

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 Kuipers^{3,4*}, Olivier B. Bakker⁵, Martin Jaeger^{1,2},
5 Cheng-Jian Xu^{1,3,4}, Jos W.M. van der Meer¹, Mattias Jakobsson^{6,7}, Jaume Bertranpetit⁸,
6 Leo A.B. Joosten^{1,2}, Yang Li^{1,2,3,4,#}, Mihai G. Netea^{1,2,9,10,#}

7 Affiliations:

8 ¹Department 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 ²Radboud Institute for Molecular Life Sciences (RIMLS), Radboud University Medical
12 Center, 6525, GA, Nijmegen, the Netherlands

13 ³Department 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 ⁴TWINCORE, 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 ⁵Department of Genetics, University Medical Centre Groningen, Nijmegen, The
21 Netherlands.

22 ⁶Human Evolution, Department of Organismal Biology, Uppsala University, Uppsala,
23 Sweden.

24 ⁷Centre for Anthropological Research, Department of Anthropology and Development
25 Studies, University of Johannesburg, Auckland Park, South Africa.

26 ⁸Institut de Biologia Evolutiva (UPF-CSIC), Universitat Pompeu Fabra, Dr. Aiguader 88,
27 08003 Barcelona, Catalonia, Spain

28 ⁹Department for Genomics & Immunoregulation, Life and Medical Sciences Institute
29 (LIMES), University of Bonn, Bonn, Germany

30 ¹⁰Lead 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 autoimmune 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

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

Results and Discussion

Peripheral blood mononuclear cells from these individuals were challenged with bacterial, fungal, viral and non-microbial stimuli, and six cytokines (TNF α , IL-1 β , IL-6, IL-17, IL-22 and IFN γ) were measured at 24h or 7 days after stimulation, generating 105 cytokine-stimulation pairs (Fig. S1 and Table S1). We correlated cytokine production with genetic variant data to obtain cytokine quantitative trait loci (QTLs), which were employed to compute and compare the polygenic risk score (PRS) of the genomes of 827 individuals from different human historical eras (early upper Paleolithic, late upper Paleolithic, Mesolithic, Neolithic, post-Neolithic) which were downloaded from version 37.2 of the compiled dataset containing unimputed published ancient genotypes (<https://reich.hms.harvard.edu/downloadable-genotypes-present-day-and-ancient-dna-data-compiled-published-papers>), and 250 modern Europeans randomly selected from the European 1000G cohort (see accompanying manuscript by Kuijpers et al.). We then investigated how the PRS changes over time by constructing linear models and performing correlation analysis. In order to account for the ancient DNA samples being pseudo-haploid, ambiguous SNPs (A/T and C/G) were excluded when computing PRS to prevent errors due to strand flips. PRS was computed using the most significant QTLs that had a P value lower than our predetermined threshold for each given trait and removing all variants within a 250kb window around these variants. The dosage of these variants was multiplied by their effect size while the dosage of missing variants in a sample were supplemented with the average dosage. Finally, we scaled the PRS to a range of -1 and 1 and correlated the scores of the samples with their respective carbon dated age. In order to verify the robustness of our results we repeated the analysis at multiple threshold combinations for variant

missingness and QTL thresholds. Furthermore, an analysis-based down-sampling approach shows that the trajectories observed in our results are consistent regardless of the sample size (Fig. S2). A schematic representation of the steps performed is shown in Fig. S3.

Applying the methodology described above, several patterns were apparent (Fig. 1). The first overall observation is that the estimation of cytokine production capacity based on PRS shows significant differences between populations in various historical periods, and the strength of evolutionary pressure on cytokine responses was different before and after the Neolithic revolution. We did not observe significant changes in cytokine production capacity between individuals who lived at different historical periods before the Neolithic, whereas strong pressure is apparent after adoption of agriculture and animal domestication in Europe. This different pattern may have resulted from the more limited number of samples available for the older time periods, resulting in lower statistical power, but the presence of some evolutionary pressure also before the Neolithic argues that this is most likely not the full explanation. The development of agriculture and domestication of animals in the Neolithic increased population densities on the one hand, and the contact between humans and domesticated animals as source of pathogens on the other hand. The number of zoonoses increases dramatically (examples being tuberculosis, brucellosis, Q-fever, and influenza), which strongly increased the selective pressure and caused significant adaptations of immunity at the genetic level (Flandroy et al., 2018). Most of the genetic adaptations to pathogens took place in the period since modern humans abandoned their hunting-gathering lifestyle and developed agriculture (Deschamps et al., 2016). In this respect, the strongest changes leading to tolerance (decreased cytokine production) were exerted in the cytokine responses to intracellular zoonotic infections (tuberculosis and *Coxiella*) (Fig. 1). In contrast, responses to the extracellular

pathogens *Staphylococcus aureus* and *Candida albicans* indicate increased resistance, with high production of IL-22 and TNF α , respectively. The increased response to the important fungal pathogen *C. albicans* after the Neolithic period is validated also at transcriptional level. Overall, these patterns are reminiscent of the studies showing that human immune responses need to adapt to a new landscape of infectious agents depending on geographical location and types of microbe encountered (Ferwerda et al., 2007). Such different patterns were most likely encountered also through history.

Importantly, our results also show significant patterns in the changes of the production of specific cytokines during history. The resistance against intracellular pathogens increased after Neolithic with higher IFN γ responses (see Fig.1): indeed, it is known that Th1-IFN γ responses are crucial for the host defense against intracellular pathogens such as mycobacteria or *Coxiella* (Thakur et al., 2019). In addition, the resistance to the extracellular pathogens *C. albicans* and *S. aureus* is also increased after this Neolithic era, with TNF α and IFN γ production increasing steadily after. These two cytokines are very well known to be important for anti-*Candida* and anti-*Staphylococcus* host defense (Chan et al., 2018; Domínguez-Andrés et al., 2017). On the other hand, a different pattern emerges in relation with the IL1 β -IL6-IL17 axis: the production of these cytokines is decreasing after Neolithic (see Figs. 1a and 1b). In this context, the decrease through time of poly I:C induction of cytokines, as a model of viral stimulation, is intriguing but potentially very important: many important viruses such as influenza and coronaviruses (SARS, MERS, and SARS-CoV-2) exert life-threatening effects through induction of cytokine-mediated hyperinflammation (also termed “cytokine storm”) (Tay et al., 2020): evolutionary processes to curtail this exaggerated responses are thus likely to be protective, and tolerance against viruses become a host defense mechanism (Diard and Hardt, 2017).

These evolutionary genetic adaptations to pathogens throughout human history greatly influence the way we respond to multiple diseases in modern times as well. To assess

these effects, we calculated the PRS associated with the risk of several highly prevalent immune mediated diseases. The first focus was on common infectious diseases such as malaria, HIV-AIDS, tuberculosis and chronic viral hepatitis; we calculated the changes in susceptibility to these diseases in the last 50.000 years of human history, based on summary statistics from genome-wide association studies (GWAS) databases available from the literature (Fig. S3). Our results show that humans are becoming more resistant to these diseases, with the notable exception of tuberculosis, whose risk score remained stable along the period studied (Fig. 2). These results suggest that humans have built up a genetic makeup which made them more resistant to a variety of microbes. The pattern of this adaptation is very interesting as well, with a suggested decrease of susceptibility to malaria especially in the last 10.000 years. The reason for this accelerated resistance after Neolithic might be linked to higher disease prevalence due to increased populations density, as otherwise *Plasmodium* parasites are known to have circulated in Africa since at least the Paleogene 30 million years ago (Poinar, 2005), and we have likely inherited it from gorillas(Liu et al., 2010). Intriguingly, we also observe a strong decrease in susceptibility to HIV: this is a contemporary pathogen, therefore this signal could be due to common genetic and immune pathways with other infections that were present in human populations. The increased resistance to HIV in Europeans may be derived from selective pressures induced by other pathogens such as *Yersinia pestis* (Duncan et al., 2005). Our data suggest on the other hand that the source of this increased resistance is even older.

In contrast, the lack of genetic adaptation in the susceptibility to tuberculosis is intriguing. This surprising finding may be explained by a concept in which *M. tuberculosis* is at the same time a pathogen and a symbiont, in which latent infection enhances the resistance against other pathogens and this is why our immune system tolerates mycobacterial presence (Pai et al., 2016). In this regard, individuals with

latent TB exhibit enhanced macrophage functions that may protect against other pathogens through the induction of trained immunity (Joosten et al., 2018). In this context humanity may not be adapting to tuberculosis because increased resistance against mycobacteria is not evolutionarily advantageous. All in all, these results suggest that the risk of suffering infectious diseases has steadily decreased at least for the last 50000 years as a result of the selection of genetic variants which confer resistance to infections.

It has been proposed that the increased prevalence of inflammatory and autoimmune diseases is associated with the immune-related alleles that have been positively selected through evolutionary processes to protect against infections, hence the contrasting differences in the prevalence of autoimmune diseases between populations results from diverse selective pressures (Ramos et al., 2015). In line with this, it has been hypothesized that genetic variants associated with protection against infectious agents are behind the increased prevalence of autoimmune diseases in populations with low pathogen exposure, such as Europeans (Fumagalli et al., 2011; Raj et al., 2013). To study the changing patterns of susceptibility to autoimmune and inflammatory disease during history, we used publicly available summary statistics from GWAS of digestive tract-related autoimmune and inflammatory diseases and arthritis-related diseases (Fig. S4) and calculated the PRS for each of samples under study. Interestingly, we observed a robust increase of the genetic variants related with the development of inflammatory diseases in the digestive tract after the Neolithic revolution (Fig. 3). PRS scores associated with celiac disease, Crohn's disease, ulcerative colitis and inflammatory bowel disease, were strongly associated with the age of the samples, regardless of the P value thresholds or the missing genotype rates used for PRS calculation, showing the robustness of these results (Fig. S2). The fact that especially intestinal inflammatory pathology is increased after a historical event that fundamentally modified human diet is unlikely to be an accident. Our results are in

line with earlier research demonstrating that variants in genes important for immune responses and involved in celiac disease pathophysiology (such as IL-12, IL-18RAP, SH2B3) are under strong positive selection (Zhernakova et al., 2010). The reasons for the selection pressure on these genes are not completely understood, but an advantage for host defense has been suggested (Zhernakova et al., 2010).

In contrast to intestinal inflammation, the PRS of traits linked with juvenile-idiopathic arthritis, rheumatoid arthritis and multiple sclerosis shows a decrease in genetic susceptibility with the age of the sample after the Neolithic revolution. For pre-Neolithic periods, these patterns had little impact with decreasing PRS for digestive tract diseases and increasing PRS for ankylosing spondylitis and juvenile idiopathic arthritis. A strong decrease in susceptibility to juvenile idiopathic arthritis, rheumatoid arthritis and multiple sclerosis is seen after the Neolithic period (see Fig. 3). This is likely linked to the decreased production of the IL-1/IL-6/IL-17 axis described in Fig. 2, which is particularly important in the pathophysiology of these disorders (Akioka, 2019; Mei et al., 2011).

The significant changes in cytokine production and disease susceptibility in European populations after the Neolithic can be due to selective processes on the one hand (as described above), but also with important demographic changes due to migrations of human communities such as the Anatolians (in Neolithic) or the Yamnaya populations from the Pontic steppe (during the Bronze Age) (Racimo et al., 2020). In this regard, several loci associated with inflammatory disease displayed a group alleles linked with Crohn's disease, celiac disease and ulcerative colitis in Neolithic Aegeans, the community who spread farming across Europe (Hofmanová et al., 2016), with several of these alleles showing signs of positive selection in modern Europeans (Raj et al., 2013). In addition, the gene expression PRS of several cytokines based on the *cis*- and *trans*- eQTLs from the eQTLGen Consortium (<https://www.eqtlgen.org/>) displayed a very strong association with time for TNF α after the Neolithic revolution (Fig. 4).

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

251 **Acknowledgements**

252 MGN was supported by an ERC Advanced Grant (833247) and a Spinoza Grant of the
253 Netherlands Organization for Scientific Research. YL was supported by an ERC
254 Starting Grant (948207) and the Radboud University Medical Centre Hypatia Grant
255 (2018) for Scientific Research. JB was supported by PID2019-110933GB-I00
256 (AEI/FEDER, UE) MINECO, Spain.

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.

Declaration of Interests

The authors declare no competing interests.

Methods

Cohort selection

Ancient DNA genotype data was downloaded from version 37.2 of the published aDNA genotype database, compiled by and available on the David Reich Lab website (<https://reich.hms.harvard.edu/downloadable-genotypes-present-day-and-ancient-dna-data-compiled-published-papers>). The ancient DNA samples consisted of pseudo-haploid genotype data. This was due to the low genotyping coverage. Samples with variant missingness above 96 percent were filtered out using Plink (Purcell et al., 2007). This was done in order to remove outliers with extremely low coverage. Only samples within Europe were used for this study, these samples were selected based on their geographic location, that is latitude (within 35 and 70 degrees north) and longitude (within 10 degrees west and 40 degrees east). Samples without a carbon-dated age were also filtered out. We also selected 250 European samples from the 1000 genomes project phase 3. Only variants present in both the ancient samples and the modern samples were retained. This resulted in a dataset of 827 ancient samples and 250 modern samples containing 1233013 variants.

Carbon-dated sample origin and geographical location

Both carbon-dated age of origin as well as latitudinal and longitudinal data was available for these 827 ancient European samples. Broad time periods were assigned to these samples with the Early Upper Paleolithic era for all samples originating from

before 25000 years before the common era standardized to 1950 (BCE). The Late Upper Paleolithic era follows until 11000 BCE. The Mesolithic era ranges from 11000 to 5500 BCE. The Neolithic era ranges from 8500 to 3900 BCE, and the Post-Neolithic era ranges from 5000 BCE and more recent ages. Using the geographical data in combination with archeological clues and the genetic data, the broad time period of origin was also available for samples that were dated to a point in time with overlapping broad time periods. This allowed the samples to be classified as either Early Upper Paleolithic, Late Upper Paleolithic, Mesolithic, Neolithic, or Post-Neolithic. The sample age of the 250 modern European samples was set to 0.

Summary statistics of GWAS and cytokine QTLs

Summary statistics for complex traits were obtained from the UK Biobank (Bycroft et al., 2018) and the GWAS catalog (MacArthur et al., 2017) last accessed on 29th of March 2020. The stimulated cytokine response summary statistics from the 500FG cohort of the HFGP were used (Li et al., 2016). Some complex traits had multiple different sets of summary statistics available. In these cases, the data which was more recent and used bigger cohorts that were either of European or mixed (European and Asian) ancestry were selected. The variants of these summary statistics were then filtered by only keeping bi-allelic variants. Most aDNA genotypes available are pseudo-haploid as a consequence of their lower sample quality. We excluded ambiguous SNPs (A/T and C/G) in order to prevent errors due to strand flips present in these pseudo-haploid samples.

Polygenic Risk Scores (PRS) calculation

Polygenic risk scores were then calculated by first intersecting the filtered variants from the summary statistics with the variants present in the DNA samples. Starting at the

most significant variant, all variants within a 250kb window around that variant were excluded until no variants remained. We then multiplied the dosage of these variants with the effect size and these values were summed. If a variant is missing in a sample the dosage is substituted with the average genotyped dosage for that variant within the entire dataset. This way the PRS is not skewed in any specific direction. The formula for this is described below with the score S being the weighted sum of a variant's dosage X_n multiplied by its associated weight or beta β_n calculated using m variants.

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

Piecewise correlation analysis

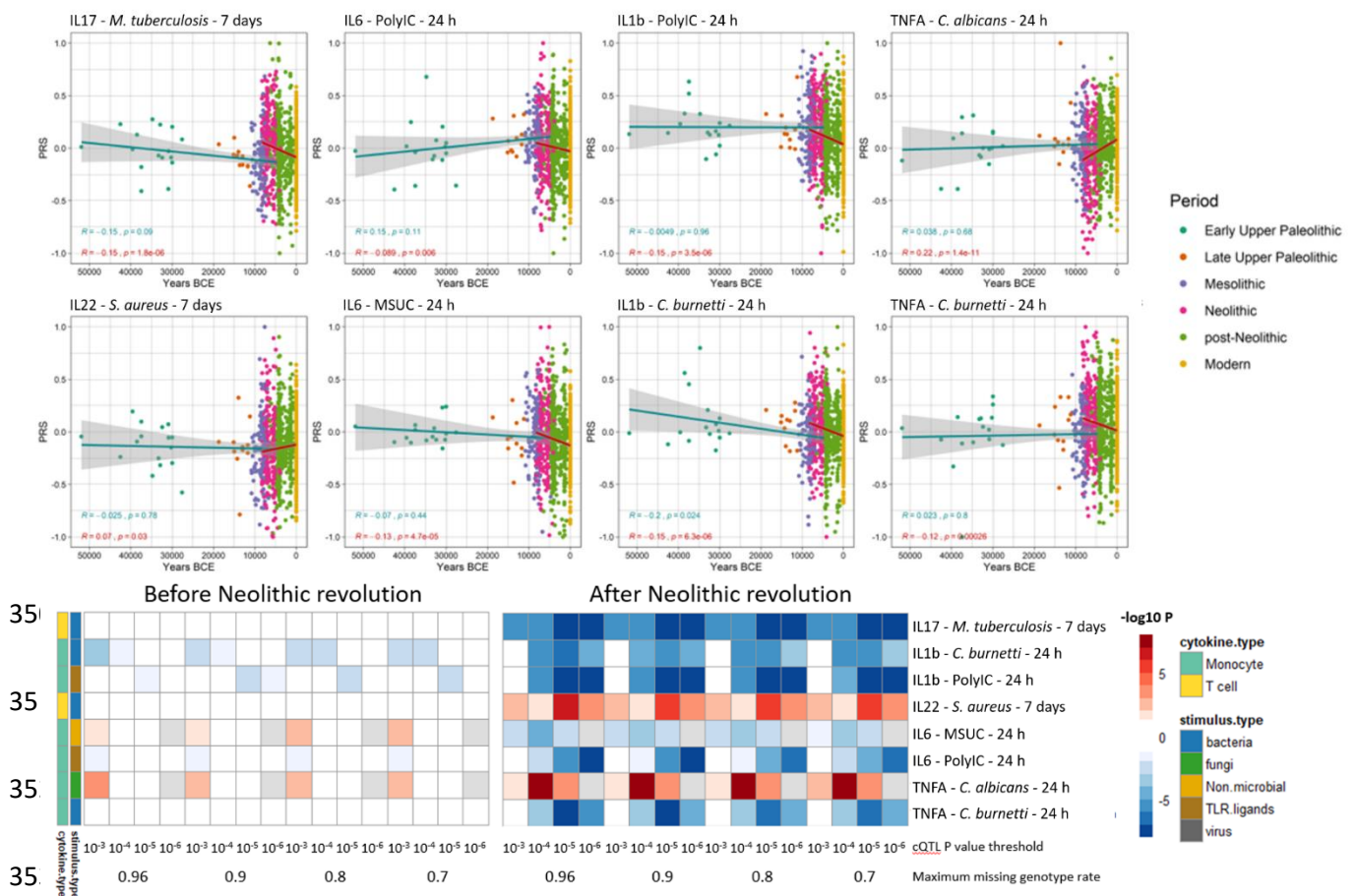
We constructed piecewise linear models for each trait by separating the samples into two groups. These two groups consisted of all samples preceding the Neolithic era and those of the Neolithic era and later respectively. We correlated PRS with the carbon dated age of our samples. We then multiplied the $-\log_{10}$ of the correlation P values with the sign of the correlation coefficients.

Robustness of results

In order to test the robustness of our results we calculated PRS using multiple different P value thresholds for QTL inclusion. We used P value thresholds from 10^{-3} to 10^{-8} for the complex traits obtained through GWAS catalog and the UK Biobank. The thresholds used for the stimulated cytokine responses ranged from 10^{-3} to 10^{-6} . We also calculated PRS using different variant missingness thresholds. This means we removed samples with a variant missingness rate higher than 96, 90, 80, or 70 percent. All of the results from the piecewise linear models were then used to create a heatmap depicting the consistency and robustness of our observed correlations.

Additionally, various window-sizes were used for clumping the QTL's and LD based clumping was also performed excluding variants with an LD greater than 0.2 compared to our lead SNP within a window. In order to see whether our observations were due to sample imbalances between the pre-Neolithic period and the later periods samples originating from the Neolithic period and later were randomly down-sampled to the same number of samples as the pre-Neolithic samples. Correlation coefficients between PRS and sample age were then recalculated for the Neolithic and younger samples and compared to the coefficients obtained using all Neolithic and younger samples.

349 Figures:



354 **Figure 1: A) Correlation between cytokine PRS and time.** Samples are colored by
355 broad age period. The blue regression lines show PRS before the Neolithic revolution
356 remained relatively constant for all traits whereas the red regression lines show the
357 correlation after the start of the Neolithic period. The threshold of max missing
358 genotype per sample was 0.96 and QTL P value cutoff was 10^{-4} . MSUC: Monosodium
359 urate crystals. B) **Correlation between cytokine PRS and time. using multiple**
360 **thresholds reveals consistent trend.** Missing genotype rate ranged from 0.96, 0.9,
361 0.8, and 0.7. QTL P value for variants included in our PRS models ranged from 10^{-3} ,
362 10^{-4} , 10^{-5} , and 10^{-6} . The color key indicates the range of -log10 P values of the Pearson
363 correlation between PRS and time. Red and blue indicate positive and negative
364 association, respectively.

365

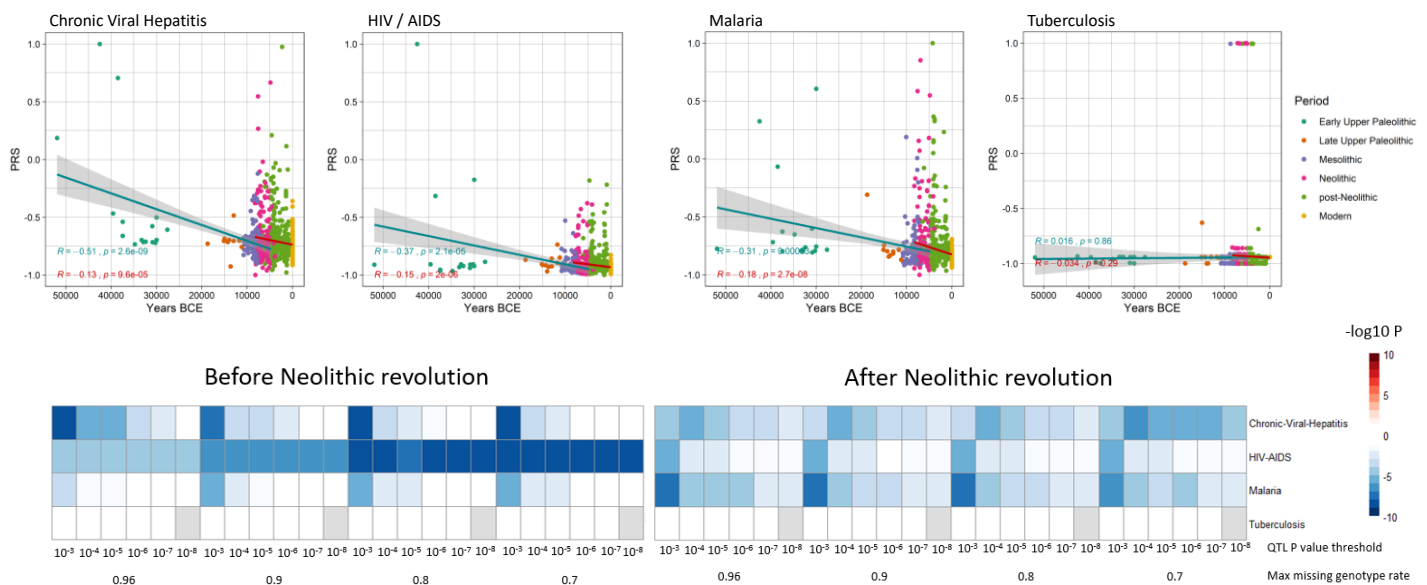


Figure 2: Infectious disease risk PRS scores decrease with time, except tuberculosis. A) Samples are colored by broad age period. The blue regression lines show PRS before the Neolithic revolution remained relatively constant for all traits whereas the red regression lines show that the correlation after 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. 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 $-\log_{10} P$ values of the Pearson correlation between PRS and time. Red and blue indicate positive and negative association, respectively.

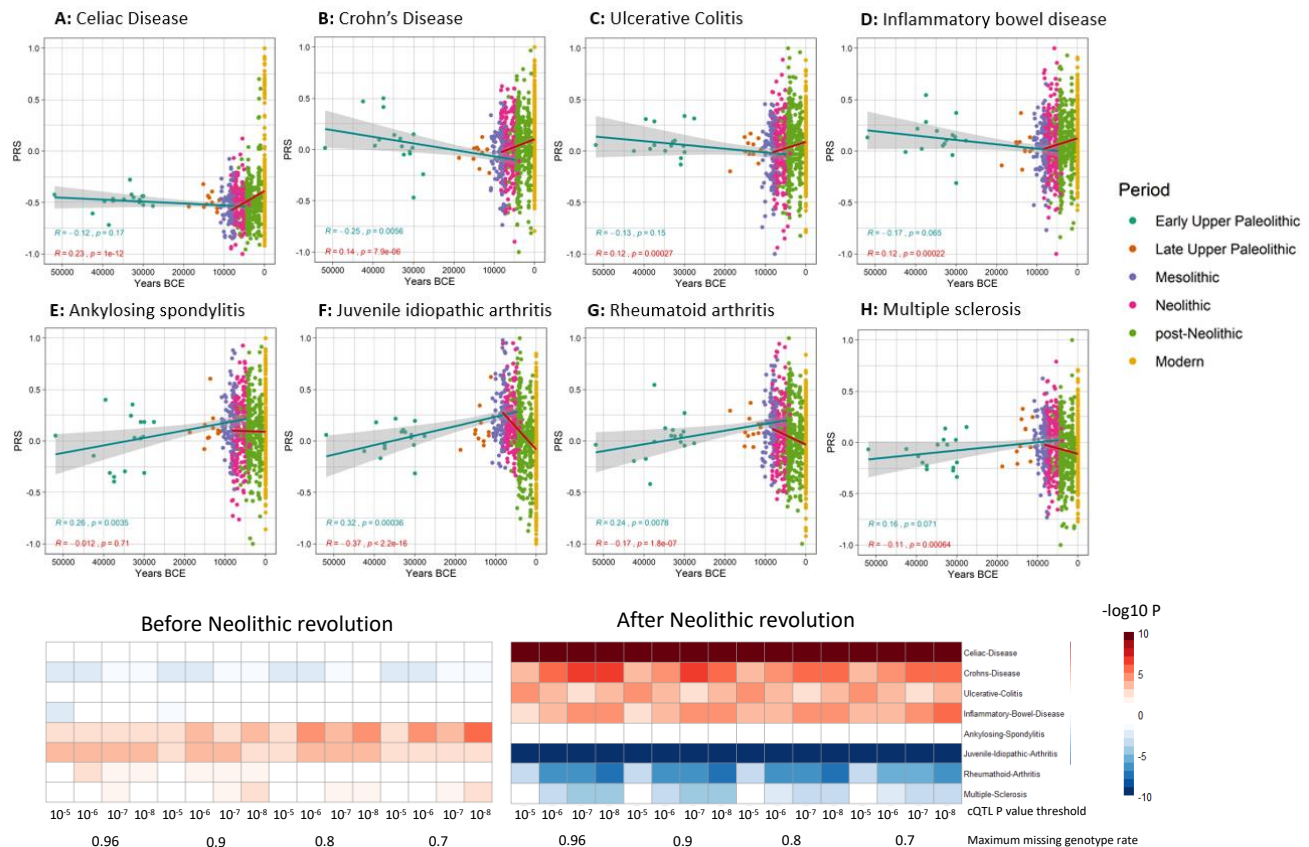


Figure 3: Correlation between auto-immune and inflammatory disease PRS and time. **A)** Samples are colored by broad age period. The blue regression lines show PRS before the Neolithic revolution remained relatively constant for all traits whereas the red regression lines show that the correlation after 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^{-4} . MSUC: Monosodium urate crystals. **B) Correlation between disease 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^{-5} , 10^{-6} , 10^{-7} , and 10^{-8} . The color key indicates the range of $-\log_{10}$ P values of the Pearson correlation between PRS and time. Red and blue indicate positive and negative association, respectively.

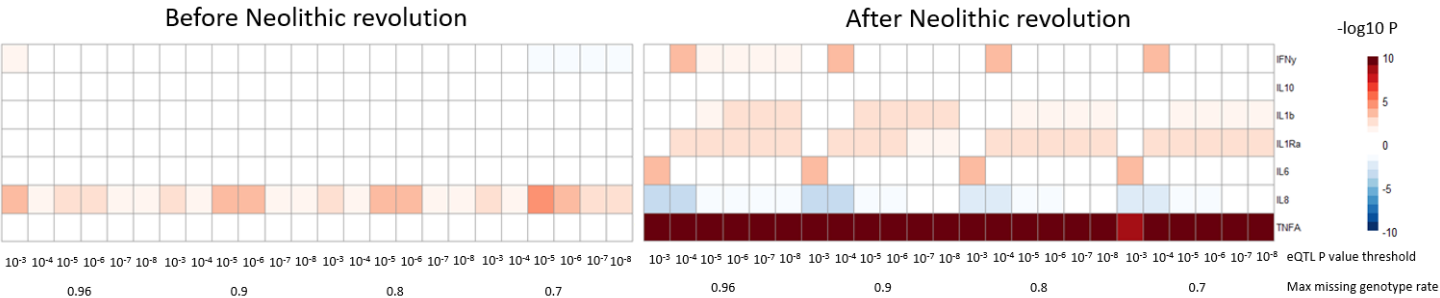
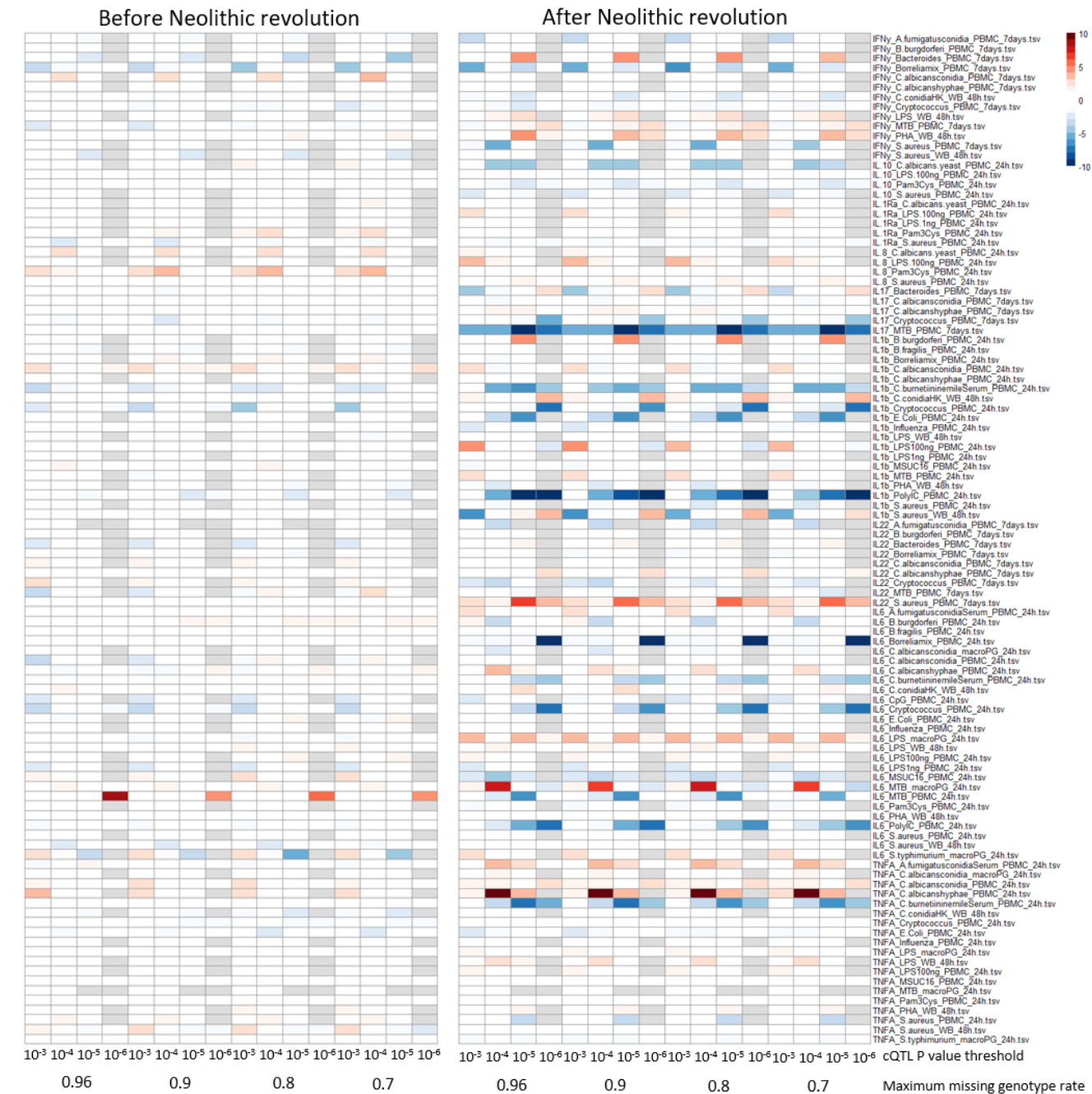


Figure 4: Cytokine gene expression PRS scores using cis- and trans-eQTLs

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.



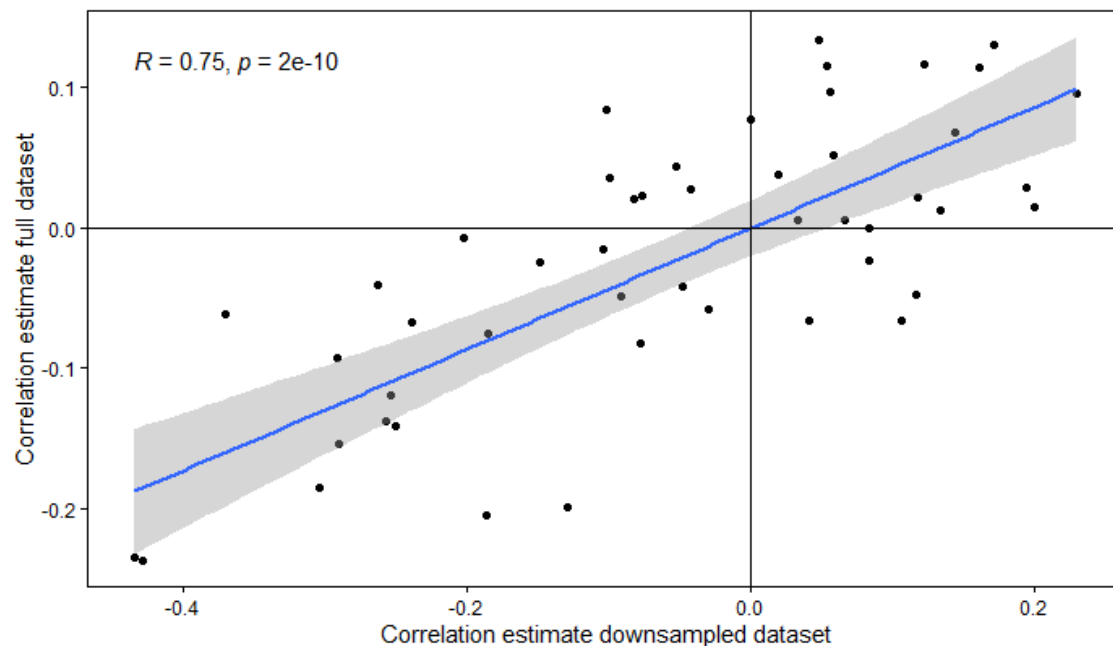


Figure S2: Robustness of correlation coefficients post Neolithic independent of sample size. Changes in PRS following the Neolithic revolution remain consistent after down-sampling samples from after the start of the Neolithic period to the same amount as samples before the Neolithic period. The lower 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.

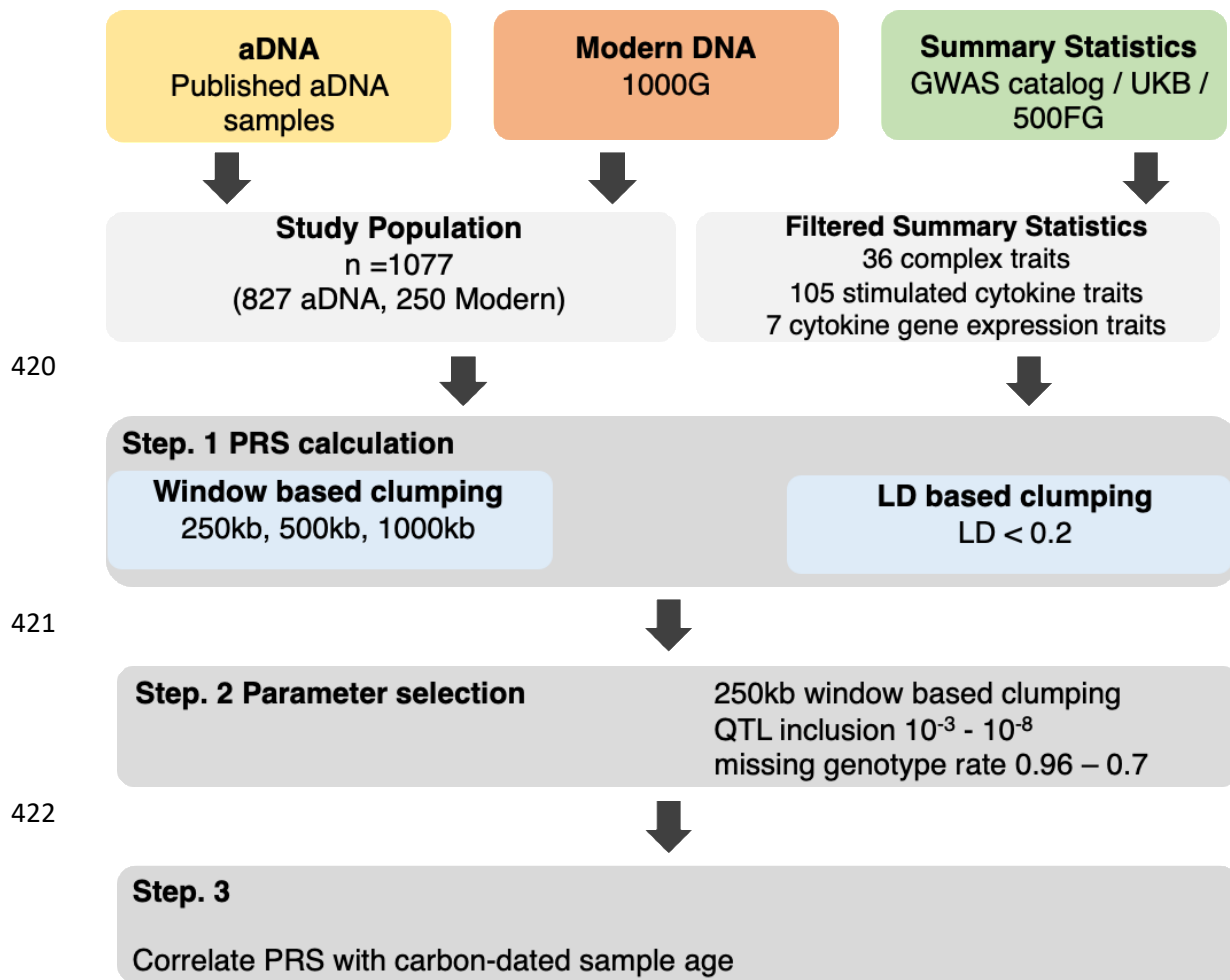


Figure S3: Both aDNA and modern DNA samples of European individuals were used in combination with summary statistics from predominantly European populations to calculate PRS of immune-related traits. This was done at various threshold combinations before correlating the scores with the sample age.

	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.

	PBMC 24h	Macrophage 24h	WB 48h	PBMC 7d
A.fumigatusconidia				IFNy, IL22
B.burgdorferi	IL1b, IL6			IFNy, IL22,
B.fragilis	IL1b, IL6			
Bacteroides				IFNy, IL17, IL22
Borrelia	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.burnetiiinileSerum	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:

- Akioka S. 2019. Interleukin-6 in juvenile idiopathic arthritis. *Mod Rheumatol* **29**:275–286. doi:10.1080/14397595.2019.1574697
- Bakker OB, Aguirre-Gamboa R, Sanna S, Oosting M, Smeekens SP, Jaeger M, Zorro M, Vősa U, Withoff S, Netea-Maier RT, Koenen HJPM, Joosten I, Xavier RJ, Franke L, Joosten LAB, Kumar V, Wijmenga C, Netea MG, Li Y. 2018. Integration of multi-omics data and deep phenotyping enables prediction of cytokine responses. *Nat Immunol* **19**:776–786. doi:10.1038/s41590-018-0121-3
- Bycroft C, Freeman C, Petkova D, Band G, Elliott LT, Sharp K, Motyer A, Vukcevic D, Delaneau O, O’Connell J, Cortes A, Welsh S, Young A, Effingham M, McVean G, Leslie S, Allen N, Donnelly P, Marchini J. 2018. The UK Biobank resource with deep phenotyping and genomic data. *Nature* **562**:203–209. doi:10.1038/s41586-018-0579-z
- Chan LC, Rossetti M, Miller LS, Filler SG, Johnson CW, Lee HK, Wang H, Gjertson D, Fowler VG, Reed EF, Yeaman MR. 2018. Protective immunity in recurrent *Staphylococcus aureus* infection reflects localized immune signatures and macrophage-conferred memory. *Proc Natl Acad Sci U S A*. doi:10.1073/pnas.1808353115
- Deschamps M, Laval G, Fagny M, Itan Y, Abel L, Casanova J-L, Patin E, Quintana-Murci L. 2016. Genomic Signatures of Selective Pressures and Introgression from Archaic Hominins at Human Innate Immunity Genes. *Am J Hum Genet* **98**:5–21. doi:10.1016/J.AJHG.2015.11.014
- Diard M, Hardt WD. 2017. Evolution of bacterial virulence. *FEMS Microbiol Rev*. doi:10.1093/femsre/fux023
- Domínguez-Andrés J, Feo-Lucas L, Minguito de la Escalera M, González L, López-

Bravo M, Ardavin C. 2017. Inflammatory Ly6C high Monocytes Protect against
Candidiasis through IL-15-Driven NK Cell/Neutrophil Activation. *Immunity*
46:1059-1072.e4. doi:10.1016/j.immuni.2017.05.009

Duncan SR, Scott S, Duncan CJ. 2005. Reappraisal of the historical selective
pressures for the CCR5-Δ32 mutation. *J Med Genet.*
doi:10.1136/jmg.2004.025346

Ferwerda B, McCall MBB, Alonso S, Giamarellos-Bourboulis EJ, Mouktaroudi M,
Izagirre N, Syafruddin D, Kibiki G, Cristea T, Hijmans A, Hamann L, Israel S,
ElGhazali G, Troye-Blomberg M, Kumpf O, Maiga B, Dolo A, Doumbo O, Hermesen
CC, Stalenhoef AFH, van Crevel R, Brunner HG, Oh D-Y, Schumann RR, de la
Rua C, Sauerwein R, Kullberg B-J, van der Ven AJAM, van der Meer JWM, Netea
MG. 2007. TLR4 polymorphisms, infectious diseases, and evolutionary pressure
during migration of modern humans. *Proc Natl Acad Sci* **104**:16645–16650.
doi:10.1073/pnas.0704828104

Flandroy L, Poutahidis T, Berg G, Clarke G, Dao M-C, Decaestecker E, Furman E,
Haahtela T, Massart S, Plovier H, Sanz Y, Rook G. 2018. The impact of human
activities and lifestyles on the interlinked microbiota and health of humans and of
ecosystems. *Sci Total Environ* **627**:1018–1038.
doi:10.1016/J.SCITOTENV.2018.01.288

Fumagalli M, Sironi M. 2014. Human genome variability, natural selection and
infectious diseases. *Curr Opin Immunol* **30**:9–16.

Fumagalli M, Sironi M, Pozzoli U, Ferrer-Admettla A, Pattini L, Nielsen R. 2011.
Signatures of Environmental Genetic Adaptation Pinpoint Pathogens as the Main
Selective Pressure through Human Evolution. *PLoS Genet* **7**:e1002355.
doi:10.1371/journal.pgen.1002355

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

Dorp L, López S, Kousathanas A, Link V, Kirsanow K, Cassidy LM, Martiniano R, Strobel M, Scheu A, Kotsakis K, Halstead P, Triantaphyllou S, Kyparissi-Apostolika N, Urem-Kotsou D, Ziota C, Adaktylou F, Gopalan S, Bobo DM, Winkelbach L, Blöcher J, Unterländer M, Leuenberger C, Çilingiroğlu Ç, Horejs B, Gerritsen F, Shennan SJ, Bradley DG, Currat M, Veeramah KR, Wegmann D, Thomas MG, Papageorgopoulou C, Burger J. 2016. Early farmers from across Europe directly descended from Neolithic Aegeans. *Proc Natl Acad Sci U S A* **113**:6886–6891. doi:10.1073/pnas.1523951113

Joosten SA, Meijgaarden KE van, Arend SM, Prins C, Oftung F, Korsvold GE, Kik S V., Arts RJW, Crevel R van, Netea MG, Ottenhoff THM. 2018. Mycobacterial growth inhibition is associated with trained innate immunity. *J Clin Invest* **128**:1837–1851. doi:10.1172/JCI97508

Karlsson EK, Kwiatkowski DP, Sabeti PC. 2014. Natural selection and infectious disease in human populations. *Nat Rev Genet* **15**:379–393. doi:10.1038/nrg3734

Li Y, Oosting M, Smeekens SP, Jaeger M, Aguirre-Gamboa R, Le KTT, Deelen P, Ricaño-Ponce I, Schoffelen T, Jansen AFM, Swertz MA, Withoff S, van de Vosse E, van Deuren M, van de Veerdonk F, Zhernakova A, van der Meer JWM, Xavier RJ, Franke L, Joosten LAB, Wijmenga C, Kumar V, Netea MG. 2016. A Functional Genomics Approach to Understand Variation in Cytokine Production in Humans. *Cell* **167**:1099-1110.e14. doi:10.1016/j.cell.2016.10.017

Liu W, Li Y, Learn GH, Rudicell RS, Robertson JD, Keele BF, Ndjanga JBN, Sanz CM, Morgan DB, Locatelli S, Gonder MK, Kranzusch PJ, Walsh PD, Delaporte E, Mpoudi-Ngole E, Georgiev A V., Muller MN, Shaw GM, Peeters M, Sharp PM, Rayner JC, Hahn BH. 2010. Origin of the human malaria parasite *Plasmodium falciparum* in gorillas. *Nature*. doi:10.1038/nature09442

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

Milano A, Morales J, MayPendlington Z, Welter D, Burdett T, Hindorff L, Flicek P, Cunningham F, Parkinson H. 2017. The new NHGRI-EBI Catalog of published genome-wide association studies (GWAS Catalog). *Nucleic Acids Res.* doi:10.1093/nar/gkw1133

Mei Y, Pan F, Gao J, Ge R, Duan Z, Zeng Z, Liao F, Xia G, Wang S, Xu S, Xu J, Zhang L, Ye D. 2011. Increased serum IL-17 and IL-23 in the patient with ankylosing spondylitis. *Clin Rheumatol.* doi:10.1007/s10067-010-1647-4

Netea MG, Joosten LAB, Li Y, Kumar V, Oosting M, Smeekens S, Jaeger M, Ter Horst R, Schirmer M, Vlamakis H, Notebaart R, Pavelka N, Aguirre-Gamboa RR, Swertz MA, Tunjungputri RN, Van De Heijden W, Franzosa EA, Ng A, Graham D, Lassen K, Schraa K, Netea-Maier R, Smit J, De Mast Q, Van De Veerdonk F, Kullberg BJ, Tack C, Van De Munckhof I, Rutten J, Van Der Graaf J, Franke L, Hofker M, Jonkers I, Platteel M, Maatman A, Fu J, Zhernakova A, Van Der Meer JWM, Dinarello CA, Van Der Ven A, Huttenhouwer C, Koenen H, Joosten I, Xavier RJ, Wijmenga C. 2016. Understanding human immune function using the resources from the Human Functional Genomics Project. *Nat Med.* doi:10.1038/nm.4140

Pai M, Behr MA, Dowdy D, Dheda K, Divangahi M, Boehme CC, Ginsberg A, Swaminathan S, Spigelman M, Getahun H, Menzies D, Raviglione M. 2016. Tuberculosis. *Nat Rev Dis Prim.* doi:10.1038/nrdp.2016.76

Poinar G. 2005. Plasmodium dominicana n. sp. (Plasmodiidae: Haemospororida) from Tertiary Dominican amber. *Syst Parasitol.* doi:10.1007/s11230-004-6354-6

Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MAR, Bender D, Maller J, Sklar P, De Bakker PIW, Daly MJ, Sham PC. 2007. PLINK: A tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet.* doi:10.1086/519795

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

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

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

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

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

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

Common Risk Alleles for Inflammatory Diseases Are Targets of Recent Positive

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

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

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

doi:10.1038/jhg.2015.94

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

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

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

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

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

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

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

020-0311-8

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

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

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

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

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

and Environmental Factors Influencing Individual Human Cytokine Responses.

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

Thakur A, Mikkelsen H, Jungersen G. 2019. Intracellular Pathogens: Host Immunity
and Microbial Persistence Strategies. *J Immunol Res* **2019**.
doi:10.1155/2019/1356540

Zhernakova A, Elbers CC, Ferwerda B, Romanos J, Trynka G, Dubois PC, de Kovel
CGF, Franke L, Oosting M, Barisani D, Bardella MT, Joosten LAB, Saavalainen P,
van Heel DA, Catassi C, Netea MG, Wijmenga C, Wijmenga C. 2010. Evolutionary
and Functional Analysis of Celiac Risk Loci Reveals SH2B3 as a Protective Factor
against Bacterial Infection. *Am J Hum Genet* **86**:970–977.
doi:10.1016/j.ajhg.2010.05.004