[22]*Committee on Genetics, Genomics, and Systems Biology, University of Chicago, Chicago, IL, USA*
[23]*Committee on Immunology, University of Chicago, Chicago, IL, USA*

Abstract

Infectious diseases have been among the strongest selective pressures during human evolution. The single greatest mortality event in recorded history is the first outbreak of the Second Pandemic of Plague, now commonly referred to as the Black Death (1346-1353 CE), which was caused by the vector-borne pathogen *Yersinia pestis*[1]. This pandemic devastated Afro-Eurasia, killing up to 30-50% of the population[2]. To identify targets of selection due to the Black Death, we characterized genetic variation around immune-related genes from 321 ancient DNA samples from two European populations before, during, and after the Black Death. Immune loci are strongly enriched for highly differentiated sites, supporting the hypothesis of genetic adaptation during the Black Death. In total, we identify 245 variants that are highly differentiated within London, four of which were replicated in an independent cohort from Denmark representing our strongest candidates for positive selection. The selected allele for one of these variants is associated with a marked increase in the expression of *ERAP2* across a large array of immune cells. This variant is also associated with increased ability to control intracellular *Y. pestis* in monocyte-derived macrophages, functionally supporting an active role of this variant in immune protection against *Y. pestis*. Finally, we show that several of our top candidate variants overlap with alleles that are today associated with increased susceptibility to autoimmune diseases, providing empirical evidence for the role played by past pandemics in shaping present-day susceptibility to disease.

Main text

Infectious diseases have presented one of the strongest selective pressures in the evolution of humans and other animals[3,4,5]. Not surprisingly, many candidates for population-

specific positive selection in humans involve immune response genes, consistent with the hypothesis that exposure to novel and or re-emerging pathogens has driven adaptation[6,7]. However, it is challenging to connect signatures of natural selection with their causative pathogens unless the underlying loci are still associated with susceptibility to the same pathogen in modern populations[8,9]. Clarifying the dynamics that have shaped the human immune system is key to understanding how historical diseases contributed to disease susceptibility today.

We sought to identify signatures of natural selection in Europeans imposed by *Yersinia pestis*, the bacterium responsible for bubonic plague[1]. The first plague pandemic began with the Plague of Justinian in 541CE[10,11]. Nearly 600 years later, the Black Death spread throughout Europe, the Middle East, and Northern Africa, reducing the population by up to 30-50%[2,12]. With no recent exposure to plague, Europeans living through the Black Death represented immunologically naïve populations with little to no adaptation to *Y. pestis*. Given the high mortality rate, it seems reasonable to assume that genetic variants conferring protection against *Y. pestis* infection be under strong selection at this time. Indeed, the nearly decadal plague outbreaks over the subsequent four hundred years of the second pandemic in Europe often (but not always) had reduced mortality[12,13,14], suggesting possible genetic adaptation to *Y. pestis* although likely complicated by pathogen evolution and changing cultural practices.

Pre- and post-Black Death populations reveal candidate loci for plague-mediated selection

Genomic targets of selection imposed by *Y. pestis* during the Black Death, if present, have remained elusive[15,16]. To better identify such loci, we characterized genetic variation from

ancient DNA extracts derived from individuals who died shortly before, during, or shortly after the Black Death in London and across Denmark. This unique sampling design isolates, to the greatest extent possible, signatures due to *Y. pestis* and not from other infectious diseases or selective processes. From London, individuals stemmed from three cemeteries close to one another, tightly dated by radiocarbon, stratigraphy and historical records to before, during and after the Black Death (**Fig 1**, **Table S1**, Supplementary Methods). From Denmark, individuals stemmed from five localities, geographically spread across the country, which were dated via archaeological means (i.e. burial arm positions), stratigraphy and historical records. We grouped all individuals into those that lived pre-Black Death (London: ~1000-1250 CE, Denmark: ~850 to ~1350 CE) and post-Black Death (London: 1350-1539 CE, Denmark: ~1350 to ~1800 CE). Within London, we also included individuals buried in the plague cemetery, East Smithfield, all of whom died during a two year window of the Black Death between 1348 and 1349[20]. Analysis of the mitogenomic diversity from these same individuals identifies solely European mitochondrial haplotypes ruling out widespread population replacement from non-European sources, which could confound the detection of natural selection[21].

In total we screened 523 samples (n=325 from London; n=198 from across Denmark) for the presence of human DNA using a modified PCR assay for the single copy nuclear cMYC gene[21,22] and identified 321 with sufficient endogenous DNA content for downstream enrichment and sequencing of additional nuclear loci (Supplementary Methods). Since many of our samples were poorly preserved and had low endogenous DNA content, we used hybridization capture to enrich for and sequence 356 immune-related genes, 496 GWAS loci previously associated with immune disorders, and 250 neutral regions (1.5kb each) based on their location >200kb from any

known gene (**Table S2**; Supplementary Methods). The targeted immune genes were manually curated based on their role in immune related processes, and include innate immune receptors, key immune transcriptions factors, cytokines and chemokines, and other effector molecules (**Table S3**). To ensure that deamination and other forms of ancient DNA damage did not lead to spurious genotype calls, we trimmed 4bp from the start and end of each sequencing read (**Fig S1**) and excluded all singleton variants (n=106,757). Our final dataset contained 33,110 biallelic, non-singleton variants within the targeted regions (2,669 near GWAS loci, 19,972 in immune genes, and 10,469 in neutral regions), with a mean coverage of 4.6x reads per site per individual (see **Table S1** for individual coverage). We further filtered our results by excluding samples with missing genotype calls at more than 50% of those sites (retaining n=206 individuals) and retaining variants with genotype calls for at least 10 individuals per time period and population. Using genotype likelihoods, we then calculated the minor allele frequency (MAF) per population at each time point, retaining only sites with a mean MAF (averaged across London and Denmark) greater than 5% (n= 22,868 sites) since our power to detect selection for variants below 5% is very low (**Fig S2**).

To detect alleles that may have conferred protection from, or susceptibility to, *Y. pestis*, we searched within candidate regions (immune genes and GWAS loci) for variants exhibiting unexpectedly large changes in allele frequency between pre- and post-Black Death. Specifically, we identified alleles for which the degree of differentiation (F_{ST}) was larger than expected by chance, when compared to the same neutral genetic regions in the same population. We used the larger sample set from London as our discovery cohort. Burials in London were also more precisely dated and better geographically controlled than those from Denmark, improving our

relative ability to detect selection in the cohort from London (Supplementary Methods). We found an enrichment of highly differentiated variants, for all frequency bins with MAF > 10%, relative to a null expectation established using our neutral loci (**Fig 2A**). Across these variants, differentiation at immune loci exceeded the 99th percentile of neutral variants at 2.4x the rate expected by chance (binomial test $p = 7.89 \times 10^{-12}$). For variants with a MAF>30%, this enrichment was even more pronounced (3.9x the rate expected by chance; binomial test $p = 1.16 \times 10^{-14}$), likely due to increased power (**Fig S2**). Simulations show that differences in recombination rate and background selection between neutral and candidate loci are insufficient to explain the observed enrichments (**Fig S3**)**.** To further validate the signatures of selection we observed among immune loci from our London sample, we performed the same analyses using the allele frequencies estimated from our Danish cohort. These samples were also enriched for highly differentiated sites relative to the expectation from neutral loci (1.6x the rate expected by chance, binomial test $p = 9.21 \times 10^{-4}$; **Fig 2B**), further supporting evidence for plague-induced selection on immune genes.

To identify specific loci that represent the strongest candidates of selection, we applied a series of stringent criteria that leveraged the time points and populations in this unique dataset. First, we identified 245 common variants (MAF > 10%) that were highly differentiated (F_{ST} >95th percentile defined using neutral sites) when comparing pre- vs post-Black Death samples in London alone (**Table S4**). Next, we reasoned that variants conferring increased susceptibility to, or protection from, *Y. pestis* should show opposing frequency patterns before, during and after the Black Death. Specifically, variants associated with susceptibility should increase in frequency among people who died during the Black Death and should decrease in frequency

among individuals sampled post-Black Death (i.e., the survivors and/or descendants of the survivors). Conversely, protective variants should show an inverse pattern. Using this reasoning, we narrowed down our list of putatively selected loci from 245 to 35 (**Table S4**). Finally, we asked if these loci were also highly differentiated before and after the Black Death in our Danish replication cohort (i.e., among the top 10% most highly differentiated sites, and in the same direction as seen in London). Four loci met our criteria, and therefore represented our strongest candidates for selection (**Fig 2C-G**). We calculated the selection coefficient (s) for each of these variants using a Hidden Markov Model (HMM) framework (based on [24], Supplementary Methods). Statistical support for non-neutral evolution ($s \neq 0$) among our four candidate loci is strong when compared to that of neutral loci ($p < 0.001$ for each locus; **Table S5**). Despite the large confidence intervals – an inherent limitation when trying to estimate s over a few generations – the absolute values for the point estimate of s range from 0.26 to 0.4, which are among the strongest selective coefficients yet reported in humans (**Fig S4**, **Table S5**).

Genes near candidate variants are differentially regulated in response to plague

None of our top candidate variants overlap (nor is in strong linkage disequilibrium with) coding variants. Thus, it is likely that the selective advantage stems from their impact on gene expression levels, particularly on immune cell types that participate in the host response to *Y. pestis* infection. Macrophages in particular are recruited to sites of infection (i.e. flea bites) where they interact with bacteria and have been shown to contribute to resistance to plague [25,26,27]. Macrophages will phagocytize *Y. pestis*, but some bacteria survive and spread to the lymph node where they replicate uncontrollably[28,29]. To test whether our variants or nearby genes are involved in the transcriptional response to *Y. pestis*, we incubated monocyte-derived

184 macrophages (MDMs) from 33 individuals with heat-killed *Y. pestis* and compared their
185 expression profiles to unstimulated control samples using RNA sequencing (Supplementary
186 Methods). As expected, macrophages responded robustly to *Y. pestis*, with principal component
187 (PC) 1 of the gene expression data explaining 56% of the total variance in gene expression levels
188 (**Fig S5**).
189
190 Seven genes within 100kb of our four candidate loci were expressed in this dataset: locus
191 1 (rs2549794): *ERAP1*, *ERAP2*, *LNPEP*; locus 3 (rs11571319): *CTLA4, ICOS;* and locus 4
192 (rs17473484): *TICAM2, TMED7*. *NFATC1*, the only gene nearby locus 2 (rs1052025), was not
193 expressed in this dataset. With the exception of *LNPEP,* all of these genes are differentially
194 expressed in response to *Y. pestis* stimulation (**Fig 3A**), supporting their putative role in the host
195 response. Macrophages from an additional panel of 8 individuals infected with live and fully
196 virulent *Y. pestis* showed similar directional changes in gene expression to those observed in
197 response to heat-killed bacteria; both genome-wide ($r = 0.88$, $p \approx 0$, **Fig S6A**) as well as at the
198 level of the genes nearby our candidate loci, with the exception of *ERAP1* that is up-regulated in
199 response to live bacteria but down-regulated in response to heat-killed bacteria (**Fig S6B**). Next,
200 we asked if changes in gene expression were specific to *Y. pestis* or shared with other infectious
201 agents. To do so, we analyzed gene expression data of macrophages infected with live *Listeria*
202 *monocytogenes* (a Gram-positive bacterium) and *Salmonella typhimurium* (a Gram-negative
203 bacterium)[30], as well as monocytes activated with bacterial and viral ligands targeting the Toll-
204 like receptor (TLR) pathways (TLR1/2, TLR4, and TLR7/8) and live influenza virus[31]. These
205 data show that all genes within our candidate loci (with the exception of *CTLA4*) respond to
206 other pathogenic agents but that the direction of change differs depending on the stimuli. For

example, *ERAP2* is downregulated in response to all live bacteria or bacterial stimuli, including *Y. pestis*, but is up-regulated in response to viral agents (**Fig S7**).

Next, we asked if different genotypes at candidate loci had any impact upon the expression levels of these genes (**Fig 3B-C**). We identified an association between rs17473484 genotypes and *TICAM2* expression where the protective allele was associated with higher expression of the gene in the unstimulated condition (**Fig 3B**; $p = 2.5 \times 10^{-6}$), although this genetic effect was smaller following *Y. pestis* stimulation ($p = 0.24$). This effect is intriguing as *TICAM2* encodes an adaptor protein for Toll-Like Receptor 4 (*TLR4*). *In vivo*, *TLR4* detects *Y. pestis* via the recognition of lipopolysaccharides (LPS) on the bacterial outer membrane[32]. *Y. pestis* attempts to circumvent this detection by deacylating surface LPS, reducing the binding affinity with *TLR4*[33,34]. *TICAM2* ushers LPS-bound *TLR4* into endosomes and activates type I interferon responses[35]. It is therefore possible that increased *TICAM2* expression confers protection against *Y. pestis* by increasing sensitivity to LPS and promoting an effective immune response.

The strongest association we identified was between rs2549794 and *ERAP2* expression, where the protective allele (C) is associated with a 5-fold increase in expression of *ERAP2* (relative to the susceptibility T allele, **Fig 3C**), in both unstimulated ($p = 4.4 \times 10^{-10}$) and *Y. pestis* challenged ($p = 8.7 \times 10^{-7}$) cells. We observed similar strong associations from macrophages and monocytes infected with other pathogens (*Salmonella, Listeria, influenza*) or stimulated with TLR-activating ligands (all $p < 1 \times 10^{-10}$; **Fig S8**). Motivated by the strong regulatory impact of rs2549794, we sought to systematically evaluate the relationship between rs2549794 genotypes and the expression of *ERAP2* in a wider variety of immune cell types before and after infection

with live and fully virulent *Y. pestis*. We generated and sequenced expression data from peripheral blood mononuclear cells from 10 individuals (5 homozygous for the protective rs2549794 C allele and 5 homozygous for the susceptible T allele), using single-cell RNA sequencing. Across all immune cell types profiled – B cells, $CD4^+$ T cells, $CD8^+$ T cells, NK cells, and monocytes (**Fig 3D**) – we identified 5,570 genes where expression level is affected by infection with *Y. pestis* (314 to 4,234 genes per cell type, 10% FDR; **Table S6**). Most genes near our candidate loci are differentially expressed in response to *Y. pestis* infection, but both the magnitude and direction of such effects is cell-type specific (**Fig S9**). For example, *ERAP2* is upregulated upon stimulation in T and B cells, but downregulated in monocyte-derived macrophages, suggesting that the regulation of *ERAP2* is likely to be controlled by a set of distinct transcription factors and enhancers across cell types. As in macrophages and other datasets, the protective rs2549794 genotype is associated with increased *ERAP2* expression across cell types for both infected and non-infected cells (**Fig 3E-F**), but we found no association between rs2549794 genotype and the magnitude of the fold change responses in *ERAP2* upon Y. pestis infection. The much lower expression of *ERAP2* in individuals homozygous for the susceptible T allele likely results from strong linkage (R^2>0.8; D'=1.0) between the T allele and a nearby alternative splicing variant which leads to a premature stop codon, nonsense mediated decay, and a non-functional protein[36].

ERAP1 and *ERAP2* are aminopeptidases that work synergistically to trim peptides for presentation to $CD8^+$ T cells by MHC class I molecules[37,38]. Given their central role in antigen presentation, it is not surprising that polymorphisms in these genes, including rs2549794, have been associated with susceptibility to a variety of infectious agents[39,40]. Interestingly, *ERAP2*

deficiency leads to a significant remodeling of the repertoire of antigens that are presented by MHC to $CD8^+$ T cells[41,42,43], including MHC ligands that have high homology to peptide sequences derived from *Yersinia* species[43]. It is therefore likely that having a functional version of *ERAP2* would help promote the presentation of a more diverse array of *Yersinia*-derived antigens to $CD8^+$ T cells, which play an important role in protection against *Yersinia* infections[44,45]. Indeed, mice depleted of $CD8^+$ T cells all die within one-week post-infection with a milder *Yersinia* spp., *Y. pseudotuberculosis*, whereas all wild-type mice survive[44]. Complementing our findings, a recent study reported positive selection on several MHC class I alleles[46], further strengthening the link between $CD8^+$ T cell responses and selection by the plague.

In addition to its canonical role in antigen presentation and $CD8^+$ T cell activation, *ERAP2* is also involved in viral clearance and cytokine responses[47]. We sought to test whether *ERAP2* genotype is associated with variation in the cytokine response to *Y. pestis* infection. To do so, we infected an additional set of MDMs from 25 individuals (9 homozygous for the protective *ERAP2* haplotype, 9 heterozygous, and 7 homozygous for the deleterious haplotype) with live and virulent *Y. pestis* and measured the protein levels of 10 cytokines involved in various aspects of the immune response, at baseline and at 24 hours post-infection. No differences in cytokine levels existed at baseline (all $p > 0.05$; **Table S7**), but four cytokines showed a significant association with *ERAP2* genotype upon stimulation. Specifically, the levels of G-CSF ($p = 0.0155$), IL1β ($p = 0.00262$), and IL10 ($p = 0.00248$) significantly decreased with the number of protective C alleles, while we observed the opposite pattern for the chemokine CCL3 levels ($p = 0.0216$), which is involved in the recruitment of neutrophils upon infection

(**Fig 4A-D**; **Table S7**)[48,49]. We also assayed the ability of macrophages to control internalized *Y. pestis* replication as a function of genotype at the *ERAP2* locus. Individuals homozygous for the protective *ERAP2* allele were better able to limit *Y. pestis* replication ($p = 0.02$; **Fig 4E**), further supporting a protective role for this variant against infection.

Discussion

Our results provide strong empirical evidence that the Black Death was an important selective force that shaped genetic diversity around some immune loci. Because our unique sampling design included individuals who died before, during, and after the Black Death, we minimized the degree to which other historical events (i.e. tuberculosis[9] or famine[50,51,52]) could affect the inference of selection. To support our ancient genomic data, we confirmed that the strongest candidates for positive selection have a direct impact on the immune response to *Y. pestis* using functional data.

We identified four loci which were strongly differentiated before and after the Black Death in London and replicated in an independent sample cohort from Denmark as the strongest candidates of selection. However, given our small sample sizes and the low sequencing coverage inherent to ancient DNA studies, our replication power is limited and some of the other 245 highly differentiated loci in London were also likely targets of selection, but did not survive our conservative, multi-step filtering criteria. Increased sample sizes coupled with additional functional data will be required to further dissect the evolutionary role played by these variants in immune protection against *Y. pestis*.

ERAP2 showed the most compelling evidence for selection, both from a genetic and functional perspective, with an estimated selection coefficient of 0.4 (95% CI 0.19,0.62), suggesting that individuals homozygous for the protective allele were approximately 40% more likely to survive the Black Death than those homozygous for the deleterious variant. The selected allele is associated with increased expression of *ERAP2*, an effect that we expect to be exacerbated at the protein level because the deleterious haplotype harbors a splicing variant that leads to a premature stop codon (ref). Thus, only individuals with the protective allele have *ERAP2* aminopeptidase activity. We hypothesize that having *ERAP2* activity increases the ability of antigen presenting cells to present *Yersinia*-derived antigens to CD8+ T cells. Furthermore, we show that the protective *ERAP2* haplotype is strongly associated with the secretion of several cytokines in macrophages infected with *Y. pestis* and the ability of macrophages to limit *Y. pestis* replication *in vitro*. In general, individuals with more copies of the protective haplotype displayed a weaker cytokine response to infection but a better ability to limit bacterial growth. For example, levels of IL1β, a key proinflammatory cytokine often associated with pyroptotic cell death[53], were 3-fold lower in individuals homozygous for the protective *ERAP2* genotype when compared to individuals homozygous for the susceptible one. Therefore, subjects with the protective haplotype are both more efficient at controlling internalized bacteria and at resisting *Y. pestis*-induced cell death than subjects with the deleterious haplotype, abilities which may help reduce bystander tissue damage during infection. On the other hand, the increased secretion of CCL3 in individuals with the protective haplotype suggests that CCL3-induced neutrophil recruitment may be beneficial[48,49]. However, since these experiments have been done *in vitro* and rodent models only possess a single *ERAP* aminopeptidase that is homologous to *ERAP1*, we were unable to directly evaluate the impact of ERAP2 genotype on tissue damage, immune cell

recruitment, and survival.

More broadly, our results highlight the contribution of natural selection to present-day susceptibility towards chronic inflammatory and autoimmune disease. We show that *ERAP2* is transcriptionally responsive to stimulation with a large array of pathogens, supporting its key role in the regulation of immune responses. Therefore, selection imposed by *Y. pestis* on *ERAP2* likely affected the immune response to other pathogens or disease traits. Consistent with this hypothesis, the protective *ERAP2* variant is a known risk factor for Crohn's disease[54], and *ERAP2* variation has also been associated with other infectious diseases[39,40]. Thus, fluctuating selection based on the presence of various pathogens (including *Y. pestis*) weighed against the costs of immune disorders likely explains why this locus appears to be evolving under long-term balancing selection[36]. Likewise, another one of our top candidate loci (rs11571319 near *CTLA4*) is associated with an increased risk of rheumatoid arthritis[55] and systemic lupus erythematosus[56], with the protective allele during the Black Death conferring increased risk for autoimmune disease. To date, most of our evidence of an association between autoimmune risk alleles and adaptation to past infectious disease remains indirect, primarily because the etiological agents driving selection remain hidden. Our ancient genomic and functional analyses suggest that *Y. pestis* has been one such agent, representing, to our knowledge, the first empirical evidence connecting the selective force of past pandemics to present-day susceptibility to disease.

Figures

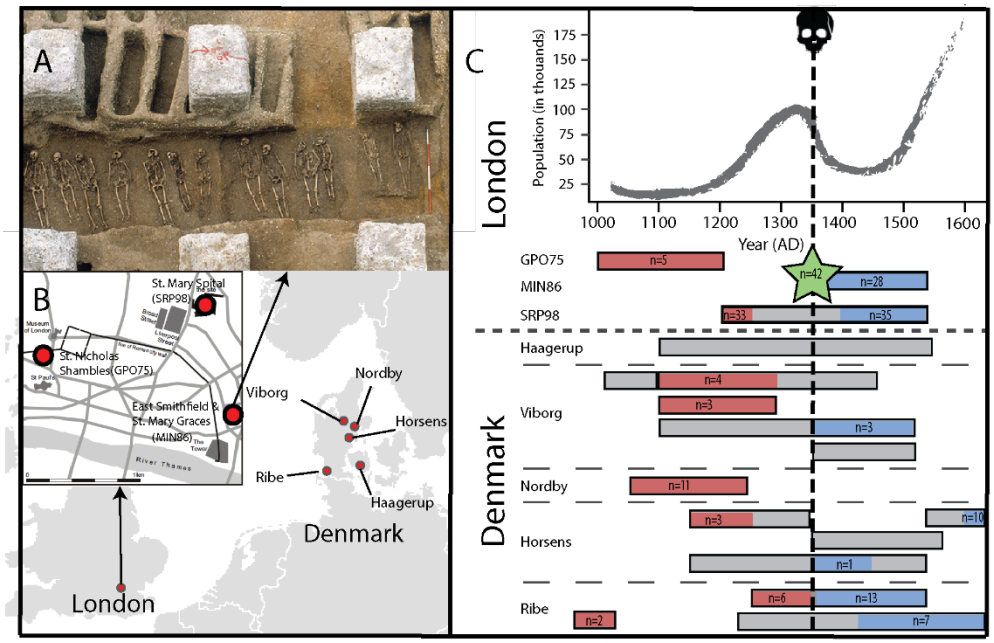

Figure 1. East Smithfield mass burial, sample locations with their date ranges along with final sample numbers used for the present study (A-C). (**A)** East Smithfield Black Death mass burial site from 1348-1349 (courtesy of the Museum of London); **(B)** Site locations for samples from London (top left, after [17] © Museum of London Archaeology) and from across Denmark (bottom right); (**C**) Top: populations size estimates for London for ~6 centuries (data taken from[15,18,19]; **Table S1**); Bottom: Site locations with associated date ranges, colored boxes indicate date range for samples, numbers in boxes indicate samples meeting all criteria for inclusion in final analyses (see main text and SI). Numbers in green star stem from East Smithfield.

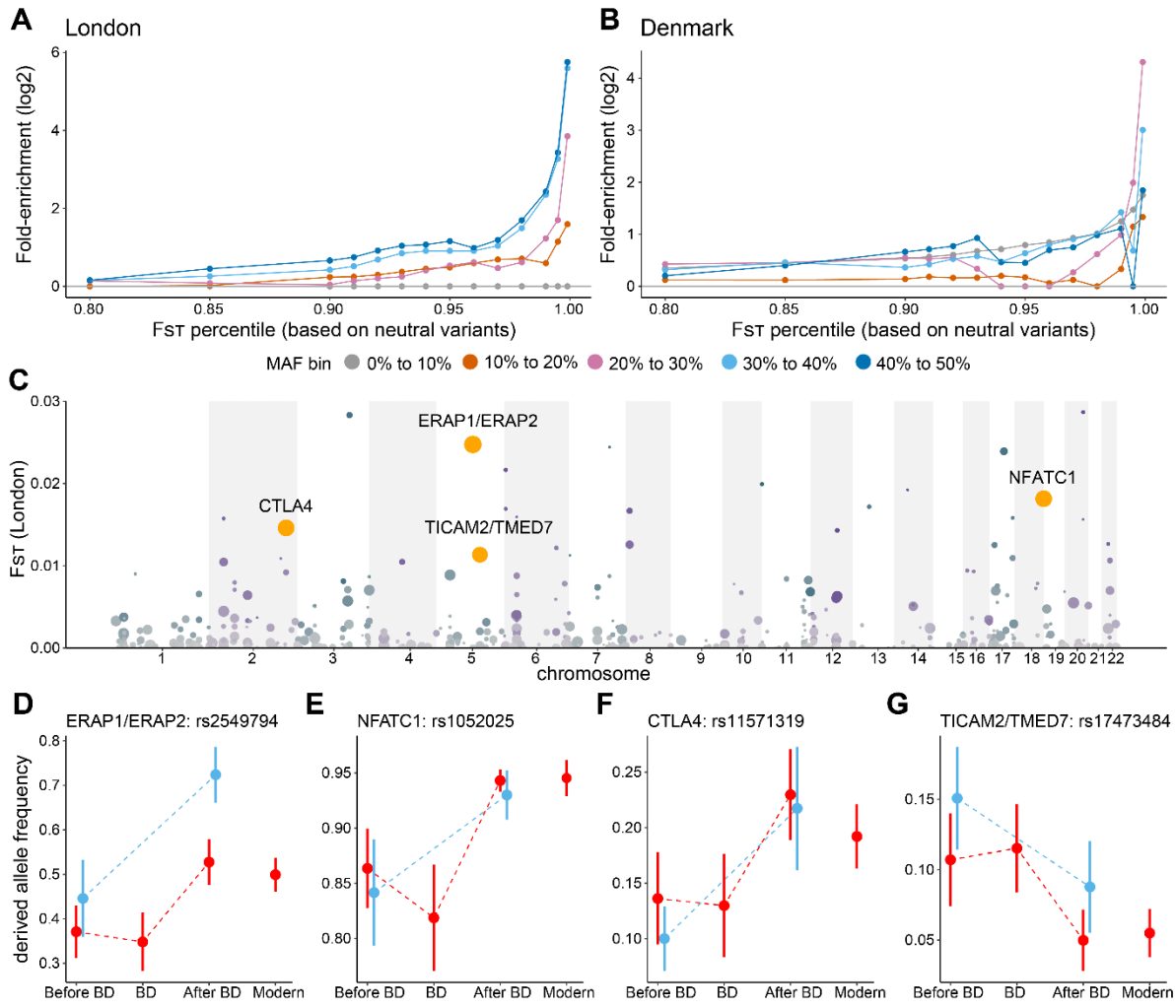

Figure 2: Positive selection at immune loci. (A-B) Enrichment of highly differentiated sites in functional regions relative to neutral regions when comparing the pre-Black Death (BD) population to the post-BD population in London **(A)** and Denmark **(B)**. **(C)** F_{ST} between London before and after the Black Death, limited to the 535 sites which show qualitative patterns consistent with natural selection (namely allele frequency changes in the same direction in both London and Denmark after the Black Death, and the opposite direction for individuals who died during the Black Death, **Table S4**). Manhattan plot showing loci with patterns indicative of positive selection. Point size and color intensity (which alternates by chromosome) represents the $-\log_{10}$ p-value comparing populations in London before and after the plague, points colored in orange represent the 4 sites which are highly differentiated in Denmark as well. **(D-G)** Patterns of allelic change over time for the four strongest candidates for positive selection. Error bars represent the standard deviation based on bootstrapping individuals from that population and each time point 10,000 times. Allele frequencies for London are shown in red and for Denmark are shown in blue. Modern allele frequencies are derived from 1000 Genomes data for Great Britain in London[23].

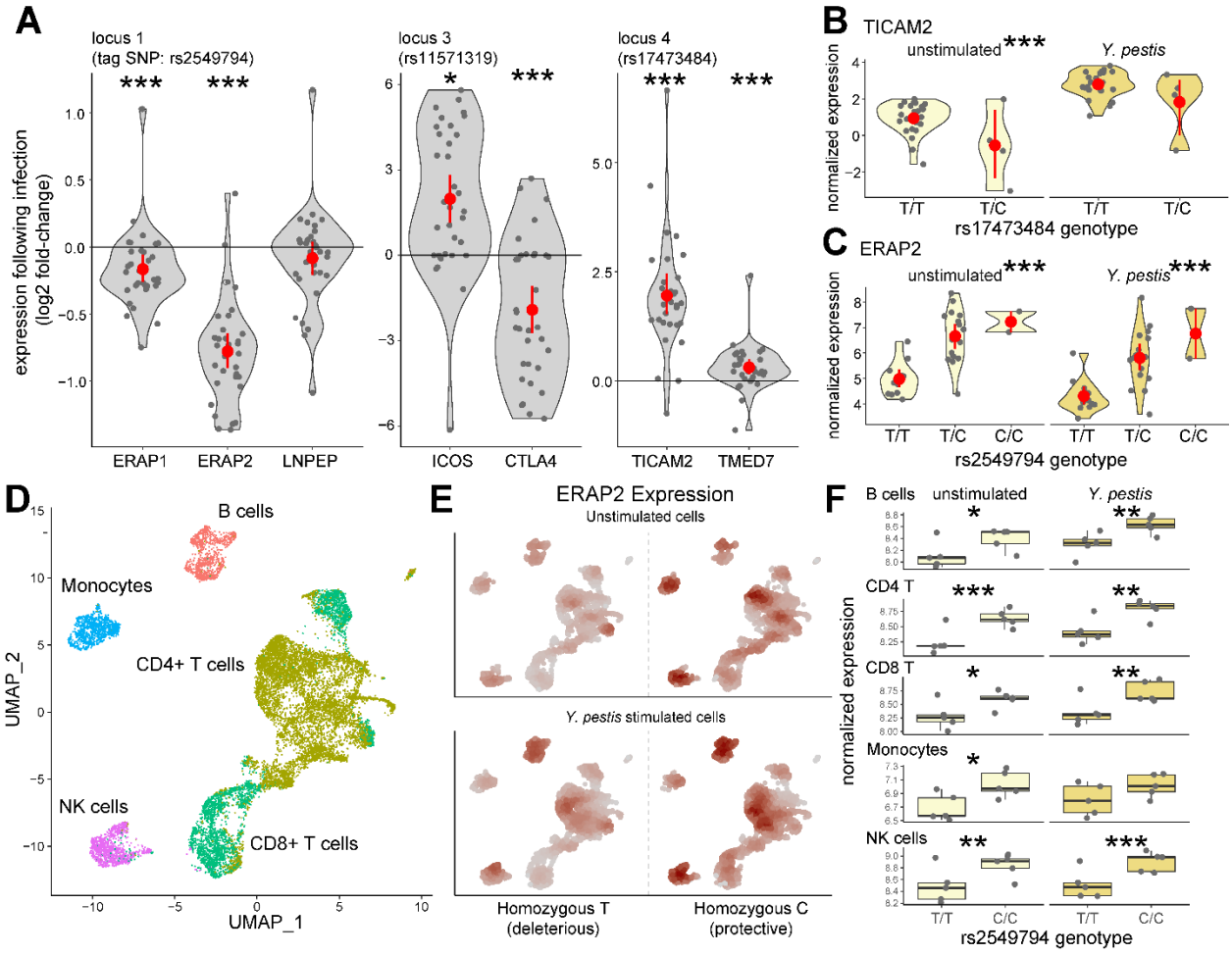

Figure 3. Positively selected loci are associated with changes in gene regulation upon *Y. pestis* stimulation. (A) Normalized log2 fold-change response to incubation of primary macrophages for 4 hours with heat killed *Y. pestis* of genes within 100kb of candidate variants that are expressed in macrophages. Gray dot corresponds to the fold-change observed for each of the 33 individuals tested; red dots and bars represent the mean ± standard deviation. With the exception of *LNPEP*, all associations are significant (10% FDR). **(B-C)** Effect of rs1747384 genotype upon *TICAM2* expression **(B)** and rs2549794 genotype upon *ERAP2* expression **(C)**, split by macrophages stimulated for 4 hours with heat killed *Y. pestis* and non-stimulated macrophages. Red dots and bars represent the mean ± standard deviation. **(D)** UMAP projection of single-cell RNA sequencing data of non-infected cells and cells infected with live *Y. pestis* for 5 hours (strain CO92, at an MOI of 1:10), after integrating samples and colored by the major immune cell types. **(E)** Comparison of *ERAP2* expression among *Y. pestis*-infected and non-infected scRNA-sequencing data in individuals with homozygous rs2549794 genotypes. Color intensity reflects the level of *ERAP2* expression, standardized for unstimulated or infected cells. **(F)** Effects of rs2549794 genotype upon *ERAP2* expression, split by infected and non-infected conditions, for each cell type. *** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$.

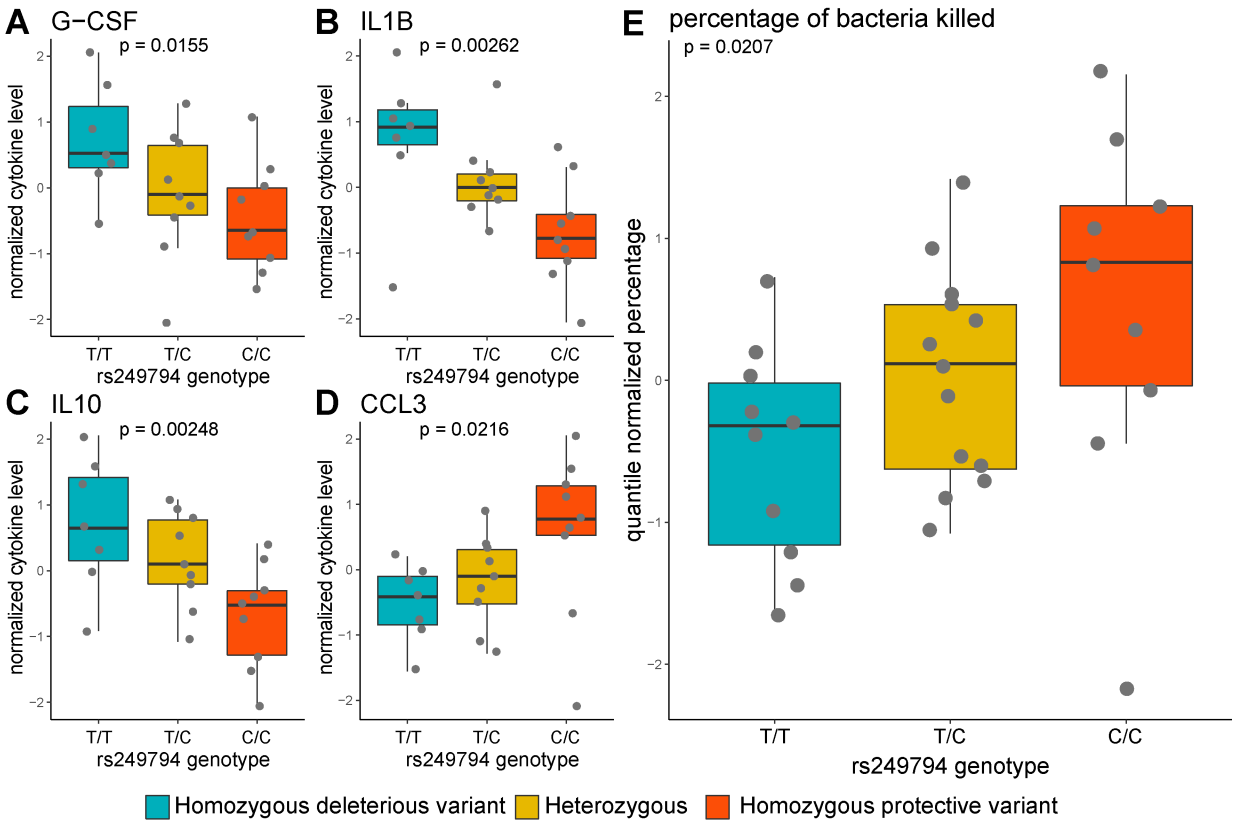

Figure 4. *ERAP2* genotype is associated with cytokine response to *Y. pestis* stimulation. (A-D) Effect of genotype upon cytokine levels for granulocyte colony-stimulating factor (G-CSF), interleukin 1β (IL1B), interleukin 10 (IL10), and C-C motif chemokine ligand 3 (CCL3). Remaining cytokines showed no significant effects and are included in **Table S7**. **(E)** Boxplots showing the quantile normalized values of the percentage of bacteria killed (y-axis) by macrophages infected for 24 hours macrophages as a function of *ERAP2* genotype (x-axis). The percentage of bacteria killed was calculated as the CFU_{2h} - CFU_{24h} / CFU_{2h}. The *p* value results from a linear model examining the association between *ERAP2* SNP genotypes (coded as the number of protective rs2549794 alleles) and the percentage of bacteria killed coded as the number of protective alleles found in each individual; 0, 1 or 2) and the percentage of bacteria killed. We reach similar conclusions to quantile normalization with either ranked CFU levels (i.e., the percentage of *Y. pestis* killed 24 hours post infection; $p = 0.0155$, $r = 0.424$) or using the Kendall-Theil Sen Siegel nonparametric linear regression ($p < 10^{-5}$) on the raw values.

References

1. Bos, K. I., et al. A draft genome of Yersinia pestis from victims of the Black Death. *Nature* **478**, 506-510 (2011).
2. Benedictow, O. J. *The Black Death, 1346-1353: The Complete History*. (Boydell Press, 2004).
3. Inhorn, M. C. & Brown, P. J. The anthropology of infectious disease. *Anuu Rev Anthrolpol* **19**, 89-117 (1990).
4. Barreiro, L. B. & Quintana-Murci, L. From evolutionary genetics to human immunology: how selection shapes host defence genes. *Nature Rev Genet* **11** (2010).
5. Fumagalli, M., et al. Signatures of environmental genetic adaptation pinpoint pathogens as the main selective pressure through human evolution. *PLoS Genet* **7**, e1002355 (2011).
6. Quintana-Murci, L. & Clark, A. G. Population genetic tools for dissecting innate immunity in humans. *Nat Rev Immun* **13**, 280 (2013).
7. Karlsson, E. K., Kwiatkowski, D. P., & Sabeti, P.C. Natural selection and infectious disease in human populations. *Nat Rev Genet* **15**, 379-393 (2014).
8. Allison, A. C. Genetic control of resistance to human malaria. *Curr Opin Immunol* **21**, 499-505. (2009).
9. Kerner, G., et al. Human ancient DNA analyses reveal the high burden of tuberculosis in Europeans over the last 2,000 years. *Am J Hum Genet* **108**, 517-524 (2021).
10. Varlık, N. New Science and Old Sources: Why the Ottoman Experience of Plague Matters. *The Medieval Globe* **1**, 9 (2014)
11. Stathakopoulos, D. C. *Famine and pestilence in the late Roman and early Byzantine empire: a systematic survey of subsistence crises and epidemics*. (Routledge, 2017).
12. Green, M. H. The Four Black Deaths. *Am Hist Rev* **125**, 1601-1631 (2021).
13. Mordechai, L., et al. The Justinianic Plague: an inconsequential pandemic? *Proc Natl Acad Sci USA* **116**, 25546-25554 (2019).
14. DeWitte, S. N. & Wood, J. W. Selectivity of Black Death mortality with respect to preexisting health. *Proc Natl Acad Sci USA* **105**, 1436-1441 (2008).
15. Earn, D. J., Ma, J., Poinar, H., Dushoff, J., & Bolker, B. M. Acceleration of plague outbreaks in the second pandemic. *Proc Natl Acad Sci USA* **117**, 27703-27711 (2020).
16. Di, D., Simon Thomas, J., Currat, M., Nunes, J. M., & Sanchez-Mazas, A. Challenging Ancient DNA Results About Putative HLA Protection or Susceptibility to *Yersinia pestis.* *Mol Bio Evol* **39**, 1537-1719 (2022).
17. Sidell, J., Thomas, C., & Bayliss, A. Validating and improving archaeological phasing at St. Mary Spital, London. *Radiocarbon*, **49,** 593-610 (2007).
18. Krylova, O., & Earn, D. J. Patterns of smallpox mortality in London, England, over three centuries. *PLoS Biol*, **18**, e3000506 (2020).
19. Greater London, Inner London & Outer London Population & Density History http://www.demographia.com/dm-lon31.htm (2001).
20. Grainger, I., Hawkins, D., Cowal, L., & Mikulski, R. *The Black Death Cemetery, East Smithfield*, *London.* (Museum of London Archaeology Service, 2008).
21. Klunk, J., et al. Genetic resiliency and the Black Death: No apparent loss of mitogenomic diversity due to the Black Death in medieval London and Denmark. *Am J Phys Anthropol* **169**, 240-252 (2019).
22. Morin, P. A., Chambers, K. E., Boesch, C., & Vigilant, L. Quantitative polymerase chain

reaction analysis of DNA from noninvasive samples for accurate microsatellite genotyping of wild chimpanzees (Pan troglodytes verus). *Mol Ecol* **10**, 1835-1844 (2001).
23. Consortium, G. P. A global reference for human genetic variation. *Nature* **526**: 68 (2015).
24. Bollback, J. P., York, T. L. & Nielsen, R. Estimation of 2Nes from temporal allele frequency data. *Genetics* **179**, 497–502 (2008).
25. Pachulec, E., et al. Enhanced macrophage M1 polarization and resistance to apoptosis enable resistance to plague. *J Infect Dis* **216**, 761-770 (2017).
26. Shannon, J. G., Bosio, C. F., & Hinnebusch, B. J. Dermal neutrophil, macrophage and dendritic cell responses to *Yersinia pestis* transmitted by fleas. *PLOS Pathog* **11**, e1004734 (2015).
27. Spinner, J. L., et al. *Yersinia pestis* survival and replication within human neutrophil phagosomes and uptake of infected neutrophils by macrophages. *J Leukocyte Biol* **95** 389-398 (2014).
28. Pujol, C. & Bliska, J. B. The ability to replicate in macrophages is conserved between Yersinia pestis and Yersinia pseudotuberculosis. *Infect Immun* **71**, 5892-5899 (2003).
29. Arifuzzaman, M., et al. Necroptosis of infiltrated macrophages drives Yersinia pestis dispersal within buboes. *JCI Insight* **3** (2018).
30. Nédélec, Y., et al. Genetic ancestry and natural selection drive population differences in immune responses to pathogens. *Cell* **167**, 657-669 (2016).
31. Quach, H., et al. Genetic adaptation and Neandertal admixture shaped the immune system of human populations. *Cell* **167**, 643-656 (2016).
32. Fitzgerald, K. A., et al. LPS-TLR4 signaling to IRF-3/7 and NF-kappaB involves the toll adapters TRAM and TRIF. *J Exp Med* **198**,1043-55. (2003).
33. Kawahara, K., Tsukano, H., Watanabe, H., Lindner, B., & Matsuura, M. Modification of the structure and activity of lipid A in *Yersinia pestis* lipopolysaccharide by growth temperature. *Infect Immun*. **70**, 4092-8. (2002).
34. Montminy, S. W., et al. Virulence factors of *Yersinia pestis* are overcome by a strong lipopolysaccharide response. *Nat Immunol*. **7**, 1066-73. (2006).
35. Kagan, J. C., et al. TRAM couples endocytosis of Toll-like receptor 4 to the induction of interferon-beta. *Nat Immunol*. **9**, 361-8. doi: (2008).
36. Andrés, A. M., et al. Balancing selection maintains a form of ERAP2 that undergoes nonsense-mediated decay and affects antigen presentation. *PLoS Genet* **6**, e1001157 (2010)
37. Saveanu, L. et al. Concerted peptide trimming by human ERAP1 and ERAP2 aminopeptidase complexes in the endoplasmic reticulum. *Nat Immun* **6**, 689-697 (2005).
38. Hattori, A. & Masafumi, T. Endoplasmic reticulum aminopeptidases: biochemistry, physiology & pathology. *J Biochem* **154**, 219-228 (2013).
39. Yao, Y., Liu, N., Zhou, Z., & Shi, L. Influence of ERAP1 and ERAP2 gene polymorphisms on disease susceptibility in different populations. *Hum Immunol* **80**, 325-334 (2019)
40. Saulle, I., Vicentini, C., Clerici, M., & Blasin, M. An Overview on ERAP Roles in Infectious Diseases. *Cells* **9**, 720 (2020).
41. Tedeschi, V. et al. The Impact of the 'Mis-Peptidome' on HLA Class I-Mediated Diseases: Contribution of ERAP1 and ERAP2 and Effects on the Immune Response. *Int J Mol Sci* **21**, 9608 (2020).

42. Lorente, E., et al. Substantial Influence of ERAP2 on the HLA-B*40:02 Peptidome: Implications for HLA-B*27-Negative Ankylosing Spondylitis*. *Mol Cell Proteomics* **18**, P2298-2309 (2019).
43. Lorente, E., et al. Modulation of Natural HLA-B*27:05 Ligandome by Ankylosing Spondylitis-associated Endoplasmic Reticulum Aminopeptidase 2 (ERAP2)*. *Mol Cell Proteomics* **19**, P994-1004 (2020).
44. Bergman, M. A., Loomis, W. P., Mecsas, J., Starnbach, M. N., & Isberg, R. R. CD8+ T Cells Restrict Yersinia pseudotuberculosis Infection: Bypass of Anti-Phagocytosis by Targeting Antigen-Presenting Cells. *PLoS Pathog* **5**, e1000573 (2009).
45. Szaba, F. M., et al. TNFα and IFNγ but Not Perforin Are Critical for CD8 T Cell-Mediated Protection against Pulmonary Yersinia pestis Infection. *PloS Pathog* **10**, e1004142 (2014).
46. Immel, A., et al. Analysis of genomic DNA from medieval plague victims suggests long-term effect of Yersinia pestis on human immunity genes. *Mol Biol Evol* **38**, 4059-4076 (2021).
47. Saulle, I., et al. ERAPs Reduce In Vitro HIV Infection by Activating Innate Immune Response. *J Immunol*. **206**, 1609-1617 (2021).
48. Lukaszewski, R. A., et al. Pathogenesis of Yersinia pestis infection in BALB/c mice: effects on host macrophages and neutrophils. *Infect Immun* **73**, 7142-7150 (2005).
49. Laws, T. R., Davey, M. S., Titball, R. W., Lukaszewski, R. Neutrophils are important in early control of lung infection by *Yersinia pestis*. *Microbes Infect*. **12**, 331-5. (2010).
50. Jordan, W. C. *The Great Famine: Northern Europe in the Early Fourteenth Century.* (Princeton University Press, 1997).
51. Hoyle, R. Britain. In G. Alfani & C. Ó Gráda (Eds.), *Famine in European History* (pp. 141-165). (Cambridge University Press, 2017).
52. DeWitte, S., & Slavin, P. Between famine and death: England on the eve of the Black Death—evidence from paleoepidemiology and manorial accounts. *J Interdiscipl Hist* **44**, 37-60. (2013).
53. Ratner, D., et al. Manipulation of Interleukin-1β and Interleukin-18 Production by *Yersinia pestis* Effectors YopJ and YopM and Redundant Impact on Virulence. *J Biol Chem*. **291**, 9894-905. (2016).
54. Di Narzo, A. F., et al. Blood and intestine eQTLs from an anti-TNF-resistant Crohn's disease cohort inform IBD genetic association loci. *Clin Transl Gastroen* **7**, e177 (2016).
55. Laufer, V. A., et al. Genetic influences on susceptibility to rheumatoid arthritis in African-Americans. *Hum Mol Genet* **28**, 858-874 (2019).
56. Wang, Y. et al. Identification of 38 novel loci for systemic lupus erythematosus and genetic heterogeneity between ancestral groups. *Nat Commun* **12**, 771 (2021).

Data Availability

Hybridization capture data from the ancient individuals have been deposited in the NCBI Sequence Read Archive (SRA) under BioProject PRJNA798381. Expression data have been deposited into the NCBI Gene Expression Omnibus (GEO) under project GSE194118. Cytokine data is available in **Table S8**.

Code Availability

Scripts for all data analyses are available at github.com/TaurVil/VilgalysKlunk_yersinia_pestis/.

Acknowledgements

We thank all members of the Barreiro lab and the Poinar lab for their constructive comments and feedback. Computational resources were provided by the University of Chicago Research Computing Center. Sequencing was performed at the Farncombe Sequencing Facility McMaster University. We thank the Cytometry and Biomarkers platform at the Institut Pasteur for support in conducting this study with a special thanks to Caroline Petitdemange for help running the Luminex assay. Finally, we thank Xinjun Zhang for assistance in simulating allele frequency changes under neutral evolution. This work was supported by grant R01-GM134376 to LBB, HP, and JP-C, a grant from the Wenner-Gren Foundation to JB (8702), and the UChicago DDRCC, Center for Interdisciplinary Study of Inflammatory Intestinal Disorders (C-IID) (NIDDK P30 DK042086). HNP was supported by an Insight Grant #20008499 from the Social Sciences and Humanities Research Council of Canada (SSHRC) and The Canadian Institute for Advanced Research under the Humans and the Microbiome program. TPV was supported by NIH F32GM140568. We also thank the University of Chicago Genomics Facility (RRID:SCR_019196), especially P Faber, for their assistance with RNA-sequencing. HP thanks Debi Poinar for continued support and manuscript suggestions and editing.

Author contributions

LBB and HP directed the study. JK designed the enrichment assays and generated ancient genomic data. TPV led all data and computational analyses, with contributions from J-CG. XC performed all analyses to estimate selection coefficients under the supervision of MSt. JFB, RS. , and VY performed challenge experiments with macrophages and heat-killed *Y. Pestis.* MS and ADu performed the infections experiments on PBMCs and generated the single cell RNA-sequencing data, with assistance from DE and DM. CED and JP-C performed and designed the infection experiments with live *Y. pestis* on macrophages and generated both cytokine and CFU data, with assistance from JM and RB. RR, SND, JG. JLB provided access to samples, archeological information, including dating, and other relevant information. AC and NV provided insights into historical context. KE and GBG provided additional sampling and bioinformatic processing and cluster maintenance. ADe and JMR provided insight on targeted enrichment and modified versions of baits used for immune enrichment. GAR provided genomic input on loci and contributed financially to the sequencing of targets. TPV, JK, HP, and LBB wrote the manuscript, with input from all authors.

Competing interest declaration

JK, ADe, and J-MR declare financial interest in Daicel Arbor Biosciences, who provided the myBaits hybridization capture kits for this work.

Additional information

Supplementary Information is available for this paper. Correspondence and requests for materials should be addressed to lbarreiro@uchicago.edu and poinarh@mcmaster.ca.

SUPPLEMENTARY INFORMATION

Black Death shaped the evolution of immune genes

Jennifer Klunk[1,2]*, Tauras P. Vilgalys[3]*, Christian E. Demeure[4], Xiaoheng Cheng [5], Mari Shiratori[3], Julien Madej[4], Rémi Beau[4], Derek Elli[6], Rebecca Redfern[7], Sharon N. DeWitte[8], Julia A. Gamble[9], Jesper L. Boldsen[10], Ann Carmichael[11], Nükhet Varlik[12], Katherine Eaton[1], Jean-Christophe Grenier[13], G. Brian Golding[1], Alison Devault[2], Jean-Marie Rouillard[2,14], Vania Yotova[15], Renata Sindeaux[15], Anne Dumaine[3], Jessica F Brinkworth[16,17], Dominique Missiakas[6], Guy A. Rouleau[18], Matthias Steinrücken[5,19], Javier Pizarro-Cerdá[4], Hendrik N. Poinar[1,20,21#], Luis B. Barreiro[3,19,22,23#]

*These authors contributed equally to this work, and are presented in alphabetical order
[#]These authors jointly supervised this work

Correspondence should be addressed to: lbarreiro@uchicago.edu; poinarh@mcmaster.ca

Author Affiliations
[1]*McMaster Ancient DNA Centre, Departments of Anthropology Biology and Biochemistry, McMaster University, Hamilton, Ontario, Canada L8S4L9*
[2]*Daicel Arbor Biosciences, Ann Arbor, MI, USA*
[3]*Section of Genetic Medicine, Department of Medicine, University of Chicago, Chicago, IL, USA*
[4]*Yersinia Research Unit, Institut Pasteur, Paris, France*
[5]*Department of Ecology and Evolution, University of Chicago, Chicago, IL, USA*
[6]*Department of Microbiology, Ricketts Laboratory, University of Chicago, Lemont, IL, USA*
[7]*Centre for Human Bioarchaeology, Museum of London, London, UK, EC2Y 5HN*
[8]*Department of Anthropology, University of South Carolina, Columbia, SC, USA*
[9]*Department of Anthropology, University of Manitoba, Winnipeg, Manitoba, R3T2N2*
[10]*Department of Forensic Medicine, Unit of Anthropology (ADBOU), University of Southern Denmark, Odense S, 5260, Denmark*
[11]*History Department, Indiana University, Bloomington, IN, USA*
[12]*Department of History, Rutgers University-Newark, NJ, USA*
[13]*Montreal Heart Institute, Faculty of Medicine, Université de Montréal, Montréal, Quebec, Canada, H1T 1C7*
[14]*Department of Chemical Engineering, University of Michigan Ann Arbor, Ann Arbor, MI, USA*
[15]*Centre Hospitalier Universitaire Sainte-Justine, Montréal, Quebec, Canada, H3T 1C5*
[16]*Department of Anthropology, University of Illinois Urbana-Champaign, Urbana, IL, USA*
[17]*Carl R Woese Institute for Genomic Biology, University of Illinois at Urbana-Champaign, Urbana, IL, USA*
[18]*Montreal Neurological Institute-Hospital, McGill University, Montréal, Quebec, Canada, H3A 2B4*
[19]*Department of Human Genetics, University of Chicago, Chicago, IL, USA*
[20]*Michael G. DeGroote Institute of Infectious Disease Research, McMaster University, Hamilton, Ontario, Canada L8S4L9*
[21]*Canadian Institute for Advanced Research, Toronto, Canada*
[22]*Committee on Genetics, Genomics, and Systems Biology, University of Chicago, Chicago, IL, USA*
[23]*Committee on Immunology, University of Chicago, Chicago, IL, USA*

TABLE OF CONTENTS

c. Table S3: hg18 positions targeted in Exon captures
d. Table S4: Highly differentiated variants
e. Table S5: Estimates of selection coefficients
f. Table S6: Changes in gene expression following *Y. pestis* stimulation
g. Table S7: Cytokine Results
h. Table S8: Cytokine Data

1. Supplementary Methods

a. Sampling locations and dates

The details of all remains used in this study, including burial locations and associated dates can be found in [21] as well as in the **Table S1**. For the purposes of our study, we split the samples from London into three time points: early or Pre-Black Death (~1000-1250), Black Death (1348-1350) and late or Post-Black Death (1350-1500) (**Table S1**). The London cemeteries are dated based on a combination of primary sources, Bayesian radiocarbon dating, and archaeological remains (primarily coins). We did the same for our Danish samples, as best we could from associated dates, which are less precise due to the lack of records associated with many of the cemeteries and the imprecision of the dating method based on arm position at burial used to date medieval burials in this region. We therefore group Danish samples into either an Early (~850-1350) or Late (1350-1800) period corresponding to before or after the Black Death period.

b. Ancient DNA sample processing, bait design, enrichment, sequencing and data processing

i. Sample processing

All sample processing was performed in the specialized cleanroom facilities of the McMaster Ancient DNA Centre (McMaster University, Hamilton, Ontario). Individuals from three cemeteries in London, England (n=325), and five localities from Denmark (n=198) were included in this study (**Fig 1**). Both reagent blanks and carrier blanks (bacteriophage lambda DNA sheared to an average of 150 bp) were processed alongside the samples at 1:20 reagent blank and 3:20 carrier blank ratio. Subsamples of 10-300 mg of long bone, tooth (root or pulp cavity), or petrous portion of the temporal bone were pulverized manually. The resulting bone or tooth fragments and powder were demineralized and digested using a previously described protocol[57,21]. The DNA was purified using either a modified phenol:chloroform:isoamyl alcohol (PCI) method or a large-volume guanidinium hydrochloride method designed to retain short DNA fragments[57,58,59]. Samples were screened for the presence of single copy nuclear DNA using an 81bp quantitative PCR assay targeting a portion of the nuclear cMYC gene[22] or a shortened 52 bp version of the assay[21]. Extracts of samples that returned positive results were prepared into double-stranded, double indexed Illumina-style sequencing libraries[60,61] for downstream enrichment.

ii. Bait design

Three sets of baits were designed and used for the enrichment of i) neutral, ii) GWAS, and iii) exon targets (**Tables S2 & S3**).

1. **Neutral.** 250 independent (non-linked) and putative neutral loci were randomly selected from a larger set of "neutral loci" defined by Gronau and colleagues[62]. Briefly, these are regions outside any known protein-coding and conserved noncoding region of the genome, and with low intralocus recombination but sufficient between regions that the genealogies are largely uncorrelated. All bait regions were initially designed to target 1,500 bp, but some of the targets required shortening to avoid non-specific hybridization. The final neutral bait set included 16,062 probes, each 80bp in length tiled every 20 bp.

2. **GWAS.** A total of 496 GWAS variants associated with immune-related disorders (**Table S2**). Initially, the targets included 50 bp on either side of the variant of interest but were redesigned to include 80 bp on either side after experiments showed better capture success. The redesigned baits outperformed the old baits in target coverage breadth and depth by a factor of 3, at approximately 4.8x lower sequencing depth. Probes for both variant alleles were designed. The final GWAS bait set included 5,800 probes, 80 bp in length tiled every 20 bp.
3. **Exons.** A total of 3,512 sequences, including all exons from 356 genes associated with immune function were included in the exon bait set. The genes were manually curated and chosen based on their known role in immune processes. 69% of all 399 genes are associated with one or more of the following GO terms ("immune function", "immune system process", "Innate Immunity", "adaptive immunity", "cytokine", "NF-kappaB", and "interleukin"). The total number of genes targeted was constrained by the library volume remaining from these precious remains. Using the maximum size of existing bait sets as our threshold we fitted the 3,512 sequences, representing a total of 2,506,004 bp into a bait set of 93,180 probes. The probes were designed to be 60 bp in length and tiled every 20 bp.

iii. Enrichment

Samples were enriched using baits targeting the neutral and GWAS loci (75ng of each probe set, resulting in a final mass of 150ng of baits per reaction) according to the myBaits v3 protocol (Daicel Arbor Biosciences), with minor modifications to result in the recommended final concentrations at a final volume of 26 µL instead of 30 µL, allowing a maximum of 10.75 µL template input instead of 7 µL template input., and using 20 µL of beads (pre- and post- wash) instead of 30 µL pre-wash and 70 µL post-wash. A volume of 5 µL template input was used. During the removal of the supernatant post-incubation with the magnetic beads (2B.5 Bead Washing, step 1), the supernatant was saved, labeled as the non-enriched fraction (NEF), and frozen at -20°C. 25 µL of the NEF was used as input to the new enrichment for the exon probes, with 150 ng baits, with incubations at 55°C. Samples were reamplified with the following conditions: Real-time PCR was per- formed in 10 µL reactions consisting of: 1X KAPA SYBR® FAST qPCR Master Mix (2X), 150 nM of each primer (IS5_long_amp.P5: AATGATACGGCGACCACCGA and IS6_long_amp.P7: CAAGCAGAAGACGGCATACGA), and 18.8 µL template. The cycling conditions were: 5 minutes of initialization at 95°C, followed by 12 cycles of a 30 second denaturation at 95°C (30 sec) and a 45 second annealing/extension at 60°C, followed by a final extension for 3 minutes at 60°C, and ending with a 1 minute hold at 4°C. A second round of enrichment was performed with the modifications above. For libraries requiring a second exon enrichment, a new aliquot of library was indexed and 5 µL was used as the input for capture using the protocol described above.

iv. Pooling, size selection, and sequencing

Sample concentration was quantified via qPCR in 0.001x dilutions against the 425bp- 525bp PhiX Control (Illumina) standard serially diluted from 1 nM to 62.5 fM. Real-time PCR was performed in 10 µL reactions consisting of: 1X KAPA SYBR® FAST qPCR Master Mix (2X),

200 nM of each primer (IS5_long_amp.P5 and IS6_long_amp.P7, as above), 4 µL template, and nuclease-free water. The cycling conditions were: 5 minutes of initialization at 95°C, followed by 35 cycles of a 30 second denaturation at 95°C (30 sec) and a 45 second annealing/extension at 60°C, ending with a 30 second hold at 8°C.

Samples were pooled at equimolar ratios. If the pool volume was greater than 12 µL, the pool was concentrated over a Minelute PCR Purification column, with above modifications. Size selection was performed on a 3% NuSieve GTG Agarose Gel with a GeneRuler 50 bp ladder (Fermentas), run for 35 minutes at 100V, and then bands 150-500 bp were excised. Gel slices were purified using a Minelute Gel Extraction kit, with the above modifications. Pools were eluted in 20 µL Qiagen Buffer EB. Pools were sequenced at the Farncombe Metagenomics Facility on the Illumina HiSeq 1500 platform, using V2 Rapid chemistry, with 2x90 paired end reads. Each sample received an average of 732,209 clusters (range: 3,549- 20,052,880).

v. Data processing

Adapters were trimmed and overlapping reads were merged using *leeHom* with aDNA specific settings[63]. Reads were mapped to each target set separately using a modified version of *BWA* which allows for bam-to-bam mapping[64] (https://github.com/udo-stenzel/network-aware-bwa). PCR duplicates were removed with biohazard software (https://github.com/udo-stenzel/biohazard) and reads with quality 30 and length below 24bp were removed with *samtools*[65]. Unique indexed libraries stemming from the same individual were merged using *samtools* and RG tags were standardized with a custom perl script.

Reads mapped strictly the targeted region of the genome produced excess heterozygosity, presumably because of slightly divergent sequences which result in off-target enrichment and are falsely interpreted as alternative alleles[66]. To correct this, we remapped reads to the entire human genome (hg19) and only retained uniquely mapped reads. Resulting bam files were sorted and indexed using *samtools*[65]. We then merged the exon, immune loci, and neutral loci enriched reads together, reasoning that off-target reads which overlap other regions of interest remain valuable and that sequencing from the same sample should share similar patterns of ancient DNA damage. Quality scores in each bam file were adjusted for DNA degradation using *mapDamage 2.0*[67], and then trimmed of the first and last 4 bases using bamUtil (trimBam –L 4 –R 4)[68], where ancient DNA damage is typically concentrated.

Analysis of damage patterns was performed with *mapDamage 2.0* and plotted in *R* using the *ggplot2* package[67,69]. No blanks yielded enough mappable sequences to be included in any downstream analyses. Our final dataset contained 33,110 biallelic, non-singleton variants within the targeted regions (2,669 near GWAS loci, 19,972 in immune genes, and 10,469 in neutral regions), with a mean coverage of 4.6x reads per site per individual (see **Table S1** for individual coverage). While low relative to sequencing of modern genomes, 4.6x coverage is quite high compared published ancient genomes (for example, a large dataset of ancient genomes compiled across multiple studies has a median autosomal coverage of 2.2x with 164 of 804 individuals have greater than 4x coverage[70]). Furthermore, to account for the uncertainty in genotype calls associated with low coverage data all our analyses used genotype likelihood (see point C. below), and we required our top candidates for selection to be replicated in both the London and Denmark ancient DNA cohorts.

c. Genotype calling

Genotypes were called for each individual using GATK's Haplotype Caller in gVCF mode (version 4.1.4.1) with default setting and a minimum base quality of 20[71]. Genotype calls across individuals were then combined using the CombineGVCFs and GenotypeGVCFs functions, and filtered for biallelic sites with minQ ≥ 30 in regions of the genome which were targeted for enrichment (n=139,867 sites). We next excluded singleton variants (n=106,757) as they are more prone to reflect DNA damage or sequencing errors, resulting in a dataset of 33,110 biallelic, non-singleton variants within the targeted regions (2,669 near GWAS loci, 19,972 in immune exons, and 10,469 in neutral regions). We then excluded samples with missing genotype calls at more than 50% of those sites (retaining n=206 individuals) and filtered for sites with at least 10 individuals per population and any given time point. Because of the low coverage intrinsic to ancient genomic studies, many genotype calls were based on a few reads. We therefore capture uncertainty in genotype calls by extracting the genotype likelihoods for each individual at each site using filtbaboon1b from *LCLAE*[72]. From these estimates, we calculated the expected number of alternate alleles as the likelihood the individual is heterozygous plus 2x the likelihood the individual is homozygous alternate. Based on these genotype likelihoods, we then calculated the minor allele frequency per population and time point, retaining only sites with a mean minor allele frequency between London and Denmark greater than 0.05 (n= 22,868 sites), as our power to detect selection for variants below 5% frequency is very low (**Fig S2**).

d. Genetic differentiation following the Black Death

We calculated F_{ST} for each variant as the loss in expected heterozygosity between time points relative to the expected heterozygosity of a panmictic population[73]. Specifically, $F_{ST} = \frac{Ehet_{panmictic} - mean(Ehet_{populations})}{Ehet_{panmictic}}$, where E_{het} (the expected heterozygosity) is calculated for each population-time point and a hypothetical panmictic population from the allele frequency assuming Hardy-Weinberg equilibria. Because high variance in estimating F_{ST} at low allele frequencies can lead to high false positive rates[74], we also filtered out variants with minor allele frequencies less than 10% (calculated as the mean across populations). After this, we compared 3,293 sites in functional regions to 760 putatively neutral loci; functional loci showed an enrichment for high levels of differentiation compared to the neutral loci (**Fig 2A-B**).

We then asked whether sites in functional regions were more differentiated than expected based upon neutral loci. Specifically, we compared the F_{ST} values in functional regions to neutral loci, calculating a p-value for each site as the proportion of neutral loci with a more extreme F_{ST}. As these proportions were generally stable when controlling for minor allele frequency, we maximized the number of neutral sites to compare against by not further restricting the set of neutral loci based on difference in minor allele frequency from the tested variant. Enrichment in the Denmark cohort is smaller than in London, although still considerably higher than expected based on forward simulations (**section 1.e**). There are several reasons why the population in London is better suited to detect positive selection relative to those in Denmark: the dataset from London is ~50% larger (which directly impacts our ability to accurately estimate allele frequencies; hence, power), comes from a smaller geographic area, and from samples which were more accurately dated and occurred over a shorter period of time. Therefore, putatively neutral sites in Denmark show larger F_{ST} values before vs after the Black Death than those in

London (78% larger on average, paired t-test $p = 3.4x10^{-10}$), reflecting increased population heterogeneity and/or our reduced ability to accurately estimate allele frequencies in Denmark given the small sample sizes. Thus, the reduced signal in Denmark (which nonetheless remains quite large: 1.6x the number of variants expected by chance exceed the 99th percentile of neutral loci) can likely reflects contributions of these technical and demographic differences.

The strongest evidence of selection would come from candidate loci with concordant evidence of high differentiation across populations and time points. In particular, we required three conditions be met for a loci to be of interest: (i) individuals who survived the Black Death in both London and Denmark, have the same direction of allele-frequency change, (ii) within London, allele frequencies of Black Death survivors are shifted in the opposite direction of individuals who died during the plague (e.g. those buried in East Smithfield), and (iii) variants of interest are more differentiated than expected based on variation at neutral loci. To address point iii, we required sites be more differentiated than 95% of neutral loci in London before vs after the plague and be more differentiated than 90% of neutral loci in Denmark before vs after the plague.

To ensure that widespread gene flow did not drive these signatures of selection, we proceeded with two additional analyses. We note that a population replacement would affect the totality of the genome and therefore the fact that our enrichments of highly-differentiated variation among immune loci is relative to putative neutral regions sequenced in the same exact samples should alleviate that concern. That said, we cannot exclude the possibility that the highly differentiated sites that we identified correspond to sites that were already unusually differentiated between the local population in and the one replacing it. Therefore, we first reasoned that if our highly differentiated loci in London when comparing pre- and post-Black Death samples were the result of a significant population replacement, those same loci should be amongst the most differentiated amongst contemporary European populations. We found no evidence that the 235 variants we identify as outliers in London before vs after the plague are enriched for large F_{ST} values among European populations, relative to variants that were not highly differentiated. Out of all pairwise comparisons of the five European 1000 Genomes populations, the only evidence for elevated genetic differentiation at candidate loci was observed when contrasting between Iberian (IBS) and Tuscany (TSI) (t-test p=0.0051), which is an implausible populations pair to explain populations replacements in London in the 14th and 15th century (especially given that Great Britain is not even included in this pair, despite the fact that among our strongest candidate variants, post-plague allele frequencies in London are most similar to modern 1000 Genomes frequencies from Great Britains in London: mean difference of 1.8% vs 5.5-6.7% for other European populations).

e. Forward simulations of neutral evolution

Our results indicate an excess of highly diverged alleles near immune-related genes and GWAS variants relative to neutral regions of the genome. However, the genomic context of candidate sites differs from neutral regions in ways that may affect the variance in F_{ST} [75,76]. We therefore contrasted the recombination rate (derived from the DeCODE project[77]) and the level of background selection (as measured by the B statistic[78]) between the neutral regions and our targeted immune loci. Recombination rates were significantly different between neutral regions and candidate loci, with an overall trend for higher recombination rates around our candidate loci

($p = 4.03\text{x}10^{-9}$; **Fig S3A**). Levels of background selection were also significantly higher around immune loci as compared to putative neutral regions ($p \approx 0$; **Fig S3B**). Importantly, neither recombination rate nor background selection was enriched for candidates of positive selection relative to other tested loci in exonic regions or near GWAS hits ($p > 0.5$, **Fig S3A-B**).

Nevertheless, since recombination rate and background selection can affect the variance in F_{ST} (albeit in opposing directions[79,80]), we used forward simulations to confirm that the differences in genetic differentiation observed between neutral and candidate regions are unlikely under neutral evolution. Specifically, we used SLiM[75,76] to simulate the neutral evolution of full genomes in a population of 10,000 individuals while controlling for recombination rates (estimated by the DeCODE project[77]) and negative selection against mutations in coding regions (defined using the UCSC table of canonical genes[81]). Deleterious mutations were introduced at a rate of $7.4\text{x}10^{-8}$ mutations per base pair per generation, and an exponential distribution of fitness effects with a mean deleterious effect size of $2.5\text{x}10^{-3}$ (inferred by [78]). We simulated 1,000 generations under this model, allowing for neutral and deleterious variation to accumulate and stabilize, before we simulated our data collection surrounding the Black Death. At 1,000 generations, we extracted "pre-BD" allele frequencies for variants within 5kb of genomic regions targeted in this study. Assuming a generation time of approximately 30 years, we simulated an additional 8 generations before the BD event (equivalent to the mean number of generations of our pre-plague London samples), the BD which led to 50% mortality in the population, and a further 7 generations after the BD event (equivalent to the mean age for our post-plague London samples), at which point we sampled "post-BD" allele frequencies.

As in our main analyses, we then asked whether sites near candidate regions were more strongly differentiated before vs after the plague than those of neutral regions. We detected a slight enrichment of highly differentiated sites that was much smaller than those observed for our empirical data (**Fig S3C-D**), suggesting background selection and recombination are insufficient to explain the enrichment of highly differentiated sites we observe. Importantly, these results are consistent across a range of parameter values including population sizes up to 50,000 (the approximate census population in London at the beginning of the Black Death) and mortality rates from 30 to 50%.

f. Gene regulatory response to *Y. pestis*

i. Isolation of monocytes, differentiation of macrophages, and challenge with heat-killed *Y. pestis*

This study has been approved by the Institutional Review Board at the University of Chicago (protocol #: IRB19-0432). All Buffy coats from 33 healthy donors were obtained from the Indiana Blood Center (Indianapolis, IN, USA). All individuals recruited in this study were males, self-identified as African-American (AA) (n = 17) or European-American (EA) (n = 16) between the age of 18 and 55 years old. Only individuals self-reported as currently healthy and not under medication were included in the study. In addition, each donor's blood was tested for Hepatitis B, Hepatitis C, Human Immunodeficiency Virus (HIV), and West Nile Virus, and only samples negative for all of the tested pathogens were used.

Blood mononuclear cells were isolated by density gradient centrifugation (Ficoll-Paque Premium, Sigma Aldrich, St. Louis, MI, USA). Monocytes were purified from peripheral blood

mononuclear cells by positive selection with magnetic CD14 MicroBeads (Miltenyi Biotech, Bergisch Gladbach, Germany) using the autoMACS Pro Separator. The purity of the isolated monocytes was verified using an antibody against CD14 (BD Biosciences) and only samples showing > 90% purity were used to differentiate into macrophages. Monocytes were then cultured for 7 days in RPMI-1640 (Fisher) supplemented with 10% heat-inactivated FBS (FBS premium, US origin, Wisent), L-glutamine (Fisher) and M-CSF (20ng/mL; R&D systems). Cell cultures were fed every 2 days with complete medium supplemented with the cytokines previously mentioned. Before inoculation with *Y. pestis*, we checked the differentiation/activation status of the monocyte-derived macrophages by flow cytometry, using antibodies against CD1a, CD14, CD83, and HLA-DR (BD Biosciences). All samples presented the expected phenotype for non-activated macrophages (CD1a+, CD14+, CD83, and HLA-DRlow). Macrophages were stimulated for 4 hours with lysate of heat-killed *Y. pestis* (strain CO92) at a value of 10 colony forming units (CFUs) per cell, or with endotoxin-free water (negative control).

ii. RNA Extraction, Library Preparation, and Sequencing

Total RNA was extracted from the non-stimulated and stimulated macrophages using the miRNeasy kit (QIAGEN). RNA quantity was evaluated spectrophotometrically, and the quality was assessed with the Agilent 2100 Bioanalyzer (Agilent Technologies). Only samples with no evidence for RNA degradation (RNA integrity number > 8) were kept for further experiments. RNA- sequencing libraries were prepared using the Illumina TruSeq protocol. Once prepared, indexed cDNA libraries were pooled (6 or 7 libraries per pool) in equimolar amounts and were sequenced with single-end 100bp reads on an Illumina HiSeq2500.

Adaptor sequences and low quality score bases (Phred score < 20) were first trimmed using Trim Galore (version 0.2.7). The resulting reads were then mapped to the human genome reference sequence (Ensembl GRCh37 release 75) using STAR (2.4.1d)[82] with a hg19 transcript annotation GTF file downloaded from Ensembl. Gene-level expression estimates were calculated using RSEM (version 1.2.21)[83] with default parameters.

After removing lowly expressed genes (mean counts per million < 1 in both stimulated and null conditions), we normalized expression data using voom and analyzed 5,874 genes using the R package *limma* (v3.48.3). We first tested for an effect of stimulation with *Y. pestis* while controlling for individual identity for an effect of stimulation on expression level, while controlling for individual identity. We then calculated a permutation-based FDR, comparing our observed results to permuted results observed from randomizing which sample was stimulated within each individual (repeated 25 times). We extracted expressed genes within 100kb of our candidate loci for positive selection, retaining *ERAP1*, *ERAP2*, *LNPEP*, *TICAM2*, *TMED7*, *CTLA4*, and *ICOS*. We then used previously published genotype data for these individuals[30] to test whether genotype at our candidate variants was associated with expression level using a linear model. Finally, we tested for an interaction effect between stimulation and genotype. Despite several trends which appear to resemble interaction effects, our sample was underpowered to identify any significant interactions. We applied an identical computational pipeline to test for effects of stimulation and genotype in previously published datasets using different bacterial and viral pathogens and ligands[30,31].

iii. Comparison with Quach et al. and Nedelec et al. data

To evaluate if the differences in gene expression identified among our candidate loci were specific to flu or shared in response to other pathogens we re-analyzed existing datasets: Nedelec and colleagues[30] measured the gene expression response of monocyte-derived macrophages to infection with two live intracellular bacteria: *Listeria monocytogenes* (a Gram-positive bacterium) and *Salmonella typhimurium* (a Gram-negative bacterium), using the same experimental design and time point as the current study. Quach and colleagues[31] characterized the transcriptional response of primary monocytes to bacterial and viral ligands activating Toll-like receptor pathways (TLR1/2, TLR4, and TLR7/8) and live influenza virus 6 hours after stimulation. Processing of the RNAseq data from these datasets and analysis was done using the same methods described above in section **ii**.

iv. Infection of macrophages with live *Y. pestis*: cytokine and gentamycin protection assays

In addition to the stimulation of macrophages with heat-killed *Y pestis* (**section 1.f.i** above), we also performed a separate set of infections with live *Y. pestis* of a set monocyte-derived macrophages derived from 33 healthy anonymous donors (Etablissement Français du Sang). This experiment was specifically designed to *(i)* evaluate if the differences in gene expression observed in response to heat-killed Y. pestis were similar to those observed in response to live Y. pestis (this was done on a subset of 8 individuals); and *(ii)* evaluate if the *ERAP2* protective allele impacted cytokine secretion levels or intracellular bacteria levels. DNA from 200µl of blood was extracted using PureLink Genomic DNA kits (Invitrogen, Waltham, MA, USA) and allelic discrimination at the rs2248374 SNP was determined by TaqMan qPCR with Applied Biosystems Human *ERAP2* probes. CD14+ monocytes were isolated and were differentiated into macrophages as described above. Fully virulent *Y. pestis* (strain CO92) was grown on LB Agar supplemented with 0.0002% porcine hemin and cells ($25x10^4$/well) were infected with live bacteria at a multiplicity of infection of 10 bacteria/cell. Supernatants (for cytokine profiles) or cells (for RNA-sequencing) were collected after 5 hours of incubation without washing cells, or cells were washed with PBS after 1h and incubated in fresh medium for a further 23 hours in the presence of gentamycin (40µg/ml). Samples were sterilized by centrifugation over 0.22µm Spin-X filters (Costar) before deep-freezing. Levels of ten cytokines (IL-1ß, IL-1RA, TNFα, IL-6, IL-10, G-CSF, GM-CSF, CCL2/MCP-1, CCL3/MIP-1α, CXCL8/IL-8) were determined using Magnetic Luminex performance assays from R&DSystems with Curiox DropArray plates and LT washing station for 25 individuals. Data were acquired on a Luminex LX-200 reader. For 8 individuals, we performed RNA-sequencing on infected and non-infected cells. Sequencing data were mapped and analyzed for an effect of *Y. pestis* stimulation as described above in **section 1.f.ii**.

For all 33 individuals, macrophages (6 biological replicates per condition) were infected with *Y. pestis* (MOI of 10 bacteria/cell) with gentle centrifugation (300xG, 5 min) to enhance contacts. After one hour, samples were washed twice with PBS, and new medium containing Gentamycin (40 µg/ml) was added to kill extracellular bacteria. Then, another hour or 24h, macrophages were washed twice with PBS, lysed with sterile water and scratched from plastic. The lysate was serially diluted, and samples immediately cultured at 28°C on LB agar + 0.002% Hemin to count the number of colony forming units (CFU) of bacteria. Intracellular bacterial killing after 24h was calculated as 100x(CFU at 2h – CFU at 24h)/(CFU at 2 h). This strategy is the most used strategy to analyze intracellular survival or killing of bacteria inside macrophages[84]. For both

cytokine and CFU data, we quantile-normalized the data across individuals to minimize the risk of spurious association due to outliers, and to ensure that the data was normally distributed. We then examined the association between SNP genotypes (coded as the number of protective "C" alleles at rs2549794: 0, 1, or 2) and cytokine or CFU levels by using a linear regression model in which phenotype was regressed against genotype. This model returns a single *p* value that reflects the relationship between genotype (i.e., the number of protective "C" alleles) and cytokine/CFU levels.

v. Single-cell RNA-sequencing of PBMCs before and after infection with *Y. pestis.*

Experimental design

To consider the response to *Y. pestis* infection across a broader array of immune cell types, we also conducted single-cell RNA-sequencing of peripheral blood mononuclear cells (PBMCs) for 10 individuals of European ancestry, five of whom were homozygous for the reference allele at rs2549754 and five of whom were homozygous for the alternate allele. These individuals were previously used to study the response to viral infections and the full details of sample providence can be found in previous work[85]. Briefly, samples were obtained from BioIVT along with a signed, written consent from healthy participants at a collection site in Miami, Florida (United States) utilizing a standard protocol with a sodium heparin anticoagulant. PBMCs were extracted from whole blood using a density gradient, washed with HBSS, reconstituted in CryoStor CS10 to a concentration of 10 million cells/mL, and cryopreserved in sets of 5-10 million cells a vial.

PBMCs were thawed approximately 14 hours prior to infection and cultured overnight in RPMI 1640 supplemented with 10% fetal bovine serum (FBS), 2 mM L-glutamine, and 10 µg/mL gentamycin. Infection experiments were balanced for individual genotypes. The morning of the experiment 10 individuals were collected, counted, plated at 1 million cells per ml per well, and infected with *Y. pestis* (strain CO92), at an MOI of ten. Following 5 hours of infection, cells were collected, washed, and prepared for single-cell capture using the 10X workflow. Multiplexed cell pools of 6 samples were used as input for the single cell captures, and for each pool 10,000 cells were targeted using the Chromium Single Cell 3' Reagent (v3 chemistry) kit (10X Genomics #CG000183 Rev A). Post Gel Bead-in-Emulsion (GEM) generation, the reverse transcription (RT) reaction was performed in a thermal cycler as described (53°C for 45 min, 85°C 537 for 5 min), and post-RT products were temporarily stored at -20°C. After all captures were complete, post-RT reaction cleanup, cDNA amplification, and sequencing library preparation were performed as described in the Single Cell 3' Reagents Kits v3 User Guide (10X Genomics #CG000183 Rev A). Briefly, cDNA was cleaned with DynaBeads MyOne SILANE beads (ThermoFisher Scientific) and amplified in a thermal cycler using the following program: 98°C for 3 min, 11 cycles x 98°C for 15 s, 63°C for 20 s, 72°C for 1 min, and 72°C 1 min. After cleanup with the SPRIselect reagent kit (Beckman Coulter), the libraries were constructed by performing the following steps: fragmentation, end-repair, A-tailing, SPRIselect cleanup, adaptor ligation, SPRIselect cleanup, sample index PCR (98°C for 45 s, 13 or 14 cycles x 98°C for 20 s, 54°C for 30 s, 72°C for 20 s, and 72°C 1 min), and SPRIselect size selection. Libraries were sequenced 100 base pair paired-end (R1: 30 cycles, I1: 10 cycles, R2: 85 cycles) on an Illumina NovaSeq. Cell pools contained infected or uninfected samples, to limit the possibility for cross-contamination.

Mapping, demultiplexing, cell filtering, and UMAP analysis

FASTQ files were mapped to a custom reference containing GRCh38 and the *Y. pestis* reference genome (downloaded from NCBI, created using cellranger mkref) using the cellranger (v3.0.2) (10X Genomics) count function. Souporcell (v2.0, Singularity v3.4.0)[86] in --skip_remap mode was used to demultiplex cells based on genotypes from a common variants file (1000GP samples filtered to SNPs with >= 2% allele frequency in the population, downloaded from https://github.com/wheaton5/souporcell). Briefly, souporcell clusters cells based on cell allele counts in common variants, assigning all cells with similar allele counts to a single cluster corresponding to one individual, while also estimating singlet/doublet/negative status for that cell. Hierarchical clustering of the true genotypes known for each individual (obtained from low-pass whole-genome-sequencing) and the cluster genotypes estimated from souporcell was used to assign individual IDs to souporcell cell clusters. All 10 individuals were successfully assigned to a single cluster.

After demultiplexing cells into samples, Seurat (v4.0.4, R v4.1.1)[87] was used to perform quality control filtering of cells. Cells were considered "high quality" and retained for downstream analysis if they had: 1) a "singlet" status called by souporcell, 2) between 200 – 2500 genes detected (nFeature_RNA), and 3) a mitochondrial reads percentage < 20%. In total, we captured 14,174 high quality cells (range of cells per individual: 780 to 2030).

We split the cells by infection status (uninfected or *Y. pestis*) and individual, then ran SCTransform (Seurat v4.0.4) to normalize and scale the UMI counts within condition. In this step, we simultaneously regressed out variables corresponding to experiment batch and percent mitochondrial reads per cell. Cells were assigned to cell types using Seurat's Multimodal Reference Mapping workflow. Briefly, we applied SCTransform, FindTransferAnchors, and MaqQuery to assign each cell a predicted cell type based on published data from PBMCs[88]. Data were then integrated on infection status and individual using reciprocal PCA on SCTransformed data (https://satijalab.org/seurat/articles/integration_rpca.html) prior to dimensionality reduction (RunUMAP in Seurat, with the first 20 dimensions). Reassuringly, predicted cell types reflected major structure in the UMAP plot.

Differential expression analysis

To model expression data for each cell type, we converted raw counts per cell into pseudobulk mean counts for each cell type and sample (as in [85]). Within each cell type for each sample, raw counts were summed across all cells assigned to that cell type and sample for each gene using the sparse_Sums function in textTinyR (v1.1.3) (https://cran.rproject.org/web/packages/textTinyR/textTinyR.pdf), yielding an n x m expression matrix, where n is the number of samples (n = 20) and m is the number of genes detected. For each cell type, we removed lowly expressed genes (cpm < 1) and then voom-normalized the count data to control for library. Using lmFit (R *limma* package v3.48.3), we modeled effects of stimulation and an effect of rs2549754 genotype within each condition, while controlling for the number of cells. False discovery rates were estimated using the q-value approach[89] relative to an empirical null distribution constructed by randomly assigning which sample condition within each individual and by randomizing genotypes between individuals (n=10 iterations). Results for stimulation of each gene and each cell type are available on in **Table S6**. Finally, we find no evidence for widespread *trans*-effects of *ERAP2* haplotype on the gene expression levels of other

genes across the genome after correction for multiple testing. However, we urge caution in interpreting this results as this particular data set is severely underpowered to identify interaction effects of this type. Moreover, we are only able to capture a snapshot of the immune response *in vitro* (i.e., measures of gene expression 5 hours post-infection). It is possible that differential gene regulation between the protective and deleterious haplotypes may occur at later time points, or in an *in vivo* setting where many different cell types interact with each other.

g. Hidden-Markov Model-based inference of selection coefficients

We analyzed the four SNPs identified in the main text (rs2549794, rs17473484, rs1052025, and rs11571319) further to quantify the evolutionary advantage that the respective alleles conferred. To this end, we aimed to infer a selection coefficient from the temporal genetic variation at each of these SNPs. Assuming the allelic effects are additive, we define s (the selection coefficient) as the expected number of offspring of an individual with genotype aa, Aa, or AA is 1, $1 + s/2$, or $1 + s$, respectively (where a is the ancestral allele and A is the derived variant).

To infer this selection coefficient s from the temporal data, we used a common Hidden Markov Model (HMM) framework[24]. Briefly, in this framework, the underlying population allele frequency is treated as the hidden state, and the observed genetic variation at the respective times present the emissions, which are binomially distributed, given the underlying allele frequency. To compute the likelihood of the data, the framework integrates over all possible hidden allele frequency trajectories under the Wright-Fisher model. Since the selection coefficient s affects the likelihood of these trajectories (larger s makes an increase in frequency more likely), this allows us to infer s using a maximum likelihood approach. Here, we use an implementation of this HMM framework available at https://github.com/steinrue/diplo_locus which allows efficient computation of the likelihood by discretizing the allele frequency space and using a Gaussian approximation to compute the transition probabilities of the Wright-Fisher model[90].

With this implementation, we computed joint likelihoods of observing samples in both populations across time, which yields maximum-likelihood point estimates of s. In addition, we applied this implementation to simulated data to provide confidence intervals and assess the statistical power of our inference scheme. Because the sampling times across both populations span no more than 20 generations, we set the mutation rate to zero for convenience. All scripts to perform the analyses are available at https://github.com/TaurVil/VilgalysKlunk_yersinia_pestis/tree/main/part1_aDNA/infer_selection_coefficients.

i. Maximum likelihood estimation of selection coefficients from data

We follow the analysis in the main text and group the samples into before, during, and after the Black Death pandemic. We use the expected derived allele frequency in each group for each candidate SNP, computed as $\mathcal{P}_{Aa} + 2\mathcal{P}_{AA}$ for each individual and the mean across individuals in each group, where $\mathcal{P}$ is the genotype likelihood per individual inferred in section 1.c.

For samples from London (blue lines and boxes in **Fig S10**), we assume an effective population size of 5,000 and that the individuals before, during, and after the pandemic are sampled in generation 1, 9, and 16, respectively. Considering that multiple generations of people have been impacted by the circulating disease, we assume selection acted during generation 6, 7, and 8, the

three generations preceding the mid-pandemic sample. Because the mid-pandemic samples in London are known to be from individuals who died from the Black Death, we assume that the lineages leading to mid-pandemic samples experience opposite selective pressure from those who survived the pandemic (the post-pandemic samples).

For samples from Denmark (brown lines and orange boxes in **Fig S10**), we only consider samples taken before and after the pandemic. We assume an effective population size of 5,000, that the samples were taken at generation 0 and 17, and that selection occurred during generation 6, 7, and 8 (same as in London). The generation times were chosen to represent approximate mid-points of the pre- and post-Black Death samples. For each candidate SNP, we computed log-likelihoods on a grid of s values in the interval [-0.9,0.9], with a higher density of grid-points around zero, for three different scenarios (see **Fig S10**, straight lines in the lower half): before-to-during in London, before-to-after in London, and before-to-after in Denmark. To reduce the number of free parameters, we assumed the selection coefficient s in all three scenarios has the same absolute value, but the coefficient for before-to-during in London has inverted sign. The rationale for this is that the samples before-to-after represent the survivors, whereas before-to-during samples in London represent the ones who did not survive. The final composite log-likelihood for each locus is then obtained by summing the log-likelihoods from the three scenarios.

To estimate $\hat{s}_{\mathrm{MLE}}$, we interpolate the composite log-likelihood at each locus as a function of s using cubic splines (with the scipy.interpolate module in Python3), and infer the maximum in the [-0.9,0.9] interval (with the scipy.optimize module) from this interpolated function. The log-likelihood surfaces and the resulting maximum-likelihood estimates (MLEs), $\hat{s}_{\mathrm{MLE}}$, are depicted in **Fig S11** for the four SNPs. We also performed the analysis with an effective population size of 10,000 in both populations, but the results did not change substantially.

ii. Simulation study to assess confidence in selection coefficients

To assess the statistical properties of our estimation procedure, we simulated allele frequency trajectories under the Wright-Fisher model, and simulated samples at the respective times given the underlying allele frequency. In the simulations of the London population, the three sampling times are set to be generation 1, 9, and 16. In simulations for Denmark, sampling times are set as generations 0 and 17. As samples sizes, we used the sample sizes in the empirical data at the respective SNPs. The effective population sizes for London and Denmark populations are set to be 5,000, and the mutation rate is set to zero. As in the likelihood-computations, selection was assumed to act in generations 6, 7, and 8.

To reflect that the mid-pandemic samples from London were taken from people who died of the disease, we simulate this set of samples by assuming a parallel population where the allele frequency experiences reverse changes during the pandemic. In particular, for each replicate, we track the allele frequency throughout the period of selection (pandemic), take the logit-transformed difference between the start and end, and deduct it from the logit-transformed pre-pandemic frequency. The use of logit here ensures that all values are bounded within (0,1). This then gives the frequency in the parallel population who died of Black Death, from which mid-pandemic samples are generated. To reflect uncertainty of the initial frequency in the empirical data, at each of the four target SNPs, we re-sample each pool of pre-pandemic samples with

replacement and use the resulting allele frequencies as the initial frequencies for the simulated replicates.

We simulate 5,000 replicates for a grid of *s*-values ranging from -0.35 to 0.5 with increments of 0.05, and then applied the same procedure as described in the previous section to each replicate to obtain $\hat{s}_{\mathrm{MLE}}$. **Fig S12** shows the empirical distribution of these estimates, highlighting the mean, median, and 2.5-th and 97.5-th percentiles. We observed a slight bias, albeit very small, in the estimates. Thus, we performed an empirical bootstrap bias-correction of the estimates as follows: We interpolate the values of the median of the empirical $\hat{s}_{\mathrm{MLE}}$ distributions as functions of true *s* value simulated. With these interpolated functions, we use the $\hat{s}_{\mathrm{MLE}}$ directly computed from the likelihood for the four SNPs to identify which underling value of *s* would result in such an estimate. We denote the latter bootstrap corrected MLE by $\tilde{s}_{\mathrm{MLE}}$.

To provide confidence intervals, for each of the four SNPs, we simulated 5,000 replicates using the bootstrap-corrected $\tilde{s}_{\mathrm{MLE}}$ as the underlying true selection coefficient and re-sampled initial frequencies, as before. We then adjust the MLEs from all replicates by the same amount that $\hat{s}_{\mathrm{MLE}}$ was adjusted to obtain $\tilde{s}_{\mathrm{MLE}}$, for each SNP. After this adjustment, we report the 2.5-th and 97.5-th percentiles of the empirical distribution of the MLEs as confidence intervals for our estimates in **Fig S4A** and **Table S5**. We also simulated sets of 5,000 replicates with the same initial conditions but under $s = 0$, applied the same inference procedure to assess the power, and report Receiver Operating Characteristics (ROC) curves and Precision-Recall curves in **Fig S4B-C**. Note that the power is affected by the selection coefficient and the initial frequencies, and thus varies between the SNPs.

To provide further evidence for selection acting at the four focal SNPs, we applied the same inference procedure that was used to obtain $\hat{s}_{\mathrm{MLE}}$ for the focal SNPs to all SNPs in the neutral regions (described in **section 1.b**), since these reflect the empirical neutral distribution. For each SNP, we then computed the log-likelihood-ratio by subtracting the log-likelihood at $s = 0$ from value of the log-likelihood at $\hat{s}_{\mathrm{MLE}}$. We report the percentile of the log-likelihood-ratio for the four focal SNPs in this empirical neutral distribution in **Table S5**, with the rationale that low percentiles indicate strong evidence that non-neutral processes were driving the allele frequency change at the respective SNP. We note for neutral SNPs where the ancestral allele is unknown, we polarized for the reference allele. Since this only changes the sign of the estimated s, the polarization does not affect the value of the log-likelihood-ratio.

Last, we note that $1 + s$ yields the expected number of viable offspring of an individual that is homozygous for the derived allele at the respective locus. Thus, we can loosely interpret $1 + s$ as reflecting the chance of an individual to survive the Black Death pandemic. With this interpretation, we can quantify the increase in chance of survival of individuals homozygous for the derived allele over individuals homozygous for to ancestral allele by $p_{surv} \coloneqq (1 + s)/1$.

h. Supplementary References

57. Schwarz, C., et al. New insights from old bones: DNA preservation and degradation in permafrost preserved mammoth remains. *Nucleic Acids Res* **37**, 3215-3229 (2009).

58. Dabney, J., et al. Complete mitochondrial genome sequence of a Middle Pleistocene cave bear reconstructed from ultrashort DNA fragments. *P Natl Acad Sci USA* **110**, 15758-15763 (2013).
59. Glocke, I. & Meyer, M. Extending the spectrum of DNA sequences retrieved from ancient bones and teeth. *Genome Res* **27**, 1230-1237 (2017).
60. Meyer, M. & Kircher, M. Illumina sequencing library preparation for highly multiplexed target capture and sequencing. *CSH Protocols* **2010.6**, pdb-prot5448 (2010).
61. Kircher, M., Sawyer, S., & Meyer, M. Double indexing overcomes inaccuracies in multiplex sequencing on the Illumina platform. *Nucleic Acids Res* **40**, e3-e3 (2012).
62. Gronau, I., Hubisz, M. J., Gulko, B., Danko, C. G., & Siepel, A. Bayesian inference of ancient human demography from individual genome sequences. *Nat Genet* **43**, 1031-1034 (2011).
63. Renaud, G., Stenzel, U., & Kelso, J. leeHom: adaptor trimming and merging for Illumina sequencing reads. *Nucleic Acids Res* **42**, e141-e141 (2014).
64. Li, H. & Durbin, R.. Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics* **25**, 1 (2009).
65. Li, H., et al., The sequence alignment/map format and SAMtools. *Bioinformatics* **25**, 2078-2079 (2009).
66. Graffelman, J., Jain, D., & Weir, B. A genome-wide study of Hardy–Weinberg equilibrium with next generation sequence data. *Hum Genet* **136**, 727-741 (2017).
67. Jónsson, H., Ginolhac, A., Schubert, M., Johnson, P.L., & Orlando, L. mapDamage2. 0: fast approximate Bayesian estimates of ancient DNA damage parameters. *Bioinformatics* **29**, 1682-1684 (2013).
68. Jun, G., Wing, M. K., Abecasis, G. R., & Kang, H. M. An efficient and scalable analysis framework for variant extraction and refinement from population-scale DNA sequence data. *Genome Res* **25**, 918-925 (2015).
69. Wickham, H. & Chang, W. Package 'ggplot2'. Create Elegant Data Visualisations Using the Grammar of Graphics. Version 2.1 (2016).
70. Allen Ancient DNA Resource https://reich.hms.harvard.edu/allen-ancient-dna-resource-aadr-downloadable-genotypes-present-day-and-ancient-dna-data, version 50.0
71. McKenna, A., et al., The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res* **20**, 1297-1303 (2010).
72. Wall, J. D., et al. Genome-wide ancestry and divergence patterns from low-coverage sequencing data reveal a complex history of admixture in wild baboons. *Mol Ecol* **25**, 3469-3483 (2016).
73. Weir, B. S. & Cockerham, C. C. Estimating F-statistics for the analysis of population structure. *Evolution* **38**, 1358-1370 (1984).
74. Whitlock, M. C. & Lotterhos, K. E. Reliable detection of loci responsible for local adaptation: Inference of a null model through trimming the distribution of F_{ST}. *Am Nat* **186**, S24-S36 (2015).

75. Haller, B.C. & Messer, P.W. SLiM 2: flexible, interactive forward genetic simulations. *Mol Biol Evol* **34**, 230-240 (2017).
76. Messer, P.W. SLiM: Simulating Evolution with Selection and Linkage. *Genetics* **194**, 1037-1039 (2013).
77. Kong, A., et al. A high-resolution recombination map of the human genome. *Nat Genet* **31**, 241–247 (2002).
78. McVicker, G., Gordon, D., Davis, C. & Green, P. Widespread genomic signatures of natural selection in hominid evolution. *PLOS Genet* **5**, e1000471 (2009).
79. Matthey-Doret, R. & Whitlock, M.C., Background selection and FST: consequences for detecting local adaptation. *Mol Ecol* **28**, 3902-3914 (2019).
80. Keinan, A. & Reich, D. Human population differentiation is strongly correlated with local recombination rate. *PLOS Genet*, **6**, e1000886 (2010).
81. Hsu, F., et al. The UCSC known genes. *Bioinformatics*, **22**, 1036-1046 (2006).
82. Dobin, A., et al. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15-21 (2013).
83. Li, B., & Dewey, C.N. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinformatics* **12**, 1-16 (2011).
84. Kaneko, M., Emoto, Y. & Emoto, M. A simple, reproducible, inexpensive, yet old-fashioned method for determining phagocytic and bactericidal activities of macrophages. *Yonsei Med J* **57**, 283-290 (2016).
85. Randolph, H. E., et al. Genetic ancestry effects on the response to viral infection are pervasive but cell type specific. *Science* **374**, 1127-1133 (2021).
86. Heaton, H., et al. Souporcell: robust clustering of single-cell RNA-seq data by genotype without reference genotypes. *Nat Methods* **17**, 615-620 (2020).
87. Stuart, T., et al. Comprehensive integration of single-cell data. *Cell* **177**, 1888-1902 (2019).
88. Hao, Y. et al. Integrated analysis of multimodal single-cell data. *Cell* (2021).
89. Storey, J.D. & Tibshirani, R. Statistical significance for genome wide studies. *P Natl Acad Sci USA* **100**, 9440-9445 (2003).
90. Mathieson, I. & McVean, G. Estimating selection coefficients in spatially structured populations from time series data of allele frequencies. *Genetics* **193**, 973–984 (2013).

2. Supplementary Figures

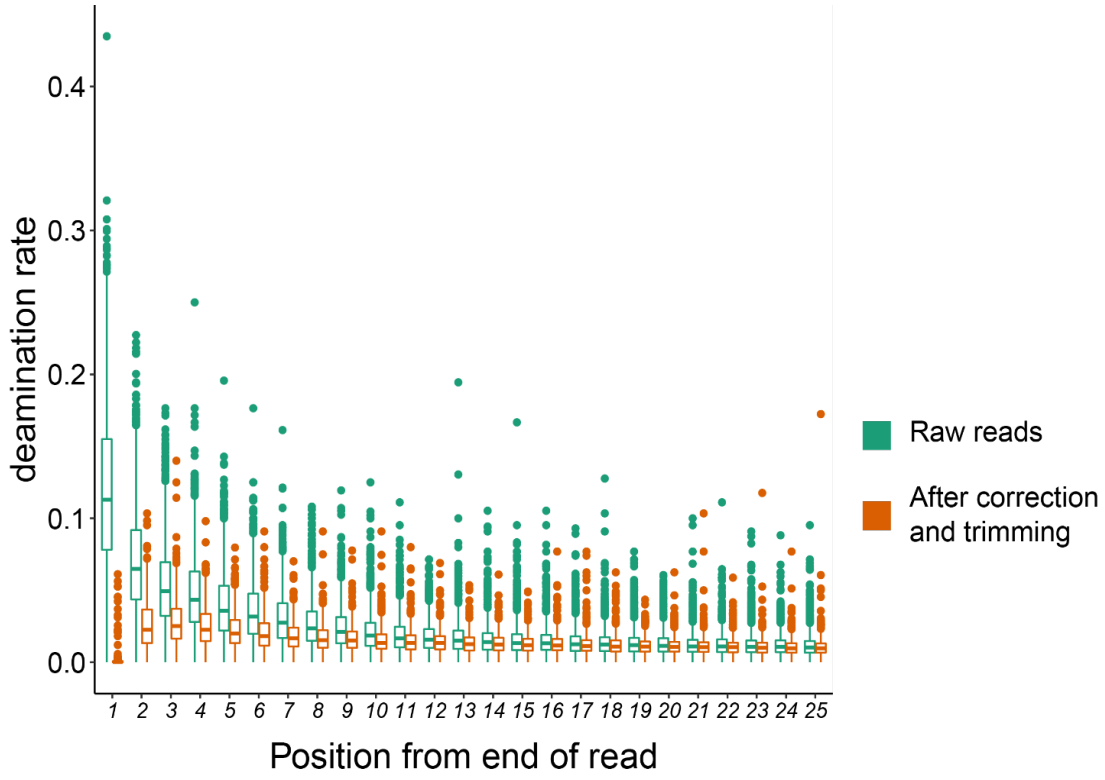

Figure S1: Correcting for ancient genomic damage. Damage and degradation in ancient DNA samples result in an increased deamination rate near the end of reads (green). We correct for this source of DNA damage by using *mapDamage*[59] and trimming the terminal 4 bases of each read, which eliminates the bias due to DNA damage (orange). After correcting and trimming reads, error rates are lower for the first base pair because *mapDamage* is more liberal in masking possible DNA damage near the ends of reads.

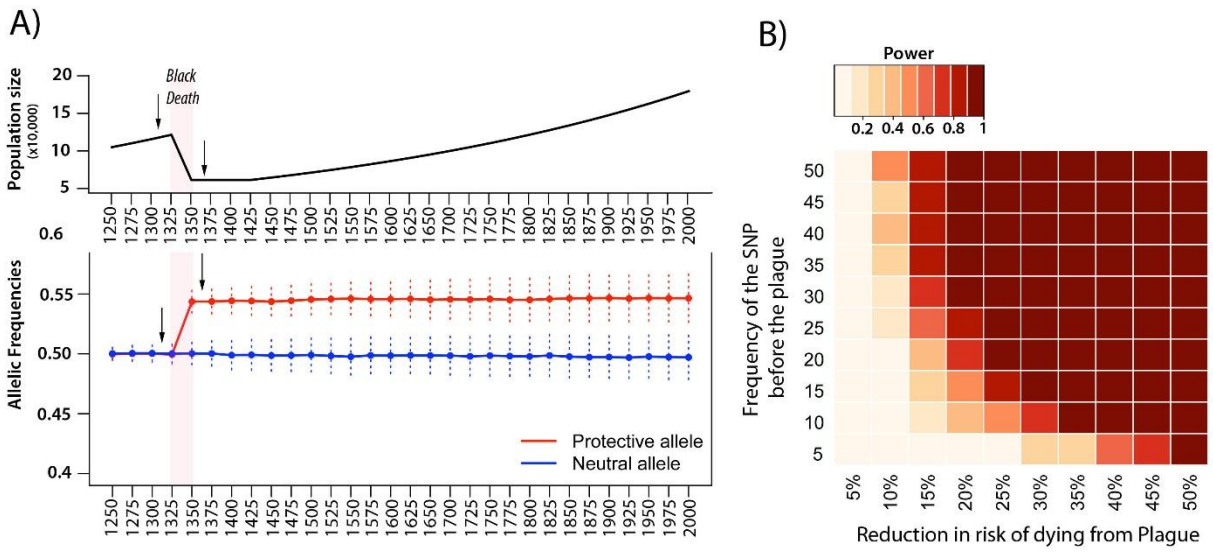

Figure S2: Power to identify protective alleles against Black Death. (A) Simulations of the expected changes in allele frequencies across generations for a protective allele (red) that decreases the risk of mortality from plague by 30% in homozygous individuals and 15% in heterozygous individuals (i.e., additive effect) as compared to a neutral allele (blue), with a starting allele frequency of the protective or neutral allele of 50%. Dashed lines - standard deviation for 100 simulated alleles. Under the simulated scenario the protective allele increases by ~5% in a single generation (in red), which is significantly more than what would be expected by a neutral allele (in blue). **(B)** Power to detect protective alleles as a function of their frequency prior to *Black Death* and their level of protection against the disease (the figure shows the level of protection for homozygous individuals) by calculating *Fst* values between pre- and post-Black Death populations. Power was calculated by comparing *Fst* values for a simulated set of 100 protective alleles between the pre- and post-Black Death populations with respect to *Fst* values observed among 1,000 simulated neutral alleles sampled from the same populations.

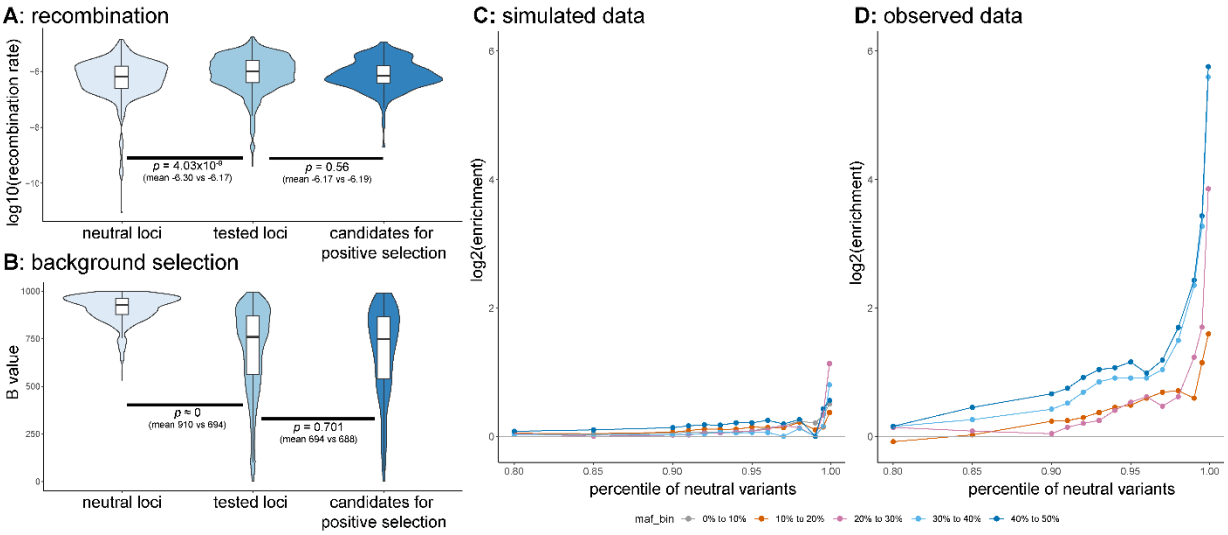

Figure S3: Differences in recombination rate and background selection are insufficient to explain the marked enrichment of high F_{ST} values among immune loci. Comparison of recombination rate (A) and background selection levels (B) between neutral loci and our candidate regions. Candidate regions were stratified into those which were tested and those which were candidates for positive selection based on high differentiation in London pre- vs post-BD. (C) Forward simulations matched for the rates of recombination and background selection of the regions targeted in our study show a slight enrichment of highly differentiated sites in candidate regions, but far from the level of enrichment observed in the real data (D), replicated from Fig 2A for comparison. For example, whereas in the real data differentiation at immune loci exceeded the 99th percentile of neutral variants at 2.4x the rate expected by chance (among variants with a MAF > 10%), the same enrichment is less than 1.2x in the simulated data.

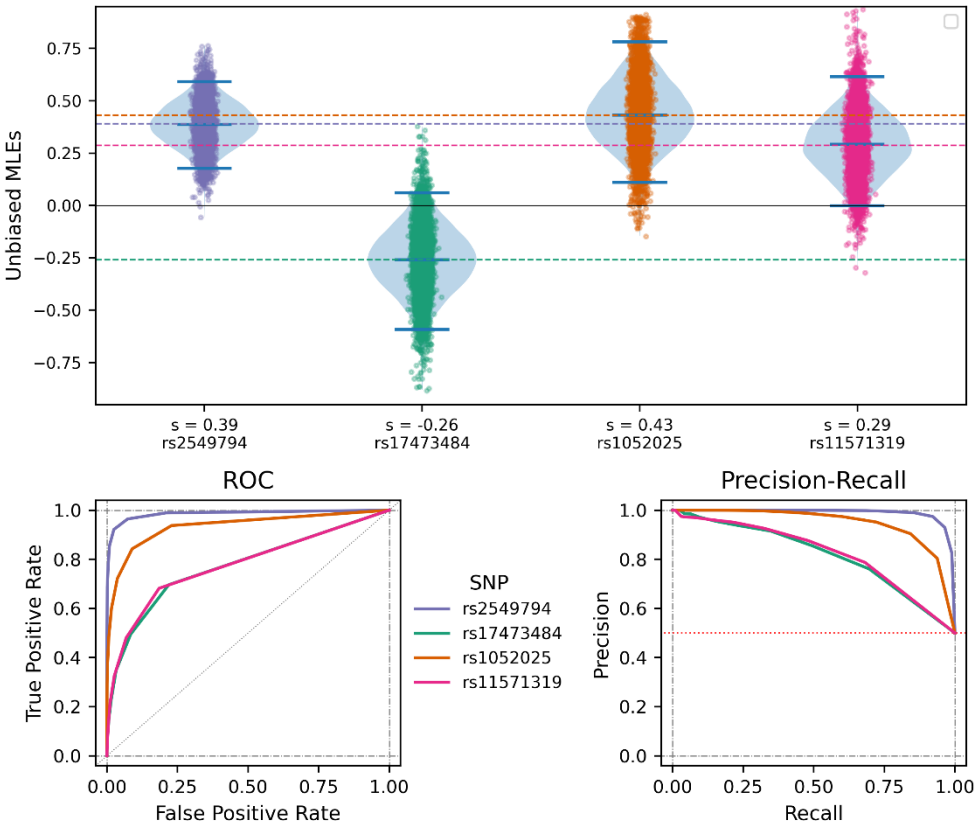

Figure S4: Estimates of the selection coefficients for the four SNPs of interest and power of the inference procedure. (A) Distributions of $\hat{s}_{\mathrm{MLE}}$ for the four SNPs when replicates are simulated with the corresponding bootstrapped allele frequency distributions as initial conditions and bootstrap-corrected estimates $\tilde{s}_{\mathrm{MLE}}$. Whiskers on the violin plots label the 2.5-, 50-, and 97.5-percentiles of their respective distributions. **(B)** ROC and **(C)** Precision-Recall curves for the estimation procedure to distinguish replicates under selection from those under neutrality.

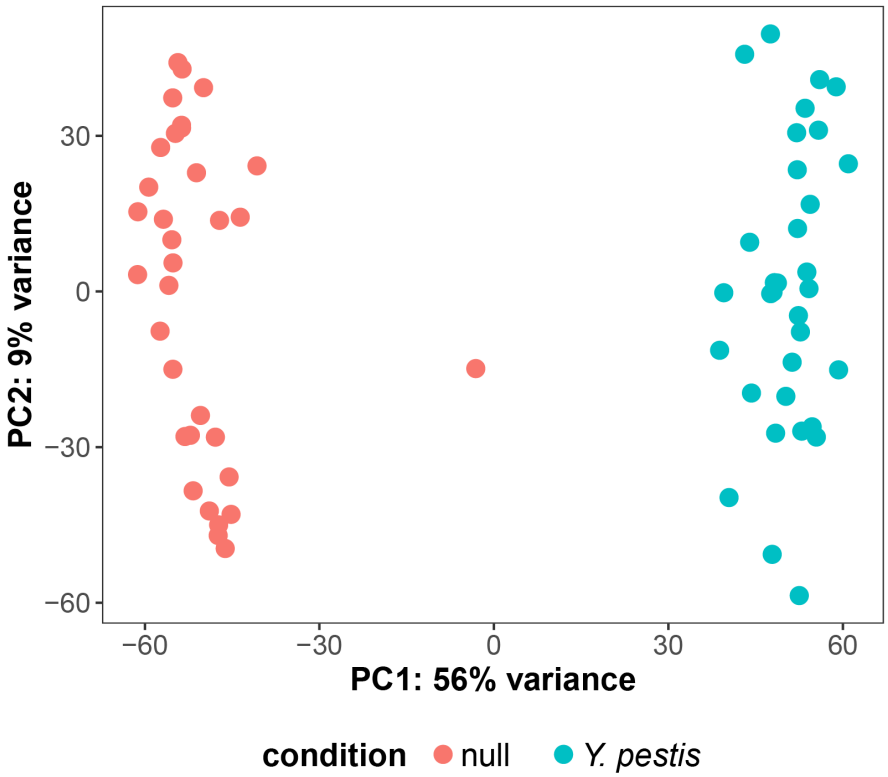

Figure S5: Principal components of gene expression for macrophages stimulated with heat-killed *Y. pestis*. The first principal component clearly separates stimulated samples from matched controls.

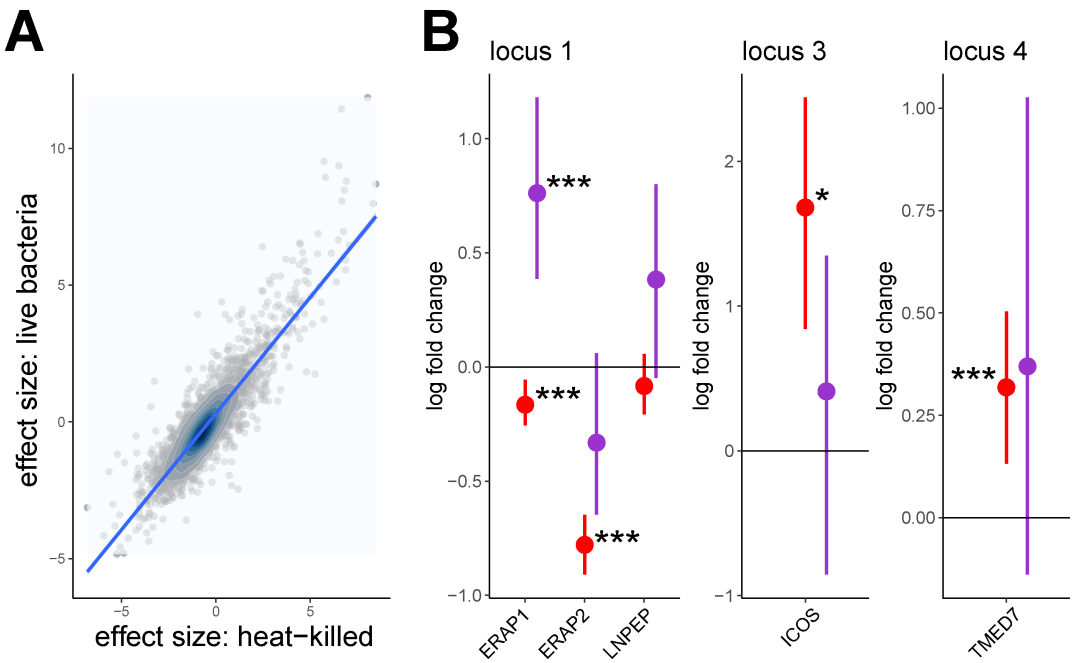

Figure S6. Response to *Y. pestis* is similar between macrophages stimulated with heat-killed and live *Y. pestis*. Effect size of *Y. pestis* stimulation compared between heat-killed *Y. pestis* (x-axis, n=33 individuals) and live *Y. pestis* (y-axis, n=8 individuals). **(A)** shows all genes, with a blue line representing the best fit line (r = 0.88). **(B)** compares effect sizes at genes near candidates for positive selection profiled in both expression datasets (red: heat-killed; purple: live bacteria). Error bars represent the standard error in estimating the effect size. The direction of effect is consistent except for *LNPEP* (which is not significant in either analysis) and *ERAP1*. *CTLA4* and *TICAM2* are now shown because there were not expressed at sufficiently highly levels in the macrophages from the 8 individuals infected with live *Y. pestis*. Asterisks placed near the point estimate of each value represent the significance: *** p < 0.001; ** p < 0.01; * p < 0.05.

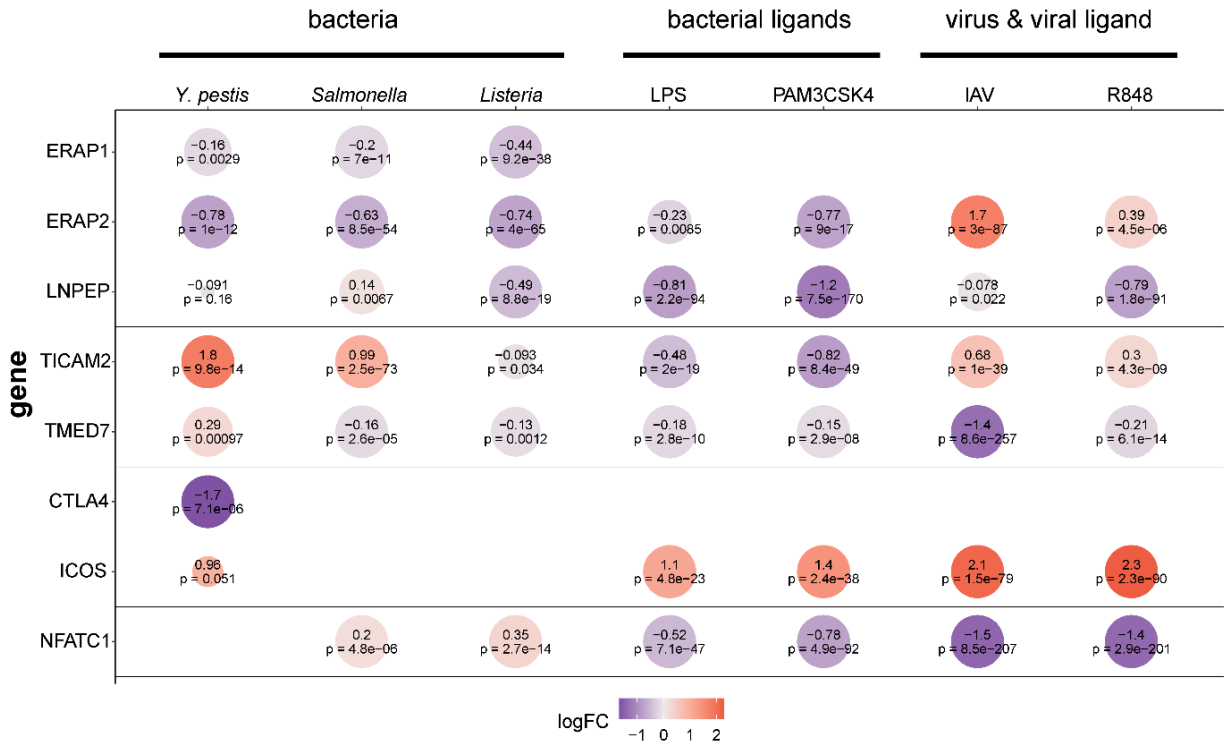

Figure S7. Transcriptional changes of genes nearby candidate loci in response to bacterial and viral stimuli. Data are derived from Nedelec *et al*.[30] and Quach *et al*. Nedelec *et al*.[31] measured the gene expression response of monocyte-derived macrophages to infection with two live intracellular bacteria: *Listeria monocytogenes* (a Gram-positive bacterium) and *Salmonella typhimurium* (a Gram-negative bacterium). Quach *et al*. characterized the transcriptional response at 6 hours of primary monocytes to bacterial and viral stimuli ligands activating Toll-like receptor pathways (TLR1/2, TLR4, and TLR7/8) and live influenza virus. The data for *Y. pestis* are the fold change responses observed in response to heat-killed bacteria. A negative estimate in plot (purple) indicates that the gene is downregulated and a positive value (red) indicates that the gene is up-regulated. The statistical support for the reported changes is given by the associated *p* values. Larger circle sizes represent smaller *p* values and empty circles refer to cases where that the gene was not expressed in that dataset.

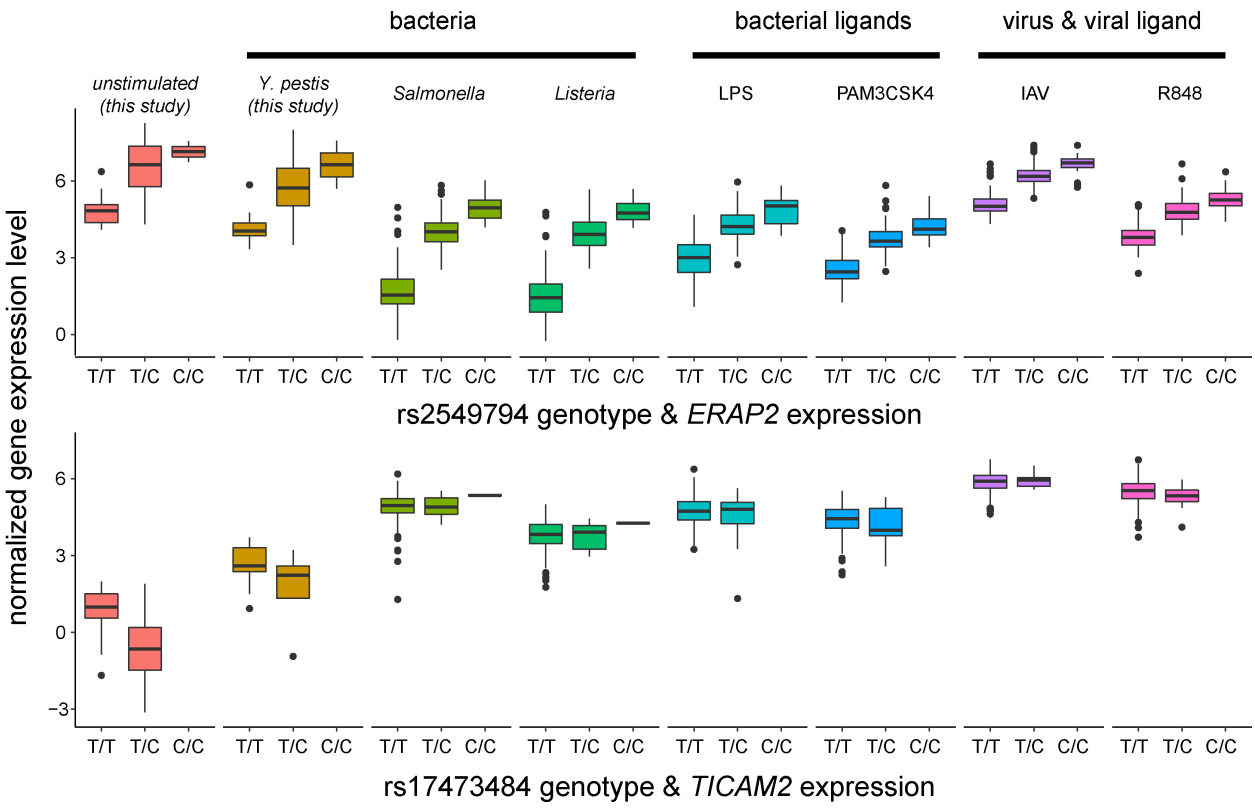

Figure S8. Genotype effects on transcription at candidate loci. Effect of genotype at nearby loci on the expression of *ERAP2* (top) and *TICAM2* (bottom), across experimental conditions in this study and previously published[30,31]. For *ERAP2*, the protective "T" haplotype increases expression in all conditions ($p < 0.001$). For *TICAM2*, the protective reference haplotype decreases expression only in the unstimulated condition ($p = 2.5 \times 10^{-6}$; $p > 0.05$ in all other conditions).

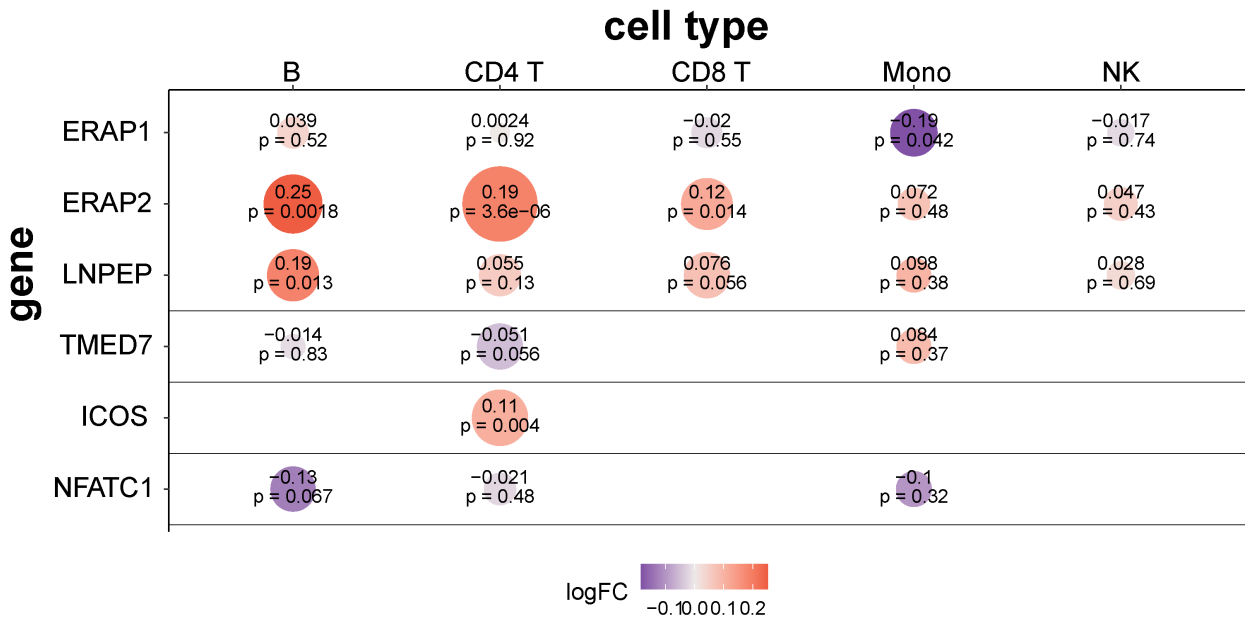

Figure S9. Transcriptional changes of genes nearby candidate loci in response to *Y. pestis* infection across cell types. For each cell type profiled using single-cell RNA sequencing, we show the effect of *Y. pestis* infection upon gene expression. A negative estimate (purple) indicates that the gene is downregulated and a positive value (red) indicates that the gene is up-regulated. The statistical support for the reported changes is given by the associated *p* values. Larger circle sizes represent smaller *p* values and empty circles refer to cases where that the gene was not expressed in that cell type.

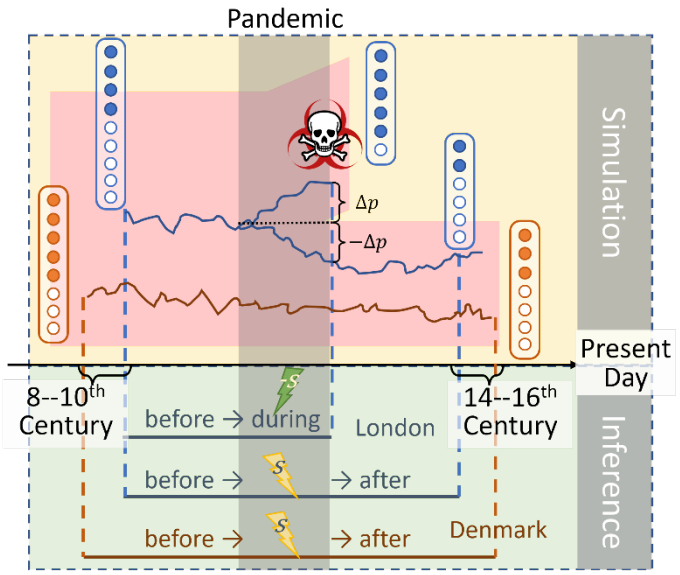

Figure S10: Schematics for estimating selection coefficients. The time axis serves as an approximate reference of the relative sampling times for the empirical samples. Dashed vertical lines indicate the relative sampling time for each group of samples considered in the analysis, and the floating boxes represent pools of samples from a bi-allelic locus. Above and below the time axis are sketches that respectively correspond to the simulation scheme and the likelihood computations. The shaded red horizontal tree represents the population-continuity along approximate time (x-axis), with the Black Death pandemic occurring in the dark shaded period. The shortened branch with a skull at the end represents people who died of the disease. In each simulated replicate, Δp and $-\Delta p$ mark the respective changes of allele frequency during the pandemic in the mid-pandemic and post-pandemic sample pools. In the inference schematics, each horizontal straight line represents a sampling scheme from which a likelihood was computed. Lightning bolts labeled with s or −s represent the selection coefficients.

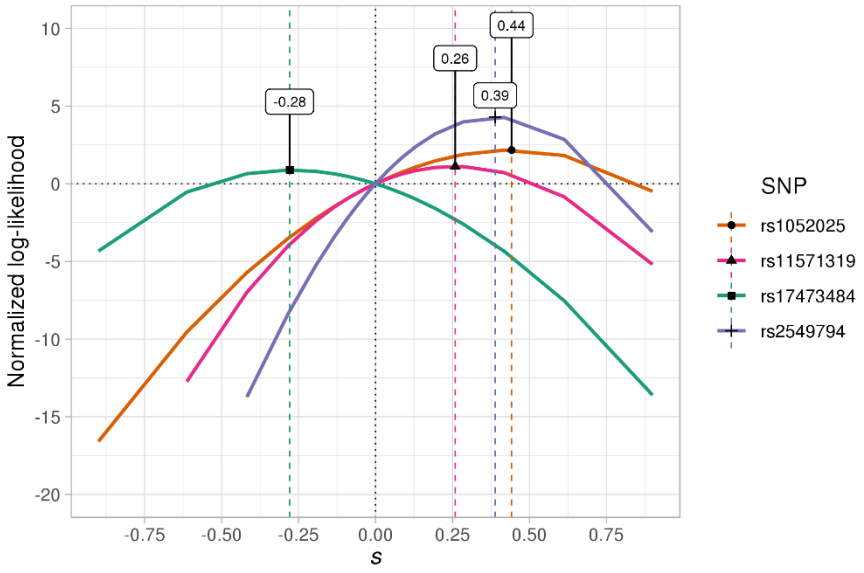

Figure S11: Log-likelihood as a function of the selection coefficient *s* for each target SNP. Labeled numbers show the approximate values of respective interpolated $\hat{s}_{\mathrm{MLE}}$. The log-likelihood is centered to 0 for *s* = 0.

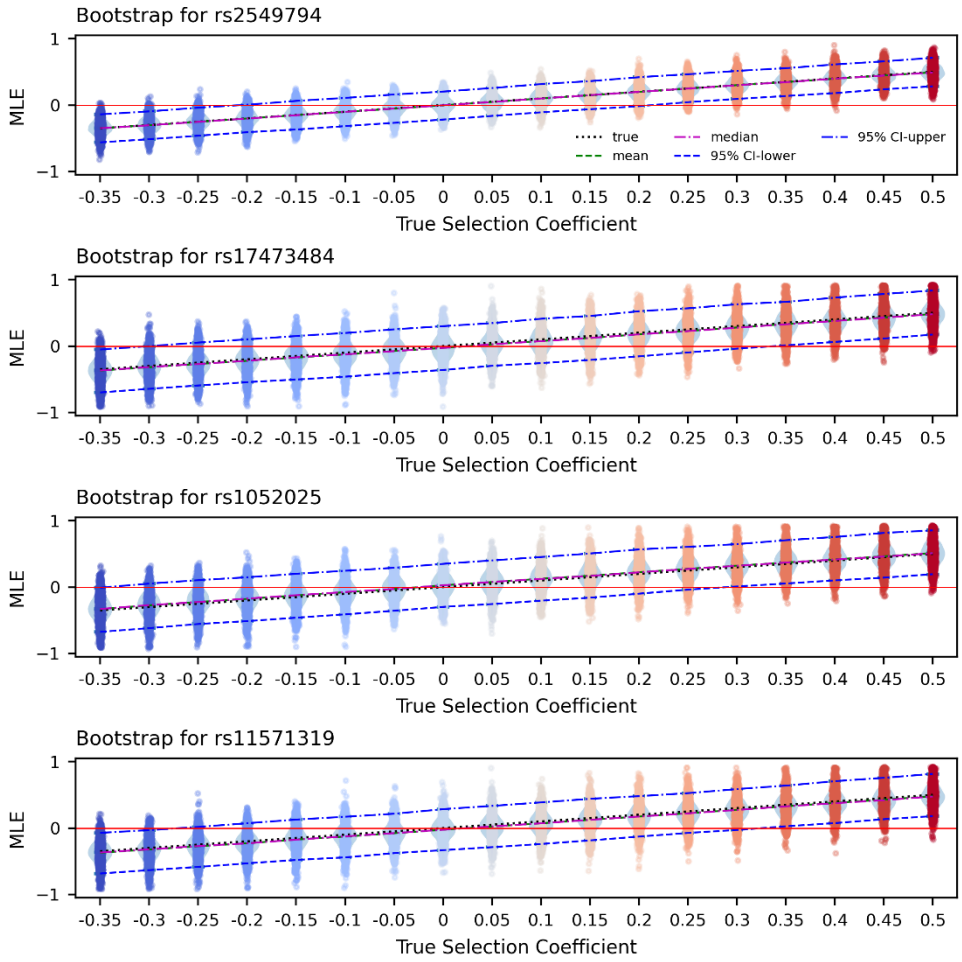

Figure S12: Distributions of $\hat{s}_{\mathrm{MLE}}$ for simulated data starting with the pre-pandemic frequencies of the four target SNPs. X-axes label the values of the true s used in the simulations, and y-axes show $\hat{s}_{\mathrm{MLE}}$. The black doted lines are the y = x reference lines. The green dashed lines and red dash-dotted lines indicate the means and medians of the $\hat{s}_{\mathrm{MLE}}$. The blue dashed and dot-dashed lines represent the 2.5-th and 97.5-th percentiles of the $\hat{s}_{\mathrm{MLE}}$ distributions.