Analysis of genomic DNA from medieval plague victims suggests long-term effect of *Yersinia pestis* on human immunity genes

Alexander Immel[1,2,3], Felix M. Key[1,4*], András Szolek[5*], Rodrigo Barquera[1*], Madeline K. Robinson[6], Genelle F. Harrison[6], William H. Palmer[6], Maria A. Spyrou[1,3], Julian Susat[2], Ben Krause-Kyora[2], Kirsten I. Bos[1,3], Stephen Forrest[3], Diana I. Hernández-Zaragoza[1,7], Jürgen Sauter[8], Ute Solloch[8], Alexander H. Schmidt[8], Verena J. Schuenemann[3,9], Ella Reiter[3,9], Madita S. Kairies[10], Rainer Weiß[11], Susanne Arnold[11], Joachim Wahl[10,11], Jill A. Hollenbach[12], Oliver Kohlbacher[5,13,14,15,16], Alexander Herbig[1,3], Paul J. Norman[6#], and Johannes Krause[1,17,3#]

[1] Max Planck Institute for the Science of Human History, Kahlaische Strasse 10, 07745 Jena, Germany
[2] Institute of Clinical Molecular Biology, Kiel University, Rosalind-Franklin-Strasse 12, 24105 Kiel, Germany
[3] Institute of Archaeological Sciences, University of Tübingen, Rümelinstrasse 23, 72070 Tübingen, Germany
[4] Max Planck Institute for Infection Biology, Charitéplatz 1, 10117 Berlin, Germany
[5] Applied Bioinformatics, Dept. for Computer Science, University of Tübingen, Sand 14, 72076 Tübingen, Germany
[6] Division of Biomedical Informatics and Personalized Medicine, and Department of Immunology & Microbiology, University of Colorado, CO 80045, USA
[7] Immunogenetics Unit, Técnicas Genéticas Aplicadas a la Clínica (TGAC), Mexico City, Mexico
[8] DKMS, Kressbach 1, 72072 Tübingen, Germany
[9] Institute of Evolutionary Medicine, University of Zurich, Winterthurerstrasse 190, 8057 Zurich, Switzerland
[10] Institute for Archaeological Sciences, WG Palaeoanthropology, University of Tübingen, Rümelinstrasse 23, 72070 Tübingen, Germany
[11] State Office for Cultural Heritage Management, Stuttgart Regional Council, Berliner Strasse 12, 73728 Esslingen, Germany
[12] UCSF Weill Institute for Neurosciences, Department of Neurology, University of California, San Francisco, USA
[13] Institute for Bioinformatics and Medical Informatics, University of Tübingen, Sand 14, 72076 Tübingen, Germany
[14] Quantitative Biology Center, University of Tübingen, Auf der Morgenstelle 10, 72076 Tübingen, Germany
[15] Translational Bioinformatics, University Hospital Tübingen, Sand 14, 72076 Tübingen, Germany
[16] Biomolecular Interactions, Max Planck Institute for Developmental Biology, Max-Planck-Ring 5, 72076 Tübingen, Germany
[17] Max Planck Institute for Evolutionary Anthropology, Deutscher Platz 6, 04103 Leipzig, Germany

*equal contribution

#corresponding authors

Abstract

Pathogens and associated outbreaks of infectious disease exert selective pressure on human populations, and any changes in allele frequencies that result may be especially evident for genes involved in immunity. In this regard, the 1346-1353 *Yersinia pestis*-caused Black Death pandemic, with continued plague outbreaks spanning several hundred years, is one of the most devastating recorded in human history. To investigate the potential impact of *Y. pestis* on human immunity genes we extracted DNA from 36 plague victims buried in a mass grave in Ellwangen, Germany in the 16th century. We targeted 488 immune-related genes, including *HLA*, using a novel in-solution hybridization capture approach. In comparison with 50 modern native inhabitants of Ellwangen, we find differences in allele frequencies for variants of the innate immunity proteins Ficolin-2 and NLRP14 at sites involved in determining specificity. We also observed that *HLA-DRB1*13* is more than twice as frequent in the modern population, whereas *HLA-B* alleles encoding an isoleucine at position 80 (I-80+), *HLA C*06:02* and *HLA-DPB1* alleles encoding histidine at position 9 are half as frequent in the modern population. Simulations show that natural selection has likely driven these allele frequency changes. Thus, our data suggests that allele frequencies of *HLA* genes involved in innate and adaptive immunity responsible for extracellular and intracellular responses to pathogenic bacteria, such as *Y. pestis*, could have been affected by the historical epidemics that occurred in Europe.

Introduction

Throughout evolution, humans have likely experienced multiple major episodes of infectious disease. Of exceptional virulence and lethality, *Yersinia pestis* has been responsible for at least three major plague pandemics during the last few millennia. Studies of ancient DNA have confirmed *Yersinia pestis* caused widespread infections in Europe from the Late Neolithic period, nearly 5,000 years ago, until the 18th century AD (Andrades Valtuena et al., 2016; Bos et al., 2016; Bos et al., 2011; Feldman et al., 2016; Keller et al., 2019; Namouchi et al., 2018; Rascovan et al., 2019; Rasmussen et al., 2015; Spyrou et al., 2018; Wagner et al., 2014). Historical records show the first pandemic began with the Justinianic Plague in the 6th century AD and lasted until the 8th century, the second began with the 1346-1353 Black Death and continued with thousands of local plague outbreaks until the 18th century (Biraben, 1976; Büntgen, Ginzler, Esper, Tegel, & McMichael, 2012) and the third pandemic started in China in the 19th century AD and spread the pathogen worldwide lasting up until the mid-20th century (Morelli et al., 2010; Politzer & WHO, 1954). Of the three recorded pandemics, the Black Death claimed up to half of the European population during its five-year period (Benedictow, 2004). Although *Y. pestis* is now absent from most of Europe, it still causes sporadic infections among humans in the Americas, Africa and Asia, usually transmitted by fleas from rodent populations that serve as plague reservoirs (Drummond et al., 2014; WHO, 2017). Although the lethality of plague is very high without treatment (WHO, 2017), it remains likely that specific individuals are protected from, or more susceptible to, severe disease through polymorphism in the determinants of natural immunity. In this case, any changes in allele frequencies that occurred during a given epidemic crisis could be evident as genetic adaptation and detectable in modern day individuals.

There are multiple examples of natural selection affecting human immunity-related genes that can be attributed to challenge by pathogens. These examples include specific pathogens causing malaria, cholera or Lassa fever, or to wider differences in pathogen exposure between geographically discrete populations (Harrison et al., 2019; Karlsson et al., 2013; F. M. Key et al., 2014; Kwiatkowski, 2005; McManus et al., 2017; Sabeti et al., 2007; Voight, Kudaravalli, Wen, & Pritchard, 2006; Wang, Kodama, Baldi, & Moyzis, 2006). The toll like receptors (TLRs) are innate immune proteins that detect the presence of specific pathogens to initiate an immune

3

90 response. Signatures of purifying selection have been identified within specific *TLR* genes that

91 correlate with distinct pattern specificities of the encoded allotypes (Barreiro et al., 2009). In

92 another example, recent signatures of positive selection in the *IFITM3* gene accompany

93 differential abilities of the alternative variants to control pandemic H1N1 influenza A virus

94 infection (Albright, Orlando, Pavia, Jackson, & Cannon Albright, 2008; Everitt et al., 2012). A

95 final example of human genetic adaptation to pathogens is the 32 base pair deletion in the

96 chemokine receptor *CCR5* (*CCR5-Δ32*), which prevents HIV from entering and infecting human

97 T-cells (Dean et al., 1996). Although once postulated as a plague resistance allele (Stephens et

98 al., 1998), there is little evidence for positive selection acting on *CCR5-Δ32* (Sabeti et al., 2006).

99 By contrast, the Major Histocompatibility Complex (*MHC*), which encodes multiple immunity-

100 related genes including the Human Leukocyte Antigen (*HLA*) molecules, does show evidence for

101 recent positive and balancing selection and has established roles initiating and directing the

102 immune response to infection (Klebanov, 2018; Parham & Moffett, 2013; Prugnolle et al., 2005;

103 Trowsdale & Knight, 2013).

104

105 Here, we extracted genomic DNA from 36 individuals who apparently died from plague (*Y.*

106 *pestis*) in Ellwangen in Southern Germany during the 16th century. We also extracted DNA from

107 50 modern day Ellwangen inhabitants. We then compared their frequency spectra for a large

108 panel of immunity-related genes. We observed evidence for pathogen-induced changes in allele

109 distributions for two innate pattern-recognition receptors and four *HLA* molecules. We propose

110 that these frequency changes could have resulted from *Y. pestis* plague exposure during the 16th

111 century.

112

113 **Results**

114 *Archaeological and Anthropological Findings*

115 Ellwangen is a small town of 27,000 inhabitants situated in South Germany near the border of

116 Baden-Wuerttemberg and Bavaria. Ellwangen was founded in the 7th century AD, with only a

117 few hundred inhabitants until modern times. The town was affected by multiple plague outbreaks

118 during the 16th and 17th centuries (Ellwangen, 2007). From 2013 to 2015 an excavation took

119 place in Ellwangen during the restoration of the town's market square (Figure 1). Three mass

120 graves were discovered with a total of 101 inhumated remains (Supplementary Figure 1A).

4

121 Consistent with 16th century bubonic plague predominantly affecting children (Bowsky, 1971;
122 Clouse, 2002; Cohn, 2003) only 23 of the individuals had reached adult age (Supplementary
123 Note 1). The individuals were buried close to each other and there was little sediment between
124 the distinct layers. The proximity, as well as radiocarbon dating, suggests that all three mass
125 burials were created during the same epidemic crisis event during the 16th century
126 (Supplementary Figure 1B). Genomic DNA from *Y. pestis* was identified previously from 13 of
127 the individuals, and the complete genome of a strain consistent with the era was reconstructed
128 from one of them (Spyrou et al., 2019; Spyrou et al., 2016). We also performed shotgun
129 sequencing directly on DNA libraries prepared from tooth samples of 30 distinct individuals, and
130 pathogen screening using the metagenomic alignment tool MALT (Vågene et al., 2018)
131 identified reads matching to *Y. pestis* in 25 of them, with aDNA characteristic terminal
132 substitutions in samples with sufficient coverage (Supplementary Table 1), confirming that the
133 reads are of ancient DNA origin. With exception of one sequence read of Hepatitis B Virus in
134 one of the petrous bone samples (ELW012), no evidence of other pathogens was detected.
135 Additional archaeological and anthropological findings suggest little physical trauma, albeit poor
136 health condition prior to death, which can likely be considered as normal health and nutrition
137 status for people living in that time period (Supplementary Note 1). Taken together, these
138 findings strongly suggest that these individuals were victims of a single *Y. pestis* plague outbreak
139 that occurred during the 16th century.

140

141 *The 16th century Ellwangen plague victims display genetic similarity with modern inhabitants*
142 From the 16th century mass grave site in Ellwangen, we successfully extracted DNA from 40
143 petrous bones (Supplementary Figure 1C) and four teeth. DNA of sufficient quality and quantity
144 for genome wide sequence analysis was obtained from all samples. An average of 1.76 million
145 unique, human genome reads per individual was generated by shotgun sequencing
146 (Supplementary Data 1A). Kinship analysis revealed three pairs of individuals to be first-degree
147 relatives (Supplementary Data 1B, Supplementary Figure 2). In order to obtain the most accurate
148 frequency distributions, one in each pair of the directly related individuals was removed from the
149 allele frequency calculations. In these cases, the individual having the lowest yield of sequence
150 reads was excluded (Material and Methods). In addition, one individual who was second-degree
151 related to two of the other individuals was also excluded (ELW030). We also obtained genomic

DNA samples from 51 contemporary inhabitants of Ellwangen and shotgun-sequenced them with an average of 2.74 million unique human genome reads per individual (Supplementary Data 1A). Here, we identified a single pair of first-degree relatives and removed one individual. In order to test whether the two cohorts derive from a single continuous population, we tested for population genetic similarity using Principle Component Analysis (PCA) (Patterson, Price, & Reich, 2006) and ADMIXTURE analysis (Alexander, Novembre, & Lange, 2009). Showing that the 16th century and modern groups indeed are genetically very similar, we found that the 16th century Ellwangen plague victims form a tight cluster in PCA space, which overlaps with the modern inhabitants (Figure 2A). This finding is bolstered by the highly similar genetic ancestry composition of the two groups as illustrated by their population admixture proportions (Figure 2B). This latter finding is important because recent demographic changes could alter allele frequencies of the modern compared with the 16th century group (Hellenthal et al., 2014).

Two immunity-related genes harbor strongly differentiated SNPs

In order to compare the allele spectra of immunity-related genes in the 16th century *Y. pestis* plague victims with modern-day inhabitants of Ellwangen, we developed an in-solution hybridization capture approach to enrich for 488 human genes implicated in immunity (Supplementary Table 2). This approach allowed us to specifically target the genes of interest while reducing the amount of sequencing required, leading to an average of 308 times more reads on target compared to undirected genome wide sequencing (Supplementary Figure 3). We applied this 'immunity capture' method to all 16th century and modern DNA samples. The targeted genes were covered with a mean read depth of 55.8 (Supplementary Data 2). We investigated the allele spectra of the 488 immunity-related genes by leveraging a branching statistic, *Differentiation with Ancestral* (*DAnc*) (F.M. Key, Fu, Romagné, Lachmann, & Andrés, 2016). *DAnc* is calculated per site and uses derived allele frequency estimates across three populations; 16th century and modern Ellwangen, and a non-European outgroup (we used Han Chinese from Beijing (Abecasis et al., 2012)). *DAnc* scores can range from -1 to +1, and those in the respective far tails of the distribution identify candidates for simulation studies that could indicate positive selection has occured. We established the expected distribution of *DAnc* scores (Supplementary Data 3) under neutrality through simulations using a human demographic model (Gravel et al., 2011).

183

In our analysis, the distribution of *DAnc* scores closely matches between the simulated and test data (Supplementary Figure 4, Supplementary Table 3). In the far tail of the distribution (>99.9%) we observed three SNPs, two in the Ficolin-2 (FCN2) gene and one in the NOD-like receptor purine domain containing 14 (NLRP14) gene (Table 1) also corresponding to the greatest F_{ST} values among the 488 genes for the same three SNPs (Supplementary Data 4). F_{ST} is an established measure for population differentiation and corrects for expected heterozygosity and sampling error (Weir & Cockerham, 1984). However, due to the ascertainment of SNPs, which are not representative of the whole genome, the far tail of the observed DAnc or F_{ST} distribution is no evidence alone for positive selection. Alternatively, we compared the fraction of SNPs observed in the far tail of the simulated and test distribution, which suggest no enrichment of SNPs in our test data and thus no evidence for positive selection using the data at hand (Supplementary Table 3). Further analyses using a larger sample size and whole genome data is necessary in order to understand the role of positive selection due to historic epidemics.

197

The identified SNPs of FCN2 are a 5' UTR promoter variant [rs17514136 (-4 A to G)] and one coding change variant [rs17549193 (717 C to T; 236 Thr to Met)]. The UTR and coding change variants occur in complete linkage disequilibrium (Δ'= 1.0, R^2=0.9), and appear to represent a single haplotype that has risen in frequency in the modern population. Interestingly, FCN2 binds to specific molecules on the surface of bacteria, triggering the complement pathway to neutralize the pathogen (Hoang et al., 2011; Luo et al., 2013). The promoter variant is associated with increased serum concentration of FCN2 (Cedzynski et al., 2007), whereas polymorphism at residue 236 (rs17549193) affects binding to the target bacteria (Hummelshoj et al., 2005). Similarly, NLRP14 belongs to inflammasome complex proteins, which are intracellular pattern recognition receptors that trigger local and systemic responses to microbial invasion (Martinon, Burns, & Tschopp, 2002). Inflammasomes are implicated in the immune response to *Yersinia* infection, amongst other pathogens (Philip, Zwack, & Brodsky, 2016; Vladimer, Marty-Roix, Ghosh, Weng, & Lien, 2013). The NLRP14 SNP is a coding change variant (rs10839708 [2745 G to A: 808 Glu-Lys]) that occurs in the leucine-rich repeat (LRR) domain, which in related molecules controls the ligand specificity (Inohara, Chamaillard, McDonald, & Nunez, 2005). Thus, in summary we show immune-related genes have no significant frequency changes

214   between 16th century *Y. pestis* victims and modern Ellwangen inhabitants.

215

216   *No evidence for role of CCR5-Δ32 in protection from Y. pestis infection*

217   We investigated the Δ32 deletion in the CCR5 locus (chr3:46414947-46414978), which was

218   included in our target regions because this mutation has previously been suggested as protective

219   from the plague. We found that CCR5-Δ32 has a frequency of 16.6% in the 16th century

220   compared to 10.8% in the modern individuals (p=0.27) and 11.2% in Germany (Supplementary

221   Data 5A and 5B, Supplementary Data 6A, Table 2). Consistent with epidemiological modeling

222   and lack of evidence that CCR5 can serve as a *Y. pestis* receptor (Galvani & Slatkin, 2003) this

223   finding suggests that the CCR5-Δ32 mutation provided no protection from *Y. pestis*. Similarly

224   we also investigated SNPs *rs4986790*, *rs4986791* within the gene *TLR4* previously suggested to

225   be associated with resistance to *Y. pestis* (Al Nabhani, Dietrich, Hugot, & Barreau; Laayouni et

226   al., 2014). However, we did not find any significant differences in their respective frequencies

227   (Supplementary Table 4).

228

229   *Natural selection has increased HLA-DRB*13 and reduced HLA-B*51 and -C*06 frequencies in*

230   *modern individuals*

231   With more than 28,000 distinct alleles described (Robinson et al., 2015), HLA molecules are

232   encoded by the most polymorphic gene complex in humans. When human populations are

233   exposed to novel diseases through contact with populations or environments they had not

234   encountered previously, changes in *HLA* allele frequencies can occur rapidly (Lindo et al., 2016;

235   Patin et al., 2017). Consequently, the signatures of balancing selection in the genomic region that

236   contains *HLA* are consistently the strongest in the genome (Quintana-Murci, 2019; Sabeti et al.,

237   2006), and specifically correspond to amino acid residues that bind peptide fragments derived

238   from pathogens (Bjorkman & Parham, 1990). Significant shifts in *HLA* allele frequencies can

239   thus reveal evidence of natural selection for specific pathogen resistance. We were able to

240   identify *HLA class I (-A, -B, -C)* and *HLA class II (-DPA1, -DPB1, -DQA1, -DQB1* and *-DRB1)*

241   genotypes from all of the 16th century and modern inhabitants of Ellwangen. We observed a total

242   of 86 distinct *HLA class I* alleles, 66 distinct *HLA class II* alleles and 168 distinct *HLA*

243   haplotypes (Supplementary Data 6B). The most frequent haplotype (*HLA-*

244   *A*01:01~B*08:01~C*07:01~DRB1*03:01*) is the same in both groups and is also the most

common and widespread across Europe today (Darke et al., 1998; Dunne, Crowley, Hagan, Rooney, & Lawlor, 2008; Johansson, Ingman, Mack, Ehrlich, & Gyllensten, 2008; Nowak et al., 2008; Pingel et al., 2013). Thus the diversity and composition of *HLA* haplotypes appears as expected for Northern European populations (Alfirevic et al., 2012), and we did not observe any significant differences in their frequencies between the 16th century and modern individuals.

By contrast to the haplotype distributions, on examining the individual *HLA class I* genes, we observed that the *B*51:01* allele of *HLA-B* decreased from 15.3% in the 16th century Ellwangen plague victims to only 6.0% (p=0.04 (p-corrected = NS); DANc= -0.093) in the modern Ellwangen population (Table 3, Supplementary Data 4, Supplementary Data 5A). Similarly, the *C*06:02* allele of *HLA-C* decreased from 13.9% to 5% (p=0.04 (p-corrected = NS); DANc= 0.053). *HLA-B*51:01* and *-C*06:02* are not in linkage disequilibrium in either population (Supplementary Data 6B), and so these two observations are independent. In addition, although there were no significant frequency differences observed for any *HLA class II* alleles as determined at two-field resolution, we observed that all allotypes present representing the DR13 serological group (Holdsworth et al., 2009) were at substantially lower frequency in the 16th century than modern Ellwangen population. Accordingly, by considering them together, there was an increase in *DR13* frequency from 5.6% in the 16th century to 17.0% in the modern individuals (p=0.026, Table 3, Supplementary Data 5A). Repeating this analysis for all the major DRB1 lineages present (Holdsworth et al., 2009), showed DRB1*13 as the only allotype differing in frequency between the two groups (Table 3). We used Wilson Score Interval estimation of the 95% binomial confidence interval. The 95% CI of HLA-B*51:01 was 0.09 – 0.25 (observed = 0.06), the 95% CI of HLA-C*06:02 was 0.08 – 0.24 (observed = 0.05), and DRB1*13 was 0.02 – 0.13 (observed = 0.16). Thus, for each of the three HLA allotypes showing distinctions between modern and 16th century inhabitants of Ellwangen, the observed modern allele frequencies are outside the 95% binomial confidence intervals surrounding sampling of the 16th century allele frequencies. We further validated these findings by comparing the *HLA* allele frequencies observed in the Ellwangen individuals with a large panel (N=8,862) of unrelated bone marrow donor registry volunteers gathered from all of Germany (Supplementary Data 5B). Whereas there were no significant allele frequency differences when comparing modern inhabitants of Ellwangen with modern Germany as a whole, we observed significantly lower

276 frequencies of *B*51:01* in modern Germany (5.5%) than the Ellwangen plague victims (15.3%),
277 when applying a pairwise proportion test (p=0.005; DANc= -0.098). We also observed
278 differences in *HLA-C*06* and *DRB1*13* between the plague victims and modern Germany, but
279 these were not statistically significant (Supplementary Data 5B).

281 To distinguish if the changes in frequencies of *B*51:01*, *C*06:02* and *DRB1*13:01* were more
282 likely to be due either to natural selection or genetic drift we performed forward time simulations
283 by starting from the observed polymorphisms in the 16th century Ellwangen and modelling
284 neutrality for the last 500 years. This way it was possible to start from reasonable levels of
285 genetic variation without the necessity to determine the impact of ancient selection on HLA and
286 episodic turnover of HLA alleles. Moreover, this way each allotype could be tested individually.
287 Again, we observed an overall concordance between median frequencies of the simulated neutral
288 alleles and the modern Ellwangen allele frequencies, as is expected under genetic drift. By
289 contrast, the allele frequencies of *B*51:01*, *C*06:02* and *DRB1*13:01* observed for modern
290 inhabitants of Ellwangen were in the extreme tails of their respective distributions (p^{sim}=0.006,
291 *0.004*, *<0.001*, respectively, Figure 3), suggesting natural selection likely drove the change in
292 these allele frequencies. A similar significant shift was observed when we considered DR13
293 broadly (p^{sim}<0.001). To quantify the selection coefficient (*s*) responsible for these changes, we
294 performed the simulations incorporating selection, mirroring the timeline of the plague, across a
295 range of *s* values. We identified an *s* equal to -0.25 was most likely to produce the observed
296 decrease in *B*51:01* alleles as well as an *s* equal to -0.27 in case of *C*06:02*. An *s* of 0.37 was
297 most likely to cause the increase in *DRB1*13:01* (Supplementary Figure 5). Notably, these
298 values are within the range of previously reported values of *s* acting on MHC (Radwan, Babik,
299 Kaufman, Lenz, & Winternitz, 2020).

301 *Higher incidence of KIR3DL1 interaction with HLA-B in plague victims than modern inhabitants*
302 *of Ellwangen*
303 The binding specificity of HLA allotypes, and thus their function and distinctiveness, is
304 determined by specific amino acid residues in the alpha-helix of the molecule. Polymorphism of
305 these amino acid residues is associated with autoimmune diseases and response to pathogens
306 (Achkar et al., 2012; Hammer et al., 2015; Hollenbach et al., 2019; Sun et al., 2018). We

10

307 identified three of these residues having significant (p<0.05) differences in frequency between
308 the 16th century victims and modern individuals (Supplementary Data 7A). We observed
309 histidine (H) at position 9 of HLA-DPB1 to be approximately three times more frequent in the
310 16th century (13%) than the modern (4%) individuals (p=0.03 (p-corrected = NS); DANc= 0.13);
311 Supplementary Data 7A). We also observed isoleucine (I) at position 80 (I-80) in HLA-B twice
312 as frequently in the 16th century (28%) than in the modern individuals (15%) (p=0.04 (p-
313 corrected = NS); DANc= 0.28), and aspartic acid (D) at position 114 in HLA-C more frequently
314 in the 16th century (85%) than the modern (70%) individuals (p=0.02 (p-corrected = NS);
315 DANc= -0.062); Supplementary Data 7A). Residue D-114 is located in the outward-facing
316 groove of HLA-C and its variation can directly affect the sequence of endogenous peptides able
317 to bind (Di Marco et al., 2017). Since *HLA-B* alleles encoding I-80 are most commonly observed
318 on haplotypes that also have *HLA-C* alleles that encode D-114 (Cao et al., 2001), it is likely that
319 the observed frequency difference at this position is driven by linkage disequilibrium with *HLA-*
320 *B*. HLA-B*27:02, -B*38:01, -B*49:01, -B*51:01, -B*52:01, -B*57:01 and -B*58:01 are all I-
321 80[+] allotypes that are more frequent in the 16th century than the modern inhabitants of Ellwangen
322 (Supplementary Data 7B and 7C), together accounting for the observed difference in I-80
323 frequency (Table 3). Thus, it is likely that the significant difference in frequencies we observed
324 for HLA-B*51:01 can be attributed to the fact that it possesses an isoleucine at position 80. We
325 next tested whether the observed frequency changes in HLA-DPB1 H-9 and HLA-B I-80 were
326 more likely due to genetic drift or natural selection, using neutral forward genetic simulations as
327 above. In both cases, we found these allele frequency shifts were unlikely to be observed unless
328 natural selection was included in the model (HLA-DPB1 H-9 $p^{sim}=0.014$, HLA-B I-80
329 $p^{sim}=0.002$).

330

331 *KIR* genes encode surface proteins on natural killer (NK) cells whose interaction with HLA class
332 I molecules can determine the outcome of NK cell responses (Guethlein, Norman, Hilton, &
333 Parham, 2015). For example, polymorphism of residue 80 in HLA-B controls its ability to bind
334 to KIR3DL1, with I-80 defining ligand specificity and permitting the strongest interaction
335 (Saunders et al., 2015). We therefore sought to determine whether the observed high frequency
336 of *HLA-B I-80*[+] alleles in the 16th century samples affects the frequency of HLA-B interaction
337 with KIR3DL1. The *KIR* region varies by gene content (Uhrberg et al., 1997), and we were able

338 to determine this diversity across all the 16th century individuals (Supplementary Figure 6). We

339 observed that 97% of the 16th century individuals possess at least one copy of *KIR3DL1*,

340 compared to 88% of the modern Ellwangen samples. These values are within the range observed

341 in modern European populations (Hollenbach, Nocedal, Ladner, Single, & Trachtenberg, 2012)

342 as well as those predicted in our neutral forward genetic simulations ($p^{sim}=0.258$). Thus, we

343 observe no statistically significant differences in *KIR3DL1* gene frequencies between the 16th

344 century and modern samples (Table 4A). Similarly, we observed no difference in *KIR3DL1*

345 allele frequencies (Supplementary Data 6C) between the 16th century and modern individuals

346 (Table 4A). By contrast, we found that *KIR3DL1* and *HLA-B I-80$^+$*, and thus their combined

347 genotype, is more frequent in 16th century (53%) than modern (26%) individuals (p=0.011 (p-

348 corrected = NS), Table 4B). Using simulations, we found that genetic drift was unlikely to

349 produce the modern day observed frequencies of *HLA-B I-80* and *KIR3DL1*, but that selection

350 against *HLA-B I-80* likely drove the decreased *HLA-B I-80$^+$/KIR3DL1$^+$* joint genotype

351 frequencies ($p^{sim}=0.002$).

352

353 **Discussion**

354 In this study we investigate a large panel of immunity-related genes from 36 individuals

355 discovered in three 16th century plague mass graves in Ellwangen, Southern Germany, and

356 compare them to 50 present-day inhabitants of Ellwangen. For this purpose, we developed a

357 targeted DNA capture protocol comprising 488 human immune system genes including the six

358 major *HLA class I* and *class II* genes and the *KIR* locus. We also compared the 16th century *HLA*

359 allele frequencies with sequence data of 8,862 potential stem cell donors registered with DKMS

360 (German Bone Marrow Donor Registry). Although we observe a predominant genetic stability of

361 human immune genes over at least five centuries in Central Europe, we find distinct allele

362 frequency changes in the *HLA* and in two other genes that encode components of innate

363 immunity.

364

365 Given its devastating effect, the *Y. pestis*-driven 2nd plague pandemic is a strong candidate for

366 exerting selection pressure on the human immune response (Laayouni et al., 2014; Lenski,

367 1988). We observed strong allele frequency differences at SNPs located in the *FCN2* and

368 *NLRP14* genes, albeit we find no clear evidence that positive selection has contributed to the

12

369   observed allele frequency differentiation. Both of these molecules are pattern recognition
370   receptors that bind specific pathogen-derived components to initiate the inflammation response;
371   Ficolin-2 does this extracellularly, and NLRP14 intracellularly. Ficolin-2 promotes phagocytosis
372   of pathogenic bacteria (Hoang et al., 2011; Luo et al., 2013). Interestingly, we observed two
373   SNPs of known direct functional effect to be in strong LD, forming a single haplotype that is
374   elevated in frequency in the modern compared to the 16th century individuals. This haplotype
375   both increases serum concentration and alters the binding properties of Ficolin-2 (Cedzynski et
376   al., 2007; Hummelshoj et al., 2005), which makes it a good candidate for providing improved
377   resistance to *Y. pestis* infection. On the other hand, less is known about NLRP14, which has
378   similar domain organization to other inflammasome proteins. Inflammasomes act to trigger
379   inflammation as well as self-destruction of infected cells (Lamkanfi & Dixit, 2014) and have
380   been identified recently as important mediators of the immune response to *Y. pestis* (Park et al.,
381   2020). Interestingly, the same variant we observed at lower frequency in modern individuals than
382   plague victims (K-808) was identified at high frequency due to positive selection in the Human
383   Genome Diversity-Project populations from East-Asia (Vasseur et al., 2012). Similar
384   inflammasome molecules, including NLRP3 and NLRP12, are known to respond to *Y. pestis*
385   (Vladimer et al., 2013; Vladimer et al., 2012), but may also be exploited by bacteria to inhibit
386   immunity (Anand et al., 2012; Philip et al., 2016; Zaki, Man, Vogel, Lamkanfi, & Kanneganti,
387   2014). K-808 is located in the LRR domain of NLRP14 and influences ligand specificity.
388   Therefore, the fluctuating frequencies of the variants at this position point to an evolutionary
389   battle between host and pathogen (Abi-Rached, Dorighi, Norman, Yawata, & Parham, 2007).
390   Functional tests are thus required to determine if mutation at residue 808 permits recognition of
391   any components of *Y. pestis*.

392   On examination of *HLA* alleles we observed candidates for natural selection of human adaptive
393   immune responses. HLA class I and II are cell surface molecules that bind to peptides derived
394   from intracellular or extracellular proteins, respectively. To trigger and drive the adaptive
395   immune response, these peptides are presented by the HLA molecules to T-cells. Antibody
396   production is elicited through highly polymorphic HLA class II molecules, HLA-DP, -DQ, and -
397   DR, presenting pathogen-derived peptides to CD4+ T-cells (Neefjes, Jongsma, Paul, & Bakke,
398   2011). Direct killing of infected cells can occur when any of three highly polymorphic HLA
399   class I molecules, HLA-A, -B or -C, presents pathogen-derived peptides to cytotoxic CD8+ T-

cells (Doherty & Zinkernagel, 1975). The *HLA-DRB1*13* allelic group increased in frequency from 5.6% in the plague victims to 17% in the modern Ellwangen individuals and 12% in the German bone marrow donors, potentially indicating antibody-driven protection from the plague for individuals having this allotype. *HLA-DRB1*13* is associated with resistance to *M. tuberculosis* (Dubaniewicz, Lewko, Moszkowska, Zamorska, & Stepinski, 2000). Similar to *Y. pestis*, *M. tuberculosis* can invade and survive within macrophages (Pieters, 2008). Macrophages express high levels of HLA class II and are cells that are specialized for presenting peptides to CD4+ T cells to initiate antibody production. These HLA class II molecules can present antigens from intracellular pathogens, such as *M. tuberculosis* (Ankley, Thomas, & Olive, 2020). Thus, the same adaptive immune pathway triggered by HLA-DRB1*13 that provides resistance to *M. tuberculosis*, might also provide resistance to *Y. pestis*.

Some HLA class I allotypes interact with killer-cell immunoglobulin-like receptors (KIRs) to modulate the function of Natural Killer (NK) cells, which are essential components of innate immunity, providing front-line defense against infection (Guethlein et al., 2015; Long, Kim, Liu, Peterson, & Rajagopalan, 2013). The *KIR* locus varies by gene content (Uhrberg et al., 1997; Wilson et al., 2000) and is located on a separate chromosome (chr19) to HLA (chr6). Combinatorial diversity of HLA class I and KIR allotypes directly impacts NK cell responses to infection (Bashirova, Martin, McVicar, & Carrington, 2006; Parham & Moffett, 2013). We observed a lower frequency of the HLA-B allotypes that can interact with KIR3DL1 in the modern individuals than we did in the plague victims, suggesting this combination could have been disadvantageous for individuals infected with *Y. pestis*. KIR3DL1 is an inhibitory receptor that enables NK cells to respond strongly to changes in HLA expression by infected cells (Boudreau & Hsu, 2018; Saunders et al., 2015) (Gumperz, Litwin, Phillips, Lanier, & Parham, 1995). Finally, we observed a marked decrease in the frequency of *HLA-C*06:02* when comparing the 16th century and modern Ellwangen populations. *HLA-C*06:02*, which also interacts strongly with KIR (Hilton et al., 2015), is strongly associated with psoriasis (Ogawa & Okada, 2020), an immune-mediated disease. These observations implicate excess collateral damage caused by NK cells responding to infection (Guo, Patil, Luan, Bohannon, & Sherwood, 2018; Kim et al., 2008), as a potential mechanism of pathology.

431   A limitation of this study is the relatively small sample size of 36 plague victims from the 16th

432   century. As suggested by our effect size analyses (Supplementary Figure 7), with the given

433   sample size only large effects (w = 0.40 - 0.45) can be detected, and therefore, the observed

434   frequency changes do not withstand multiple testing correction. Thus, it also remains possible

435   that the signals of selection we detected for some variants are caused by drift and/or sampling

436   biases, and, on the other hand, some other variants under selection were potentially not targeted

437   through this approach. Increasing the sample size in future studies will allow addressing this

438   caveat. Moreover, all tested individuals are 16th century late plague victims, and it remains

439   possible that stronger selection signatures could be observed when analyzing individuals who

440   died of plague in earlier pandemics. Importantly, further cohorts of *Y. pestis* victims are required

441   to verify the observations in this study in different geographic contexts, and also whether the

442   associations with the above-mentioned immunity genes are specific to the plague or might be

443   caused by other pathogens (Galvani & Slatkin, 2003). Furthermore, the demographic model we

444   used for simulation of natural selection is fitted to the CEU population (Central Europeans from

445   Utah) (Gravel et al., 2011) and assumes an exponential population growth. However, the CEU

446   population might have had a different demographic history than Ellwangen. It cannot, therefore,

447   be ruled out that the results from our analysis of natural selection may be inaccurate, if

448   Ellwangen has undergone stronger genetic drift than CEU. Nevertheless, our simulation results

449   provide preliminary evidence for natural selection as the main driving agent for the decrease of

450   frequencies in HLA-C*06:02, HLA-B*51:01 (and other HLA-B I-80[+] alleles), and HLA-DPB1-

451   H9 on the one hand, and the frequency-increase in HLA-DRB1*13 on the other hand. We note

452   that the frequency changes we observed are based on simulations of episodic selection and could

453   also be derived through alternative scenarios, including constant selection pressure (e.g. *s* of -

454   0.012 B*51, -0.014 C*06, 0.06 DRB1*13); (data not shown), or other epidemic challenges, such

455   as smallpox or tuberculosis, occurring since the 16th century. However, we did not find evidence

456   of smallpox or tuberculosis in the plague victims' DNA. Comparison with non-plague victims

457   from the same time period will be necessary to definitively answer this question. Our results do

458   not provide support for the proposition that evolution of human immunity drove reduction of *Y.*

459   *pestis* virulence and its disappearance from Europe (Ell, 1984). Instead, we provide first evidence

460   for evolutionary adaptive processes that might be driven by *Y. pestis* and may have been shaping

461   certain human immunity-relevant genes in Ellwangen and possibly also in Europe. As the earliest

15

462 victims of *Y. pestis* in Europe were already present in the Late Neolithic (Andrades Valtuena et
463 al., 2016; Rascovan et al., 2019; Rasmussen et al., 2015) and Europeans were intermittently
464 exposed to plague for almost 5,000 years, it is possible that relevant immunity alleles had already
465 been pre-selected in the European population long ago and maintained by standing variation
466 (Ralph & Coop, 2015) but recently became selected through epidemic events. Whilst *Y. pestis*
467 seems a likely culprit, this remains to be determined through replication cohorts and further
468 functional analyses.

469
470
471
472
473
474

475 **Material and Methods**

476 *Anthropological analyses*
477 Anthropological analyses on the skeletal remains were conducted in the Institute of
478 Paleoanthropology, University of Tübingen. Diseases of the periodontium and the teeth,
479 nonspecific stress markers and deficiencies, degenerative transformations, inflammatory bone
480 changes, and trauma were recorded (Supplementary Note 1). The body height of the adult
481 individuals was reconstructed and the growth course of the sub adult individuals was analyzed
482 (Kairies, 2015).

483

484 *C14-Dating of the archaeological remains from Ellwangen*
485 Acceleration Mass Spectrometry Radiocarbon (AMS-C14) dating was conducted at the Curt-
486 Engelhorn Center for Archaeometry in Mannheim. Calibration was performed based on the
487 INTCAL13 and the SwissCal 1.0 calibration curves.

488

489 *DNA extraction*
490 Petrous pyramids were cut longitudinally in order to enable access to the bony labyrinth
491 (Supplementary Figure 1C), which is the densest part of the mammalian body (Frisch, Sorensen,
492 Overgaard, Lind, & Bretlau, 1998) and provides the highest endogenous DNA yields (Pinhasi et

16

493   al., 2015). After cleaning the surface on one side of the bony labyrinth with the drill bit,
494   sampling was performed along the semicircular canal, which yielded 80-120 mg bone powder.
495   DNA extraction was performed by guanidinium-silica based extraction (Rohland & Hofreiter,
496   2007) using all the bone powder obtained. Tooth samples were cut in the middle, thus separating
497   the crown from the root, followed by drilling into the dental pulp to produce bone powder (ca.
498   100 mg). Saliva samples were obtained from 51 living Ellwangen citizens using *Whatman*
499   *OmniSwab* cheek swabs. Samples were obtained only from individuals whose families have been
500   resident in Ellwangen for at least four generations. Consent was given by the contributing
501   persons and their samples were anonymized. Approval for the study was granted by the Ethics
502   Committee of the Faculty of Medicine of the Eberhard Karls University and the University
503   Hospital Tübingen. Isolation of genomic DNA was performed using the *QIAamp DNA Blood*
504   *Mini Kit* following the *Qiagen* protocol.

505

506   *Preparation of libraries and sequencing*
507   Overall strategy: Indexed libraries were generated from all samples and the sequencing was then
508   performed in two stages. First, an aliquot of the full DNA library was subjected to whole genome
509   sequencing. Then, a second aliquot was subjected to enrichment for selected immunity-related
510   genes and subsequently sequenced.

511

512   Since our protocols for DNA extraction and library preparation are optimized for short-length
513   ancient DNA, and in order to avoid potential bias through laboratory methods, we sheared the
514   DNA extracted from modern individuals using ultrasonic DNA shearing to the same average
515   length as the ancient DNA. Therefore, the modern DNA was sheared to an average fragment
516   length of 75 bp using a *Covaris M220 Focused ultrasonicator*. DNA libraries, including sample-
517   specific indices, were prepared using 20 µl of each extract following published protocols
518   (Kircher, Sawyer, & Meyer, 2012; Meyer & Kircher, 2010). For the ancient samples, partial
519   uracil-DNA-glycosylase treatment was first applied (Rohland, Harney, Mallick, Nordenfelt, &
520   Reich, 2015). Sequencing was performed using an Illumina *Hiseq 4000* instrument with 75+8
521   cycles in single-end mode.

522

523

17

524 *Screening for pathogens*

525 DNA samples were screened for their metagenome content using the alignment tool MALT

526 version 0.3.8 (Vågene et al., 2018) and the metagenome analyzer MEGAN V6.11.4 (Huson,

527 Auch, Qi, & Schuster, 2007) (Supplementary Table 1).

528

529 Since petrous bone samples are not ideal for pathogen screening, we additionally accessed well-

530 preserved tooth samples from 30 distinct 16th century plague victims. Teeth were not available

531 from all individuals from whom we had obtained petrous bones, nor could they be

532 unambiguously attributed to specific individuals. Sequencing libraries and shotgun sequencing

533 were performed on the teeth following published protocols as described above.

534 MALT was used to align all pre-processed reads against a collection of all complete bacterial

535 genomes obtained from NCBI (ftp.ncbi.nlm.nih.gov/genomes/refseq/bacteria, access

536 12.03.2018). MALT was executed in BLASTN mode for bacteria using the following command:

537 malt-run --mode BlastN --e 0.001 --id 95 --alignmentType SemiGlobal --index $REF --inFile

538 $IN --output $OUT (where $REF is the MALT index). The e-value (--e) is a parameter that

539 describes the number of hits that are expected to be found just by chance. The --id parameter

540 describes the minimum percent identity that is needed for a hit to be reported. As the screening

541 with MALT was performed on aDNA data the applied filters are not very stringent since we

542 expect substitutions in organisms from ancient samples.

543 Reads assigned to the *Yersinia pestis* node and reads assigned to the nodes below, were extracted

544 using the extract reads function in MEGAN. For subsequent verification blastn (version 2.7.1)

545 was used to blast the extracted reads against *Yersinia pestis* (NC_003143.1) and *Yersinia*

546 *pseudotuberculosis* (NC_010634.1). The following custom blast command was used:

547 blastn -db $REF -query $IN -outfmt "6 qseqid sseqid pident length mismatch gapopen qstart

548 qend sstart send evalue bitscore gaps" (where $REF is the reference genome and $IN are the

549 extracted reads from MEGAN).

550

551 *Targeted sequencing of immunity-related genes*

552 Indexed libraries containing 20 µl DNA each were amplified in 100 µl reactions in a variable

553 number of one to seven cycles to reach the required concentration of 200 ng/µl for enrichment,

554 followed by purification using *Qiagen MinElute* columns. Using an in-solution capture-by-

18

555 hybridization approach (Fu et al., 2013), DNA molecule fragments originating from immunity
556 genes were enriched from the total DNA. The design and manufacture of the capture probes are
557 described below. Sequencing was performed as above.

558

559 *DNA damage estimation*

560 We performed an initial analysis of the merged data using the *EAGER* pipeline (Peltzer et al.,
561 2016) as follows: reads were mapped to *hg19* (The Genome Sequencing Consortium, 2001)
562 using the *aln* algorithm in *BWA 0.7.12* (H. Li & Durbin, 2010) with a seed length (k) of 32, the
563 *samtools* mapping quality parameter "q" set to 30 and a reduced mapping stringency parameter
564 "-n 0.01" to account for damage in ancient DNA. On average 2.2 million reads (51%) from the
565 plague victims with an average length of 59 bp, and 4.2 million reads (91%) from the modern
566 individuals with an average length of 68 bp, mapped uniquely to hg19 (Supplementary Data 1A).
567 To assess the authenticity of the ancient DNA fragments, C to T misincorporation frequencies
568 (Briggs et al., 2007) were obtained using *mapDamage 2.0* (Jonsson, Ginolhac, Schubert,
569 Johnson, & Orlando, 2013). As expected from partial UDG-treatment (Rohland et al., 2015),
570 ancient DNA sequences showed C to T substitutions at the first two positions of their 5' ends and
571 G to A substitutions at the 3' ends (Supplementary Data 1A). The first two positions from the 5'
572 end of the fastq-reads were trimmed off. The modern sample DNA sequence reads were not
573 subjected to this trimming.

574

575 *Sex determination*

576 Genetic sex was determined based on the ratio of sequences aligning to the X and Y
577 chromosomes compared to the autosomes (Skoglund, Storå, Götherström, & Jakobsson, 2013).

578

579 *Final data collation*

580 Contamination was estimated through examination of mitochondria sequences using the software
581 *Schmutzi* (Renaud, Slon, Duggan, & Kelso, 2015), and in males additionally on the X-
582 chromosomal level by applying *ANGSD* (Korneliussen, Albrechtsen, & Nielsen, 2014).
583 Contamination estimates ranged between 1 and 3% on mitochondrial and between 0.2 and 2.9%
584 on X-chromosomal level (Supplementary Data 1A). Datasets showing >8% contamination were
585 excluded from further analyses.

19

586

Genotyping

SNPs were drawn at random at each position from a previously published dataset of 1,233,013 SNPs (Haak et al., 2015; Lazaridis et al., 2014; Mathieson et al., 2015) in a pseudo-haploid manner using *pileupcaller* from the *sequenceTools* package (Lamnidis et al., 2018). Samples having fewer than 10,000 calls from a set of 1,233,013 SNPs were excluded. Forty-four datasets from ancient samples (40 from petrous bones and four from teeth) and 52 datasets from modern saliva samples remained.

594

Population genetic analyses

The genotype data from both Ellwangen populations were merged with a dataset of previously published West Eurasian populations genotyped on the aforementioned 1,233,013 SNPs (Mathieson et al., 2015) using the program *mergeit* from the *EIGENSOFT* package (Patterson et al., 2006). Principle Component Analysis (PCA) was performed using the software *smartpca* (Patterson et al., 2006). Admixture modelling was performed using the software ADMIXTURE (Alexander et al., 2009) with 65 West Eurasian populations from the *Affymetrix Human Origin* dataset, and the number of ancestral components ranging from K=3 to K=12. Cross-validation was performed for every admixture model and the model with the highest accuracy was determined by the lowest cross-validation error.

605

Kinship analysis

Kinship was assessed using three different software packages: *READ* (Monroy, Jose, Jakobsson, & Günther, 2017), *lcMLkin* (Lipatov, Sanjeev, Patro, & Veeramah, 2015) and *outgroup f3* statistics (Patterson et al., 2012). *READ* identifies relatives based on the proportion of non-matching alleles. *lcMLkin* infers individual kinship from calculated genotype likelihoods, and *f3* statistics can be used to identify relatives based on the amount of shared genetic drift. A pair of individuals was regarded related if evidence of relatedness was independently provided by at least two programs. For the modern population a first-degree relationship (parent-child or siblings) was detected by all three programs for EL1 and EL57. For the plague victims, evidence of a first-degree kinship was provided from all three programs for three pairs of individuals: ELW015 and ELW037, ELW016 and ELW017, and ELW036 and ELW039. Support from at

least two programs was given for second degree relatedness (grandparent-grandchild, uncle-nephew or first cousins) for two pairs: ELW021 and ELW030, and ELW007 and ELW039. Second- or higher-degree relatedness was suggested for the pair ELW030 and ELW034. Nine further observations of second-degree relationships were observed, but supported by only one program at a time and therefore not regarded as reliable kinship estimates (Supplementary Figure 2). Individuals EL57, ELW017, ELW030, ELW037 and ELW039 were excluded from allele frequency calculations, since they constitute kinship "nodes" or "leaves" that would bias allele frequencies as they do not contribute to the total allele diversity.

Effect size analysis

Effect sizes were estimated and plotted in *G*Power* 3.1.9.2 (Faul, Erdfelder, Lang, & Buchner, 2007) based on the given sample size and a power of 0.8. Effect size analysis has shown that with the current sample size large to medium effects (*w*=0.45-0.4) could be detected (Supplementary Figure 7).

Probe design for immune-capture

Enrichment of selected target genomic regions prior to sequencing can save sequencing costs and significantly reduce microbial DNA contaminants (Fu et al., 2016; Gnirke et al., 2009; Haak et al., 2015; Lazaridis et al., 2014; Mathieson et al., 2015). We therefore selected a set of 488 different human genes representative of the innate and adaptive immune system (Supplementary Table 2). Exon sequences were extracted from the human genome build *hg19* (The Genome Sequencing Consortium, 2001) using the *RefSeqGene* records from the *NCBI/Nucleotide* database and then selecting "Highlight Sequence Features" and "Exon". We added alternative alleles for *HLA, MIC, TAP* and *KIR*, which were obtained from the IMGT/HLA database (Robinson et al., 2015). For *HLA class I* and *KIR* genes the intronic regions were also included. For the *HLA* and *MIC* genes a set of 83 representative alleles with full-length gene sequences was chosen that encompasses the major serologically defined subclasses (Holdsworth et al., 2009) and covers 95% of the known polymorphism. To capture the remaining 5%, a set of 162 x 160 bp consensus sequences was designed.

A 60 bp probe was designed at every 5 bp interval along the target sequence. The last (3') 8 bp of each generated sequence was replaced by a custom primer sequence, so that probes could be

21

648   amplified. The final 52 bp probe sequences were mapped to *hg19* using *RazerS3* (Weese,

649   Holtgrewe, & Reinert, 2012) with minimum threshold of 95% identity. Duplicates and probe

650   sequences that mapped more than 20 times were removed. This process resulted in a final set of

651   322,667 unique probe sequences of 52 bp length. The probe set was tripled to complete capacity

652   of the *Agilent* 1-million feature array. The probes were cleaved from the array and amplified

653   using PCR (Fu et al., 2016). In summary, we generated 322,667 unique probes of 52 bp length

654   using stepwise 5 bp tiling to cover a total of 3,355,736 bp. The final set of probe sequences is

655   available in Supplementary Data 8. To validate the capture protocol, we used seven cell lines

656   from the Immunogenetics and Histocompatibility Workshop (IHW) chosen to represent

657   divergent HLA alleles that we had previously sequenced to full resolution (P. J. Norman et al.,

658   2017). The results are shown in Supplementary Data 2.

659

660   *Analysis of the CCR5-Δ32 frequency*

661   The CCR5 locus (chr3:46414947-46414978) was included in the target regions. To genotype

662   CCR5 for wildtype (wt) and Δ32 alleles, the sequence data was remapped to *hg19* using BWA-

663   *mem* with the mapping quality filter turned off. To generate genotypes, the CCR5 locus was

664   visually inspected using the Integrative Genomics Viewer (Robinson et al., 2011).

665

666   *1000 Genome data for selection scan*

667   We obtained the "low coverage" and "exome" aligned data for a set of 50 unrelated individuals

668   for the East Asian population CHB (Han Chinese in Beijing, China) from the 1000 Genomes

669   Phase 3 dataset (ftp://ftp.1000genomes.ebi.ac.uk/vol1/ftp/phase3/data/). The bam files were

670   converted into *fastq* format using the *bamtofastq* option from the software *bedtools 2.28.0*

671   (Quinlan, 2014).

672

673   *Variant detection*

674   We used *samtools mpileup* (Heng Li et al., 2009) for variant detection with a minimum mapping

675   and base quality of 30 while ignoring indels (-q 30 -Q 30 -C 50 -t DP,SP -g --skip-indels), and

676   used bcftools (Heng Li, 2011) for variant calling (-m -f GQ -O b). We considered only variants

677   that were within the captured regions +/- 1,000 bp. Variants were kept when at least 10

678   individuals had a genotype quality of 30 or higher as measured using vcftools (Danecek et al.,

679   2011). Resulting vcf-files were further annotated by adding the ancestral allele and dbSNP IDs

680   version 147. The ancestral allele was called as the most parsimonious based on 1000 Genomes

681   data (Abecasis et al., 2012) and a multiple species alignment (F.M. Key et al., 2016).

682

683   *DAnc calculation*

684   For all variants shared across the Ellwangen data (modern and ancient) and 1000 Genomes CHB

685   population we calculated a *Differentiation with Ancestral* (*DAnc*) score (F.M. Key et al., 2016).

686   *DAnc* is calculated per site, and uses derived allele frequency estimates to infer population-

687   specific allele frequency changes. Therefore, we inferred the derived allele frequency (DAF)

688   using the annotated ancestral allele for every site. Using the DAF we calculated *DAnc* scores per

689   site:

690

$$DAnc = |(ELW_{MOD} - CHB)| - |(ELW_{ANC} - CHB)|$$

691

692   For every site the resulting *DAnc* scores can range from -1 to +1. Invariable sites have a score of

693   0. A positive *DAnc* score indicates that the modern Ellwangen population has a different allele

694   frequency compared to the ancient Ellwangen population and the outgroup population, e.g. due

695   to recent positive selection. A negative *DAnc* score indicates that the ancient Ellwangen

696   population differentiates from both modern Ellwangen and the outgroup CHB.

697

698   *Estimating* F_{ST} *values*

699   We calculated F_{ST} values for all variants using the (Weir & Cockerham, 1984) estimator

700   implemented in *vcf-tools* (Danecek et al., 2011).We report the empirical p-values, which were

701   obtained by comparing the F_{ST} of all three candidate SNPs to the empirical distribution of F_{ST}

702   scores from all other variants.

703

704   *Simulation of neutral evolution*

705   In order to estimate the expected distribution of *DAnc* scores under neutral evolution, we

706   simulated the European demographic history, using a published model (Gravel et al., 2011) and

707   the simulation software *slim2* (Haller & Messer, 2017). The demographic model is based on

708   genome-wide data; we however had predominantly capture data from coding regions. To account

23

709 for increased drift in coding regions due to background selection, we reduced the effective
710 population sizes using background selection coefficients (*B scores*) (F.M. Key et al., 2016). We
711 estimated background selection for every genomic region captured using a published genome-
712 wide map (McVicker, Gordon, Davis, & Green, 2009). The complete model including all
713 parameters is available in Supplementary Data 9. We ran 100,000 simulations of genomic loci
714 matched in length to the captured region and used the resulting variants to calculate the neutral
715 expectation of the *DAnc* score distribution.

716

717 *Simulation of natural selection*
718
719 We performed 10,000 forward genetic simulations using *slim3* (Haller & Messer, 2019) to
720 determine null distributions for neutral frequency changes over 500 years in Ellwangen for each
721 HLA and KIR allotype. We used the sampled ancestral Ellwangen HLA and KIR allotype
722 frequencies as input and simulated 20 generations, assuming a 25 year generation time (e.g.
723 Gravel et al, 2011). We assumed a constant population growth rate of 1.085 per generation,
724 resulting in growth from 5000 to approximately 25000 in Ellwangen. We explicitly modelled
725 HLA and KIR allotypes, using linked binary identifiers to differentiate between alleles of a gene,
726 and therefore assumed no new mutations or intragenic recombination. We calculated intergenic
727 recombination rates per generation between *HLA* genes using a recent sex-averaged refined
728 genetic map (Bhérer, Campbell, & Auton, 2017). We allowed free recombination between *HLA*
729 and *KIR* regions (because they are on separate chromosomes). Specifically, we assumed the
730 following number of crossovers per generation between each HLA gene. HLA-A/HLA-C: $5.3e^{-3}$,
731 HLA-C/HLA-B: $1e^{-8}$, HLA-B/HLA-DRB345: $5.4e^{-3}$, HLA-DRB345/HLA-DRB1: $1e^{-8}$, HLA-
732 DRB1/HLA-DQA1: $3.2e^{-5}$, HLA-DQA1/HLA-DQB1: $3.9e^{-7}$, HLA-DQB1/HLA-DPA1: $6.5e^{-3}$,
733 and HLA-DPA1/HLA-DPB1: $1e^{-8}$. We calculated neutral frequency changes of each allotype.
734 We conclude the frequency changes of an allotype is due to natural selection if the sampled
735 modern day allotype frequency falls within the 0.5% extremes of the respective neutral
736 distribution (p < 0.01).
737 We re-implemented the neutral slim3 models as described above, but included a non-zero *s*
738 parameter for each HLA-allotype in question. For each selected allotype, we ran 100 simulations
739 with a positive or negative *s* with an absolute value of 0.001, 0.01, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6 or
740 0.7. Mirroring the timeline of the European plague outbreak, allotypes were selected for 7

24

741 generations and then returned to neutrality. The reported *s* values were consistent with previous
742 reported values of *s* acting on *MHC* genes (Radwan et al., 2020). We estimated the strength of
743 natural selection by fitting a LOESS curve to the simulated relationship between *s* and allotype
744 frequency and mapping the observed modern Ellwangen allotype frequency.

745

746 *HLA typing of the Ellwangen individuals*
747 We applied the *OptiType* algorithm, which is a program that enables HLA genotyping from high-
748 throughput sequence data. *OptiType* requires a minimum of a 12-fold coverage to reliably
749 determine the HLA alleles present at two-field (distinct polypeptide sequences) resolution. We
750 applied *OptiType* (Szolek et al., 2014) to identify HLA class I alleles, using a reference set of
751 present-day HLA allele sequences and a required sequence identity of at least 97% for every
752 alignment. We set no limit on the number of potential best matches during read mapping. We
753 manually verified the results obtained by a development version of the upcoming OptiType 2.0
754 package in order to determine HLA class II alleles. For every sample, the *OptiType* call having
755 highest confidence was used.

756

757 *Reconstruction of HLA haplotypes*
758 Haplotypes were assigned based on previously reported frequencies and linkage disequilibrium
759 (LD) (Cao et al., 2001; González-Neira et al., 2004). Maximum-likelihood haplotype frequencies
760 for alleles and two-point, three-point and four-point associations were estimated using an
761 Expectation-Maximization (EM) algorithm provided by the computer program *Arlequin* ver. 3.5
762 (Excoffier & Lischer, 2010).

763

764 *Comparing allele and haplotype frequencies*
765 HLA allele frequencies were calculated from HLA-A, -B, -C, and -DRB1 sequence data of 8862
766 potential stem cell donors registered with DKMS (German Bone Marrow Donor Registry) until
767 June 2014. Donors were of self-assessed German origin. Allele frequencies were calculated to
768 the two-field level (polypeptide sequence) (Schmidt et al., 2009). For allele frequency
769 comparisons, chi-squared tests (Pearson, 1900) were applied in R (R Development Core Team,
770 2011). Pairwise proportion tests were made between the allele or haplotype frequencies, where *p*
771 < 0.05 was considered significant. Omnibus tests for association with specific amino acid

772   positions, as well as pairwise tests for specific residues, were computed using the BIGDAWG R

773   package (Pappas, Marin, Hollenbach, & Mack, 2016). Linear mixed effects modeling was

774   performed using the Gaston package in R (https://cran.r-project.org/web/packages/gaston/) and

775   the *lcMLkin* kinship matrix generated as above. Wilson Score Interval estimation was performed

776   using the 'Hmisc' package of R (https://CRAN.R-project.org/package=Hmisc).

777

778   *Reconstruction of KIR genotypes and KIR allele frequency analyses*

779   Sequence reads specific to the *KIR* locus were identified by alignment to the human genome

780   reference hg19 using BWA *mem*, and then extracting those mapping to chr19:55,228,188-

781   55,383,188 or chr19_gl000209_random. The presence or absence and copy number of each *KIR*

782   gene were determined using the PING pipeline (P.J. Norman et al., 2016), modified for single-

783   end (SE; i.e. non-paired) reads. The alleles of *KIR3DL1/S1* were also determined using an SE

784   modified version of PING, and further validated by determining the alleles of the genes flanking

785   *KIR3DL1/S1* in the telomeric portion of the *KIR* locus. As an additional step, virtual sequence

786   probes were used to identify specific alleles directly from FASTQ data files, with a threshold of

787   ten reads used to positively identify a given allele. The scripts and probe sequences are available

788   at https://github.com/n0rmski/ThePlague/.

789

790

791

792   **References**

793   Abecasis, G. R., Auton, A., Brooks, L. D., DePristo, M. A., Durbin, R. M., Handsaker, R.
794   E., . . . McVean, G. A. (2012). An integrated map of genetic variation from 1,092
795   human genomes. *Nature, 491*(7422), 56–65. doi:10.1038/nature11632
796   Abi-Rached, L., Dorighi, K., Norman, P. J., Yawata, M., & Parham, P. (2007). Episodes of
797   natural selection shaped the interactions of IgA-Fc with FcalphaRI and bacterial decoy
798   proteins. *J Immunol, 178*(12), 7943–7954. doi:10.4049/jimmunol.178.12.7943
799   Achkar, J. P., Klei, L., de Bakker, P. I. W., Bellone, G., Rebert, N., Scott, R., . . . Duerr, R.
800   H. (2012). Amino acid position 11 of HLA-DR β 1 is a major determinant of
801   chromosome 6p association with ulcerative colitis. *Genes and immunity, 13*(3), 245–
802   252. doi:10.1038/gene.2011.79
803   Al Nabhani, Z., Dietrich, G., Hugot, J. P., & Barreau, F. Nod2: The intestinal gate keeper.

804   (1553–7374 (Electronic)).

805   Albright, F. S., Orlando, P., Pavia, A. T., Jackson, G. G., & Cannon Albright, L. A. (2008).
806   Evidence for a heritable predisposition to death due to influenza. *J Infect Dis.,*
807   *197*(0022–1899 (Print)).

808   Alexander, D. H., Novembre, J., & Lange, K. (2009). Fast model–based estimation of
809   ancestry in unrelated individuals. *Genome Res., 19*(9), 1655–1664.

810   Alfirevic, A., Gonzalez–Galarza, F., Bell, C., Martinsson, K., Platt, V., Bretland, G., . . .
811   Pirmohamed, M. (2012). In silico analysis of HLA associations with drug–induced liver
812   injury: use of a HLA–genotyped DNA archive from healthy volunteers. *Genome Med,*
813   *4*(6), 51.

814   Anand, P. K., Malireddi, R. K., Lukens, J. R., Vogel, P., Bertin, J., Lamkanfi, M., &
815   Kanneganti, T. D. (2012). NLRP6 negatively regulates innate immunity and host
816   defence against bacterial pathogens. *Nature, 488*(7411), 389–393.
817   doi:10.1038/nature11250

818   Andrades Valtuena, A., Mittnik, A., Massy, K., Allmae, R., Daubaras, M., Jankauskas,
819   R., . . . Krause, J. (2016). The Stone Age Plague: 1000 years of Persistence in
820   Eurasia. *bioRxiv.*

821   Ankley, L., Thomas, S., & Olive, A. A.–O. (2020). Fighting Persistence: How Chronic
822   Infections with Mycobacterium tuberculosis Evade T Cell–Mediated Clearance and
823   New Strategies To Defeat Them. LID – e00916–19 [pii] LID – 10.1128/IAI.00916–19
824   [doi]. *Infect. Immun., 88*(1098–5522 (Electronic)).

825   Barreiro, L. B., Ben–Ali M Fau – Quach, H., Quach H Fau – Laval, G., Laval G Fau – Patin,
826   E., Patin E Fau – Pickrell, J. K., Pickrell Jk Fau – Bouchier, C., . . . Quintana–Murci,
827   L. (2009). Evolutionary dynamics of human Toll–like receptors and their different
828   contributions to host defense. *PLoS Genet, 5*(1553–7404 (Electronic)).

829   Bashirova, A. A., Martin, M. P., McVicar, D. W., & Carrington, M. (2006). The killer
830   immunoglobulin–like receptor gene cluster: tuning the genome for defense. *Annu Rev*
831   *Genomics Hum Genet, 7*, 277–300. doi:10.1146/annurev.genom.7.080505.115726

832   Benedictow, O. J. (2004). *The Black Death, 1346–1353: The Complete History, 383*: Boydell
833   & Brewer.

834   Bhérer, C., Campbell, C. L., & Auton, A. (2017). Refined genetic maps reveal sexual
835   dimorphism in human meiotic recombination at multiple scales. *Nature*
836   *Communications, 8*(1), 14994. doi:10.1038/ncomms14994

837   Biraben, J.–N. (1976). *Les hommes et la peste en France et dans les pays européens et*
838   *méditerranéens.* Mouton: Paris–La Haye.

839   Bjorkman, P. J., & Parham, P. (1990). Structure, function, and diversity of class I major
840   histocompatibility complex molecules. *Annu Rev Biochem, 59*, 253–288.
841   doi:10.1146/annurev.bi.59.070190.001345

842   Bos, K. I., Herbig, A., Sahl, J., Waglechner, N., Fourment, M., Forrest, S. A., . . . Poinar, H.
843   N. (2016). Eighteenth century Yersinia pestis genomes reveal the long–term
844   persistence of an historical plague focus. *Elife, 5*, e12994. doi:10.7554/eLife.12994

845 Bos, K. I., Schuenemann, V. J., Golding, G. B., Burbano, H. A., Waglechner, N., Coombes,
846   B. K., . . . Krause, J. (2011). A draft genome of Yersinia pestis from victims of the
847   Black Death. *Nature, 478*(7370), 506–510. doi:10.1038/nature10549
848 Boudreau, J. E., & Hsu, K. C. (2018). Natural Killer Cell Education and the Response to
849   Infection and Cancer Therapy: Stay Tuned. *Trends Immunol, 39*(3), 222–239.
850   doi:10.1016/j.it.2017.12.001
851 Bowsky, W. M. (1971). *The Black Death: a turning point in history?* : New York, HOlt,
852   Rinhart and Winston.
853 Briggs, A. W., Stenzel, U., Johnson, P. L., Green, R. E., Kelso, J., Prufer, K., . . . Paabo, S.
854   (2007). Patterns of damage in genomic DNA sequences from a Neandertal. *Proceedings*
855   *of the National Academy of Sciences of the United States of America, 104*(37), 14616–
856   14621.
857 Büntgen, U., Ginzler, C., Esper, J., Tegel, W., & McMichael, A. J. (2012). Digitizing
858   historical plague. *Clin. Infect. Dis., 55*(11), 1586–1588.
859 Cao, K., Hollenbach, J., Shi, X., Shi, W., Chopek, M., & Fernández-Viña, M. A. (2001).
860   Analysis of the frequencies of HLA-A, B, and C alleles and haplotypes in the five
861   major ethnic groups of the United States reveals high levels of diversity in these loci
862   and contrasting distribution patterns in these populations. *Hum. Immunol., 62*(9),
863   1009–1030.
864 Cedzynski, M., Nuytinck, L., Atkinson, A. P. M., St Swierzko, A., Zeman, K., Szemraj,
865   J., . . . Kilpatrick, D. C. (2007). Extremes of L-ficolin concentration in children with
866   recurrent infections are associated with single nucleotide polymorphisms in the FCN2
867   gene. *Clin. Exp. Immunol., 150*(1), 99–104.
868 Clouse, M. (2002). The Black Death Transformed: Disease and Culture in Early Renaissance
869   Europe.: Samuel K Cohn Jr. London and New York: Arnold and Oxford University
870   Press, 2002, pp. 318, US$65.00 (HB) ISBN: 0-340-70646-5. *International Journal of*
871   *Epidemiology, 31*(6), 1280–1281. doi:10.1093/ije/31.6.1280
872 Cohn, S. K., Jr. (2003). *The Black Death Transformed: Disease and Culture in Early*
873   *Renaissance Europe.* London: Edward Arnold.
874 Danecek, P., Auton, A., Abecasis, G., Albers, C. A., Banks, E., DePristo, M. A., . . .
875   Genomes Project Analysis, G. (2011). The variant call format and VCFtools.
876   *Bioinformatics, 27*(15), 2156–2158.
877 Darke, C., Guttridge, M. G., Thompson, J., MacNamara, S., Street, J., & Thomas, M. (1998).
878   HLA class I (A, B) and II (DR, DQ) gene and haplotype frequencies in blood donors
879   from Wales. *Exp Clin Immunogenet, 15*, 69–83.
880 Dean, M., Carrington, M., Winkler, C., Huttley, G. A., Smith, M. W., Allikmets, R., . . .
881   O'Brien, S. J. (1996). Genetic restriction of HIV-1 infection and progression to AIDS
882   by a deletion allele of the CKR5 structural gene. Hemophilia Growth and Development
883   Study, Multicenter AIDS Cohort Study, Multicenter Hemophilia Cohort Study, San
884   Francisco City Cohort, ALIVE Study. *Science, 273*(5283), 1856–1862.
885 Di Marco, M., Schuster, H., Backert, L., Ghosh, M., Rammensee, H. G., & Stevanovic, S.

28

886   (2017). Unveiling the Peptide Motifs of HLA-C and HLA-G from Naturally Presented
887   Peptides and Generation of Binding Prediction Matrices. *J Immunol, 199*(8), 2639-
888   2651. doi:10.4049/jimmunol.1700938

889 Doherty, P. C., & Zinkernagel, R. M. (1975). A biological role for the major histocompatibility
890   antigens. *Lancet, 1*(7922), 1406-1409.

891 Drummond, W. K., Nelson, C. A., Fowler, J., Epson, E. E., Mead, P. S., & Lawaczeck, E. W.
892   (2014). Plague in a Pediatric Patient: Case Report and Use of Polymerase Chain
893   Reaction as a Diagnostic Aid. *J Pediatric Infect Dis Soc, 3*(4), e38-41.
894   doi:10.1093/jpids/piu001

895 Dubaniewicz, A., Lewko, B., Moszkowska, G., Zamorska, B., & Stepinski, J. (2000).
896   Molecular subtypes of the HLA-DR antigens in pulmonary tuberculosis. *Int. J. Infect.*
897   *Dis., 4*(3), 129-133.

898 Dunne, C., Crowley, J., Hagan, R., Rooney, G., & Lawlor, E. (2008). HLA-A, B, Cw, DRB1,
899   DQB1 and DPB1 alleles and haplotypes in the genetically homogenous Irish
900   population. *Int. J. Immunogenet., 35*(4-5), 295-302.

901 Ell, S. R. (1984). Immunity as a Factor in the Epidemiology of Medieval Plague. *Reviews of*
902   *Infectious Diseases, 6*(6), 866-879. doi:10.1093/clinids/6.6.866

903 Ellwangen, S. (2007). *Die dunkle Zeit. Hexenverfolgung in der Stadt und Fürstpropstei*
904   *Ellwangen.* Ellwangen: Stadtverwaltung Ellwangen.

905 Everitt, A. R., Clare S Fau - Pertel, T., Pertel T Fau - John, S. P., John Sp Fau - Wash, R.
906   S., Wash Rs Fau - Smith, S. E., Smith Se Fau - Chin, C. R., . . . Kellam, P. (2012).
907   IFITM3 restricts the morbidity and mortality associated with influenza. *Nature,*
908   *484*(1476-4687 (Electronic)), 519-523. doi:D - NLM: EMS41068 FIR - Everingham, K

909 Excoffier, L., & Lischer, H. E. L. (2010). Arlequin suite ver 3.5: a new series of programs to
910   perform population genetics analyses under Linux and Windows. *Mol Ecol Resour,*
911   *10*(3), 564-567.

912 Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: a flexible statistical
913   power analysis program for the social, behavioral, and biomedical sciences. *Behav Res*
914   *Methods, 39*(2), 175-191.

915 Feldman, M., Harbeck, M., Keller, M., Spyrou, M. A., Rott, A., Trautmann, B., . . . Krause,
916   J. (2016). A High-Coverage Yersinia pestis Genome from a Sixth-Century Justinianic
917   Plague Victim. *Mol. Biol. Evol., 33*(11), 2911-2923.

918 Frisch, T., Sorensen, M. S., Overgaard, S., Lind, M., & Bretlau, P. (1998). Volume-Referent
919   Bone Turnover Estimated From the Interlabel Area Fraction After Sequential Labeling.
920   *Bone, 22*, 677-682. doi:10.1016/s8756-3282(98)00050-7

921 Fu, Q., Meyer, M., Gao, X., Stenzel, U., Burbano, H. A., Kelso, J., & Pääbo, S. (2013). DNA
922   analysis of an early modern human from Tianyuan Cave, China. *Proc. Natl. Acad. Sci.*
923   *U.S.A., 110*(6), 2223-2227.

924 Fu, Q., Posth, C., Hajdinjak, M., Petr, M., Mallick, S., Fernandes, D., . . . Reich, D. (2016).
925   The genetic history of Ice Age Europe. *Nature, 534*(7606), 200-205.

926 Galvani, A. P., & Slatkin, M. (2003). Evaluating plague and smallpox as historical selective

927   pressures for the CCR5-Delta 32 HIV-resistance allele. *Proceedings of the National*
928   *Academy of Sciences of the United States of America, 100*(25), 15276-15279.
929   doi:10.1073/pnas.2435085100

930 Gnirke, A., Melnikov, A., Maguire, J., Rogov, P., LeProust, E. M., Brockman, W., . . .
931   Nusbaum, C. (2009). Solution hybrid selection with ultra-long oligonucleotides for
932   massively parallel targeted sequencing. *Nat Biotech, 27*(2), 182-189.
933   doi:http://www.nature.com/nbt/journal/v27/n2/suppinfo/nbt.1523_S1.html

934 González-Neira, A., Calafell, F., Navarro, A., Lao, O., Cann, H., Comas, D., & Bertranpetit,
935   J. (2004). Geographic stratification of linkage disequilibrium: a worldwide population
936   study in a region of chromosome 22. *Hum. Genomics, 1*(6), 399-409.

937 Gravel, Henn, B., Gutenkunst, R., Indap, A., Marth, G., Clark, A., . . . Project, T. G.
938   (2011). Demographic history and rare allele sharing among human populations.
939   *Proceedings of the National Academy of Sciences of the United States of America,*
940   *108*(29), 11983-11988. doi:10.1073/pnas.1019276108

941 Guethlein, L. A., Norman, P. J., Hilton, H. G., & Parham, P. (2015). Co-evolution of MHC
942   class I and variable NK cell receptors in placental mammals. *Immunol. Rev., 267*(1),
943   259-282.

944 Gumperz, J. E., Litwin, V., Phillips, J. H., Lanier, L. L., & Parham, P. (1995). The Bw4 public
945   epitope of HLA-B molecules confers reactivity with natural killer cell clones that
946   express NKB1, a putative HLA receptor. *J. Exp. Med., 181*(3), 1133-1144.

947 Guo, Y., Patil, N. K., Luan, L., Bohannon, J. K., & Sherwood, E. R. (2018). The biology of
948   natural killer cells during sepsis. *Immunology, 153*(2), 190-202.
949   doi:10.1111/imm.12854

950 Haak, W., Lazaridis, I., Patterson, N., Rohland, N., Mallick, S., Llamas, B., . . . Reich, D.
951   (2015). Massive migration from the steppe was a source for Indo-European languages
952   in Europe. *Nature, 522*(7555), 207-211.

953 Haller, B. C., & Messer, P. W. (2017). SLiM 2: Flexible, Interactive Forward Genetic
954   Simulations. *Mol. Biol. Evol., 34*(1), 230-240.

955 Haller, B. C., & Messer, P. W. (2019). SLiM 3: Forward Genetic Simulations Beyond the
956   Wright-Fisher Model. *Molecular Biology and Evolution, 36*(3), 632-637.
957   doi:10.1093/molbev/msy228

958 Hammer, C., Begemann, M., McLaren, P. J., Bartha, I., Michel, A., Klose, B., . . . Fellay, J.
959   (2015). Amino Acid Variation in HLA Class II Proteins Is a Major Determinant of
960   Humoral Response to Common Viruses. *American journal of human genetics, 97*(5),
961   738-743. doi:10.1016/j.ajhg.2015.09.008

962 Harrison, G. F., Sanz, J., Boulais, J., Mina, M. A.-O., Grenier, J. A.-O., Leng, Y., . . .
963   Barreiro, L. B. (2019). Natural selection contributed to immunological differences
964   between hunter-gatherers and agriculturalists. *Nature Ecology and Evolution, 3*(2397-
965   334X (Electronic)), 1253-1264.

966 Hellenthal, G., Busby, G. B., Band, G., Wilson, J. F., Capelli, C., Falush, D., & Myers, S.
967   (2014). A genetic atlas of human admixture history. *Science, 343*(6172), 747-751.

968 Hilton, H. G., Guethlein, L. A., Goyos, A. A.-O. X., Nemat-Gorgani, N., Bushnell, D. A.,
969   Norman, P. J., & Parham, P. (2015). Polymorphic HLA-C Receptors Balance the
970   Functional Characteristics of KIR Haplotypes. *J Immunol, 195*(1550-6606
971   (Electronic)), 160-170.
972 Hoang, T. V., Toan, N. L., Song, L. H., Ouf, E. A., Bock, C. T., Kremsner, P. G., . . .
973   Velavan, T. P. (2011). Ficolin-2 levels and FCN2 haplotypes influence hepatitis B
974   infection outcome in Vietnamese patients. *PLoS ONE, 6*(11), e28113.
975 Holdsworth, R., Hurley, C. K., Marsh, S. G., Lau, M., Noreen, H. J., Kempenich, J. H., . . .
976   Maiers, M. (2009). The HLA dictionary 2008: a summary of HLA-A, -B, -C, -
977   DRB1/3/4/5, and -DQB1 alleles and their association with serologically defined HLA-
978   A, -B, -C, -DR, and -DQ antigens. *Tissue Antigens, 73*(2), 95-170.
979   doi:10.1111/j.1399-0039.2008.01183.x
980 Hollenbach, J., Nocedal, I., Ladner, M. B., Single, R. M., & Trachtenberg, E. A. (2012).
981   Killer cell immunoglobulin-like receptor (KIR) gene content variation in the HGDP-
982   CEPH populations. *Immunogenetics*. doi:10.1007/s00251-012-0629-x
983 Hollenbach, J., Norman, P. J., Creary, L. E., Damotte, V. A.-O., Montero-Martin, G.,
984   Caillier, S., . . . Oksenberg, J. R. (2019). A specific amino acid motif of HLA-DRB1
985   mediates risk and Interacts with smoking history in Parkinson's Disease. *Proc. Natl.
986   Acad. Sci. U.S.A., 116*(1091-6490 (Electronic)), 7419-7424.
987 Hummelshoj, T., Munthe-Fog, L., Madsen, H. O., Fujita, T., Matsushita, M., & Garred, P.
988   (2005). Polymorphisms in the FCN2 gene determine serum variation and function of
989   Ficolin-2. *Hum. Mol. Genet., 14*(12), 1651-1658.
990 Huson, D. H., Auch, A. F., Qi, J., & Schuster, S. C. (2007). MEGAN analysis of
991   metagenomic data. *Genome Res., 17*(3), 377-386.
992 Inohara, Chamaillard, McDonald, C., & Nunez, G. (2005). NOD-LRR proteins: role in host-
993   microbial interactions and inflammatory disease. *Annu Rev Biochem, 74*, 355-383.
994   doi:10.1146/annurev.biochem.74.082803.133347
995 Johansson, A., Ingman, M., Mack, S. J., Ehrlich, H., & Gyllensten, U. (2008). Genetic origin
996   of the Swedish Sami inferred from HLA class I and class II allele frequencies. *Eur J
997   Hum Genet., 16*, 1341-1349.
998 Jonsson, H., Ginolhac, A., Schubert, M., Johnson, P. L., & Orlando, L. (2013).
999   mapDamage2.0: fast approximate Bayesian estimates of ancient DNA damage
1000   parameters. *Bioinformatics, 29*(13), 1682-1684.
1001 Kairies, M.-S. (2015). *Drei frühneuzeitliche Massengräber aus Ellwangen (Jagst) -
1002   Paläopathologie und demographische Struktur.* (Master of Science), University of
1003   Tübingen, Germany.
1004 Karlsson, E. K., Harris Jb Fau - Tabrizi, S., Tabrizi S Fau - Rahman, A., Rahman A Fau -
1005   Shlyakhter, I., Shlyakhter I Fau - Patterson, N., Patterson N Fau - O'Dushlaine,
1006   C., . . . Larocque, R. C. (2013). Natural selection in a bangladeshi population from the
1007   cholera-endemic ganges river delta. *Sci Transl Med., 5*(1946-6242 (Electronic)).
1008 Keller, M., Spyrou, M. A., Scheib, C. L., Neumann, G. U., Kröpelin, A., Haas-Gebhard,

31

1009   B., . . . Krause, J. (2019). Ancient genomes from across Western Europe reveal early
1010   diversification during the First Pandemic (541–750). *Proc. Natl. Acad. Sci. U.S.A.,*
1011   *116*(25), 12363–12372.

1012   Key, F. M., Fu, Q., Romagné, F., Lachmann, M., & Andrés, A. M. (2016). Human adaptation
1013   and population differentiation in the light of ancient genomes. *Nat Commun, 7*, 10775.

1014   Key, F. M., Peter, B., Dennis, M. Y., Huerta-Sánchez, E., Tang, W., Prokunina-Olsson,
1015   L., . . . Andrés, A. M. (2014). Selection on a variant associated with improved viral
1016   clearance drives local, adaptive pseudogenization of interferon lambda 4 (IFNL4).
1017   (1553–7404 (Electronic)).

1018   Kim, S., Sunwoo, J. B., Yang, L., Choi, T., Song, Y.-J., French, A. R., . . . Yokoyama, W. M.
1019   (2008). HLA alleles determine differences in human natural killer cell responsiveness
1020   and potency. *Proceedings of the National Academy of Sciences of the United States of*
1021   *America, 105*(8), 3053–3058. doi:10.1073/pnas.0712229105

1022   Kircher, M., Sawyer, S., & Meyer, M. (2012). Double indexing overcomes inaccuracies in
1023   multiplex sequencing on the Illumina platform. *Nucleic Acids Res, 40*(1), e3.

1024   Klebanov, N. (2018). Genetic Predisposition to Infectious Disease. *Cureus, 10*(8), e3210.

1025   Korneliussen, T. S., Albrechtsen, A., & Nielsen, R. (2014). ANGSD: Analysis of Next
1026   Generation Sequencing Data. *BMC Bioinformatics, 15*, 356.

1027   Kwiatkowski, D. P. (2005). How malaria has affected the human genome and what human
1028   genetics can teach us about malaria. *Am. J. Hum. Genet., 77*(0002–9297 (Print)), 171–
1029   192.

1030   Laayouni, H., Oosting, M., Luisi, P., Ioana, M., Alonso, S., Ricaño-Ponce, I., . . . Netea, M.
1031   G. (2014). Convergent evolution in European and Rroma populations reveals pressure
1032   exerted by plague on Toll-like receptors. *Proc. Natl. Acad. Sci. U.S.A., 111*(7), 2668–
1033   2673.

1034   Lamkanfi, M., & Dixit, V. M. (2014). Mechanisms and functions of inflammasomes. *Cell,*
1035   *157*(5), 1013–1022. doi:10.1016/j.cell.2014.04.007

1036   Lamnidis, T. C., Majander, K., Jeong, C., Salmela, E., Wessman, A., Moiseyev, V., . . .
1037   Schiffels, S. (2018). Ancient Fennoscandian genomes reveal origin and spread of
1038   Siberian ancestry in Europe. *Nat Commun, 9*(1), 5018. doi:10.1038/s41467-018-
1039   07483-5

1040   Lazaridis, I., Patterson, N., Mittnik, A., Renaud, G., Mallick, S., Kirsanow, K., . . . Krause,
1041   J. (2014). Ancient human genomes suggest three ancestral populations for present-day
1042   Europeans. *Nature, 513*(7518), 409–413.

1043   Lenski, R. E. (1988). Evolution of plague virulence. *Nature, 334*(6182), 473–474.

1044   Li, H. (2011). A statistical framework for SNP calling, mutation discovery, association
1045   mapping and population genetical parameter estimation from sequencing data.
1046   *Bioinformatics, 27*(21), 2987–2993.

1047   Li, H., & Durbin, R. (2010). Fast and accurate long-read alignment with Burrows-Wheeler
1048   transform. *Bioinformatics, 26*(5), 589–595.

1049   Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., . . . Genome Project

1050   Data Processing, S. (2009). The Sequence Alignment/Map format and SAMtools.
1051   *Bioinformatics, 25*(16), 2078-2079.

1052 Lindo, J., Huerta-Sánchez, E., Nakagome, S., Rasmussen, M., Petzelt, B., Mitchell, J., . . .
1053   Malhi, R. S. (2016). A time transect of exomes from a Native American population
1054   before and after European contact. *Nat Commun, 7*, 13175.

1055 Lipatov, M., Sanjeev, K., Patro, R., & Veeramah, K. (2015). Maximum Likelihood Estimation
1056   of biological Relatedness from low Coverage Sequencing Data. *bioRxiv*.

1057 Long, E. O., Kim, H. S., Liu, D., Peterson, M. E., & Rajagopalan, S. (2013). Controlling
1058   natural killer cell responses: integration of signals for activation and inhibition. *Annu*
1059   *Rev Immunol, 31*, 227-258. doi:10.1146/annurev-immunol-020711-075005

1060 Luo, F., Sun, X., Wang, Y., Wang, Q., Wu, Y., Pan, Q., . . . Zhang, X.-L. (2013). Ficolin-2
1061   defends against virulent Mycobacteria tuberculosis infection in vivo, and its
1062   insufficiency is associated with infection in humans. *PLoS ONE, 8*(9), e73859.

1063 Martinon, F., Burns, K., & Tschopp, J. (2002). The inflammasome: a molecular platform
1064   triggering activation of inflammatory caspases and processing of proIL-beta. *Mol Cell,*
1065   *10*(2), 417-426.

1066 Mathieson, I., Lazaridis, I., Rohland, N., Mallick, S., Patterson, N., Roodenberg, S. A., . . .
1067   Reich, D. (2015). Genome-wide patterns of selection in 230 ancient Eurasians. *Nature,*
1068   *528*(7583), 499-503.

1069 McManus, K. A.-O., Taravella, A. M., Henn, B. A.-O. X., Bustamante, C. D., Sikora, M., &
1070   Cornejo, O. A.-O. (2017). Population genetic analysis of the DARC locus (Duffy)
1071   reveals adaptation from standing variation associated with malaria resistance in
1072   humans. *PLoS Genet, 13*(1553-7404 (Electronic)).

1073 McVicker, G., Gordon, D., Davis, C., & Green, P. (2009). Widespread genomic signatures of
1074   natural selection in hominid evolution. *PLoS Genet., 5*(5), e1000471.

1075 Meyer, M., & Kircher, M. (2010). Illumina sequencing library preparation for highly
1076   multiplexed target capture and sequencing. *Cold Spring Harb Protoc, 2010*(6),
1077   pdb.prot5448.

1078 Monroy, K., Jose, M., Jakobsson, M., & Günther, T. (2017). Estimating genetic Kin
1079   Relationships in prehistoric Populations. *bioRxiv*.

1080 Morelli, G., Song, Y., Mazzoni, C. J., Eppinger, M., Roumagnac, P., Wagner, D. M., . . .
1081   Achtman, M. (2010). Yersinia pestis genome sequencing identifies patterns of global
1082   phylogenetic diversity. *Nat. Genet., 42*(12), 1140-1143.

1083 Namouchi, A., Guellil, M., Kersten, O., Hänsch, S., Ottoni, C., Schmid, B. V., . . . Bramanti,
1084   B. (2018). Integrative approach using Yersinia pestis genomes to revisit the historical
1085   landscape of plague during the Medieval Period. *Proceedings of the National Academy*
1086   *of Sciences, 115*(50), E11790. doi:10.1073/pnas.1812865115

1087 Neefjes, J., Jongsma, M. L. M., Paul, P., & Bakke, O. (2011). Towards a systems
1088   understanding of MHC class I and MHC class II antigen presentation. *Nat. Rev.*
1089   *Immunol., 11*(12), 823-836.

1090 Norman, P. J., Hollenbach, J., Nemat-Gorgani, N., Marin, W., Norberg, S., Ashouri, E., . . .

33

Parham, P. (2016). Defining KIR and HLA Class I Genotypes at Highest Resolution via High-Throughput Sequencing. *Am. J. Hum. Genet., 99*(2), 375-391.

Norman, P. J., Norberg, S. J., Guethlein, L. A., Nemat-Gorgani, N., Royce, T., Wroblewski, E. E., . . . Parham, P. (2017). Sequences of 95 human MHC haplotypes reveal extreme coding variation in genes other than highly polymorphic HLA class I and II. *Genome Res, 27*(5), 813-823. doi:10.1101/gr.213538.116

Nowak, J., Mika-Witkowska, R., Polak, M., Zajko, M., Rogatko-Koros, M., Graczyk-Pol, E., & Lange, A. (2008). Allele and extended haplotype polymorphism of HLA-A, -C, -B, -DRB1 and -DQB1 loci in Polish population and genetic affinities to other populations. *Tissue Antigens, 71*(3), 193-205. doi:10.1111/j.1399-0039.2007.00991.x

Ogawa, K., & Okada, Y. (2020). The current landscape of psoriasis genetics in 2020. *J Dermatol Sci., S0923-1811*(1873-569X (Electronic)). doi:10.1016/j.dermsci.2020.05.008

Pappas, D. J., Marin, W., Hollenbach, J. A., & Mack, S. J. (2016). Bridging ImmunoGenomic Data Analysis Workflow Gaps (BIGDAWG): An integrated case-control analysis pipeline. *Hum. Immunol., 77*(3), 283-287.

Parham, P., & Moffett, A. (2013). Variable NK cell receptors and their MHC class I ligands in immunity, reproduction and human evolution. *Nat Rev Immunol, 13*(2), 133-144. doi:10.1038/nri3370

Park, Y. H., Remmers, E. A.-O., Lee, W. A.-O., Ombrello, A. K., Chung, L. K., Shilei, Z., . . . Chae, J. A.-O. (2020). Ancient familial Mediterranean fever mutations in human pyrin and resistance to Yersinia pestis. *Nat Immunol., 21*(1529-2916 (Electronic)), 857-867.

Patin, E., Lopez, M., Grollemund, R., Verdu, P., Harmant, C., Quach, H., . . . Quintana-Murci, L. (2017). Dispersals and genetic adaptation of Bantu-speaking populations in Africa and North America. *Science, 356*(6337), 543-546. doi:10.1126/science.aal1988

Patterson, N., Moorjani, P., Luo, Y., Mallick, S., Rohland, N., Zhan, Y., . . . Reich, D. (2012). Ancient admixture in human history. *Genetics, 192*(3), 1065-1093.

Patterson, N., Price, A. L., & Reich, D. (2006). Population structure and eigenanalysis. *PLoS Genet., 2*(12), e190.

Pearson, K. (1900). On the criterion that a given system of deviations from the probable in the case of a correlated system of variables is such that it can be reasonably supposed to have arisen from random sampling. *Philosophical Magazine Series 5, 50*, 157-175. doi:doi:10.1080/14786440009463897

Peltzer, A., Jäger, G., Herbig, A., Seitz, A., Kniep, C., Krause, J., & Nieselt, K. (2016). EAGER: efficient ancient genome reconstruction. *Genome Biol., 17*(60). doi:10.1186/s13059-016-0918-z

Philip, N. H., Zwack, E. E., & Brodsky, I. E. (2016). Activation and Evasion of Inflammasomes by Yersinia. *Curr Top Microbiol Immunol, 397*, 69-90. doi:10.1007/978-3-319-41171-2_4

Pieters, J. (2008). Mycobacterium tuberculosis and the macrophage: maintaining a balance.

1132   *Cell Host Microbe, 3*(6), 399–407.

1133 Pingel, J., Solloch, U. V., Hofmann, J. A., Lange, V., Ehninger, G., & Schmidt, A. H. (2013).
1134   High-resolution HLA haplotype frequencies of stem cell donors in Germany with
1135   foreign parentage: how can they be used to improve unrelated donor searches? *Hum.*
1136   *Immunol., 74*(3), 330–340.

1137 Pinhasi, R., Fernandes, D., Sirak, K., Novak, M., Connell, S., Alpaslan-Roodenberg, S., . . .
1138   Hofreiter, M. (2015). Optimal Ancient DNA Yields from the Inner Ear Part of the
1139   Human Petrous Bone. *PLoS ONE, 10*(6), e0129102. doi:10.1371/journal.pone.0129102

1140 Politzer, R., & WHO. (1954). *Plague.*

1141 Prugnolle, F., Manica, A., Charpentier, M., Guégan, J., V., G., & Balloux, F. (2005).
1142   Pathogen-driven selection and worldwide HLA class I diversity. *Current Biology, 15*,
1143   1022–1027.

1144 Quinlan, A. R. (2014). BEDTools: The Swiss-Army Tool for Genome Feature Analysis. *Curr*
1145   *Protoc Bioinformatics, 47*, 11.12.11–34. doi:10.1002/0471250953.bi1112s47

1146 Quintana-Murci, L. (2019). Human Immunology through the Lens of Evolutionary Genetics.
1147   *Cell, 177*(1), 184–199. doi:10.1016/j.cell.2019.02.033

1148 R Development Core Team. (2011). *R: A Language and Environment for Statistical*
1149   *Computing. Vienna, Austria : the R Foundation for Statistical Computing*: R
1150   Foundation for Statistical Computing.

1151 Radwan, J., Babik, W., Kaufman, J., Lenz, T. L., & Winternitz, J. (2020). Advances in the
1152   Evolutionary Understanding of MHC Polymorphism. *Trends in Genetics, 36*, 298–311.

1153 Ralph, P. L., & Coop, G. (2015). Convergent Evolution During Local Adaptation to Patchy
1154   Landscapes. *PLoS Genet., 11*(11), e1005630.

1155 Rascovan, N., Sjögren, K.-G., Kristiansen, K., Nielsen, R., Willerslev, E., Desnues, C., &
1156   Rasmussen, S. (2019). Emergence and Spread of Basal Lineages of Yersinia pestis
1157   during the Neolithic Decline. *Cell, 176*(1-2), 295–305.e210.

1158 Rasmussen, S., Allentoft, M. E., Nielsen, K., Orlando, L., Sikora, M., Sjogren, K. G., . . .
1159   Willerslev, E. (2015). Early divergent strains of Yersinia pestis in Eurasia 5,000 years
1160   ago. *Cell, 163*(3), 571–582. doi:10.1016/j.cell.2015.10.009

1161 Renaud, G., Slon, V., Duggan, A. T., & Kelso, J. (2015). Schmutzi: estimation of
1162   contamination and endogenous mitochondrial consensus calling for ancient DNA.
1163   *Genome Biology, 16*(1), 224. doi:10.1186/s13059-015-0776-0

1164 Robinson, J., Halliwell, J. A., Hayhurst, J. D., Flicek, P., Parham, P., & Marsh, S. G. E.
1165   (2015). The IPD and IMGT/HLA database: allele variant databases. *Nucleic Acids*
1166   *Res., 43*(Database issue), D423–D431.

1167 Robinson, J., Thorvaldsdóttir, H., Winckler, W., Guttman, M., Lander, E. S., Getz, G., &
1168   Mesirov, J. P. (2011). Integrative genomics viewer. *Nat. Biotechnol., 29*(1), 24–26.

1169 Rohland, N., Harney, E., Mallick, S., Nordenfelt, S., & Reich, D. (2015). Partial uracil-DNA-
1170   glycosylase treatment for screening of ancient DNA. *Philos Trans R Soc Lond B Biol*
1171   *Sci, 370*(1660), 20130624. doi:10.1098/rstb.2013.0624

1172 Rohland, N., & Hofreiter, M. (2007). Ancient DNA extraction from bones and teeth. *Nat*

Protoc, 2(7), 1756–1762.

Sabeti, Schaffner, S., Fry, B., Lohmueller, J., Varilly, P., Shamovsky, O., . . . Lander, E. S. (2006). Positive natural selection in the human lineage. *Science, 312*(5780), 1614–1620. doi:10.1126/science.1124309

Sabeti, Varilly, P., Fry, B., Lohmueller, J., Hostetter, E., Cotsapas, C., . . . Stewart, J. (2007). Genome-wide detection and characterization of positive selection in human populations. *Nature, 449*(1476-4687 (Electronic)), 913–918. doi:D – NLM: UKMS4416 EDAT- 2007/10/19 09:00 MHDA- 2007/12/06 09:00 CRDT- 2007/10/19 09:00 PHST- 2007/08/08 00:00 [received] PHST- 2007/09/13 00:00 [accepted] PHST- 2007/10/19 09:00 [pubmed] PHST- 2007/12/06 09:00 [medline] PHST- 2007/10/19 09:00 [entrez] AID – nature06250 [pii] AID – 10.1038/nature06250 [doi] PST – ppublish

Saunders, P. M., Vivian, J. P., O'Connor, G. M., Sullivan, L. C., Pymm, P., Rossjohn, J., & Brooks, A. G. (2015). A bird's eye view of NK cell receptor interactions with their MHC class I ligands. *Immunol Rev, 267*(1), 148–166. doi:10.1111/imr.12319

Schmidt, A. H., Baier, D., Solloch, U. V., Stahr, A., Cereb, N., Wassmuth, R., . . . Rutt, C. (2009). Estimation of high-resolution HLA-A, -B, -C, -DRB1 allele and haplotype frequencies based on 8862 German stem cell donors and implications for strategic donor registry planning. *Hum. Immunol., 70*(11), 895–902.

Skoglund, P., Storå, J., Götherström, A., & Jakobsson, M. (2013). Accurate sex identification of ancient human remains using DNA shotgun sequencing. *Journal of Archaeological Science, 40*(12), 4477–4482.

Solloch, U. V., Lang, K., Lange, V., Böhme, I., Schmidt, A. H., & Sauter, J. (2017). Frequencies of gene variant CCR5-Δ32 in 87 countries based on next-generation sequencing of 1.3 million individuals sampled from 3 national DKMS donor centers. *Hum. Immunol., 78*(11-12), 710–717.

Spyrou, M. A., Keller, M., Tukhbatova, R. I., Scheib, C. L., Nelson, E. A., Andrades Valtueña, A., . . . Krause, J. (2019). Phylogeography of the second plague pandemic revealed through analysis of historical Yersinia pestis genomes. *Nature Communications, 10*(1), 4470–4470. doi:10.1038/s41467-019-12154-0

Spyrou, M. A., Tukhbatova, R. I., Feldman, M., Drath, J., Kacki, S., Beltran de Heredia, J., . . . Krause, J. (2016). Historical Y. pestis Genomes Reveal the European Black Death as the Source of Ancient and Modern Plague Pandemics. *Cell Host Microbe, 19*(6), 874–881. doi:10.1016/j.chom.2016.05.012

Spyrou, M. A., Tukhbatova, R. I., Wang, C.-C., Valtueña, A. A., Lankapalli, A. K., Kondrashin, V. V., . . . Krause, J. (2018). Analysis of 3800-year-old Yersinia pestis genomes suggests Bronze Age origin for bubonic plague. *Nat Commun, 9*(1), 2234.

Stephens, J. C., Reich, D. E., Goldstein, D. B., Shin, H. D., Smith, M. W., Carrington, M., . . . Dean, M. (1998). Dating the origin of the CCR5-Delta32 AIDS-resistance allele by the coalescence of haplotypes. *Am. J. Hum. Genet., 62*(6), 1507–1515.

Sun, J., Yang, C., Fei, W., Zhang, X., Sheng, Y., Zheng, X., . . . Zhang, X. (2018). HLA-

1214   DQ β 1 amino acid position 87 and DQB1*0301 are associated with Chinese Han SLE.
1215   *Molecular genetics & genomic medicine, 6*(4), 541–546. doi:10.1002/mgg3.403

1216 Szolek, A., Schubert, B., Mohr, C., Sturm, M., Feldhahn, M., & Kohlbacher, O. (2014).
1217   OptiType: precision HLA typing from next-generation sequencing data.
1218   *Bioinformatics, 30*(23), 3310–3316. doi:10.1093/bioinformatics/btu548

1219 The Genome Sequencing Consortium. (2001). Initial sequencing and analysis of the human
1220   genome. *Nature, 409*(6822), 860–921.

1221 Trowsdale, J., & Knight, J. C. (2013). Major histocompatibility complex genomics and human
1222   disease. *Annu Rev Genomics Hum Genet, 14*, 301–323.

1223 Uhrberg, M., Valiante, N. M., Shum, B. P., Shilling, H. G., Lienert-Weidenbach, K., Corliss,
1224   B., . . . Parham, P. (1997). Human diversity in killer cell inhibitory receptor genes.
1225   *Immunity, 7*(6), 753–763.

1226 Vågene, Å. J., Herbig, A., Campana, M. G., Robles García, N. M., Warriner, C., Sabin,
1227   S., . . . Krause, J. (2018). Salmonella enterica genomes from victims of a major
1228   sixteenth-century epidemic in Mexico. *Nature Ecology & Evolution, 2*, 520–528.

1229 Vasseur, E., Boniotto, M., Patin, E., Laval, G., Quach, H., Manry, J., . . . Quintana-Murci,
1230   L. (2012). The evolutionary landscape of cytosolic microbial sensors in humans.
1231   *American journal of human genetics, 91*(1), 27–37. doi:10.1016/j.ajhg.2012.05.008

1232 Vladimer, G. I., Marty-Roix, R., Ghosh, S., Weng, D., & Lien, E. (2013). Inflammasomes and
1233   host defenses against bacterial infections. *Curr Opin Microbiol, 16*(1), 23–31.
1234   doi:10.1016/j.mib.2012.11.008

1235 Vladimer, G. I., Weng, D., Paquette, S. W., Vanaja, S. K., Rathinam, V. A., Aune, M.
1236   H., . . . Lien, E. (2012). The NLRP12 inflammasome recognizes Yersinia pestis.
1237   *Immunity, 37*(1), 96–107. doi:10.1016/j.immuni.2012.07.006

1238 Voight, B. F., Kudaravalli, S., Wen, X., & Pritchard, J. K. (2006). A map of recent positive
1239   selection in the human genome. *PLoS Biol., 4*(1545-7885 (Electronic)).

1240 Wagner, D. M., Klunk, J., Harbeck, M., Devault, A., Waglechner, N., Sahl, J. W., . . .
1241   Poinar, H. (2014). Yersinia pestis and the plague of Justinian 541–543 AD: a genomic
1242   analysis. *Lancet Infect Dis, 14*(4), 319–326. doi:10.1016/s1473-3099(13)70323-2

1243 Wang, E. T., Kodama, G., Baldi, P., & Moyzis, R. K. (2006). Global landscape of recent
1244   inferred Darwinian selection for Homo sapiens. *Proc. Natl. Acad. Sci. U.S.A.,*
1245   *103*(0027-8424 (Print)), 135–140.

1246 Weese, D., Holtgrewe, Manuel, & Reinert, K. (2012). RazerS 3: Faster, fully sensitive read
1247   mapping. *Bioinformatics*. doi:10.1093/bioinformatics/bts505

1248 Weir, B. S., & Cockerham, C. C. (1984). ESTIMATING F-STATISTICS FOR THE
1249   ANALYSIS OF POPULATION STRUCTURE. (1558-5646 (Electronic)).

1250 WHO. (2017). Plague in Madagascar. *WHO*.

1251 Wilson, M. J., Torkar, M., Haude, A., Milne, S., Jones, T., Sheer, D., . . . Trowsdale, J.
1252   (2000). Plasticity in the organization and sequences of human KIR/ILT gene families.
1253   *Proceedings of the National Academy of Sciences of the United States of America,*
1254   *97*(9), 4778–4783. doi:10.1073/pnas.080588597

1255 Zaki, M. H., Man, S. M., Vogel, P., Lamkanfi, M., & Kanneganti, T. D. (2014). Salmonella
1256   exploits NLRP12-dependent innate immune signaling to suppress host defenses during
1257   infection. *Proceedings of the National Academy of Sciences of the United States of*
1258   *America, 111*(1), 385–390. doi:10.1073/pnas.1317643111
1259

Acknowledgements

Ethical Approval for modern and ancient Samples

Saliva samples were obtained from the contributing persons after their consent was given. The samples were obtained in a completely anonymised way, so that no personal information was registered. Ethical Approval was granted by the Ethics Committee of the Faculty of Medicine of the Eberhard Karls University and the University Hospital of Tübingen. According to ethical guidelines, sampling of 16th century teeth and petrous bones was performed with archaeological / anthropological supervision to minimize the destructive effect. Each sample was sampled only once and photographs were taken before and after the sampling procedure. Every sample has been documented and archived in the aDNA facility of the Max Planck Institute for the Science of Human History in Jena, Germany.

Author Contributions

AI, AS, PJN, AH, KIB, SF, OK and JK designed the immunity capture. AI, ER and JK collected modern saliva samples. AI, MAS, VJS and ER conducted lab work. AI, RB, AS and DIHZ performed HLA genotyping, and FMK conducted *DAnc* and F_{st} analysis. JS and BKK performed pathogen screening. PJN and MKR performed KIR genotyping. JH and PJN performed data

1287   analyses. PJN, WHP and GFH performed simulations for natural selection. JS, US and AHS

1288   contributed the DKMS allele frequencies. RW and SA led the Excavation in Ellwangen. JW and

1289   RW provided the 16th century Ellwangen samples. MSK and JW conducted anthropological

1290   analyses. AI, JH, PJN and JK wrote the manuscript. JK led the study.

1291

1292   **Conflicts of Interests**

1293   The authors declare no conflicts of interests.

1294

1295   **Data Availability**

1296   The data underlying this article is available on the European Nucleotide Archive under the

1297   accession number PRJEB44124 (ERP128137).

1298

1299

1300

1301

1302

1303

1304

1305

1306

1307

1308

1309

1310

1311

1312

1313

1314

1315

1316

Figures

Figure 1. Mass burials discovered at Ellwangen. A. Location of Ellwangen in Germany. **B.** Location of the marketplace, where the mass burials were discovered during an excavation in 2013-2015. **C.** Mass grave 549 showing several individuals being buried together.
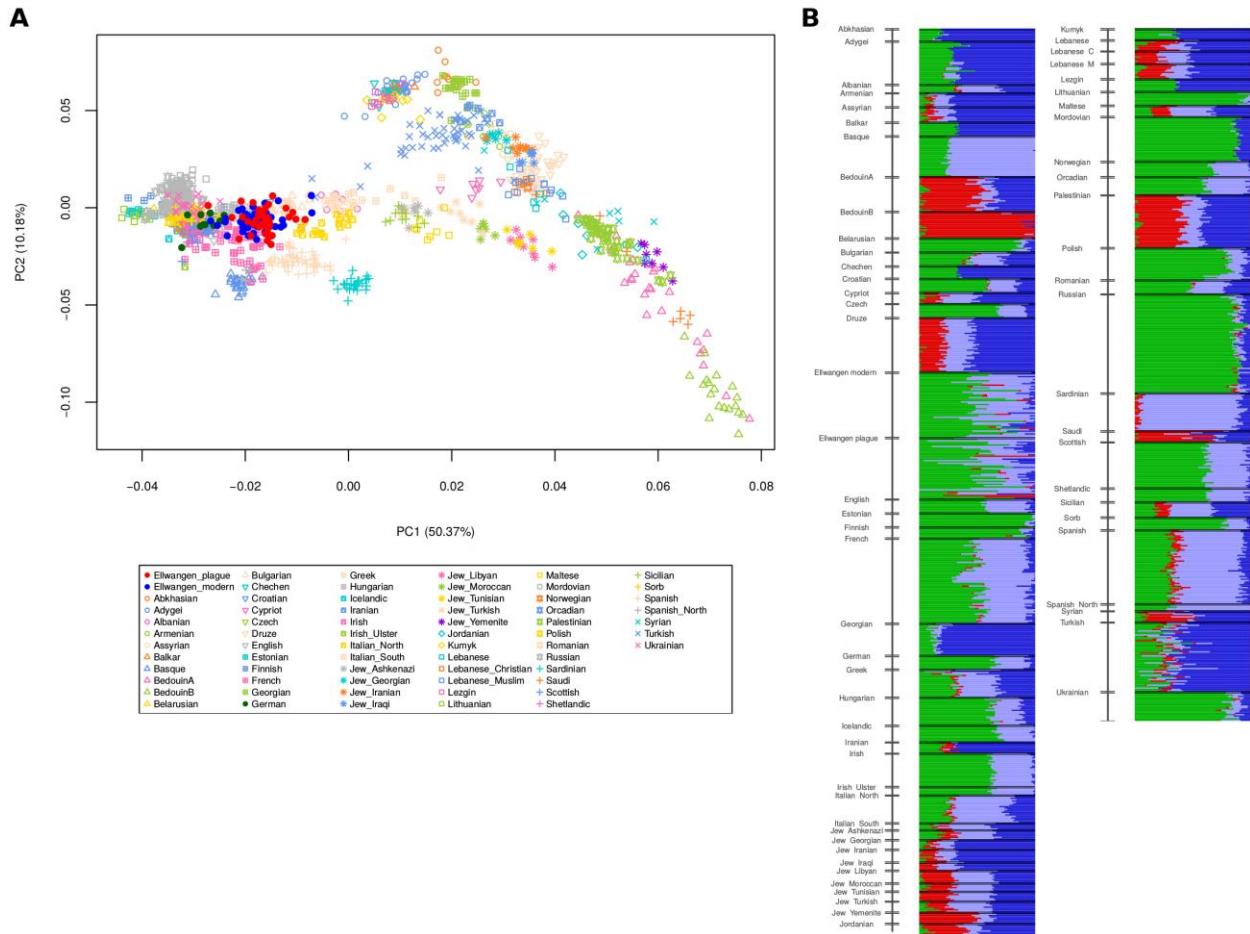
Figure 2. The 16th century plague victims and modern inhabitants of Ellwangen form a continuous population

A. PCA showing the 16th century (red) and modern (blue) Ellwangen populations in the context of 65 modern day populations from West-Eurasia based on 1,233,013 genome wide SNPs (Lazaridis et al. 2014; Haak et al. 2015; Fu et al. 2016). **B**. Admixture modeling based on four ancestral components (K=4) of the same 65 modern West Eurasian populations including 16th century (Ellwangen plague) and modern Ellwangen (Ellwangen modern) populations. The K=4 model was chosen due to the lowest cross-validation error.
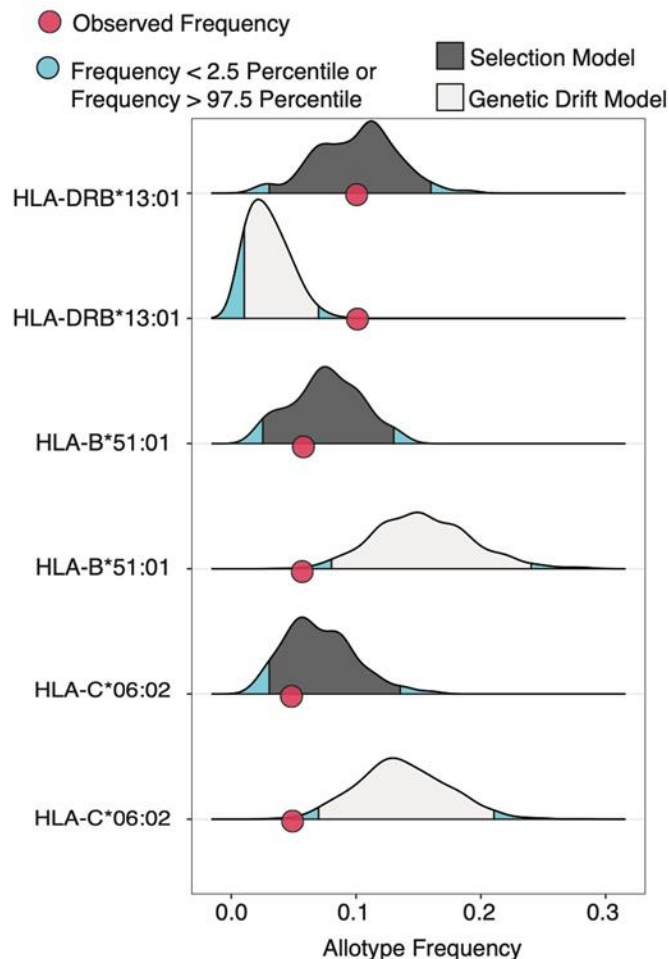
1347

1348 **Figure 3. Natural selection drives HLA allele frequency changes**

1349 Density plots showing the distributions of allele frequencies from SLIM3 model simulations with

1350 (dark grey) or without (light grey) natural selection. The starting frequency for simulations was

1351 the observed frequency in the 16th century population. Selection coefficients for the models with

1352 natural selection were -0.1 for *HLA-B*51:01* and *HLA-C*06:02* and 0.2 for *DRB1*13:01*

1353 (Supplementary Figure 5). The 2.5% extremes are shown in blue illustrating where the p-value

1354 cut-off of 0.05 would occur. Red points represent the frequency in the modern-day population.

1355

1356

1357 **Table 1.** Genes identified in the 0.01% tail of distribution following *DAnc* analysis. **Ref** =

1358 reference, **Alt** = alternative, **Der** = derived, **Anc** = derived allele frequency in Ellwangen 16th

1359 century, **Mod** = derived allele frequency in Ellwangen modern, **CEU** = derived allele frequency

1360 in Central Europeans from Utah, **FIN** = derived allele frequency in Finnish, **GBR** = derived

42

allele frequency in Great Britains (obtained from the 1000 Genomes Project Phase 3 data). Red text indicates alleles that have significantly (p<0.05) increased in frequency in the modern individuals and blue text indicates that allele frequency has significantly (p<0.05) decreased in frequency in the modern individuals. F_{ST} empirical p-value refers to the empirical distribution of F_{ST} calculated between the 16th century and the modern Ellwangen population.

Chromosome			Allele			Ellwangen						F_{ST} empirical p-value	Variant	Gene	Function
No.	Position	SNP ID	Ref	Alt	Der	Anc	Mod	CEU	FIN	GBR	DAnc				
9	137772664	rs17514136	A	G	G	0.17	0.37	0.28	0.24	0.21	0.20	0.99	5_prime_UTR	FCN2	activation of the lectin complement pathway
9	137779026	rs17549193	C	T	T	0.19	0.39	0.29	0.25	0.23	0.20	0.99	missense	FCN2	
11	7079038	rs10839708	G	A	A	0.69	0.51	0.60	0.65	0.63	0.18	0.97	missense	NLRP14	activation of pro-inflammatory caspases

Table 2. Genotype and allele frequencies of *CCR5*-wildtype (wt), and *CCR5-Δ32* (Δ32), among the plague victims and modern individuals from Ellwangen. The individual genotypes are given in Supplementary Data 6A. The frequencies for Germany were obtained from a study of German bone marrow donor registry volunteers (Solloch et al., 2017) for comparison.

	Genotype frequency (%)			**Allele Frequency (%)**	
	wt/wt	**wt/Δ32**	**Δ32/Δ32**	**wt**	**Δ32**
Ellwangen plague	71.4	23.8	4.8	83.4	16.6
Ellwangen modern	78.4	21.6	0	89.2	10.8
Germany	79.2	19.4	1.4	88.8	11.2

1374 **Table 3.** *HLA-B*, *-C* and *-DRB1* allele frequencies in 16th century plague victims and modern

1375 inhabitants of Ellwangen. Frequency differences with p<0.05 are highlighted in red. No

1376 significance could be obtained after multiple testing correction (for details see Supplementary

1377 Data 5A).

Locus	Frequency (%) Plague	Frequency (%) Modern	P-value
B*07:02	13.89	14	0.984
B*08:01	9.72	13	0.508
B*13:02	1.39	3	0.490
B*14:02	1.39	4	0.315
B*15:01	5.56	10	0.294
B*18:01	5.56	3	0.402
B*27:05	4.17	3	0.680
B*35:01	4.17	6	0.595
B*35:03	4.17	2	0.403
B*38:01	2.78	1	0.379
B*40:01	4.17	2	0.403
B*44:02	4.17	6	0.595
B*49:01	1.39	1	0.814
B*50:01	4.17	1	0.174
B*51:01	15.28	6	0.044
B*52:01	1.39	1	0.814
B*57:01	4.17	2	0.403
C*01:02	4.17	2	0.403
C*02:02	4.17	3	0.680
C*03:03	4.17	11	0.106
C*03:04	6.94	4	0.393
C*04:01	11.11	15	0.460
C*05:01	2.78	5	0.467
C*06:02	13.89	5	0.041
C*07:01	13.89	17	0.580
C*07:02	15.28	12	0.533
C*07:04	1.39	1	0.814
C*08:02	1.39	4	0.315
C*12:02	1.39	1	0.814
C*12:03	5.56	5	0.871
C*15:02	5.56	4	0.632
DRB1*01:01	9.72	7	0.520
DRB1*01:02	1.39	2	0.763
DRB1*03:01	11.11	11	0.982
DRB1*04:01	9.72	3	0.063
DRB1*04:07	1.39	1	0.814
DRB1*04:08	1.39	2	0.763
DRB1*07:01	11.11	12	0.857
DRB1*09:01	1.39	2	0.763
DRB1*11:01	9.72	6	0.363
DRB1*11:03	2.78	2	0.738
DRB1*11:04	1.39	4	0.315
DRB1*12:01	1.39	1	0.814
DRB1*13:01	2.78	10	0.067
DRB1*13:02	2.78	6	0.323
DRB1*15:01	18.06	15	0.592
DRB1*15:02	1.39	1	0.814
DRB1*16:01	1.39	3	0.490
DRB1*01	11.11	9	0.625
DRB1*03	11.11	11	0.956
DRB1*04	15.28	10	0.281
DRB1*07	11.11	12	0.883
DRB1*11	13.89	13	0.838
DRB1*13	5.56	17	0.026
DRB1*15	19.44	16	0.529

1378

1379

1380 **Table 4. A.** Allele frequencies for (top) presence of *KIR3DL1* gene, (bottom) *KIR3DL1* alleles

1381 (‡) indicates allele not expressed at the cell surface (Guethlein et al., 2015).

1382 **B.** Genotype frequencies for *KIR3DL1* and *I80⁺HLA-B* in 16th century (plague) and modern

1383 inhabitants of Ellwangen. Shown are p values from (chi) chi square, and (lme) logistic mixed

1384 effects model.

1385

1386

A

		frequency plague	frequency modern
Gene	*KIR3DL1*	0.81	0.74
Alleles	*00101	0.25	0.18
	*002	0.17	0.13
	*00401‡	0.11	0.08
	*00501	0.13	0.18
	*007	0.03	0.02
	*008	0.07	0.04
	*01502	0.06	0.11

B

	Number Observed				p=	
Genotype	**plague (N=36)**		**modern (N=50)**		**(chi)**	**(lme)**
I-80⁺HLA-B	20	(55.6%)	16	(32%)	0.029	0.07
KIR3DL1	35	(97.2%)	44	(88%)	0.099	0.29
KIR3DL1 + I-80⁺HLA-B	19	(52.8%)	13	(26%)	0.011	0.01

1387

45