1 **Evolution of cytokine production capacity in ancient and modern European**

2 **populations**

3 *Short title: Evolution of cytokine production in Europeans*

4 Jorge Domínguez-Andrés[1,2,*], Yunus Kuijpers[3,4*], Olivier B. Bakker[5], Martin Jaeger[1,2],

5 Cheng-Jian Xu[1,3,4], Jos W.M. van der Meer[1], Mattias Jakobsson[6,7], Jaume Bertranpetit[8],

6 Leo A.B. Joosten[1,2], Yang Li[1,2,3,4,#], Mihai G. Netea[1,2,9,10,#]

7 Affiliations:

8 [1]Department of Internal Medicine and Radboud Center for Infectious diseases (RCI),

9 Radboud University Nijmegen Medical Centre, Geert Grooteplein 8, 6500HB Nijmegen,

10 the Netherlands.

11 [2]Radboud Institute for Molecular Life Sciences (RIMLS), RadboudUniversity Medical

12 Center, 6525, GA, Nijmegen, the Netherlands

13 [3]Department of Computational Biology for Individualised Infection Medicine, Centre for

14 Individualised Infection Medicine (CiiM), a joint venture between Helmholtz-Centre for

15 Infection Research (HZI) and the Hannover Medical School (MHH), Hannover,

16 Germany

17 [4]TWINCORE, Centre for Experimental and Clinical Infection Research, a joint venture

18 between Helmholtz-Centre for Infection Research (HZI) and the Hannover Medical

19 School (MHH), Hannover, Germany

20 [5]Department of Genetics, University Medical Centre Groningen, Nijmegen, The

21 Netherlands.

22 [6]Human Evolution, Department of Organismal Biology, Uppsala University, Uppsala,

23 Sweden.

24 [7]Centre for Anthropological Research, Department of Anthropology and Development

25 Studies, University of Johannesburg, Auckland Park, South Africa.

26   [8]Institut de Biologia Evolutiva (UPF-CSIC), Universitat Pompeu Fabra, Dr. Aiguader 88,

27   08003 Barcelona, Catalonia, Spain

28   [9]Department for Genomics & Immunoregulation, Life and Medical Sciences Institute

29   (LIMES), University of Bonn, Bonn, Germany

30   [10]Lead contact

31   *These authors contributed equally to the manuscript

32   #These authors share senior authorship

33   Correspondence: mihai.netea@radboudumc.nl, Yang.Li@helmholtz-hzi.de

34

Abstract

As our ancestors migrated throughout the different continents, natural selection increased the presence of alleles advantageous in the new environments. Heritable variations that alter the susceptibility to diseases vary with the historical period, the virulence of the infections, and their geographical spread. In this study we built polygenic scores for heritable traits influencing the genetic adaptation in the production of cytokines and immune-mediated disorders, including infectious, inflammatory, and autoimmune diseases, and applied them to the genomes of several ancient European populations. We observed that the advent of the Neolithic was a turning point for immune-mediated traits in Europeans, favoring those alleles linked with the development of tolerance against intracellular pathogens and promoting inflammatory responses against extracellular microbes. These evolutionary patterns are also associated with an increased presence of traits related to inflammatory and auto-immune diseases.

Introduction

Human history has been shaped by infectious diseases. Human genes, especially host defense genes, have been constantly influenced by the pathogens encountered (Fumagalli and Sironi, 2014; Karlsson et al., 2014; Quintana-Murci and Clark, 2013). Pathogens drive the selection of genetic variants affecting resistance or tolerance to the infection, and heritable variations that increase survival to diseases with high morbidity and mortality will be naturally selected in people before reproductive age (Karlsson et al., 2014). These selection signatures vary with historical period, virulence of the pathogen, and the geographical spread.

Here we investigated the historical evolutionary patterns leading to genetic adaptation in cytokine production and immune-mediated diseases, including infectious, inflammatory, and autoimmune diseases. Cytokine production capacity is a key component of the host defense mechanisms: it induces inflammation, activates phagocytes to eliminate the pathogens and present antigens, and controls induction of T-helper adaptive immune responses. We have therefore chosen to investigate the evolutionary trajectories of cytokine production capacity in modern human populations during history. To determine the difference in polygenic regulation of diseases and cytokine production capacity, we used data derived from the 500 Functional Genomics (500FG) cohort of the Human Functional Genomics Project (HFGP; http://www.humanfunctionalgenomics.org). The HFGP is an international collaboration aiming to identify the host and environmental factors responsible for the variability of human immune responses in health and disease (Netea et al., 2016). Within the HFGP project, the 500FG study generated a large database of immunological, phenotypic and multi-omics data from a cohort of 534 individuals of Western-European ancestry, which have been used to integrate the impact of genetic and environmental factors on cytokine production and immune parameters. We subsequently deciphered the factors that influence inter-individual variation in the immune responses against different

77　stimuli (Bakker et al., 2018; Li et al., 2016; Schirmer et al., 2016; Ter Horst et al.,

78　2016).

79

80　**Results and Discussion**

81　Peripheral blood mononuclear cells from these individuals were challenged with

82　bacterial, fungal, viral and non-microbial stimuli, and six cytokines (TNFα, IL-1β, IL-6,

83　IL-17, IL-22 and IFNγ) were measured at 24h or 7 days after stimulation, generating

84　105 cytokine-stimulation pairs (Fig. S1 and Table S1). We correlated cytokine

85　production with genetic variant data to obtain cytokine quantitative trait loci (QTLs),

86　which were employed to compute and compare the polygenic risk score (PRS) of the

87　genomes of 827 individuals from different human historical eras (early upper

88　Paleolithic, late upper Paleolithic, Mesolithic, Neolithic, post-Neolithic) which were

89　downloaded from version 37.2 of the compiled dataset containing unimputed published

90　ancient genotypes (https://reich.hms.harvard.edu/downloadable-genotypes-present-

91　day-and-ancient-dna-data-compiled-published-papers), and 250 modern Europeans

92　randomly selected from the European 1000G cohort (see accompanying manuscript by

93　Kuijpers et al.). We then investigated how the PRS changes over time by constructing

94　linear models and performing correlation analysis. In order to account for the ancient

95　DNA samples being pseudo-haploid, ambiguous SNPs (A/T and C/G) were excluded

96　when computing PRS to prevent errors due to strand flips. PRS was computed using

97　the most significant QTLs that had a P value lower than our predetermined threshold

98　for each given trait and removing all variants within a 250kb window around these

99　variants. The dosage of these variants was multiplied by their effect size while the

100　dosage of missing variants in a sample were supplemented with the average dosage.

101　Finally, we scaled the PRS to a range of -1 and 1 and correlated the scores of the

102　samples with their respective carbon dated age. In order to verify the robustness of our

103　results we repeated the analysis at multiple threshold combinations for variant

104   missingness and QTL thresholds. Furthermore, an analysis-based down-sampling

105   approach shows that the trajectories observed in our results are consistent regardless

106   of the sample size (Fig. S2). A schematic representation of the steps performed is

107   shown in Fig. S3.

108

109   Applying the methodology described above, several patterns were apparent (Fig. 1).

110   The first overall observation is that the estimation of cytokine production capacity

111   based on PRS shows significant differences between populations in various historical

112   periods, and the strength of evolutionary pressure on cytokine responses was different

113   before and after the Neolithic revolution. We did not observe significant changes in

114   cytokine production capacity between individuals who lived at different historical

115   periods before the Neolithic, whereas strong pressure is apparent after adoption of

116   agriculture and animal domestication in Europe. This different pattern may have

117   resulted from the more limited number of samples available for the older time periods,

118   resulting in lower statistical power, but the presence of some evolutionary pressure

119   also before the Neolithic argues that this is most likely not the full explanation. The

120   development of agriculture and domestication of animals in the Neolithic increased

121   population densities on the one hand, and the contact between humans and

122   domesticated animals as source of pathogens on the other hand. The number of

123   zoonoses increases dramatically (examples being tuberculosis, brucellosis, Q-fever,

124   and influenza), which strongly increased the selective pressure and caused significant

125   adaptations of immunity at the genetic level (Flandroy et al., 2018). Most of the genetic

126   adaptations to pathogens took place in the period since modern humans abandoned

127   their hunting-gathering lifestyle and developed agriculture (Deschamps et al., 2016). In

128   this respect, the strongest changes leading to tolerance (decreased cytokine

129   production) were exerted in the cytokine responses to intracellular zoonotic infections

130   (tuberculosis and *Coxiella*) (Fig. 1). In contrast, responses to the extracellular

131 pathogens *Staphylococcus aureus* and *Candida albicans* indicate increased resistance,

132 with high production of IL-22 and TNFα, respectively. The increased response to the

133 important fungal pathogen *C. albicans* after the Neolithic period is validated also at

134 transcriptional level. Overall, these patterns are reminiscent of the studies showing that

135 human immune responses need to adapt to a new landscape of infectious agents

136 depending on geographical location and types of microbe encountered (Ferwerda et

137 al., 2007). Such different patterns were most likely encountered also through history.

138 Importantly, our results also show significant patterns in the changes of the production

139 of specific cytokines during history. The resistance against intracellular pathogens

140 increased after Neolithic with higher IFNγ responses (see Fig.1): indeed, it is known

141 that Th1-IFNγ responses are crucial for the host defense against intracellular

142 pathogens such as mycobacteria or *Coxiella* (Thakur et al., 2019). In addition, the

143 resistance to the extracellular pathogens *C. albicans* and *S. aureus* is also increased

144 after this Neolithic era, with TNFα and IFNγ production increasing steadily after. These

145 two cytokines are very well known to be important for anti-*Candida* and anti-

146 *Staphylococcus* host defense (Chan et al., 2018; Domínguez-Andrés et al., 2017). On

147 the other hand, a different pattern emerges in relation with the IL1β-IL6-IL17 axis: the

148 production of these cytokines is decreasing after Neolithic (see Figs. 1a and 1b). In this

149 context, the decrease through time of poly I:C induction of cytokines, as a model of

150 viral stimulation, is intriguing but potentially very important: many important viruses

151 such as influenza and coronaviruses (SARS, MERS, and SARS-CoV-2) exert life-

152 threatening effects through induction of cytokine-mediated hyperinflammation (also

153 termed "cytokine storm') (Tay et al., 2020): evolutionary processes to curtail this

154 exaggerated responses are thus likely to be protective, and tolerance against viruses

155 become a host defense mechanism (Diard and Hardt, 2017).

156 These evolutionary genetic adaptations to pathogens throughout human history greatly

157 influence the way we respond to multiple diseases in modern times as well. To assess

158 these effects, we calculated the PRS associated with the risk of several highly

159 prevalent immune mediated diseases. The first focus was on common infectious

160 diseases such as malaria, HIV-AIDS, tuberculosis and chronic viral hepatitis; we

161 calculated the changes in susceptibility to these diseases in the last 50.000 years of

162 human history, based on summary statistics from genome-wide association studies

163 (GWAS) databases available from the literature (Fig. S3). Our results show that

164 humans are becoming more resistant to these diseases, with the notable exception of

165 tuberculosis, whose risk score remained stable along the period studied (Fig. 2). These

166 results suggest that humans have built up a genetic makeup which made them more

167 resistant to a variety of microbes. The pattern of this adaptation is very interesting as

168 well, with a suggested decrease of susceptibility to malaria especially in the last 10.000

169 years. The reason for this accelerated resistance after Neolithic might be linked to

170 higher disease prevalence due to increased populations density, as otherwise

171 *Plasmodium* parasites are known to have circulated in Africa since at least the

172 Paleogene 30 million years ago (Poinar, 2005), and we have likely inherited it from

173 gorillas(Liu et al., 2010). Intriguingly, we also observe a strong decrease in

174 susceptibility to HIV: this is a contemporary pathogen, therefore this signal could be

175 due to common genetic and immune pathways with other infections that were present

176 in human populations. The increased resistance to HIV in Europeans may be derived

177 from selective pressures induced by other pathogens such as *Yersinia pestis* (Duncan

178 et al., 2005). Our data suggest on the other hand that the source of this increased

179 resistance is even older.

180 In contrast, the lack of genetic adaptation in the susceptibility to tuberculosis is

181 intriguing. This surprising finding may be explained by a concept in which *M.*

182 *tuberculosis* is at the same time a pathogen and a symbiont, in which latent infection

183 enhances the resistance against other pathogens and this is why our immune system

184 tolerates mycobacterial presence (Pai et al., 2016). In this regard, individuals with

185   latent TB exhibit enhanced macrophage functions that may protect against other

186   pathogens through the induction of trained immunity (Joosten et al., 2018). In this

187   context humanity may not be adapting to tuberculosis because increased resistance

188   against mycobacteria is not evolutionarily advantageous. All in all, these results

189   suggest that the risk of suffering infectious diseases has steadily decreased at least for

190   the last 50000 years as a result of the selection of genetic variants which confer

191   resistance to infections.

192   It has been proposed that the increased prevalence of inflammatory and autoimmune

193   diseases is associated with the immune-related alleles that have been positively

194   selected through evolutionary processes to protect against infections, hence the

195   contrasting differences in the prevalence of autoimmune diseases between populations

196   results from diverse selective pressures (Ramos et al., 2015). In line with this, it has

197   been hypothesized that genetic variants associated with protection against infectious

198   agents are behind the increased prevalence of autoimmune diseases in populations

199   with low pathogen exposure, such as Europeans (Fumagalli et al., 2011; Raj et al.,

200   2013). To study the changing patterns of susceptibility to autoimmune and

201   inflammatory disease during history, we used publicly available summary statistics from

202   GWAS of digestive tract-related autoimmune and inflammatory diseases and arthritis-

203   related diseases (Fig. S4) and calculated the PRS for each of samples under study.

204   Interestingly, we observed a robust increase of the genetic variants related with the

205   development of inflammatory diseases in the digestive tract after the Neolithic

206   revolution (Fig. 3). PRS scores associated with celiac disease, Crohn's disease,

207   ulcerative colitis and inflammatory bowel disease, were strongly associated with the

208   age of the samples, regardless of the P value thresholds or the missing genotype rates

209   used for PRS calculation, showing the robustness of these results (Fig. S2). The fact

210   that especially intestinal inflammatory pathology is increased after a historical event

211   that fundamentally modified human diet is unlikely to be an accident. Our results are in

212   line with earlier research demonstrating that variants in genes important for immune

213   responses and involved in celiac disease pathophysiology (such as IL-12, IL-18RAP,

214   SH2B3) are under strong positive selection (Zhernakova et al., 2010). The reasons for

215   the selection pressure on these genes are not completely understood, but an

216   advantage for host defense has been suggested (Zhernakova et al., 2010).

217   In contrast to intestinal inflammation, the PRS of traits linked with juvenile-idiopathic

218   arthritis, rheumatoid arthritis and multiple sclerosis shows a decrease in genetic

219   susceptibility with the age of the sample after the Neolithic revolution. For pre-Neolithic

220   periods, these patterns had little impact with decreasing PRS for digestive tract

221   diseases and increasing PRS for ankylosing spondylitis and juvenile idiopathic arthritis.

222   A strong decrease in susceptibility to juvenile idiopathic arthritis, rheumatoid arthritis

223   and multiple sclerosis is seen after the Neolithic period (see Fig. 3). This is likely linked

224   to the decreased production of the IL-1/IL-6/IL-17 axis described in Fig. 2, which is

225   particularly important in the pathophysiology of these disorders (Akioka, 2019; Mei et

226   al., 2011).

227   The significant changes in cytokine production and disease susceptibility in European

228   populations after the Neolithic can be due to selective processes on the one hand (as

229   described above), but also with important demographic changes due to migrations of

230   human communities such as the Anatolians (in Neolithic) or the Yamnaya populations

231   from the Pontic steppe (during the Bronze Age) (Racimo et al., 2020). In this regard,

232   several loci associated with inflammatory disease displayed a group alleles linked with

233   Crohn's disease, celiac disease and ulcerative colitis in Neolithic Aegeans, the

234   community who spread farming across Europe (Hofmanová et al., 2016), with several

235   of these alleles showing signs of positive selection in modern Europeans (Raj et al.,

236   2013). In addition, the gene expression PRS of several cytokines based on the *cis-* and

237   *trans-* eQTLs from the eQTLGen Consortium (https://www.eqtlgen.org/) displayed a

238   very strong association with time for TNFα after the Neolithic revolution (Fig. 4).

10

239

240 Collectively, our results show that the advent of the Neolithic era was a turning point for

241 the evolution of immune-mediated traits in European populations, driving the expansion

242 of alleles that favor the development of tolerance against intracellular pathogens and

243 promote inflammatory responses against extracellular microbes. This is associated with

244 a higher presence of genetic traits related with inflammatory and auto-immune

245 diseases of the digestive tract and a lower number of alleles linked with the

246 development of arthritis. Further research should compare the trends in different

247 populations that have been exposed to different environments across the planet and

248 clarify the influence of ancestry, time, rural vs. urban lifestyle to shed light on the

249 influence of the infectious environment in genetics and human evolution.

250

257

258 **Author contributions**

259 JDA, YK, OBB and MJ designed and performed experiments and analysed the data.

260 JDA and YK wrote the first draft of the manuscript with all authors contributing to writing

261 and providing feedback. MJ, CJX, JWMvdM, LABJ and JB provided guidance and

262 advice. YL and MGN conceived ideas, designed experiments, offered supervision and

263 oversaw the research program.

264

Declaration of Interests

266   The authors declare no competing interests.

267

Methods

Cohort selection

270   Ancient DNA genotype data was downloaded from version 37.2 of the published aDNA

271   genotype database, compiled by and available on the David Reich Lab website

272   (https://reich.hms.harvard.edu/downloadable-genotypes-present-day-and-ancient-dna-

273   data-compiled-published-papers). The ancient DNA samples consisted of pseudo-

274   haploid genotype data. This was due to the low genotyping coverage. Samples with

275   variant missingness above 96 percent were filtered out using Plink (Purcell et al.,

276   2007). This was done in order to remove outliers with extremely low coverage. Only

277   samples within Europe were used for this study, these samples were selected based

278   on their geographic location, that is latitude (within 35 and 70 degrees north) and

279   longitude (within 10 degrees west and 40 degrees east). Samples without a carbon-

280   dated age were also filtered out. We also selected 250 European samples from the

281   1000 genomes project phase 3. Only variants present in both the ancient samples and

282   the modern samples were retained. This resulted in a dataset of 827 ancient samples

283   and 250 modern samples containing 1233013 variants.

284

Carbon-dated sample origin and geographical location

286   Both carbon-dated age of origin as well as latitudinal and longitudinal data was

287   available for these 827 ancient European samples. Broad time periods were assigned

288   to these samples with the Early Upper Paleolithic era for all samples originating from

12

289 before 25000 years before the common era standardized to 1950 (BCE). The Late

290 Upper Paleolithic era follows until 11000 BCE. The Mesolithic era ranges from 11000 to

291 5500 BCE. The Neolithic era ranges from 8500 to 3900 BCE, and the Post-Neolithic

292 era ranges from 5000 BCE and more recent ages. Using the geographical data in

293 combination with archeological clues and the genetic data, the broad time period of

294 origin was also available for samples that were dated to a point in time with overlapping

295 broad time periods. This allowed the samples to be classified as either Early Upper

296 Paleolithic, Late Upper Paleolithic, Mesolithic, Neolithic, or Post-Neolithic. The sample

297 age of the 250 modern European samples was set to 0.

298

299 <u>Summary statistics of GWAS and cytokine QTLs</u>

300 Summary statistics for complex traits were obtained from the UK Biobank (Bycroft et

301 al., 2018) and the GWAS catalog (MacArthur et al., 2017) last accessed on 29th of

302 March 2020. The stimulated cytokine response summary statistics from the 500FG

303 cohort of the HFGP were used (Li et al., 2016). Some complex traits had multiple

304 different sets of summary statistics available. In these cases, the data which was more

305 recent and used bigger cohorts that were either of European or mixed (European and

306 Asian) ancestry were selected. The variants of these summary statistics were then

307 filtered by only keeping bi-allelic variants. Most aDNA genotypes available are pseudo-

308 haploid as a consequence of their lower sample quality. We excluded ambiguous SNPs

309 (A/T and C/G) in order to prevent errors due to strand flips present in these pseudo-

310 haploid samples.

311

312 <u>Polygenic Risk Scores (PRS) calculation</u>

313 Polygenic risk scores were then calculated by first intersecting the filtered variants from

314 the summary statistics with the variants present in the DNA samples. Starting at the

13

315   most significant variant, all variants within a 250kb window around that variant were

316   excluded until no variants remained. We then multiplied the dosage of these variants

317   with the effect size and these values were summed. If a variant is missing in a sample

318   the dosage is substituted with the average genotyped dosage for that variant within the

319   entire dataset. This way the PRS is not skewed in any specific direction. The formula

320   for this is described below with the score S being the weighted sum of a variant's

321   dosage X_n multiplied by its associated weight or beta β_n calculated using m variants.

$$S = \sum_{n=1}^{m} X_n \, \beta_n$$

322

323   <u>Piecewise correlation analysis</u>

324   We constructed piecewise linear models for each trait by separating the samples into

325   two groups. These two groups consisted of all samples preceding the Neolithic era and

326   those of the Neolithic era and later respectively. We correlated PRS with the carbon

327   dated age of our samples. We then multiplied the -log10 of the correlation P values

328   with the sign of the correlation coefficients.

329

330   <u>Robustness of results</u>

331   In order to test the robustness of our results we calculated PRS using multiple different

332   P value thresholds for QTL inclusion. We used P value thresholds from 10^{-3} to 10^{-8} for

333   the complex traits obtained through GWAS catalog and the UK Biobank. The

334   thresholds used for the stimulated cytokine responses ranged from 10^{-3} to 10^{-6}. We

335   also calculated PRS using different variant missingness thresholds. This means we

336   removed samples with a variant missingness rate higher than 96, 90, 80, or 70 percent.

337   All of the results from the piecewise linear models were then used to create a heatmap

338   depicting the consistency and robustness of our observed correlations.

339   Additionally, various window-sizes were used for clumping the QTL's and LD based

340   clumping was also performed excluding variants with an LD greater than 0.2 compared

341   to our lead SNP within a window. In order to see whether our observations were due to

342   sample imbalances between the pre-Neolithic period and the later periods samples

343   originating from the Neolithic period and later were randomly down-sampled to the

344   same number of samples as the pre-Neolithic samples. Correlation coefficients

345   between PRS and sample age were then recalculated for the Neolithic and younger

346   samples and compared to the coefficients obtained using all Neolithic and younger
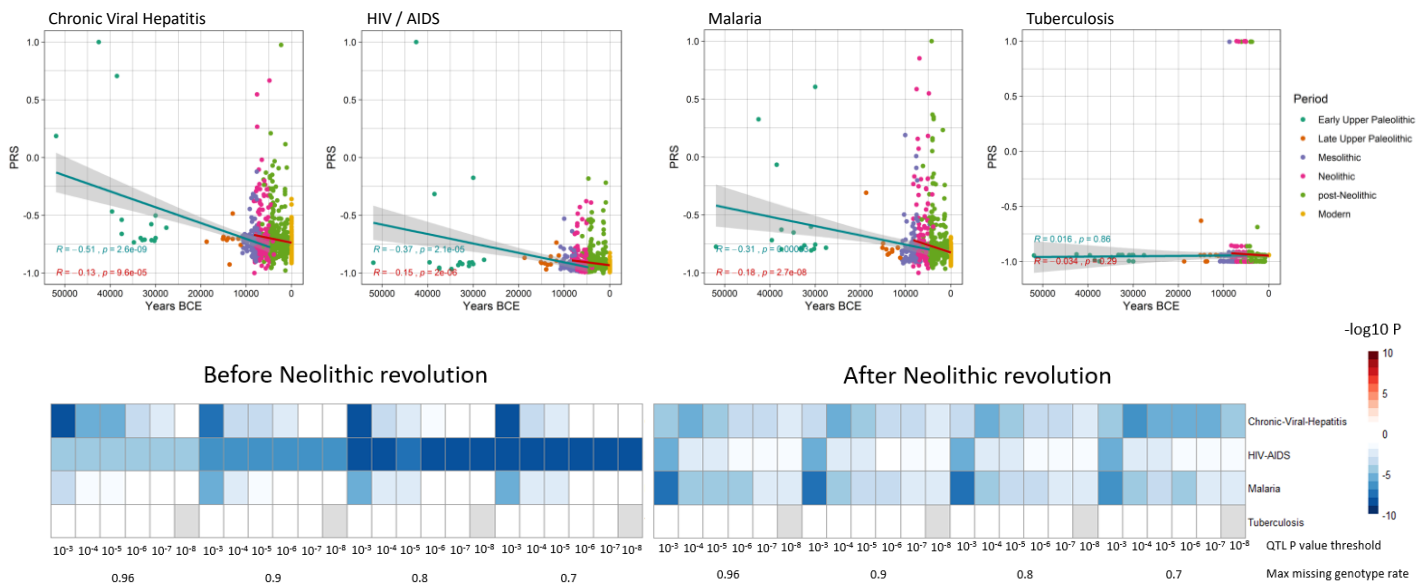
347   samples.

348
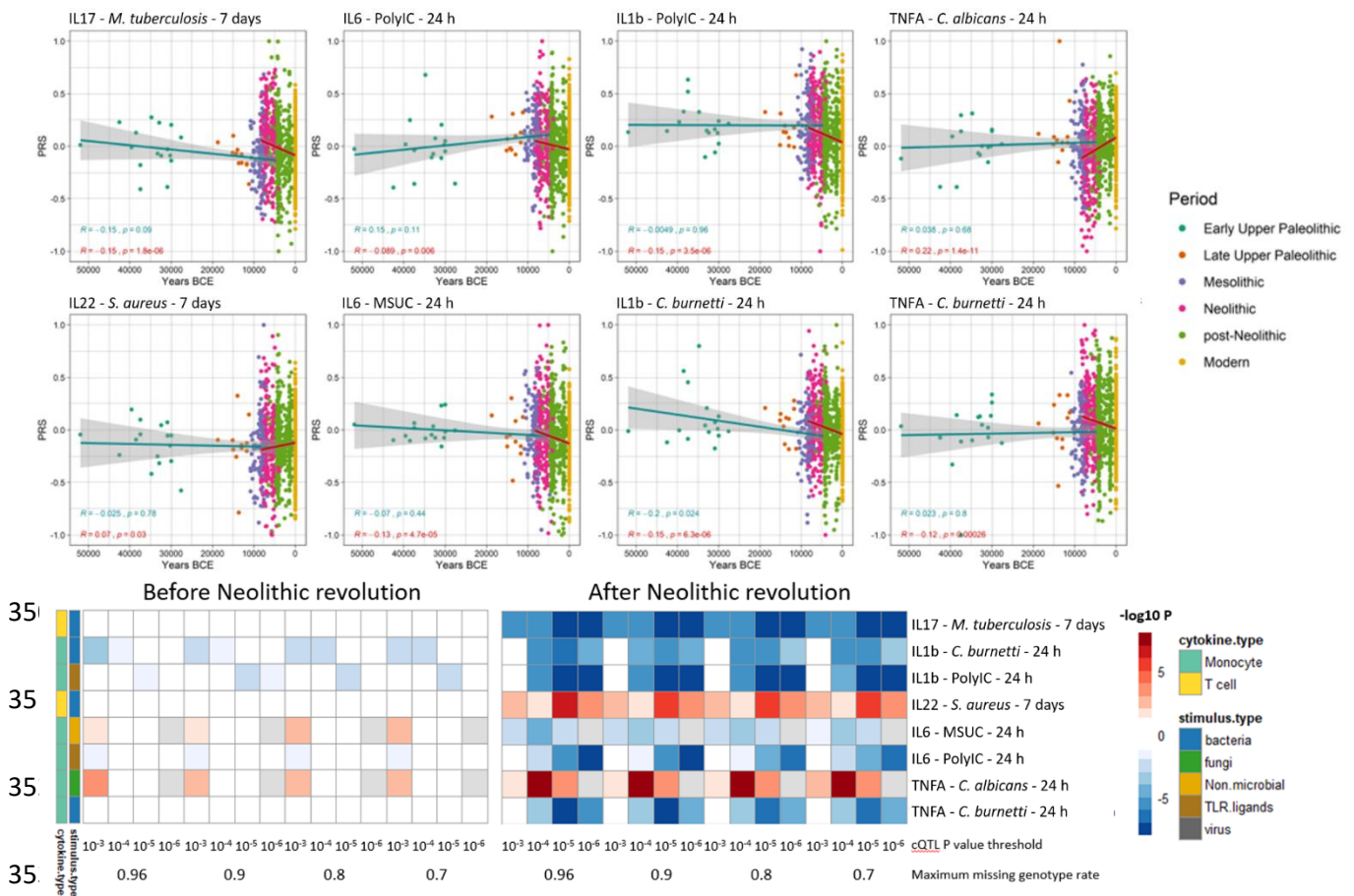
Figures:

Figure 1: A) Correlation between cytokine PRS and time. Samples are colored by broad age period. The blue regression lines show PRS before the Neolithic revolution remained relatively constant for all traits whereas the red regression lines show the correlation after the start of the Neolithic period. The threshold of max missing genotype per sample was 0.96 and QTL P value cutoff was 10^{-4}. MSUC: Monosodium urate crystals. B) **Correlation between cytokine PRS and time. using multiple thresholds reveals consistent trend.** Missing genotype rate ranged from 0.96, 0.9, 0.8, and 0.7. QTL P value for variants included in our PRS models ranged from 10^{-3}, 10^{-4}, 10^{-5}, and 10^{-6}. The color key indicates the range of -log10 P values of the Pearson correlation between PRS and time. Red and blue indicate positive and negative association, respectively.

Figure 2: Infectious disease risk PRS scores decrease with time, except tuberculosis. A) Samples are colored by broad age period. The blue regression lines show PRS before the Neolithic revolution remained relatively constant for all traits whereas the red regression lines show that the correlation after the start of the Neolithic period changed significantly. The threshold of max missing genotype per sample was 0.96 and QTL P value cutoff was 10^{-3}. MSUC: Monosodium urate crystals. **B) Correlation between disease PRS and time using multiple thresholds reveals consistent trend.** Miss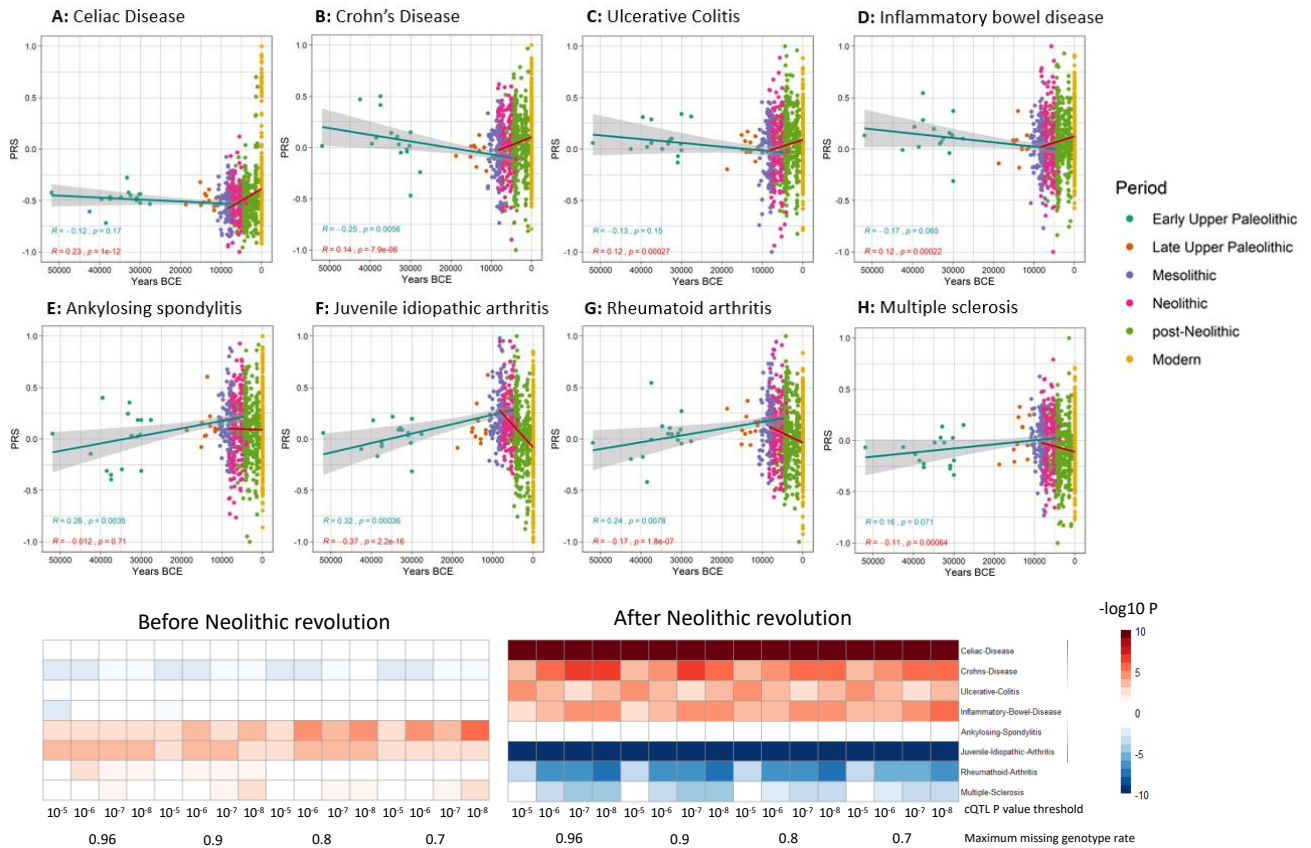ing genotype rate ranged from 0.96, 0.9, 0.8, and 0.7. QTL P value for variants included in our PRS models ranged from 10^{-3}, 10^{-4}, 10^{-5}, 10^{-6}, 10^{-7}, and 10^{-8}. The color key indicates the range of -log10 P values of the Pearson correlation between PRS and time. Red and blue indicate positive and negative association, respectively.

380   **Figure 3: Correlation between auto-immune and inflammatory disease PRS and**

381   **time. A)** Samples are colored by broad age period. The blue regression lines show

382   PRS before the Neolithic revolution remained relatively constant for all traits whereas

383   the red regression lines show that the correlation after the start of the Neolithic period

384   changed significantly. The threshold of max missing genotype per sample was 0.96

385   and QTL P value cutoff was 10^{-4}. MSUC: Monosodium urate crystals. **B) Correlation**

386   **between disease PRS and time using multiple thresholds reveals consistent**

387   **trend.** Missing genotype rate ranged from 0.96, 0.9, 0.8, and 0.7. QTL P value for

388   variants included in our PRS models ranged from 10^{-5}, 10^{-6}, 10^{-7}, and 10^{-8}. The color

389   key indicates the range of -log10 P values of the Pearson correlation between PRS and

390   time. Red and blue indicate positive and negative association, respectively.

391

392
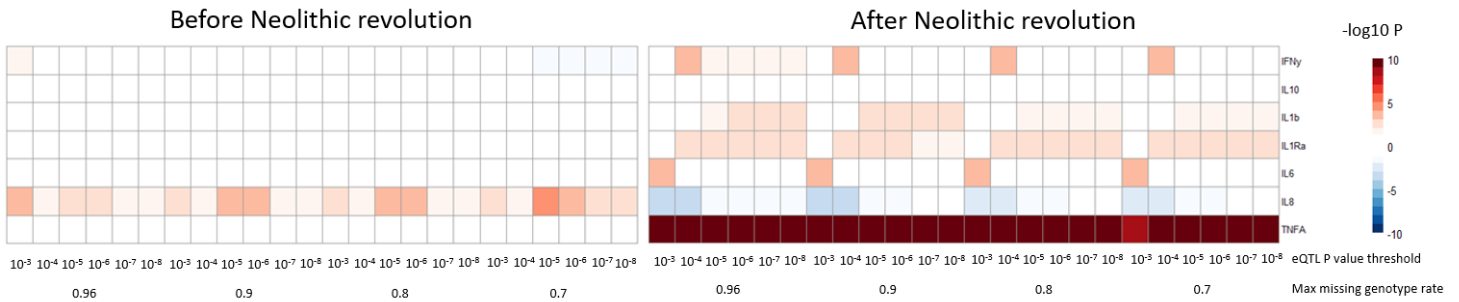
Figure 4: Cytokine gene expression PRS scores using cis- and trans-eQTLs correlated with time. Most notably is the highly significant increase in *TNFA* gene expression PRS over time following the Neolithic revolution. Prior to the Neolithic revolution an increase in *IL8* gene expression PRS can be observed which shifts to a decreasing trend after the Neolithic revolution. Both *IL1B* and *IL1RN* gene expression show a slight increase in PRS over time after the start of the Neolithic revolution. Missing genotype rate ranged from 0.96, 0.9, 0.8, and 0.7. QTL P value for variants included in our PRS models ranged from 10^{-3}, 10^{-4}, 10^{-5}, 10^{-6}, 10^{-7}, and 10^{-8}. The color key indicates the range of -log10 P values of the Pearson correlation between PRS and time. Red and blue indicate positive and negative association, respectively.
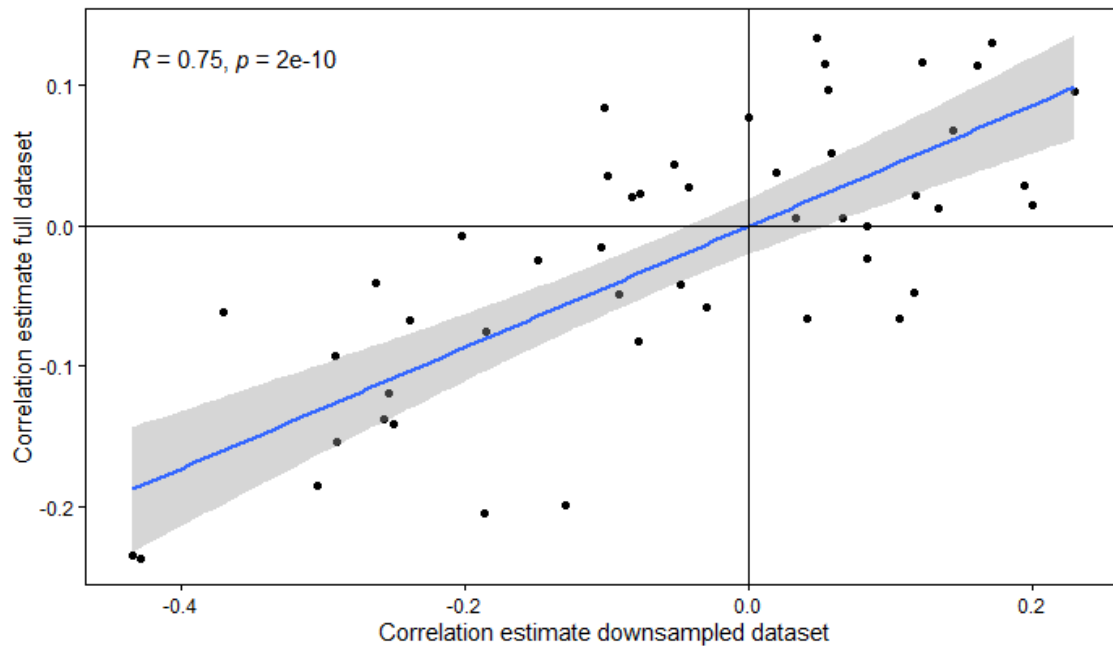
19
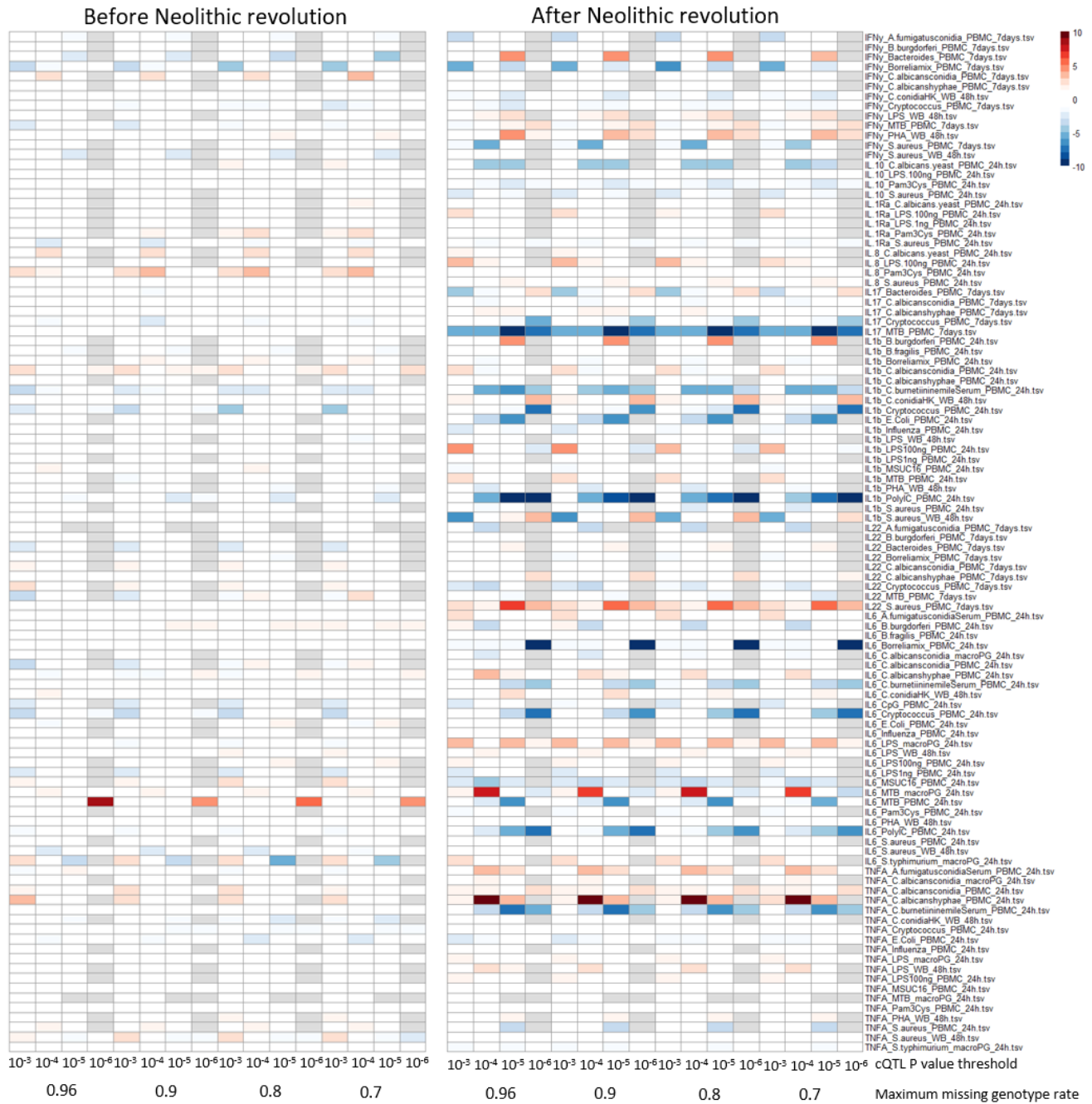
Figure S1: Correlation between cytokine PRS and time. Missing genotype rate ranged from 0.96, 0.9, 0.8, and 0.7. QTL P value for variants included in our PRS models ranged from 10^{-3}, 10^{-4}, 10^{-5}, and 10^{-6}. The color key indicates the range of -log10 P values of the Pearson correlation between PRS and time. Red and blue indicate positive and negative association, respectively.

412

Figure S2: Robustness of correlation coefficients post Neolithic independent of sample size. Changes in PRS following the Neolithic revolution remain consistent after down-sampling samples from after the start of the Neolithic period to the same amount as samples before the Neolithic period. The lower number of samples reduces the power which reduces the amount of significant correlations but does not influence the direction of changes in PRS which were previously identified as significant.

419
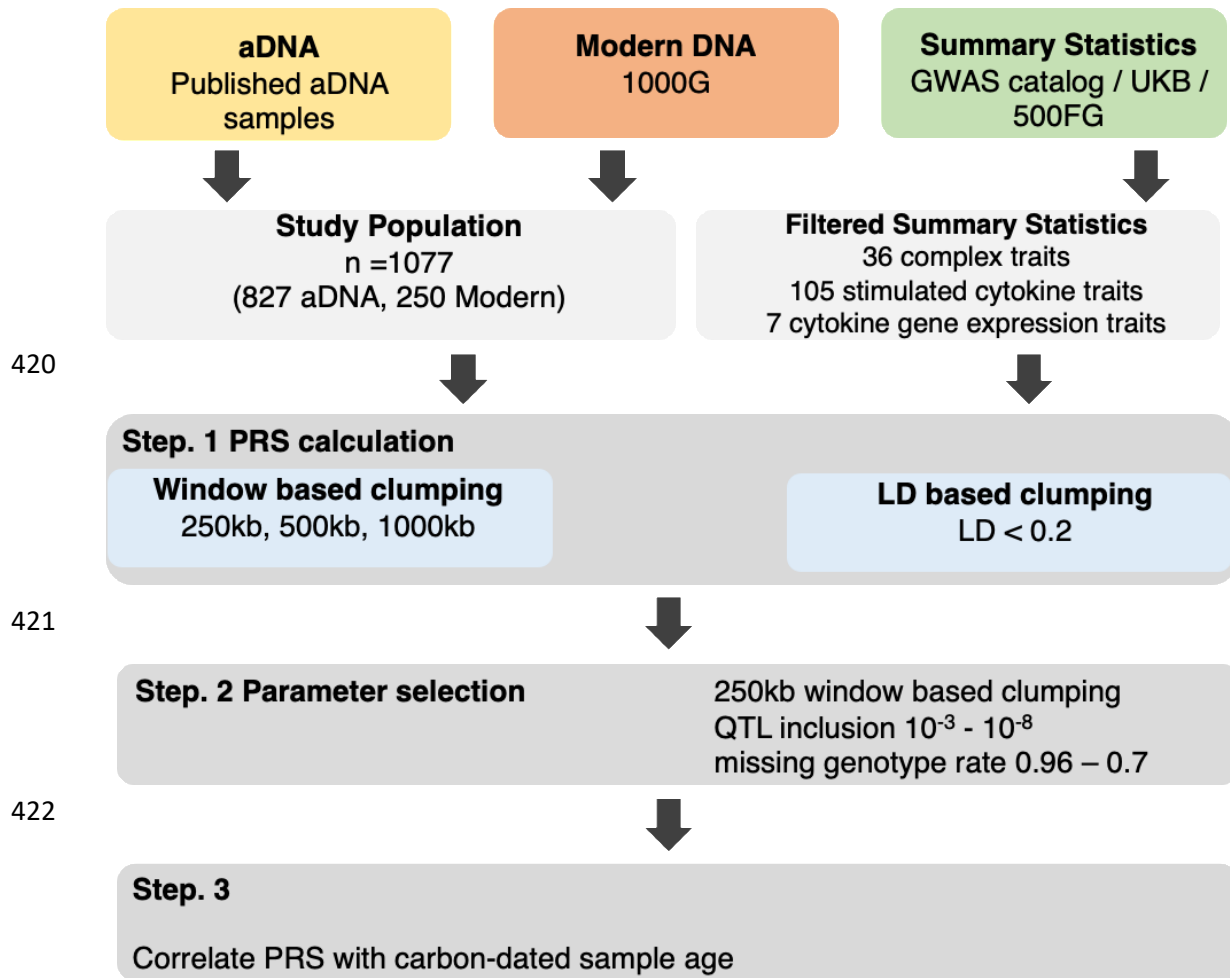
Figure S3: Both aDNA and modern DNA samples of European individuals were used in combination with summary statistics from predominantly European populations to calculate PRS of immune-related traits. This was done at various threshold combinations before correlating the scores with the sample age.

	Trait	Source/PMID	Population	Cohort Size
Auto-immune and inflammatory diseases	Celiac Disease	22057235	European Ancestry	24269
	Crohns Disease	26192919	European Ancestry: 86640 Asian Ancestry: 9846	96486
	Ulcerative Colitis	26192919	European Ancestry: 86640 Asian Ancestry: 9846	96486
	Inflammatory Bowel Disease	26192919	European Ancestry: 86640 Asian Ancestry: 9846	96486
	Ankylosing Spondylitis	23749187	European Ancestry	25764
	Juvenile Idiopathic Arthritis	23603761	European Ancestry	2816
	Rheumatoid Arthritis	24390342	European and Asian Ancestry	103638
	Multiple Sclerosis	24076602	European Ancestry	80094
Infectious diseases	Chronic Viral Hepatitis	UKB round 2	European Ancestry	361194
	HIV / AIDS	UKB round 2	European Ancestry	361194
	Malaria	UKB round 2	European Ancestry	361194
	Tuberculosis	UKB round 2	European Ancestry	361194

Figure S4: GWAS summary statistics and cohorts used for PRS calculation. Traits were separated into two categories: Auto-immune and inflammatory diseases-related traits, and infectious diseases-related trait. GWAS summary statistics from predominantly European populations were selected.

23

	PBMC	Macrophage	WB	PBMC
	24h	24h	48h	7d
A.fumigatusconidia				IFNy, IL22
B.burgdorferi	IL1b, IL6			IFNy, IL22,
B.fragilis	IL1b, IL6			
Bacteroides				IFNy, IL17, IL22
Borreliamix	IL1b, IL6			IFNy, IL22
C.albicans.yeast	IL1Ra, IL8, IL10			
C.albicansconidia	IL1b, IL6, TNFa	IL6, TNFa		IFNy, IL17, IL22
C.albicanshyphae	IL1b, IL6, TNFa			IFNy, IL17, IL22
C.burnetiininemileSerum	IL1b, IL6, TNFa			
C.conidiaHK			IFNy, IL1b, IL6, TNFa	
CpG	IL6			
Cryptococcus	IL1b, IL6, TNFa			IFNy, IL17, IL22
E.Coli	IL1b, IL6, TNFa			
A.fumigatusconidia Serum	IL6, TNFa			
Influenza	IL1b, IL6, TNFa			
LPS		IL6, TNFa	IFNy, IL1b, IL6, TNFa	
LPS 100ng	IL1b, IL1Ra, IL6, IL8, IL10, TNFa			
LPS 1ng	IL1b, IL1Ra, IL6, IL8			
MSUC16	IL1b, IL6, TNFa			
MTB	IL1b, IL6,	IL6, TNFa		IFNy, IL17, IL22
Pam3Cys	IL1Ra, IL6, IL8, IL10, TNFa			
PHA			IFNy, IL1b, IL6, TNFa	
Poly IC	IL1b, IL6			
S.aureus	IL1b, IL1Ra, IL6, IL8, IL10, TNFa		IFNy, IL1b, IL6, TNFa	IFNy, IL22
S.typhimurium		IL6, TNFa		
Total N	58	8	16	23

Table 1: Overview of the stimulus, cytokine, and timepoint combinations. In total 105 unique stimulated cytokine traits were available using various types of stimuli measuring both the innate and adaptive immune response.

References:

448

449 Akioka S. 2019. Interleukin-6 in juvenile idiopathic arthritis. *Mod Rheumatol* **29**:275–

450   286. doi:10.1080/14397595.2019.1574697

451 Bakker OB, Aguirre-Gamboa R, Sanna S, Oosting M, Smeekens SP, Jaeger M, Zorro

452   M, Võsa U, Withoff S, Netea-Maier RT, Koenen HJPM, Joosten I, Xavier RJ,

453   Franke L, Joosten LAB, Kumar V, Wijmenga C, Netea MG, Li Y. 2018. Integration

454   of multi-omics data and deep phenotyping enables prediction of cytokine

455   responses. *Nat Immunol* **19**:776–786. doi:10.1038/s41590-018-0121-3

456 Bycroft C, Freeman C, Petkova D, Band G, Elliott LT, Sharp K, Motyer A, Vukcevic D,

457   Delaneau O, O'Connell J, Cortes A, Welsh S, Young A, Effingham M, McVean G,

458   Leslie S, Allen N, Donnelly P, Marchini J. 2018. The UK Biobank resource with

459   deep phenotyping and genomic data. *Nature* **562**:203–209. doi:10.1038/s41586-

460   018-0579-z

461 Chan LC, Rossetti M, Miller LS, Filler SG, Johnson CW, Lee HK, Wang H, Gjertson D,

462   Fowler VG, Reed EF, Yeaman MR. 2018. Protective immunity in recurrent

463   Staphylococcus aureus infection reflects localized immune signatures and

464   macrophage-conferred memory. *Proc Natl Acad Sci U S A*.

465   doi:10.1073/pnas.1808353115

466 Deschamps M, Laval G, Fagny M, Itan Y, Abel L, Casanova J-L, Patin E, Quintana-

467   Murci L. 2016. Genomic Signatures of Selective Pressures and Introgression from

468   Archaic Hominins at Human Innate Immunity Genes. *Am J Hum Genet* **98**:5–21.

469   doi:10.1016/J.AJHG.2015.11.014

470 Diard M, Hardt WD. 2017. Evolution of bacterial virulence. *FEMS Microbiol Rev*.

471   doi:10.1093/femsre/fux023

472 Domínguez-Andrés J, Feo-Lucas L, Minguito de la Escalera M, González L, López-

473   Bravo M, Ardavín C. 2017. Inflammatory Ly6C high Monocytes Protect against

474   Candidiasis through IL-15-Driven NK Cell/Neutrophil Activation. *Immunity*

475   **46**:1059-1072.e4. doi:10.1016/j.immuni.2017.05.009

476   Duncan SR, Scott S, Duncan CJ. 2005. Reappraisal of the historical selective

477   pressures for the CCR5-Δ32 mutation. *J Med Genet*.

478   doi:10.1136/jmg.2004.025346

479   Ferwerda B, McCall MBB, Alonso S, Giamarellos-Bourboulis EJ, Mouktaroudi M,

480   Izagirre N, Syafruddin D, Kibiki G, Cristea T, Hijmans A, Hamann L, Israel S,

481   ElGhazali G, Troye-Blomberg M, Kumpf O, Maiga B, Dolo A, Doumbo O, Hermsen

482   CC, Stalenhoef AFH, van Crevel R, Brunner HG, Oh D-Y, Schumann RR, de la

483   Rua C, Sauerwein R, Kullberg B-J, van der Ven AJAM, van der Meer JWM, Netea

484   MG. 2007. TLR4 polymorphisms, infectious diseases, and evolutionary pressure

485   during migration of modern humans. *Proc Natl Acad Sci* **104**:16645–16650.

486   doi:10.1073/pnas.0704828104

487   Flandroy L, Poutahidis T, Berg G, Clarke G, Dao M-C, Decaestecker E, Furman E,

488   Haahtela T, Massart S, Plovier H, Sanz Y, Rook G. 2018. The impact of human

489   activities and lifestyles on the interlinked microbiota and health of humans and of

490   ecosystems. *Sci Total Environ* **627**:1018–1038.

491   doi:10.1016/J.SCITOTENV.2018.01.288

492   Fumagalli M, Sironi M. 2014. Human genome variability, natural selection and

493   infectious diseases. *Curr Opin Immunol* **30**:9–16.

494   Fumagalli M, Sironi M, Pozzoli U, Ferrer-Admettla A, Pattini L, Nielsen R. 2011.

495   Signatures of Environmental Genetic Adaptation Pinpoint Pathogens as the Main

496   Selective Pressure through Human Evolution. *PLoS Genet* **7**:e1002355.

497   doi:10.1371/journal.pgen.1002355

498   Hofmanová Z, Kreutzer S, Hellenthal G, Sell C, Diekmann Y, Díez-Del-Molino D, Van

499   Dorp L, López S, Kousathanas A, Link V, Kirsanow K, Cassidy LM, Martiniano R,

500   Strobel M, Scheu A, Kotsakis K, Halstead P, Triantaphyllou S, Kyparissi-

501   Apostolika N, Urem-Kotsou D, Ziota C, Adaktylou F, Gopalan S, Bobo DM,

502   Winkelbach L, Blöcher J, Unterländer M, Leuenberger C, Çilingiroğlu Ç, Horejs B,

503   Gerritsen F, Shennan SJ, Bradley DG, Currat M, Veeramah KR, Wegmann D,

504   Thomas MG, Papageorgopoulou C, Burger J. 2016. Early farmers from across

505   Europe directly descended from Neolithic Aegeans. *Proc Natl Acad Sci U S A*

506   **113**:6886–6891. doi:10.1073/pnas.1523951113

507   Joosten SA, Meijgaarden KE van, Arend SM, Prins C, Oftung F, Korsvold GE, Kik S V.,

508   Arts RJW, Crevel R van, Netea MG, Ottenhoff THM. 2018. Mycobacterial growth

509   inhibition is associated with trained innate immunity. *J Clin Invest* **128**:1837–1851.

510   doi:10.1172/JCI97508

511   Karlsson EK, Kwiatkowski DP, Sabeti PC. 2014. Natural selection and infectious

512   disease in human populations. *Nat Rev Genet* **15**:379–393. doi:10.1038/nrg3734

513   Li Y, Oosting M, Smeekens SP, Jaeger M, Aguirre-Gamboa R, Le KTT, Deelen P,

514   Ricaño-Ponce I, Schoffelen T, Jansen AFM, Swertz MA, Withoff S, van de Vosse

515   E, van Deuren M, van de Veerdonk F, Zhernakova A, van der Meer JWM, Xavier

516   RJ, Franke L, Joosten LAB, Wijmenga C, Kumar V, Netea MG. 2016. A Functional

517   Genomics Approach to Understand Variation in Cytokine Production in Humans.

518   *Cell* **167**:1099-1110.e14. doi:10.1016/j.cell.2016.10.017

519   Liu W, Li Y, Learn GH, Rudicell RS, Robertson JD, Keele BF, Ndjango JBN, Sanz CM,

520   Morgan DB, Locatelli S, Gonder MK, Kranzusch PJ, Walsh PD, Delaporte E,

521   Mpoudi-Ngole E, Georgiev A V., Muller MN, Shaw GM, Peeters M, Sharp PM,

522   Rayner JC, Hahn BH. 2010. Origin of the human malaria parasite Plasmodium

523   falciparum in gorillas. *Nature*. doi:10.1038/nature09442

524   MacArthur J, Bowler E, Cerezo M, Gil L, Hall P, Hastings E, Junkins H, McMahon A,

525   Milano A, Morales J, MayPendlington Z, Welter D, Burdett T, Hindorff L, Flicek P,

526   Cunningham F, Parkinson H. 2017. The new NHGRI-EBI Catalog of published

527   genome-wide association studies (GWAS Catalog). *Nucleic Acids Res*.

528   doi:10.1093/nar/gkw1133

529   Mei Y, Pan F, Gao J, Ge R, Duan Z, Zeng Z, Liao F, Xia G, Wang S, Xu S, Xu J, Zhang

530   L, Ye D. 2011. Increased serum IL-17 and IL-23 in the patient with ankylosing

531   spondylitis. *Clin Rheumatol*. doi:10.1007/s10067-010-1647-4

532   Netea MG, Joosten LAB, Li Y, Kumar V, Oosting M, Smeekens S, Jaeger M, Ter Horst

533   R, Schirmer M, Vlamakis H, Notebaart R, Pavelka N, Aguirre-Gamboa RR, Swertz

534   MA, Tunjungputri RN, Van De Heijden W, Franzosa EA, Ng A, Graham D, Lassen

535   K, Schraa K, Netea-Maier R, Smit J, De Mast Q, Van De Veerdonk F, Kullberg BJ,

536   Tack C, Van De Munckhof I, Rutten J, Van Der Graaf J, Franke L, Hofker M,

537   Jonkers I, Platteel M, Maatman A, Fu J, Zhernakova A, Van Der Meer JWM,

538   Dinarello CA, Van Der Ven A, Huttenhouwer C, Koenen H, Joosten I, Xavier RJ,

539   Wijmenga C. 2016. Understanding human immune function using the resources

540   from the Human Functional Genomics Project. *Nat Med*. doi:10.1038/nm.4140

541   Pai M, Behr MA, Dowdy D, Dheda K, Divangahi M, Boehme CC, Ginsberg A,

542   Swaminathan S, Spigelman M, Getahun H, Menzies D, Raviglione M. 2016.

543   Tuberculosis. *Nat Rev Dis Prim*. doi:10.1038/nrdp.2016.76

544   Poinar G. 2005. Plasmodium dominicana n. sp. (Plasmodiidae: Haemospororida) from

545   Tertiary Dominican amber. *Syst Parasitol*. doi:10.1007/s11230-004-6354-6

546   Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MAR, Bender D, Maller J, Sklar

547   P, De Bakker PIW, Daly MJ, Sham PC. 2007. PLINK: A tool set for whole-genome

548   association and population-based linkage analyses. *Am J Hum Genet*.

549   doi:10.1086/519795

550   Quintana-Murci L, Clark AG. 2013. Population genetic tools for dissecting innate

551   immunity in humans. *Nat Rev Immunol* **13**:280–293. doi:10.1038/nri3421

552   Racimo F, Woodbridge J, Fyfe RM, Sikora M, Sjögren KG, Kristiansen K, Linden M

553   Vander. 2020. The spatiotemporal spread of human migrations during the

554   European Holocene. *Proc Natl Acad Sci U S A*. doi:10.1073/pnas.1920051117

555   Raj T, Kuchroo M, Replogle JM, Raychaudhuri S, Stranger BE, De Jager PL. 2013.

556   Common Risk Alleles for Inflammatory Diseases Are Targets of Recent Positive

557   Selection. *Am J Hum Genet* **92**:517–529. doi:10.1016/j.ajhg.2013.03.001

558   Ramos PS, Shedlock AM, Langefeld CD. 2015. Genetics of autoimmune diseases:

559   insights from population genetics. *J Hum Genet* **60**:657–664.

560   doi:10.1038/jhg.2015.94

561   Schirmer M, Smeekens SP, Vlamakis H, Jaeger M, Oosting M, Franzosa EA, Jansen

562   T, Jacobs L, Bonder MJ, Kurilshikov A, Fu J, Joosten LAB, Zhernakova A,

563   Huttenhower C, Wijmenga C, Netea MG, Xavier RJ. 2016. Linking the Human Gut

564   Microbiome to Inflammatory Cytokine Production Capacity. *Cell* **167**:1125-

565   1136.e8. doi:10.1016/j.cell.2016.10.020

566   Tay MZ, Poh CM, Rénia L, MacAry PA, Ng LFP. 2020. The trinity of COVID-19:

567   immunity, inflammation and intervention. *Nat Rev Immunol*. doi:10.1038/s41577-

568   020-0311-8

569   Ter Horst R, Jaeger M, Smeekens SP, Oosting M, Swertz MA, Li Y, Kumar V,

570   Diavatopoulos DA, Jansen AFM, Lemmers H, Toenhake-Dijkstra H, van

571   Herwaarden AE, Janssen M, van der Molen RG, Joosten I, Sweep FCGJ, Smit

572   JW, Netea-Maier RT, Koenders MMJF, Xavier RJ, van der Meer JWM, Dinarello

573   CA, Pavelka N, Wijmenga C, Notebaart RA, Joosten LAB, Netea MG. 2016. Host

574   and Environmental Factors Influencing Individual Human Cytokine Responses.

575   *Cell* **167**:1111-1124.e13. doi:10.1016/j.cell.2016.10.018

576   Thakur A, Mikkelsen H, Jungersen G. 2019. Intracellular Pathogens: Host Immunity

577   and Microbial Persistence Strategies. *J Immunol Res* **2019**.

578   doi:10.1155/2019/1356540

579   Zhernakova A, Elbers CC, Ferwerda B, Romanos J, Trynka G, Dubois PC, de Kovel

580   CGF, Franke L, Oosting M, Barisani D, Bardella MT, Joosten LAB, Saavalainen P,

581   van Heel DA, Catassi C, Netea MG, Wijmenga C, Wijmenga C. 2010. Evolutionary

582   and Functional Analysis of Celiac Risk Loci Reveals SH2B3 as a Protective Factor

583   against Bacterial Infection. *Am J Hum Genet* **86**:970–977.

584   doi:10.1016/j.ajhg.2010.05.004

585

586