

# Ancient DNA reveals pervasive directional selection across West Eurasia

<https://doi.org/10.1038/s41586-026-10358-1>

Received: 14 September 2024

Accepted: 4 March 2026

Published online: 15 April 2026

 Check for updates

Ali Akbari<sup>1,2,3</sup>, Annabel Perry<sup>2,3</sup>, Alison R. Barton<sup>2,3</sup>, Mohammadreza Kariminejad<sup>4</sup>, Steven Gazal<sup>5,6,7</sup>, Zheng Li<sup>8</sup>, Yating Zeng<sup>5,9</sup>, Alissa Mittnik<sup>2,10</sup>, Nick Patterson<sup>2,3</sup>, Matthew Mah<sup>1,11</sup>, Xiang Zhou<sup>12</sup>, Alkes L. Price<sup>3,13,14</sup>, Eric S. Lander<sup>3,15,16</sup>, Ron Pinhasi<sup>17,18</sup>, Nadin Rohland<sup>1,2,3</sup>, Swapan Mallick<sup>1,2,3,11</sup> & David Reich<sup>1,2,3,11</sup>

Ancient DNA has transformed our understanding of population history<sup>1</sup>, but its potential to reveal as much about human evolutionary biology has not been realized because of limited sample sizes and the difficulty of distinguishing sustained rises in allele frequency increasing fitness—directional selection—from shifts due to migrations, population structure, or non-adaptive purifying or stabilizing selection<sup>2–7</sup>. Here we present a method for detecting directional selection in ancient DNA time-series data that tests for consistent trends in allele frequency change over time, and apply it to 15,836 West Eurasians (10,016 with new data). Previous work has shown that classic hard sweeps driving advantageous mutations to fixation have been rare over the broad span of human evolution<sup>8,9</sup>. By contrast, in the past ten millennia, we find that many hundreds of alleles have been affected by strong directional selection. We also document one-standard-deviation changes on the scale of modern variation in combinations of alleles that today predict complex traits. This includes decreases in predicted body fat and schizophrenia, and increases in measures of cognitive performance. These effects were measured in industrialized societies, and it remains unclear how these relate to phenotypes that were adaptive in the past. We estimate selection coefficients at 9.7 million variants, enabling study of how Darwinian forces couple to allelic effects and shape the genetic architecture of complex traits.

Ancient DNA data hold extraordinary promise for revealing adaptation, making it possible to track effects across time and obtain direct measurements of selection coefficients<sup>10,11</sup>. Rather than being trapped in the present and studying the scars left by selection on the genomes of descendants<sup>12,13</sup>, ancient DNA makes it possible to test directly whether frequencies of variants shifted more than could be expected by chance<sup>2–5</sup>. Such data also make it easier to measure selection on variants not of recent mutational origin (standing variation), which is challenging to detect using retrospective methods<sup>14</sup>. Previous ancient DNA selection studies in West Eurasia<sup>2–5</sup> (Europe and its neighbours in the Near East) have identified dozens of alleles influenced by selection, but despite growth in reported ancient individuals to more than 10,000 today<sup>15</sup>, the number of genome-wide significant loci reported in a single study grew only from 12 in the first genome scan in 2015 (ref. 2) to 21 in 2024 (ref. 4), raising the concern that this approach might not deliver broad insights into human adaptation. Here we increase

the yield of discoveries by 20-fold by adding more statistical power (due to a qualitatively new method and larger sample size) and reducing artefactual signals (through intensive data cleaning).

First, we increased power by testing for a consistent trend in allele frequency change over time. Several past studies have dealt with the challenge posed by population mixture by treating more recent populations as linear combinations of more ancient ones, then searching for alleles whose frequencies were outliers compared with what would be expected from this history<sup>2,3</sup>. However, changes in frequency due to selection are often less than what can be expected from gene flow and random genetic drift<sup>6</sup>, and in this context, increasing sample size helps little. We used a qualitatively different approach, using the genetic similarity of each individual to every other, and testing whether the date when they lived provides additional predictive power for the allele frequencies of their population beyond what is expected from the empirical population structure. Our test is simple: at each variant

<sup>1</sup>Department of Genetics, Harvard Medical School, Boston, MA, USA. <sup>2</sup>Department of Human Evolutionary Biology, Harvard University, Cambridge, MA, USA. <sup>3</sup>Broad Institute of MIT and Harvard, Cambridge, MA, USA. <sup>4</sup>Shamsipour Technical and Vocational College, Tehran, Iran. <sup>5</sup>Department of Population and Public Health Sciences, Keck School of Medicine, University of Southern California, Los Angeles, CA, USA. <sup>6</sup>Center for Genetic Epidemiology, Keck School of Medicine, University of Southern California, Los Angeles, CA, USA. <sup>7</sup>Department of Quantitative and Computational Biology, University of Southern California, Los Angeles, CA, USA. <sup>8</sup>Department of Biostatistics, School of Public Health, University of Michigan, Ann Arbor, MI, USA. <sup>9</sup>Department of Biostatistics and Data Science, School of Public Health, The University of Texas Health Science Center at Houston, Houston, TX, USA. <sup>10</sup>Department of Archaeogenetics, Max Planck Institute for Evolutionary Anthropology, Leipzig, Germany. <sup>11</sup>Howard Hughes Medical Institute, Harvard Medical School, Boston, MA, USA. <sup>12</sup>Department of Statistics and Data Science, Yale University, New Haven, CT, USA. <sup>13</sup>Department of Epidemiology, Harvard T.H. Chan School of Public Health, Boston, MA, USA. <sup>14</sup>Department of Biostatistics, Harvard T.H. Chan School of Public Health, Boston, MA, USA. <sup>15</sup>Department of Systems Biology, Harvard Medical School, Boston, MA, USA. <sup>16</sup>Department of Biology, Massachusetts Institute of Technology (MIT), Cambridge, MA, USA. <sup>17</sup>Department of Evolutionary Anthropology, University of Vienna, Vienna, Austria. <sup>18</sup>Human Evolution and Archaeological Sciences, University of Vienna, Vienna, Austria. <sup>✉</sup>e-mail: Ali\_Akbari@fas.harvard.edu; reich@genetics.med.harvard.edu

we asked whether hypothesizing a non-zero selection coefficient  $s$ —causing allele frequency to trend in the same direction over all times and places—predicts frequency differences across populations significantly better than empirically measured population structure alone.

Second, we increased power through a 14-fold increase in sample size, driven by 10,016 ancient individuals for whom we report new data, which combined with previously reported data yields a dataset of 15,836 people spanning 18,000 years (Supplementary Tables 1–3). We co-analysed with 6,438 modern people, sub-sampled so that their countries of origin were spread evenly over West Eurasia (Extended Data Fig. 1a,b).

Third, we increased power by imputing diploid genotypes<sup>16</sup>, leveraging known patterns of allelic correlation in a modern reference panel to fill in missing data in all individuals. We also carried out intensive data cleaning (Supplementary Section 1), filtering out single-nucleotide polymorphisms (SNPs) susceptible to false positives. The final dataset included 8,074,573 SNPs and 1,665,051 insertions or deletions (indels) on chromosomes 1–22.

### Test of selection on single variants

For each SNP in the genome, we estimated a selection coefficient,  $s$ , which we found has a standard error of typically 0.17% for common variants (Extended Data Fig. 1c,d). In theory, a valid test for selection should be  $Z$ , the number of standard errors this quantity is from zero, and we should be able to use a normal distribution to identify scores that pass the standard threshold of genome-wide significance ( $P < 5 \times 10^{-8}$ ). In practice, the median  $\chi^2$  statistic (squared  $Z$ -score for the number of standard errors  $s$  is from zero) is inflated by 4.75-fold relative to a  $\chi^2$  distribution with 1 d.f. In human genetics studies, such inflation can arise due to uncorrected population structure or non-normality of the null distribution due to case–control sample imbalance<sup>17</sup>. Inflation like this is often addressed by rescaling  $\chi^2$  statistics by the median inflation across the genome<sup>18,19</sup>. But if the null is not normally distributed, or a substantial fraction of the genome carries real signal<sup>18</sup>, random sites will not provide an appropriate neutral baseline. These concerns apply here: the null distribution is not normal due to frequency-dependent biases from imputation and the non-linear transformation of allele frequencies (Supplementary Section 2), and most of the genome is in linkage disequilibrium with sites showing evidence of directional selection (Extended Data Fig. 2b and Supplementary Section 3).

Instead, we calibrated our test by taking advantage of a finding about the connection between selection coefficients and associations to phenotypes in living people. The proportion of SNPs showing significant association to a phenotype in genome-wide association studies (GWAS) increases with our selection statistic, plateauing at around 5.2 times the rate for random SNPs (Fig. 1a; this analysis relies on 1,317,022 SNPs with a genome-wide significant association to at least one phenotype for 452 traits in the UK Biobank, and conditions on minor allele frequency (MAF) to remove bias due to signals being easier to detect for higher MAFs). The plateau occurs at the same place when we control for ‘background selection’, the loss of neutral variation linked to deleterious mutations removed by purifying selection, although the enrichment in GWAS variants is reduced (approximately 2.4 times; Extended Data Fig. 3a). This is the pattern expected for a true threshold for genome-wide significance: if SNPs beyond this threshold reflect a combination of true signal and false discoveries, we would expect enrichment to continue beyond it. This threshold occurs at  $Z = 8.23$ , larger than the standard threshold (5.45) for genome-wide significance for a normal distribution, so we rescaled the naive score by this quantity to obtain an  $X$ -statistic ( $Z/1.51$ ) whose significance threshold matches the standard threshold.

To validate our procedure, we conducted realistic forward-in-time simulations (Fig. 1b, Extended Data Fig. 4 and Supplementary Section 2). The simulations specify a demographic history inspired by

that of Europe and produce patterns of structure matching important features of real data. The simulations include coding, functional non-coding and neutral regions, recapitulating varying degrees of background selection. The simulations also permit exploration of the relationship between selected alleles and traits, in particular the phenomenon of ‘stabilizing’ selection to remove genetic variation when the average genetic predictor of a trait is at a fitness optimum. We verified that these simulations produce a deficiency of common polymorphisms influencing traits, matching patterns in real data<sup>20</sup>.

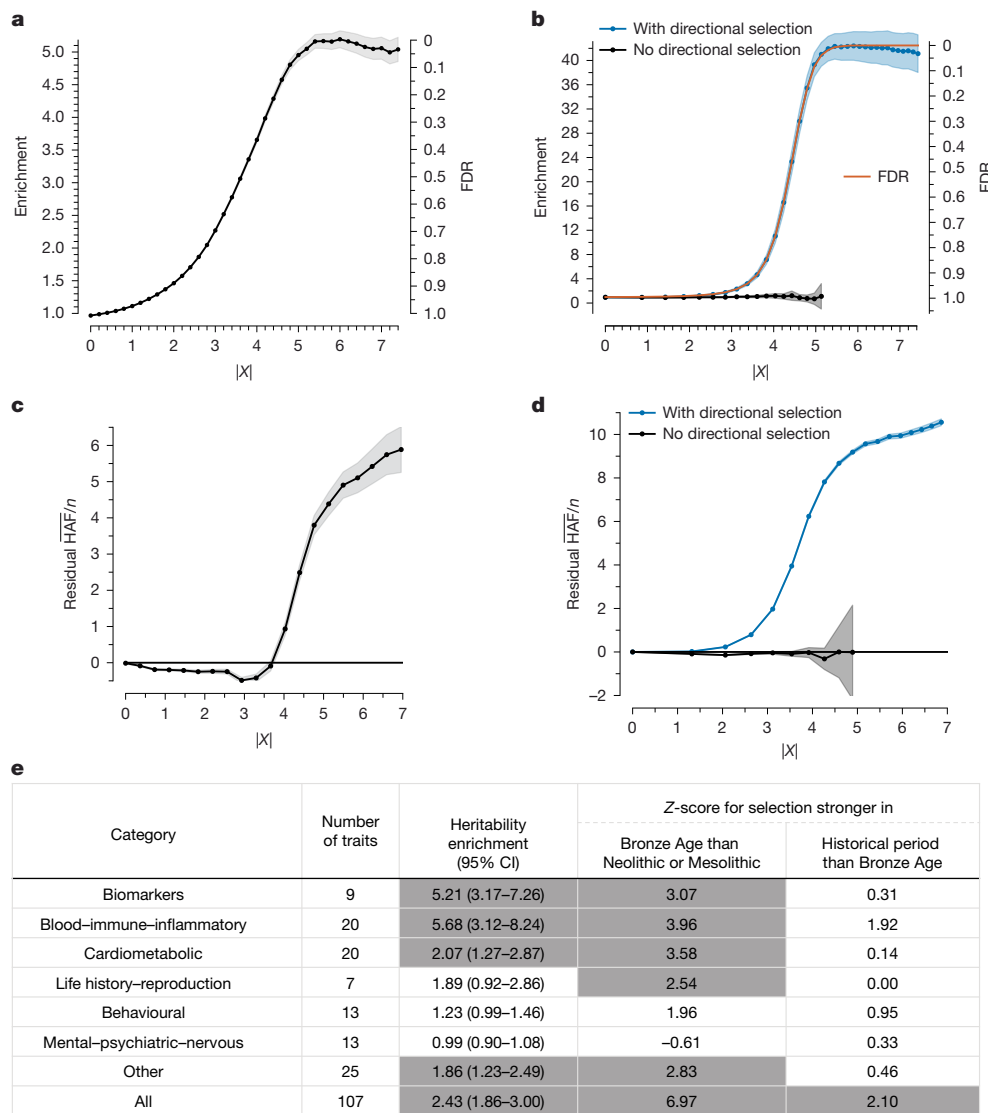
Our test for directional selection is robust to every confounding factor that we considered, including population structure, background selection due to purifying or stabilizing selection, sampling bias and selection in the ancestral population but not during the time transect itself (Supplementary Section 2). None of these produces even a hint of the enrichment in GWAS signals that we observe in data. By contrast, in simulations with directional selection, including where we shift the trait optimum, leading to deterministic changes in allele frequency to bring the population closer to that optimum<sup>21</sup>, we observed exactly this signal and could use it to provide a valid calibration of genome-wide significance (Fig. 1b, Extended Data Fig. 4 and Supplementary Section 2). To translate  $X$  to a posterior probability of selection  $\pi$ , we used a false discovery rate (FDR) approach (Fig. 1a,b and Supplementary Section 3). We fit a smooth function to the enrichment curve for GWAS signals and estimate that at  $X$  greater than our threshold of 5.45,  $\pi > 99\%$ .

A second line of evidence that these signals are real comes from computing the mean haplotype allele frequency (HAF) score, the sum of squared derived allele frequencies scaled by sample size in 200-kb windows surrounding each tested variant. Previous work<sup>22,23</sup> has shown that directional positive selection on derived alleles is expected to increase HAF scores, whereas background selection is expected to decrease it (Extended Data Fig. 3b and Supplementary Section 4). After computing the residual HAF score for each variant controlling for background selection<sup>24,25</sup>, we found that it increases with the  $X$ -statistic and begins to rise before 5.45, the standard threshold for genome-wide significance in GWAS (Fig. 1c and Extended Data Fig. 3c). Our forward-in-time simulations (Supplementary Section 2) confirm that the rise in HAF score occurs only for directional selection, with no hint of this pattern for other scenarios (Fig. 1d and Extended Data Fig. 4).

We obtained a third line of evidence that we are detecting real signals of adaptation by showing that  $|X| > 5.45$  variants are associated with some classes of traits much more than others. We found enrichment for SNPs contributing to blood–immune–inflammatory traits (95% confidence interval (CI) 3.12–8.24)<sup>3,5</sup>, compared with random SNPs with matched characteristics defining the baseline. For mental–psychiatric–nervous and behavioural traits, we did not detect enrichment (95% CIs of 0.90–1.08 and 0.99–1.46, respectively; Fig. 1e and Extended Data Fig. 5a). This cannot be explained by differences in allele frequencies or background selection as we controlled for these factors. The intensity of selection on both blood–immune–inflammatory and cardiometabolic traits increased significantly in the Bronze Age relative to the pre-farming period (Fig. 1e and Extended Data Fig. 5b), plausibly reflecting adaptation to new diets, larger population or living closer to domesticated animals<sup>26</sup>.

### Hundreds of cases of directional selection

We found evidence of 479 independent loci (410 excluding the HLA region) with  $|X| > 5.45$ , corresponding to a  $\pi > 99\%$  probability of selection (Fig. 2a). To produce this list, we identified the strongest signal in the genome and considered all SNPs in linkage disequilibrium with it in unrelated Western European individuals ( $r^2 > 0.05$ ) to potentially reflect the same signal. We then found the second-strongest signal excluding these positions, and so on until no more variants passed this threshold (Extended Data Fig. 2b). Visualizations of the trajectories for these loci (Supplementary Section 5) and summary statistics for



**Fig. 1 | Multiple lines of evidence show that we are detecting genuine directional selection.** **a**, Enrichment of SNPs significant in at least one of 452 UK Biobank GWAS, for SNPs with  $|X|$  above the value on the x-axis, controlling for allele frequency. **b**, Simulations reveal enrichment of GWAS hits (blue; left axis) closely aligns with  $1 - \text{FDR}$  (orange; right axis) until a plateau as a function of  $|X|$ : 800 simulations with directional selection (200 each for selection coefficients of 0.01, 0.02 or 0.03 for model 2.1, soft sweep; and 0.05 for model 2.2, hard sweep) and 800 simulations without it. **c**, Residual mean HAF score ( $(\text{HAF})/n$ ) for SNPs with  $|X|$  above the value on the x-axis, computed as observed minus expected, with  $n$  haploid sample size, from a linear regression correcting for background selection using as explanatory variables McVicker-B, Murphy-phastCons, Murphy-CADD, number of SNPs and heterozygosity in a 200-kb window. **d**, Residual mean HAF score in simulation for SNPs with  $|X|$

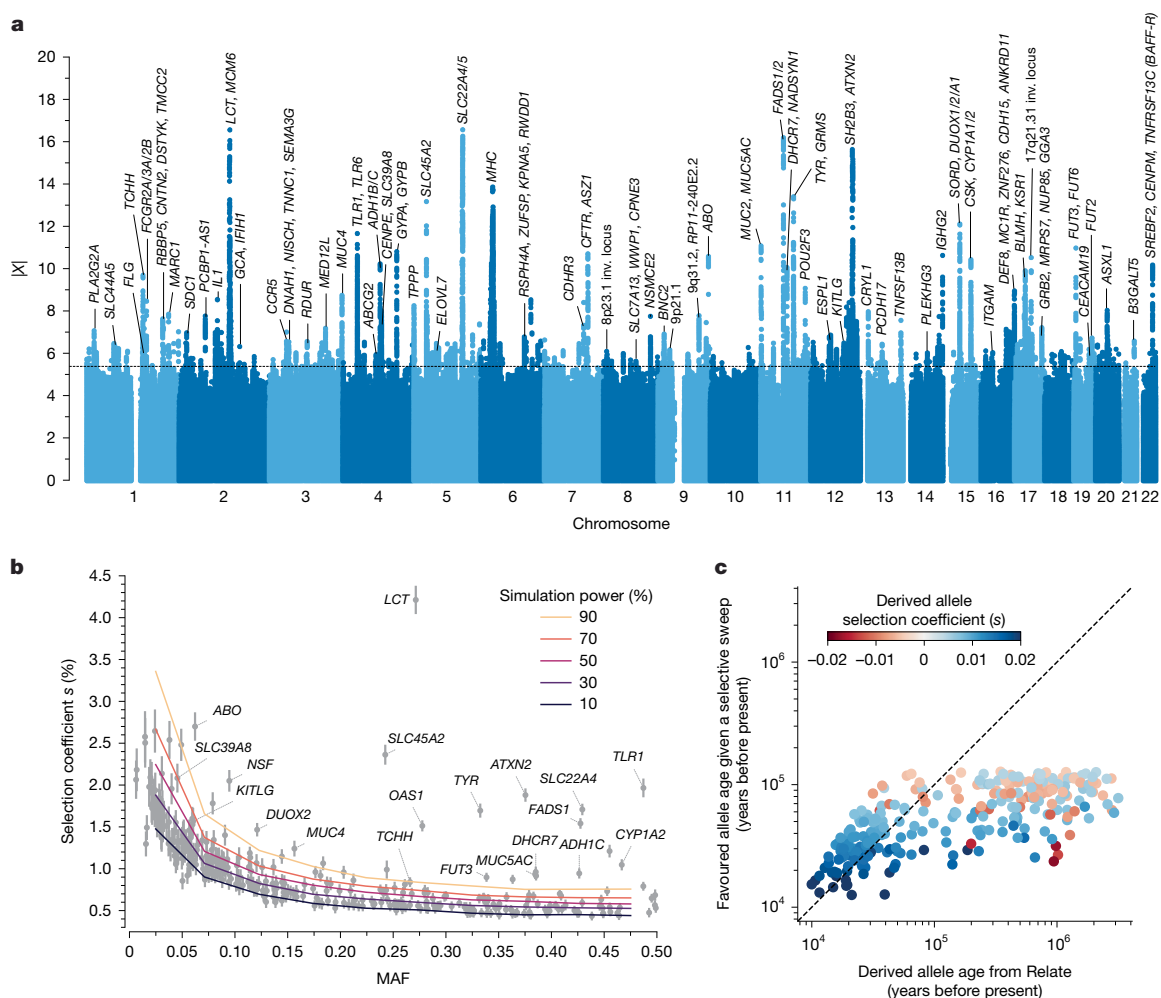
above the value on the x-axis, computed as observed minus expected, with  $n$  haploid sample size, from a linear regression correcting for background selection using as explanatory variables McVicker-B, coding region annotation, functional non-coding region annotation, number of SNPs and heterozygosity in a 200-kb window. The shaded areas show 95% CIs (**a–d**). **e**, The heritability enrichment is a meta-analysis for annotations based on a binary selection annotation, with FDR either below 1% (1) or above 1% (0). The Z-score for change in selection intensity over time is based on a meta-analysis of heritability enrichment comparing key cultural transitions: Mesolithic–Neolithic to Bronze Age, and Bronze Age to historical era. We annotated each SNP according to whether it is among the top 5% with the highest probability of a stronger magnitude of selection coefficient in one time transect versus another. Cells shaded grey denote  $P < 0.05$ .

9.7 million variants (Supplementary Table 4) can be cross-referenced with GWAS and with their frequency trajectories at the AGES browser (<https://reich-ages.rc.hms.harvard.edu>).

The actual number of loci under selection is likely to be much larger. Using a threshold of  $|X| > 3.61$  (FDR = 50%), we identified 7,689 non-HLA loci, implying more than 3,800 independent episodes of selection. The exact number of distinct signals is difficult to count: residual linkage disequilibrium exists among candidate loci, which at some loci may cause some overestimation of signals (Supplementary Section 5), whereas truly selected alleles in linkage disequilibrium with nearby stronger alleles will be undercounted. Downsampling analyses showed that

further increases in sample size are expected to increase the number of detected loci further, with people living more than 8,000 years ago providing the most added power (Extended Data Fig. 1d,e).

To obtain insight into the phenotypes shaped by directional selection, we took advantage of the fact that a high proportion (62%) of the variants with genome-wide evidence of selection are independently associated to a phenotype in at least one UK Biobank GWAS. However, biological interpretation is complicated as the allele that was the target of selection may differ from the tag SNP that we are using to represent the locus (and may even be in a neighbouring gene); because some alleles affect multiple phenotypes, because the relevant modern trait



**Fig. 2 | Genomescan for directional selection.** **a**, The x axis is the chromosomal position, and the y axis is the selection signal for each variant. The dotted line indicates our genome-wide significance threshold of  $|X| = 5.45$ . For clarity, only a subset of loci are annotated. **b**, Selection coefficient ( $s$ ) estimated from our scan plotted against MAF of tagging SNPs at independent loci with  $\pi > 99\%$ . Overlaid grids are simulation-based power estimates (90%, 70%, 50%, 30% and 10% probability of detection). The error bars indicate standard errors from

20,374 unrelated individuals. **c**, The estimated age of the favoured allele in a selective sweep versus the date of origin of the mutation inferred from Relate<sup>58</sup>, for tagging SNPs at independent loci with  $\pi > 99\%$ . The age of the sweep is defined as the time before present when the frequency of the favoured allele is expected to have been 0.0001 given the present-day frequency in 1000 Genomes Project European populations and assuming that the selection coefficient has been constant over time.

may not be measured in one of the GWAS that we are analysing, or because the phenotype in modern societies may not have existed in the ancient societies where selection acted. The median selection magnitude  $|s|$  at the tag SNPs for non-HLA loci is 0.86% (range of 0.43–4.3%), and the median MAF is 13%. Standard errors in our estimates of  $|s|$  for common alleles are 0.17% (Fig. 2b and Extended Data Fig. 1c,d).

We compared our results to those of five previous scans for selection in Holocene West Eurasia (four based on ancient DNA)<sup>2–5,13</sup> (Extended Data Table 1 and Supplementary Section 6). Of 38 unique non-HLA loci that met the formal threshold for genome-wide significance in at least one previous study, 18 passed our  $\pi > 99\%$  threshold. The other 20 did not replicate, in most cases due to what appears to be incompletely controlled population structure driven by mixtures of populations with different allele frequencies before they came together and in a few cases due to failing quality control or fluctuating selection in space or time.

We present a gallery of 36 single-allele trajectories of particular interest (Fig. 3) as well as estimates of how their selection coefficients changed over time (Extended Data Fig. 6). These loci are not necessarily those with the largest  $X$  scores, but are highlighted as they address long-standing debates. They include 27 passing the  $\pi > 99\%$  threshold,

5 with probable evidence of selection ( $57\% < \pi < 93\%$ ) and 4 with surprising negative findings.

**HLA-DQB1: selection in favour of the major risk factor for coeliac disease**

At the HLA region of chromosome 6, densely packed genes have key roles in microorganism recognition. rs391176 (C>A, meaning that the ancestral allele is C and the newly arising mutation is A) is an excellent tag for *HLA-DBQ1\*02/DQ2* *HLA-DQB1\*02/DQ2*, with individuals carrying two A alleles having a 19-fold higher susceptibility for coeliac disease or gluten sensitivity (Extended Data Fig. 7a,b). The A allele has a selection coefficient of  $s = 4.4\%$  ( $\pi > 99\%$ ), rising from approximately 0% to approximately 20% in the past 4,000 years (Fig. 3a). These findings suggest that the pathogenic exposures that drove its rise were not a phenomenon only or largely of the rise of agriculture<sup>27</sup>.

**ABO: positive selection for B at the expense of A**

The *ABO* gene modifies oligosaccharides in glycoproteins on the surface of red blood cells; its A, B and null (O) alleles encode distinct glycoprotein alterations that modulate resistance and susceptibility to diverse pathogens<sup>28</sup>. The B allele rose from approximately 0% to approximately



hypothesized to have been positively selected<sup>29</sup>, but we detected negative selection ( $s = -0.9\%$ ,  $\pi > 99\%$ ), decreasing from approximately 50% to approximately 20% in the past 7,000 years (Fig. 3c). This is expected to have decreased baldness by 1.8% compared with the no-selection expectation.

#### **TYK2: reversal of selection at a major risk factor for tuberculosis**

Individuals carrying two copies of the rs34536443 G>C allele have more than 80% prevalence of clinically significant tuberculosis<sup>30</sup>. Previous work<sup>30</sup> has found evidence of negative selection on the C allele and hypothesized that this allele was associated with the time tuberculosis became endemic in Europe. We confirmed a decrease from approximately 9% to about 3% in the past approximately 3,000 years ( $s = -1.9\%$ ,  $\pi > 99\%$ ), but also positive selection from approximately 9,000 to about 3,000 years ago, from approximately 0% to approximately 9% ( $s = 2.3\%$ ,  $\pi > 99\%$ ; Fig. 3d). This may reflect changing endemicity of different pathogens over time.

#### **HLA-DRB1: elevated multiple sclerosis risk in northern Europe is not due to selection on the steppe**

A previous study<sup>31</sup> discovered positive selection at the rs3135388 G>A tag SNP for the *HLA-DRB1\*15:01* risk factor for multiple sclerosis. Because selection was already occurring in Yamnaya steppe pastoralists, and Yamnaya ancestry is most common in northern European individuals, the authors argued that the higher risk for multiple sclerosis in northern European individuals was driven by Yamnaya ancestry and selection on the steppe. We confirmed positive selection at this allele, rising from approximately 0% to about 16% approximately 6,000–2,000 years ago ( $s = 4.2\%$ ,  $\pi > 99\%$ ; Fig. 3e). However, the primary driver of the north–south differential was not selection on the steppe (Supplementary Section 7). First, selection began south of the Caucasus mountains in people without steppe ancestry. Second, after Yamnaya ancestry spread west, selection was stronger in northern Europe at  $s = 11.1 \pm 2.5\%$  than in southwestern Europe at  $s = 6.1 \pm 2.1\%$  (more than 3,500 years before present (BP)); this was the main driver of the north–south differential. Third, we detected negative selection in the past approximately 2,000 years missed by previous work ( $s = -2.1\%$ ,  $\pi > 99\%$ ), plausibly reflecting new pathogen exposures.

#### **HFE: reversal of selection at the major risk factor for haemochromatosis**

The rs1800562 (G>A) allele predicts pathogenic iron buildup in cells in individuals with two copies, and we found evidence of positive selection from approximately 5,000–2,000 years ago, rising from approximately 1% to about 5% ( $s = 2.7\%$ ,  $\pi > 99\%$ ), then dropping to approximately 4% today ( $s = -1.3\%$ , by itself not significantly different from zero at  $\pi = 11\%$ , but the decrease relative to the earlier positive selection is significant at  $P = 2.8 \times 10^{-9}$ ; Fig. 3f). It has been hypothesized that the causal allele protected against *Yersinia pestis* (the agent of the Black Death)<sup>32</sup>, but this is unlikely as its frequency was decreasing by the time of the Justinianic and medieval pandemics<sup>33</sup>.

#### **CCR5-Δ32: positive selection at an allele conferring immunity to HIV-1 infection**

The *CCR5-Δ32* allele confers complete resistance to HIV-1 infection in people who carry two copies<sup>34</sup>. An initial study dated the rise of this allele to medieval times and hypothesized that it was selected for resistance to the Black Death<sup>35</sup>. However, improved genetic maps revised its date to more than 5,000 years ago and the signal became non-significant<sup>36</sup>. We found that the allele was positively selected approximately 6,000–2,000 years ago (Fig. 3g and Extended Data Fig. 6g), increasing from approximately 2% to about 8% ( $s = 1.2\%$ ,  $\pi > 99\%$ ). This is too early to be explained by the medieval pandemic, but ancient pathogen studies have showed that *Yersinia* was endemic in West Eurasia for the past approximately 5,000 years<sup>37</sup>, resurrecting

the possibility that it was the cause, although other pathogens are possible.

#### **Selection for light skin tone at ten loci**

We found nine loci with genome-wide signals of selection for light skin tone, one probable signal and no loci showing selection for dark skin tone (Fig. 3h–q).

#### **CFTR: no evidence of selection for the major cystic fibrosis risk allele ΔF508**

The major risk allele for this recessive disease in European individuals has been hypothesized to be an example of heterozygote advantage due to conferring resistance to cholera in carriers<sup>38</sup>, keeping its frequency substantial despite its strong deleterious effects in homozygotes. However, we found no evidence of directional selection over the time frame that cholera was probably endemic in West Eurasia ( $\pi < 1\%$ ; Fig. 3r), with the earliest direct observation approximately 6,550 years BP in Croatia and the earliest imputed one approximately 15,411 years BP in Anatolia. Another explanation is needed for the persistence of the allele, plausibly by balancing selection.

Fifteen other selection discoveries are highlighted (Fig. 3s–ag). Most are highly significant at  $\pi > 99\%$ : *TSBP1* (coeliac disease,  $s = 5.6\%$ ); a second allele at *HLA-DQB1* (coeliac disease,  $s = 1.0\%$ ); *HLA-DRB1* (rheumatoid arthritis,  $s = -0.9\%$ ); *GYP A* (increases MNS blood group N,  $s = -0.9\%$ ); *DUOX2* (increases ferritin level,  $s = 1.3\%$ ); *SLC22A4* (Crohn's disease,  $s = 1.8\%$ ); *TLRI* (leprosy resistance,  $s = 2.0\%$ ); *CYP1A2* (decreases blood pressure,  $s = 1.0\%$ ); *NADSYN1/DHCR7* (increases vitamin D levels,  $s = 0.9\%$ ); *ADHIB* (lower risk for alcoholism,  $s = 2.3\%$ ); and *ABCG2* (gout,  $s = 0.9\%$ ). Four more signals are probable: *APOE* (hyperlipidaemia,  $s = 0.9\%$ ,  $\pi = 87\%$ ); *GCKR* (hyperlipidaemia or gout,  $s = 0.5\%$ ,  $\pi = 84\%$ ); *SERPINA1* ( $\alpha 1$ -antitrypsin deficiency,  $s = 1.6\%$ ,  $\pi = 57\%$ ); and a second locus at *SERPINA1* ( $\alpha 1$ -antitrypsin deficiency), which shows positive selection from approximately 7,000 to 2,500 years ago ( $s = 2.5\%$ ,  $\pi = 90\%$ ) followed by a reversal after approximately 2,500 years ago ( $s = -2.3\%$ ,  $\pi = 93\%$ ).

Figure 3ah–aj highlights null signals at loci previously hypothesized to be selected: *PTPN22* (hypothyroidism), a second allele at *HFE* (haemochromatosis) and *IL23R* (Crohn's disease).

#### **Tests for selection on complex traits**

We searched for evidence that groups of alleles with similar influence on traits today trended in the same direction in the past, as expected if a phenotype with a similar genetic underpinning was the target of selection. We leveraged 563 GWAS results in people of European ancestry: 452 mostly quantitative traits in the UK Biobank, and 111 curated traits from studies especially of common disease (Supplementary Table 5). How phenotypes manifest today may be very different from how they manifested in past populations living in different environments with different lifestyles, so any signals discovered by this approach should not be interpreted as evidence for selection on the exact phenotype being tested.

We used three statistics to test for coordinated selection on alleles affecting the same trait (we excluded variants at the HLA locus as we were interested in polygenic signals, not those dominated by large effect variants). First, we computed a polygenic score (PGS) for each GWAS: a linear combination of allelic values, weighted by estimated effect size. We evaluated whether the change in PGS over time  $\gamma$  (scaled so one unit corresponds to a standard deviation change over ten millennia) is more than could be expected by genetic drift or population structure, using the genetic relatedness matrix to control for these effects as in our single-allele tests. To test whether the observed deviation is significant, we repeated the test 100 times with randomly flipped signs of GWAS effects, to correct for residual inflation. Second, we repeated the procedure without using magnitudes of GWAS effects,

and instead only the sign, generating a statistic  $\gamma_{\text{sign}}$  that may be less affected by concerns about transferability of PGS across groups<sup>3,39</sup>. Third, we performed a SNP-by-SNP comparison for each trait, using cross-trait linkage disequilibrium score regression (LDSC)<sup>40</sup> to estimate genetic correlation ( $r_g$ ) between selection coefficients ( $s$ ) and trait effect sizes. We computed a standard error from a block jackknife to test whether this correlation is different from zero, accounting for non-independence of SNPs. We found high Pearson's correlation for all three tests (0.78–0.92; Extended Data Fig. 8).

For 31 traits, we repeated analyses using effect sizes measured in East Asian GWAS. Population structure in East Asian individuals is uncorrelated to that in ancient West Eurasian individuals, so any validation by this test must reflect a real signal of selection<sup>3,39</sup>.

For 34 traits, we repeated analyses using effect sizes estimated from family-based GWAS. This analysis distinguished 'direct effects' of genetic variants on the biology of the individual being examined, from 'indirect effects' that affect the trait by modulating the behaviour of family members and thereby changing the environment of the individual. Despite much reduced sample sizes, standard errors are small enough for some traits to validate our findings.

Multiple lines of evidence point to the robustness of our tests of polygenic selection to the potential confounding factor of population structure. First, we carried out ordinary linear regression searching for evidence of a change in the polygenic predictor of height over time and found a significant signal without correction for structure ( $P = 7 \times 10^{-159}$ )<sup>41,42</sup>, which disappeared after our correction: ( $\gamma, P = 0.21$ ). Second, we observed strong Pearson's correlations between  $Z$ -scores for European and East Asian GWAS of 0.43 ( $\gamma$ ), 0.75 ( $\gamma_{\text{sign}}$ ) and 0.81 ( $r_s$ ; Extended Data Fig. 12a); population structure is uncorrelated for European and East Asian individuals, and thus structure artefacts cannot explain this signal. A third line of evidence comes from positive correlations between  $Z$ -scores from population GWAS and direct genetic effect GWAS from family-based studies: 0.51 ( $\gamma$ ), 0.51 ( $\gamma_{\text{sign}}$ ) and 0.76 ( $r_s$ ; Extended Data Fig. 12b). Fourth, we did not observe false positives in realistic simulations of European structure without directional selection (Supplementary Section 2).

## Directional selection on complex traits

In total, 44 of 563 GWAS showed significant signals across all three tests after correction for the number of GWAS examined ( $P < 8.9 \times 10^{-5}$ ), and 100 GWAS showed signals in at least one test after correction for multiple hypothesis testing ( $P < 3.0 \times 10^{-5}$ ; Supplementary Table 5). We focused on 12 traits of particular interest with significant signals from all three tests (Fig. 4 and Extended Data Fig. 9).

One of the strongest signals is an increase over time in the PGS for light skin pigmentation ( $\gamma = 1.80 \pm 0.10$  standard deviations increase in mean PGS in ten millennia;  $P = 5.7 \times 10^{-74}$ ; Fig. 4 and Extended Data Fig. 9). This plausibly reflects selection for increased synthesis of vitamin D in regions of low sunlight in farmers with little of it in their diets. Most of the phenotypic shift is driven by a handful of loci<sup>43</sup> (52% due to *SLC45A2* alone, and 75% to the top 10 loci); however, the signal is highly polygenic: we need to drop the top 79 loci before the signal disappears (Extended Data Fig. 10). A model in which selection for pigmentation impacted all variants in proportion to their effect size fits ( $P = 0.11$ ; Extended Data Fig. 11).

Type 2 diabetes risk factors give compelling signals of negative selection. Thus, we observed negative selection on combinations of alleles that today increase body fat percentage ( $\gamma = -1.04 \pm 0.13$ ), waist circumference ( $\gamma = -0.91 \pm 0.13$ ) and waist-to-hip ratio ( $\gamma = -0.76 \pm 0.13$ ; Fig. 4), which could be consistent with the 'Thrifty Genes' hypothesis<sup>44</sup> that a genetic adaptation to store fat to promote survival during periods of scarcity in hunter-gatherers became deleterious after the transition to food production. Skeletal evidence for nutritional stresses in early European farmers<sup>45–48</sup> could be viewed as in tension with this

hypothesis, but is consistent with it if a farming lifestyle was associated with famines that occurred on too-infrequent a timescale to be effectively buffered by higher fat. The signal of selection against the polygenic predictor of type 2 diabetes itself just misses the multiple hypothesis-testing corrected threshold of  $|Z| > 3.92$  ( $\gamma = -0.43 \pm 0.12$  ( $Z = -3.59$ ),  $\gamma_{\text{sign}} = -0.48 \pm 0.12$  ( $Z = -3.85$ ),  $r_s = -0.14 \pm 0.04$  ( $Z = -3.59$ )).

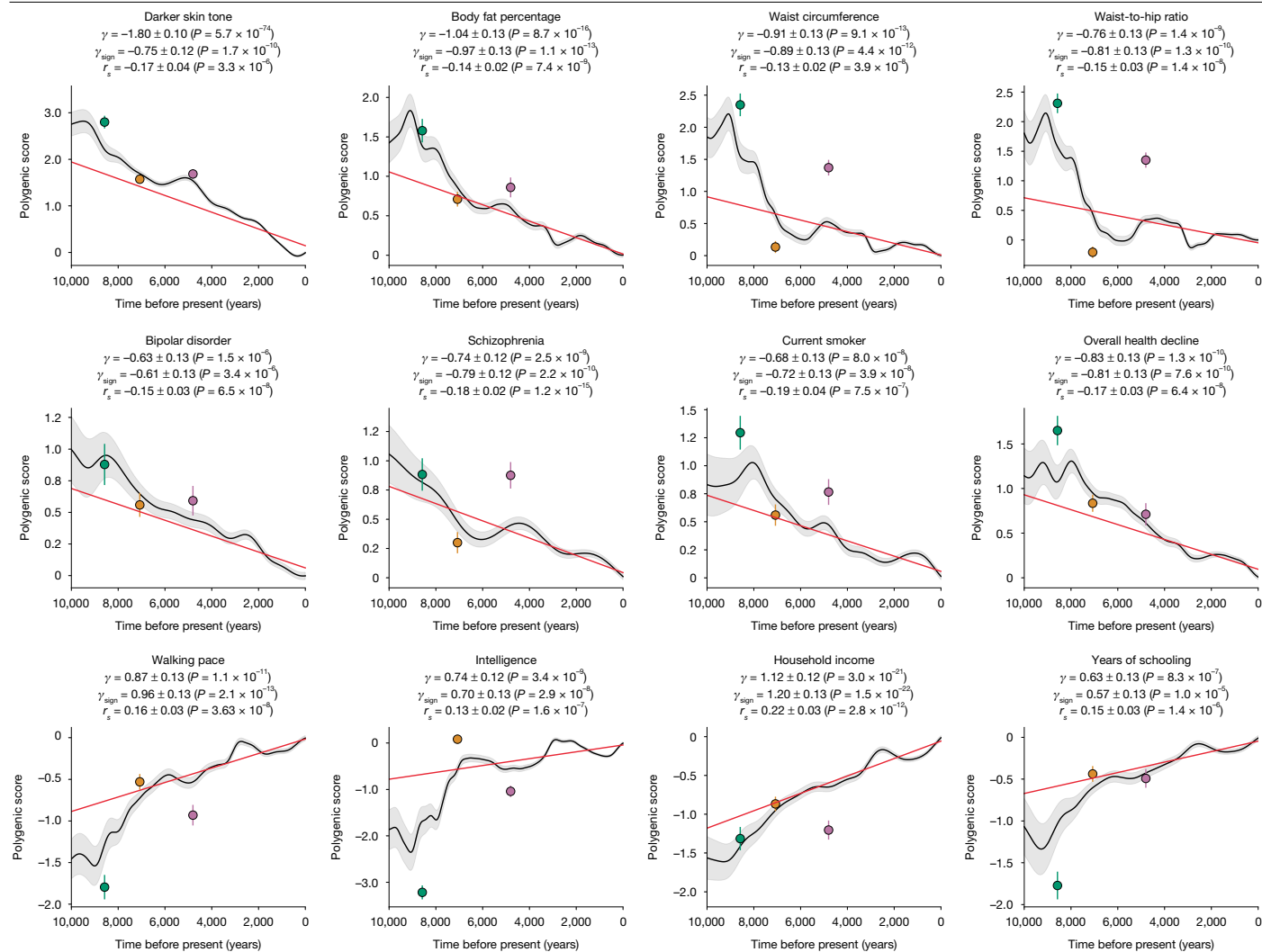
We detected negative polygenic selection against alleles associated today with psychoses such as bipolar disorder ( $\gamma = -0.63 \pm 0.13$ ) and schizophrenia ( $\gamma = -0.74 \pm 0.12$ ; Fig. 4). These results might at first seem surprising given our finding that significant signals of selection are not enriched for variants with genome-wide significant evidence of modulating psychiatric traits in GWAS (Fig. 1d). However, for variants with sub-genome-wide significant signals in GWAS, we did observe heritability enrichment<sup>49</sup> (Extended Data Fig. 5a). Brain traits have qualitatively different genetic architectures from blood-immune-inflammatory traits, with a higher proportion of sites modulating them and smaller effect sizes on average per allele<sup>50</sup>. Thus, our observations probably reflect a situation in which alleles with strong enough selection coefficients to be significant are concentrated in immune traits, whereas selection on brain traits is so polygenic that most individual contributing alleles have too-small selection coefficients to be detected. Indeed, directional selection against the genetic predictors today predisposing to psychosis is extraordinarily polygenic: we have to drop 584 loci for bipolar disorder and 825 loci for schizophrenia for the signals to disappear (Extended Data Fig. 10).

We observed signals of selection for combinations of alleles that are today associated with healthy lifestyles into old age. This includes selection against alleles today associated with smoking ( $\gamma = -0.68 \pm 0.13$ ), against alleles contributing to overall health decline ( $\gamma = -0.83 \pm 0.13$ ) and for alleles that promote faster walking pace ( $\gamma = 0.87 \pm 0.13$ ).

We finally observed signals of selection for combinations of alleles that today are associated with three correlated behavioural traits: scores on intelligence tests (increasing  $\gamma = 0.74 \pm 0.12$ ), household income (increasing  $\gamma = 1.12 \pm 0.12$ ) and years of schooling (increasing  $\gamma = 0.63 \pm 0.13$ ). These signals are all highly polygenic, and we have to drop 449–1,056 loci for the signals to become non-significant (Extended Data Fig. 10). The signals are largely driven by selection before approximately 2,000 years BP, after which  $\gamma$  tends towards zero (Extended Data Fig. 9).

We were concerned that the evidence of polygenic selection for behavioural traits might be an artefact of uncorrected population structure in West Eurasia, despite our simulations showing that our methodology corrects for such structure. To explore this, we first focused on years of schooling, for which available GWAS were largest. We tested for a correlation of East Asian GWAS effect size measurements to West Eurasian selection coefficients. Despite reduced power—driven by the smaller sample sizes of East Asian GWAS, and differences in linkage disequilibrium structure and genetic architecture across populations<sup>51–53</sup>—we replicated the signal of selection, which cannot be explained by imperfectly corrected West Eurasian population structure, which is uncorrelated to that in East Eurasia:  $\gamma$  ( $P = 7.3 \times 10^{-5}$ ),  $\gamma_{\text{sign}}$  ( $P = 2.1 \times 10^{-5}$ ) and  $r_s$  ( $P = 3.9 \times 10^{-12}$ ; Extended Data Fig. 12a). We also replicated using family-based GWAS<sup>54</sup>, which is immune to population structure (Extended Data Fig. 12b). Despite limited power of these studies due to small sample sizes, the PGS for direct effects on years of schooling showing nominal evidence of positive selection in West Eurasian individuals by all three tests:  $\gamma$  ( $P = 0.027$ ),  $\gamma_{\text{sign}}$  ( $P = 0.016$ ) and  $r_s$  ( $P = 0.029$ ). The score capturing indirect genetic effects—parental alleles not passed to offspring and thus modulating parental behaviour and family environment—showed a signal in one test:  $\gamma$  ( $P = 0.920$ ),  $\gamma_{\text{sign}}$  ( $P = 0.945$ ) and  $r_s$  ( $P = 0.009$ ).

We could not obtain insight into the phenotype that was being selected in the past through functional stratification of the years of schooling signal. Thus, when we compared GWAS for the cognitive and non-cognitive components of the years of schooling trait<sup>55</sup>, both showed significant



**Fig. 4 | Notable signals of directional polygenic selection.** We show 12 instances where tests of polygenic selection— $\gamma$ ,  $\gamma_{\text{sign}}$  and  $r_s$ —are all significant, with the relevant statistics at the top of each panel. The mean PGS of West Eurasian individuals over 10,000 years is denoted by a solid black line, with the 95% CI in grey. The red line represents the linear mixed model regression,

signals individually, but their pairwise differences were not significant: all three tests were  $P > 0.05$  (Supplementary Table 5). We also analysed brain volume<sup>56</sup>, which is correlated with years of schooling (LDSC genetic correlation ( $r_g$ ) =  $0.26 \pm 0.03$ ) and found no signals (all  $P > 0.05$ ).

There are caveats when interpreting signals of polygenic adaptation, especially for the three genetically correlated traits of scores on intelligence tests, household income and years of schooling. These traits are only relevant to modern societies and would have been unmeasurable in preliterate societies over the vast majority of the period during which selection acted. The difficulty of interpretation is enhanced by the fact that the alleles driving down the frequency of type 2 diabetes-related traits are highly correlated to those contributing to the increased scores for years of schooling, household income and intelligence (Extended Data Fig. 13). Finally, there is evidence of change in these selection pressures over time<sup>57</sup>, for example, in Iceland in the past century in which there was significant selection to decrease the predictor of years of schooling, opposite to the long-term increase that we detected.

## Discussion

Previous work has shown that classic selective sweeps driving alleles to fixation have been rare over the broad span of human evolution<sup>8,9</sup>.

adjusted for population structure, with slope  $\gamma$ . The circles represent mean PGS in Western hunter-gatherers (green;  $n = 131$ ), early European farmers (orange;  $n = 452$ ) and steppe pastoralists (pink;  $n = 293$ ); the error bars indicate 95% CIs. All  $P$  values are two-sided and evaluated against a Bonferroni-corrected threshold for multiple testing ( $P < 8.9 \times 10^{-5}$ ).

Thus, we were surprised that over the past 10,000 years in West Eurasia, there have been many hundreds of instances of directional selection with coefficients on the order of 0.5% or more (Fig. 2b). This is large enough that if a similarly dense landscape of directionally selected variants had existed tens of thousands of years ago, and if the selection coefficients had been constant since then, we would expect many fixed differences across populations, despite the fact that previous studies have shown that there are only a handful, hardly more than would be expected based on random drift<sup>9</sup>.

The simplest way to resolve this paradox is to recognize that selection coefficients are unlikely to have been constant over time, even though we make this simplifying assumption to enable detection of selection. By sliding a 2,000-year window through our time transect and re-estimating selection coefficients within each window (Fig. 3 and Extended Data Fig. 6), we can see that there have in fact been changes in selection pressures at a number of the loci that we analysed, including at *HLA-DRB1*, *TYK2* and *HFE*. By comparing the estimated age of the mutation that contributed each selected allele<sup>58</sup> to the extrapolated time to reach fixation given its estimated  $s$  value, we find that about one-third of the favoured mutations are ancestral alleles and another one-third are derived alleles with true ages an order of magnitude older than the expected sweep age, indicating that selection coefficients

must have shifted over time at these variants (Fig. 2c). The remaining up to one-third of the signals that we detect are consistent with classic sweeps where the newly arising variants have been under consistent positive selection since they arose (but they arose recently enough that the sweeps are not complete).

A second explanation for this paradox is that West Eurasian individuals have been experiencing qualitatively more and different natural selection in the Holocene than in earlier periods because of rapidly changing lifestyles and economies. This hypothesis is supported by our evidence of intensifying selection for traits including blood–immune–inflammatory traits in the Bronze Age compared with earlier periods (Fig. 1e and Extended Data Fig. 5b).

Although the exact number of independent loci cannot be estimated with high confidence, we project that there are at least 3,800 independent signals of directional selection (one-half of the 7,689 non-HLA loci at the FDR = 50% threshold) that are in linkage disequilibrium with the overwhelming majority of variants in the genome (Extended Data Fig. 2b). Superficially, this appears to be at odds with findings that there has been relatively little contribution from directional selection to allele frequency changes in genome compared with the much larger forces of gene flow, genetic drift, and purifying or stabilizing selection<sup>6</sup>. In fact, there is no conflict. Our method allows us to partition the effects of selection at each SNP into the influences of directional selection ( $s$ ) and the combined effects of fluctuating selection and drift ( $\sigma^2$ ). We estimate that only 2.16 ( $\pm 0.11$ )% (jackknife standard deviation) of allele frequency changes are due to directional selection. Allele frequency change in human populations has been so rampant that even if a small fraction of allele frequency change is due to directional selection, this corresponds to many thousands of loci.

We ruled out the possibility that our results are reflecting not directional selection but stabilizing selection—selection to reduce genetic variability even if the population is near a fitness optimum—a major force shaping patterns of variation in living people<sup>6,8,20,25,59</sup>. Previous work has shown that if there is a shift in a trait optimum, there is expected to be a deterministic directional selection phase during which the population shifts towards the optimum, followed by a stochastic stabilizing selection phase when underdominant selection pushes common variants towards fixation or loss<sup>21</sup>. Equation 7 of ref. 21 shows that the selection coefficient ( $s$ ) in the deterministic phase scales linearly with the phenotypic effect of the allele, whereas in the stochastic phase, it scales with the square. Assuming a phenotypic effect of approximately 0.01, plausible based on typical GWAS, yields a selection coefficient on the order of  $s \sim 0.0001$ . Any changes in frequency would be expected to occur at a timescale of  $1/s$  (figure 3 of ref. 21), and thus, while the deterministic phase unfolds over about 100 generations, the stochastic phase spans 10,000–100,000 generations, implying that underdominance due to stabilizing selection is effectively indistinguishable from neutral evolution in the context of our study. Confirming this, our simulations of diverse scenarios of stabilizing selection do not produce the enrichment of GWAS signals and HAF score that we observed empirically (Fig. 1b,d, Extended Data Fig. 4 and Supplementary Section 2).

We considered and ruled out the possibility that our observations are not evidence of directional selection, but instead a process noted by refs. 59,60, in which in a panmictic population, purifying selection can generate positive covariance between frequency changes in adjacent time intervals for neutral loci linked to newly arising deleterious mutations. Three theoretical reasons indicate that this cannot be explaining our observations. First, we lack the temporal resolution to detect such signals that have a timescale of 10–20 generations, on the order of the standard error of estimated dates in ancient DNA. Second, our study spans hundreds of generations, so the signals that we detected show autocovariance on a far longer timescale than can be explained by this phenomenon. Third, the population in our study is not panmictic as in ref. 60; on a timescale of 10–20 generations, it contains many effective

replicate populations isolated by minimal migration, a scenario that ref. 59 shows reduces autocorrelations, consistent with ref. 6, who analysed two parallel ancient human time transects and found no genome-wide signal of linked selection. Supporting these arguments, our simulations of purifying selection do not show enrichment of GWAS signals and HAF score similar to those we observe in real data (Fig. 1b,d, Extended Data Fig. 4 and Supplementary Section 2).

It will be of interest to apply similar approaches to ancient DNA time series over longer times and to other world regions. This would allow more generalizable insights by identifying which patterns of selection are shared and which are distinctive to Holocene West Eurasia.

## Online content

Any methods, additional references, Nature Portfolio reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at <https://doi.org/10.1038/s41586-026-10358-1>.

1. Liu, Y., Mao, X., Krause, J. & Fu, Q. Insights into human history from the first decade of ancient human genomics. *Science* **373**, 1479–1484 (2021).
2. Mathieson, I. et al. Genome-wide patterns of selection in 230 ancient Eurasians. *Nature* **528**, 499–503 (2015).
3. Le, M. K. et al. 1,000 Ancient genomes uncover 10,000 years of natural selection in Europe. Preprint at *bioRxiv* <https://doi.org/10.1101/2022.08.24.505188> (2022).
4. Irving-Pease, E. K. et al. The selection landscape and genetic legacy of ancient Eurasians. *Nature* **625**, 312–320 (2024).
5. Kerner, G. et al. Genetic adaptation to pathogens and increased risk of inflammatory disorders in post-Neolithic Europe. *Cell Genom.* **3**, 100248 (2023).
6. Simon, A. & Coop, G. The contribution of gene flow, selection, and genetic drift to five thousand years of human allele frequency change. *Proc. Natl Acad. Sci. USA* **121**, e2312377121 (2024).
7. Dehaesque, M. et al. Inference of natural selection from ancient DNA. *Evol. Lett.* **4**, 94–108 (2020).
8. Hernandez, R. D. et al. Classic selective sweeps were rare in recent human evolution. *Science* **331**, 920–924 (2011).
9. Coop, G. et al. The role of geography in human adaptation. *PLoS Genet.* **5**, e1000500 (2009).
10. Kerner, G., Choin, J. & Quintana-Murci, L. Ancient DNA as a tool for medical research. *Nat. Med.* **29**, 1048–1051 (2023).
11. Bennett, E. A. & Fu, Q. Ancient genomes and the evolutionary path of modern humans. *Cell* **187**, 1042–1046 (2024).
12. Vitti, J. J., Grossman, S. R. & Sabeti, P. C. Detecting natural selection in genomic data. *Annu. Rev. Genet.* **47**, 97–120 (2013).
13. Field, Y. et al. Detection of human adaptation during the past 2000 years. *Science* **354**, 760–764 (2016).
14. Berg, J. J. & Coop, G. A coalescent model for a sweep of a unique standing variant. *Genetics* **201**, 707–725 (2015).
15. Mallick, S. et al. The Allen Ancient DNA Resource (AADR) a curated compendium of ancient human genomes. *Sci. Data* **11**, 182 (2024).
16. Rubinacci, S., Ribeiro, D. M., Hofmeister, R. J. & Delaneau, O. Efficient phasing and imputation of low-coverage sequencing data using large reference panels. *Nat. Genet.* **53**, 120–126 (2021).
17. Zhou, W. et al. Scalable generalized linear mixed model for region-based association tests in large biobanks and cohorts. *Nat. Genet.* **52**, 634–639 (2020).
18. Yang, J. et al. Genomic inflation factors under polygenic inheritance. *Eur. J. Hum. Genet.* **19**, 807–812 (2011).
19. Devlin, B., Roeder, K. & Wasserman, L. Genomic control, a new approach to genetic-based association studies. *Theor. Popul. Biol.* **60**, 155–166 (2001).
20. Koch, E. et al. Genetic association data are broadly consistent with stabilizing selection shaping human common diseases and traits. Preprint at *bioRxiv* <https://doi.org/10.1101/2024.06.19.599789> (2024).
21. Hayward, L. K. & Sella, G. Polygenic adaptation after a sudden change in environment. *eLife* **11**, e66697 (2022).
22. Ronen, R. et al. Predicting carriers of ongoing selective sweeps without knowledge of the favored allele. *PLoS Genet.* **11**, e1005527 (2015).
23. Akbari, A. et al. Identifying the favored mutation in a positive selective sweep. *Nat. Methods* **15**, 279–282 (2018).
24. McVicker, G., Gordon, D., Davis, C. & Green, P. Widespread genomic signatures of natural selection in hominid evolution. *PLoS Genet.* **5**, e1000471 (2009).
25. Murphy, D. A., Elyashiv, E., Amster, G. & Sella, G. Broad-scale variation in human genetic diversity levels is predicted by purifying selection on coding and non-coding elements. *eLife* **12**, e76065 (2023).
26. Sikora, M. et al. The spatiotemporal distribution of human pathogens in ancient Eurasia. *Nature* **643**, 1011–1019 (2025).
27. Sams, A. & Hawks, J. Celiac disease as a model for the evolution of multifactorial disease in humans. *Hum. Biol.* **86**, 19–36 (2014).
28. Abegaz, S. B. Human ABO blood groups and their associations with different diseases. *BioMed Res. Int.* **2021**, 6629060 (2021).

29. Medland, S. E. et al. Common variants in the trichohyalin gene are associated with straight hair in Europeans. *Am. J. Hum. Genet.* **85**, 750–755 (2009).
30. Kerner, G. et al. Human ancient DNA analyses reveal the high burden of tuberculosis in Europeans over the last 2,000 years. *Am. J. Hum. Genet.* **108**, 517–524 (2021).
31. Barrie, W. et al. Elevated genetic risk for multiple sclerosis emerged in steppe pastoralist populations. *Nature* **625**, 321–328 (2024).
32. Moalem, S., Percy, M. E., Kruck, T. P. A. & Gelbart, R. R. Epidemic pathogenic selection: an explanation for hereditary hemochromatosis? *Med. Hypotheses* **59**, 325–329 (2002).
33. Wagner, D. M. et al. *Yersinia pestis* and the plague of Justinian 541–543 AD: a genomic analysis. *Lancet Infect. Dis.* **14**, 319–326 (2014).
34. Gupta, R. K. et al. HIV-1 remission following CCR5-Δ32/Δ32 haematopoietic stem-cell transplantation. *Nature* **568**, 244–248 (2019).
35. Stephens, J. C. et al. Dating the origin of the CCR5-Delta32 AIDS-resistance allele by the coalescence of haplotypes. *Am. J. Hum. Genet.* **62**, 1507–1515 (1998).
36. Sabeti, P. C. et al. The case for selection at CCR5-Δ32. *PLoS Biol.* **3**, e378 (2005).
37. Seersholm, F. V. et al. Repeated plague infections across six generations of Neolithic farmers. *Nature* **632**, 114–121 (2024).
38. Gabriel, S. E., Brigman, K. N., Koller, B. H., Boucher, R. C. & Stutts, M. J. Cystic fibrosis heterozygote resistance to cholera toxin in the cystic fibrosis mouse model. *Science* **266**, 107–109 (1994).
39. Chen, M. et al. Evidence of polygenic adaptation in Sardinia at height-associated loci ascertained from the Biobank Japan. *Am. J. Hum. Genet.* **107**, 60–71 (2020).
40. Bulik-Sullivan, B. et al. An atlas of genetic correlations across human diseases and traits. *Nat. Genet.* **47**, 1236–1241 (2015).
41. Sohail, M. et al. Polygenic adaptation on height is overestimated due to uncorrected stratification in genome-wide association studies. *eLife* **8**, e39702 (2019).
42. Berg, J. J. et al. Reduced signal for polygenic adaptation of height in UK Biobank. *eLife* **8**, e39725 (2019).
43. Ju, D. & Mathieson, I. The evolution of skin pigmentation-associated variation in West Eurasia. *Proc. Natl Acad. Sci. USA* **118**, e2009227118 (2021).
44. Neel, J. V. Diabetes mellitus: a 'thrifty' genotype rendered detrimental by 'progress'? *Am. J. Hum. Genet.* **14**, 353–362 (1962).
45. Cohen, M. N., Armelagos, G. J. & Larsen, C. S. *Paleopathology at the Origins of Agriculture* (Univ. Press of Florida, 2013).
46. Macintosh, A. A., Pinhasi, R. & Stock, J. T. Early life conditions and physiological stress following the transition to farming in Central/Southeast Europe: skeletal growth impairment and 6000 years of gradual recovery. *PLoS ONE* **11**, e0148468 (2016).
47. Theodorakopoulou, K. & Karamanou, M. Human paleopathology during the stone age. *Arch. Balk. Med. Union* **55**, 676–683 (2020).
48. Marciniak, S. et al. An integrative skeletal and paleogenomic analysis of stature variation suggests relatively reduced health for early European farmers. *Proc. Natl Acad. Sci. USA* **119**, e2106743119 (2022).
49. Finucane, H. K. et al. Partitioning heritability by functional annotation using genome-wide association summary statistics. *Nat. Genet.* **47**, 1228–1235 (2015).
50. O'Connor, L. J. et al. Extreme polygenicity of complex traits is explained by negative selection. *Am. J. Hum. Genet.* **105**, 456–476 (2019).
51. Amariuta, T. et al. Improving the trans-ancestry portability of polygenic risk scores by prioritizing variants in predicted cell-type-specific regulatory elements. *Nat. Genet.* **52**, 1346–1354 (2020).
52. Martin, A. R. et al. Clinical use of current polygenic risk scores may exacerbate health disparities. *Nat. Genet.* **51**, 584–591 (2019).
53. Wang, Y. et al. Theoretical and empirical quantification of the accuracy of polygenic scores in ancestry divergent populations. *Nat. Commun.* **11**, 3865 (2020).
54. Tan, T. et al. Family-GWAS reveals effects of environment and mating on genetic associations. Preprint at medRxiv <https://doi.org/10.1101/2024.10.01.24314703> (2024).
55. Demange, P. A. et al. Investigating the genetic architecture of noncognitive skills using GWAS-by-subtraction. *Nat. Genet.* **53**, 35–44 (2021).
56. Jansen, P. R. et al. Genome-wide meta-analysis of brain volume identifies genomic loci and genes shared with intelligence. *Nat. Commun.* **11**, 5606 (2020).
57. Kong, A. et al. Selection against variants in the genome associated with educational attainment. *Proc. Natl Acad. Sci. USA* **114**, E727–E732 (2017).
58. Speidel, L., Forest, M., Shi, S. & Myers, S. R. A method for genome-wide genealogy estimation for thousands of samples. *Nat. Genet.* **51**, 1321–1329 (2019).
59. Buffalo, V. & Coop, G. Estimating the genome-wide contribution of selection to temporal allele frequency change. *Proc. Natl Acad. Sci. USA* **117**, 20672–20680 (2020).
60. Buffalo, V. & Coop, G. The linked selection signature of rapid adaptation in temporal genomic data. *Genetics* **213**, 1007–1045 (2019).

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

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.

© The Author(s), under exclusive licence to Springer Nature Limited 2026

## Methods

### Testing for selection while correcting for population structure

We used a generalized linear mixed model (GLMM) approach to correct for population structure. Population structure is a major confounder in scans for significant changes in frequency over time, especially as migration and population mixture have been common in almost all parts of the world. Previous studies have corrected for structure in ancient DNA time transects by modelling the population history and estimating mixture proportions, which works optimally only if there are data from the true source population, which is rarely the case. It is tempting to use an unsupervised approach such as principal components to address population structure. However, after experimentation, we found that this is not effective as principal components are correlated with sample dates, which creates collinearity with the quantity that we are most interested in (the time-varying component), inflating the empirically estimated variance and reducing power (Extended Data Fig. 14).

The mixed model approach, which is often deployed in the context of genetic association studies and has also been applied in selection studies in modern populations<sup>61</sup> to address similar challenges<sup>62</sup>, offers a way to address these issues by combining structured data in an unsupervised manner and estimating fewer parameters over a wider span of time, which results in greater power than using separate regression analyses for each population or comparing the estimated means from different groups. Our simulations show that under simplifying assumptions, a GLMM is more powerful in controlling for population structure and detecting change in allele frequency than a generalized linear model using the top principal components as covariates (Extended Data Fig. 14). Thus, the method has advantages despite the fact that the model fitted by the GLMM makes the unrealistic assumption of constant selective pressure across space and time, and is thus susceptible to missing real signals of fluctuating selection such as *TYK2* rs34536443.

We used our GLMM to fit a linear time-varying component to the logit (log-odds) transformation of allele frequency at each position in the genome, and then to test whether there is evidence for a consistent trend in allele frequency change over time for all populations. We searched for evidence of such a trend beyond the prediction based on population structure and associated genetic drift relating sampled individuals in space and time (as measured by the covariance of genotypes over all the individuals, known as the genetic relationship matrix (GRM)). In our GLMM, the response variable for each tested allele  $j$  is the allele count. The allele counts for an individual  $i$  are drawn from a binomial distribution  $B(2, p_{ij})$ , where 2 is the number of chromosomes each person carries at each position, and  $p_{ij}$  is the unknown frequency of allele  $j$  in the population in which the tested individual  $i$  lives. A logit link function allows the frequency  $p_{ij}$  to be modelled as a linear combination of covariates. This is a generalization of the logistic mixed model where the response variable is binary:

$$\ln\left(\frac{p_{ij}}{1-p_{ij}}\right) = \alpha_j + s_j t_i + g_{ij}. \quad (1)$$

The logit function,  $\ln\left(\frac{p_{ij}}{1-p_{ij}}\right)$ , transforms allele frequency so that its expected change per generation is proportional to the selection coefficient  $s_j$  (regardless of  $p_{ij}$ )<sup>63,64</sup>.  $\alpha_j$  is a constant related to the average logit transformation of allele frequency in sampled individuals at time  $t = 0$  today.  $s_j$  is the per-generation selection strength at the allele, assumed constant over time and space during the period of our time transect; our test for selection is simply a test for whether the equation fits significantly better if  $s_j$  is non-zero than if it is zero.  $t_i$  is the negative sampling date in the past, in units of twice the generation interval<sup>63,64</sup> (assuming 29 years per generation).  $g_{ij}$  is a random effect, an error term capturing individual-specific variability not explained by fixed effects

( $\alpha_j + s_j t_i$ ); it differs from the error term in a generalized linear model, which is independently and identically distributed following a normal distribution. In our GLMM, the error term is drawn from the vector  $\mathbf{g}_j \sim \text{MVN}(0, \sigma_j^2 \mathbf{K})$ , following a multivariate normal (MVN) distribution, where  $\mathbf{K}$  is the covariance matrix structure (the GRM), the empirically observed relatedness of all individuals to each other, and  $\sigma_j^2$  measures the drift at that variant, constrained to be no smaller than the genome-wide expectation conditional on MAF.

$s_j$ ,  $\sigma_j^2$  and  $\alpha_j$  are independently estimated for each of 9.7 million variants. Refitting  $\sigma_j^2$  and  $\alpha_j$  at each analysed site makes the method robust to false positives due to processes that vary across SNPs such as degree of background selection, which increases the effective amount of random genetic drift or variation in MAF. These nuisance random effects are soaked up by allowing  $\sigma_j^2$  and  $\alpha_j$  to vary, allowing us to test for a time-dependent influence on allele frequency fluctuations  $s_j$  beyond what can be explained by the GRM. We placed a minimum value on  $\sigma_j^2$  constrained by genome-wide estimates conditional on MAF, as we were concerned that if we did not do this, the method might assume an unreasonably low  $\sigma_j^2$  and ascribe too much allele frequency change to the  $s_j$  term, again contributing to false positives. Our test for a non-zero  $s_j$  is thus a test for selection above and beyond what could be explained not just by structure but also other non-time-dependent processes. The penalty that we pay for estimating variance components at millions of SNPs—in contrast to the constant variance component assumption used in mixed model analysis in GWAS<sup>62</sup>—is computational load. We grouped individuals with similar ancestry and dates into 2,000 clusters; at this resolution, our analysis required approximately 130,000 CPU hours (Supplementary Section 8).

Using the GLMM, we obtained a point estimate for the selection coefficient at each variant and its standard error, and a Z-score for the number of standard errors this is from zero, a naive test for selection. In practice, the statistic needs recalibration as it is inflated due to unmodelled features of the data, so we empirically assessed significance from enrichment of signals in independent GWAS (Supplementary Section 3).

### Fitting the GLMM

We developed PQLseqPy, a faster implementation of PQLseq<sup>65</sup> for fitting the GLMM to count data (Supplementary Section 8). Despite an 82-fold speed increase, running a GLMM on approximately 20,000 individuals for about 9.7 million variants was infeasible (approximately 6,500 CPU years). To make analysis tractable, we grouped individuals into clusters with similar ancestry and coming from similar times. For the single-variant selection tests, we restricted to 13,936 ancient individuals who were unrelated up to the second degree to make computation tractable (for the polygenic tests, we included related individuals and did not apply clustering).

To identify the  $T = 2,000$  clusters that we analysed, we required there to be a maximum date gap  $G = 500$  years between any two individuals in each cluster. We initialized the interval  $l = (l = 2, r = T)$  with midpoint  $m$ . We applied hierarchical clustering on the top 30 principal components using the `sklearn.cluster.AgglomerativeClustering` function in Python with default parameters and `n_clusters = m`. For each of the  $S$  clusters from the previous step, we performed hierarchical clustering on the dates with distance threshold =  $G$  and `n_clusters = none`. If the resulting number of clusters was larger than  $T + 1$ , we repeated the process with  $l = (l, m)$ . If it was less than  $T - 1$ , we updated  $l = (m, r)$ . We repeated these steps until the final number of clusters was within  $T - 1$  to  $T + 1$ , or  $l = r$ . Across 2,000 clusters, individuals per cluster has a first quartile of 2, a median of 4, a third quartile of 10 and a maximum of 328.

We used the GLMM model in equation (1). However, the cluster can include more than one individual. The allele counts for each cluster  $i$  are drawn from a binomial distribution  $B(2n_i, p_{ij})$ , where  $n_i$  is the number

# Article

of diploid individuals in the cluster, and  $p_{ij}$  is the unknown frequency of allele  $j$  in the population where individuals in cluster  $i$  reside.

## Proportion of variance explained by directional selection

The proportion of variance in allele frequency on the logit scale for each SNP  $j$  is:

$$\text{Proportion of variance for SNP } j = \frac{s_j^2 \cdot \text{var}(t)}{s_j^2 \cdot \text{var}(t) + \sigma_j^2} \quad (2)$$

We used 1,000 independent SNPs, randomly selected across the genome with pairwise linkage disequilibrium ( $r^2$ ) less than 0.05, to estimate that directional selection explains an average of 2.16% of the variance in allele frequency, with a standard error of 0.11% based on jackknife estimation. The GLMM used for this analysis is based on the full sample size of unrelated individuals, rather than clustering individuals according to their ancestry and date.

## Covariance structure for the GLMM

The covariance structure matrix  $\mathbf{K}$  for clusters  $m$  and  $n$  is defined as:

$$\mathbf{K}_{mn} = \frac{1}{N_m N_n} \sum_{i \in c_m} \sum_{j \in c_n} A_{ij} \quad (3)$$

where  $c_m$  is the set of individuals in cluster  $m$ ,  $N_m$  is the number of individuals in cluster  $m$ , and  $A_{ij}$  is the GRM between individuals  $i$  and  $j$  and defined as<sup>17</sup>:

$$A_{ij} = \frac{1}{M} \sum_{k=1}^M \frac{(G_{ik} - 2f_k)(G_{jk} - 2f_k)}{2f_k(1-f_k)} \quad (4)$$

Here  $G_{ik}$  is the genotype for SNP  $k$  of individual  $i$ ,  $f_k$  is the allele frequency of SNP  $k$  across all individuals and  $M$  is the number of SNPs. We created a GRM using all autosomal SNPs passing quality control with MAF  $> 0$  across all individuals, and applied a leave-one-chromosome-out scheme to prevent proximal contamination<sup>66,67</sup>, creating a separate GRM for each chromosome.

## PGS computation

The PGS that we analysed is a weighted average of genotypes for  $M$  independent variants:

$$\text{PGS}_i = \sum_{j=1}^M w_j G_{ij} \quad (5)$$

Here,  $G_{ij}$  is the genotype for SNP  $j$  of individual  $i$  and  $w_j$  is the SNP weight. We generated four variations of the PGS by using the GWAS effect values ( $\beta_j$ ) or only the sign of the effects ( $\text{sign}(\beta_j)$ ) as weights, and by including or excluding the HLA region (Supplementary Table 5; the results quoted in the main text exclude the HLA region). For each phenotype, we generated an independent set of SNPs using a two-step clumping and thresholding process. To avoid ascertainment bias<sup>68</sup>, SNP selection and effect size estimation for PGS calculation were both derived from the same GWAS, independent of other GWAS or selection signals. Initially, we clumped SNPs with PLINK using a  $P < 10^{-3}$ ,  $r^2 < 0.05$  and a 500-kb window. Then, we selected the SNP with the smallest  $P$  value as the index SNP, removed SNPs with  $D' > 0.2$  within 500 kb, and repeated until no SNP remained. Consequently, all remaining SNPs had  $P < 0.001$ , and if two SNPs were within 500 kb, their  $r^2 < 0.05$  and  $D' < 0.2$ . To minimize residual population structure, we used a linear mixed model (LMM):

$$y_i = \alpha + t_i \gamma + g_i + e_i \quad (6)$$

Here  $y_i$  is the PGS of the sample  $i$ , centred and scaled by the mean and standard deviation of the PGS in the modern samples to make it

interpretable as the strength of polygenic selection and comparable across different traits<sup>21</sup>;  $t_i$  is the date scaled down by  $-10,000$  (so it is in units of ten millennia);  $\alpha$  is the intercept;  $\mathbf{g} \sim \text{MVN}(0, \sigma_g^2 \mathbf{K})$  is a vector of random effects where the covariance structure matrix  $\mathbf{K}$  is the genetic relationship matrix; and  $\mathbf{e} = \text{MVN}(0, \sigma_e^2 \mathbf{I})$  is a vector of residual errors where  $\mathbf{I}$  is the identity matrix. The coefficient  $\gamma$  is the change in the PGS over 10,000 years in units of standard deviation of the PGS in the modern samples. We used the coefficient  $\gamma$  as a proxy for directional polygenic selection.

## Fitting the LMM

We used GEMMA (v0.98.5)<sup>69</sup> to fit the LMM and estimate the polygenic selection coefficient ( $\gamma$ ). The running time was tractable, so we did not apply the clustering algorithm used in the GLMM analysis, and also included related individuals allowing the sample size of ancient individuals to increase from 13,936 for our single-allele tests to 15,836 to polygenic tests. We used the GRM as the covariance structure matrix  $\mathbf{K}$ . Here PGS is calculated over all autosomes, and so we could not use the leave-one-chromosome-out approach from single-variant GLMM to avoid influence from neighbouring positions in linkage disequilibrium. Instead, we used 112,683 high-quality, independent SNPs generated by the 'indep-pairwise 1000 1 0.05' option of PLINK2 to calculate a GRM, using this as a covariance structure in the LMM to handle population structure and reduce influence from sites whose alleles are correlated to each other.

## Adjusting for residual inflation in directional polygenic analysis

To adjust for residual inflation in the estimated  $\mathbf{Z}_i$  for each trait, we carried out 100 randomizations for each trait of interest, using the same SNPs used for calculating the PGS of that trait and randomly assigning a weight of  $+1$  or  $-1$  to each SNP for each simulation. The simulated PGS is not expected to show a signal of polygenic selection, as the weights are randomly flipped and should cancel for polygenic traits. To carry out a valid test for a signal of directional selection, we defined an inflation factor by calculating the ratio of the median  $\mathbf{Z}_i^2$  in the simulation to the median of the chi-squared distribution with 1 d.f. (0.455). This correction factor is constrained to be at least 1 and at most the across-trait median of 3.16.

## Correlation between trait effect sizes and selection coefficients

We used LDSC (v1.0.1)<sup>40,49,70</sup> to estimate the genetic correlation between trait effect sizes and selection coefficients ( $s$ ). We used the pre-calculated linkage disequilibrium scores computed with individuals of European ancestry from the 1000 Genomes Project, which are provided with the LDSC software. To compute trans-ethnic genetic correlation, we used the S-LDXR software<sup>71</sup>. We used the pre-calculated reference files for European and East Asian populations that are provided with this software.

## Heritability enrichment and standardized effect size ( $\tau^*$ )

We used stratified LDSC<sup>49</sup> to estimate the contribution of each annotation to the heritability of polygenic traits. We combined the set of annotations of interest with the baseline-LD model (v2.2), which includes 97 annotations modelling MAF, linkage disequilibrium and functional architectures including coding regions, promoters, enhancers and conserved elements<sup>49,72,73</sup>. We used heritability enrichment to quantify the effects of the annotation; it is defined as the proportion of heritability explained by SNPs in the annotation divided by the proportion of SNPs in the annotation. The standardized effect size ( $\tau^*$ ) measures the effects unique to the focal annotation after conditioning on all the other annotations in the baseline-LD model<sup>72</sup>.

## Simulations of selection for a realistic model of history

We carried out forward-in-time simulations using SLiM (v4.0.1)<sup>74</sup> to model European history, starting with the semi-realistic demographic

model of Irving-Pease et al.<sup>4</sup>. We divided the genome into coding, functional non-coding and neutral regions. We simulated both purifying selection (selection against newly arising deleterious alleles) and stabilizing selection (selection against variation), as these are two major processes known to shape patterns of human genetic variation and are not forms of directional selection (defined for our purposes as sustained rises in frequency of alleles in a direction that increases the fitness in carriers). We were concerned that our test could artefactually be detecting signals at alleles affected by these non-directional selection processes, and to explore this, we simulated various scenarios. Each model included an experimental condition with both directional and non-directional selection processes, and a corresponding negative control that differed only by the absence of directional selection. Detailed information on the simulations and our analysis of them are provided in Supplementary Section 2.

### GRM-matched simulations based on real data

For the purposes of Fig. 2b and Extended Data Fig. 1c where we wished to simulate the genotypes of individuals for a variant with a selection coefficient  $s_j$ , we used a random sample drawn from a Gaussian distribution with a covariance matrix of  $\sigma_j^2 \mathbf{K}$ . We estimated the genetic relationship matrix  $\mathbf{A}$  using real data, and randomly selected  $\sigma_j^2$  from an empirical distribution. We used equation (1) to simulate different selection coefficients and determined the initial allele frequency by drawing from an empirical distribution of allele frequencies in modern individuals. We used this value as a constraint to define the constant  $\alpha_j$ . To sample genotypes, we drew from a binomial distribution, with the probability of the alternative allele calculated using the standard logistic function applied to both sides of equation (1). All other simulations that we report in this study are based on the SLiM framework.

### Sources of data for 15,836 ancient individuals

We restricted to 15,836 ancient individuals of genetically West Eurasian-related ancestry living between longitude 50° W and 120° E and latitude greater than 24° N (Supplementary Table 1 and Supplementary Section 1). For 5,820 ancient individuals, the sequences that we analysed are published<sup>2,15,31,37,48,75-246</sup> and are reanalysed here. For 296 ancient individuals, we report new shotgun sequencing data from samples for which in-solution enrichment data from the same ancient DNA samples, extracts or libraries were previously published. The present study serves as the formal report of these new sequences, and reanalysis of the data presented here should cite both the present study and the study that originally reported data from these samples. Supplementary Table 2 lists these samples along with newly reported shotgun data for an additional 71 anonymized newly reported individuals (for a total of 367 newly reported shotgun genomes, which have a median of 4.89-fold coverage and of which 43 have at least 20-fold coverage).

For 223 ancient individuals, we report higher-coverage in-solution enrichment data based on additional extracts, libraries and sometimes recaptures of libraries for which smaller amounts of in-solution enrichment data from the same samples were previously published, obtained by adding data from 517 newly reported ancient DNA libraries (Supplementary Table 3). The present study serves as the formal report of these merges of previously published data with the newly generated data. Reanalysis of the data presented here should cite both the present study and the study that originally reported data from these individuals.

For 9,426 never-before-reported ancient individuals obtained by in-solution enrichment of 17,880 newly reported ancient DNA libraries (Supplementary Table 3) and shotgun sequencing of an additional 71 newly reported ancient DNA libraries (Supplementary Table 2), we release raw ancient DNA data with permission of sample custodians. The individuals are anonymized, with the only information provided being that used in our analysis: the full genetic data, and point estimates of their dates and broad geographical categorization into five regions of West Eurasia.

### Sources of data for contemporary individuals

We analysed data from 6,438 contemporary individuals, comprising 5,935 from the UK Biobank<sup>247</sup> and 503 from the 1000 Genomes Project<sup>248</sup>.

For the UK Biobank data, we selected individuals genotyped on the UK Biobank Axiom array, excluding those genotyped on the UK BiLEVE array to minimize batch effects. To ensure broad representation across western Eurasia, we subsampled the UK Biobank, limiting the selection to at most 300 people per 'country of birth' within western Eurasia, focusing on countries with ancient DNA in this study. This yielded 6,088 individuals.

We then performed an outlier removal step. Using their coordinates on the top 20 principal components derived from the full UK Biobank dataset, we standardized the scores and calculated Mahalanobis distances. Assuming the squared Mahalanobis distance follows a chi-squared distribution with 20 d.f., we excluded individuals with  $P$  values below the Bonferroni threshold of  $8.2 \times 10^{-6}$ . This outlier removal step yielded a final set of 5,935 individuals from 58 countries, with a median of 55 and a mean of 102 per country.

### Ancient DNA data generation

The great majority of wet laboratory work was performed in the ancient DNA laboratory at Harvard Medical School, following established protocols that evolved over time from mostly manual processing (sample preparation, DNA extraction with silica columns<sup>249,250</sup> and partial UDG-treated double-stranded library preparation<sup>251,252</sup>; capture was automated using a Perkin Elmer EP3 or Agilent Bravo NGS Workstations<sup>2,188,253</sup>) to mostly automated processing (DNA extraction<sup>254</sup>, double-stranded and single-stranded library preparation<sup>255</sup>, capture and pooling for sequencing). New libraries (if not deeply shotgun sequenced) were enriched with the Twist Ancient DNA panel<sup>242</sup>, whereas older libraries were enriched with the 1240k reagent (or its predecessors, 390k and 840k). We sequenced on an Illumina NextSeq500 instrument until 2019, when we switched to Illumina HiSeq X10 instruments, and finally to Illumina NovaSeq X instruments in 2022. Archaeologists or collaborators from other ancient DNA laboratories in some cases provided sample powder, DNA extracts or libraries. Supplementary Table 3 provides summary statistics based on in-solution enrichment for 18,397 ancient DNA libraries for which we report new capture data. Supplementary Table 2 provides summary statistics for 369 ancient DNA libraries from 367 individuals for which we report new shotgun sequencing data.

### Ancient DNA bioinformatic processing

Most of the newly reported data come from sequencing the products of in-solution enrichment targeting a set of more than 1 million known polymorphisms<sup>188,242</sup>. In-solution enrichment extracts more information per invested sequence by enriching molecules to overlap sites that are polymorphic in humans (which also helps to greatly reduce the proportion of non-endogenous bacterial and/or microbial sources that colonized the samples post-mortem). The great majority of ancient DNA libraries that we analysed are marked with identification tags (barcodes and indices) before sequencing in pools. We merged paired-end sequences, requiring that there is no more than one mismatch in the overlap between paired sequences where the base quality is at least 20 or three mismatches if the base quality is less than 20. We did not analyse sequences that we could not merge. We stripped adapters and identification tags to prepare molecules for alignment. We used a custom toolkit (ADNA-Tools v2.1.0) for all these steps. We aligned merged sequences to the hg19 version of the human reference genome with decoy sequences (hs37d5) using the single-ended aligner, BWA SAMSE (v0.7.15)<sup>256</sup> with typical ancient DNA alignment parameters -n 0.01 -o 2 and -l 16500, which disables pre-alignment seeding. We marked duplicate reads using Picard MarkDuplicates (v2.17.10)<sup>257</sup>. In addition, we mapped merged sequences to the mitochondrial DNA Reconstructed

# Article

Sapiens Reference Sequence<sup>258</sup>, which enables mitochondrial-specific metrics. Our bioinformatic processing produces data and key metrics, including estimates of authenticity based on elevated damage rates at the end of sequences (indicative of ancient DNA), contamination rates and endogenous rates. For the shotgun sequencing of a subset of libraries with a high proportion of human DNA to generate coverage throughout the genome, we processed data in the same way as for the in-solution enrichment experiments.

## Imputation

To carry out imputation, we used as input either data from ancient individuals (mapped sequences) or modern individuals (SNP array genotypes), and then used allelic correlation patterns in a haplotype reference panel<sup>248,259</sup> to predict genotypes at millions of sites.

In detail, for each sample, we used `bcftools mpileup` (v1.13)<sup>260</sup> to generate genotype likelihoods for all variants (SNPs and indels). We used the high-coverage (30×) 1000 Genomes Project<sup>248</sup> phase 3 sequences as the reference panel and converted the assembly version to GRCh37/hg19 using `CrossMap` (v0.5.2)<sup>261</sup>. We restricted to 2,504 unrelated samples and biallelic variants that pass all the quality control filters reported by `gnomAD` (v2.1.1)<sup>262</sup>. We used `GLIMPSE` (v1.0.0)<sup>16</sup> with the reference panel to impute and phase each sample individually. Owing to higher reference bias for indels, we ignored their genotype likelihood, set them to missing, and passed this to `GLIMPSE` to impute all biallelic autosomal SNPs and indels based on the genotype likelihood of SNPs and haplotype information for both SNPs and indels in the reference panel. This means that we only used reference panel data to impute indels even when we had sequences overlapping the indels. After imputation, we added the genotype caller information of all variants (SNPs and indels) to the final `bcf` file.

To minimize discrepancies between imputation of ancient DNA and UK Biobank data, we re-imputed the UK Biobank genotyping data. We used `Affymetrix` confidence files to simulate genotype likelihoods and processed these through the same imputation pipeline used for ancient DNA.

## Sample quality control

For each imputed sample, we defined the imputation quality score (IQS) as  $IQS = \text{mean}(GP_i | GT = 1)$ , where  $GT$  is the most likely genotype based on the imputed genotype posterior  $GP = (GP_0, GP_1, GP_2)$  and  $\sum_{i=0}^2 GP_i = 1$ . We only kept samples with high  $IQS > 0.9$ . We detected duplicates and related samples up to the second degree. We prioritized samples by their  $IQS$  and dropped relatives up to the second degree until there were no two that were second-degree related or closer. We also fit a linear regression model to the top 100 principal components as explanatory variables, and used the reported date of samples as the response variable to remove outliers where reported and predicted dates were very different. Full details of this procedure are presented in Supplementary Section 1.

## Variant quality control

The data analysed in this study come from multiple sources and sequencing technologies: imputed ancient DNA sequences (shotgun sequences and enrichment for more than 1 million SNPs), European ancestry individuals largely from the 1000 Genomes Project, and imputed individuals of western Eurasian ancestry from the UK Biobank genotyped using the UK Biobank Axiom Array. We applied variant quality control in a three-step procedure. Initially, we applied brute-force filtering to compute principal components, allowing for the identification of ancestry-matching samples across datasets with similar allele frequencies. We filtered out variants if their allele frequencies differed strongly between sample sets, with the goal of minimizing batch effects from combining samples from different sources. Finally, we excluded variants whose estimated selection coefficients were sensitive to data batches. This resulted in 9,739,624 variants, including 8,074,573 SNPs

and 1,665,051 indels, passing the final variant quality control out of 52,382,872 imputed variants. The step-by-step variant quality control process is detailed in Supplementary Section 1.

## Allele frequency trajectories

We computed allele frequency trajectories using all individuals in the time series. We used a moving average sliding window, with a window size of 1,000 years and a step size of 100. We used a binomial likelihood function to estimate the mean, CIs and standard error. We smoothed the mean and standard error using the `GaussianProcessRegressor` function from the `Scikit-learn` library in Python. We parameterized this function with  $\alpha = 10^{-4}$  and a `RationalQuadratic` kernel, with `length_scale_bounds` set to (10,  $10^6$ ). We clipped the resulting values to remain within the range of 0 and 1.

## GWAS data to which we correlated selection coefficients

We analysed GWAS of 452 largely quantitative phenotypes from the UK Biobank<sup>263</sup> that had the flag phenotype\_qc\_EUR set as PASS (the high-quality subset of 6,951 UK Biobank GWAS). We added to this 107 independent GWAS studies<sup>264,265</sup> mostly of complex diseases, alongside 2 brain volume GWAS from ref. 56 and 2 GWAS from ref. 55 for cognitive and non-cognitive aspects of educational attainment. These 563 = 452 + 107 + 2 + 2 GWAS of largely unrelated people of European ancestry are the ones that we refer to in the main text when discussing our primary scans for polygenic selection. For the trans-ethnic analysis, we analysed an additional 31 GWAS in East Asian individuals: 30 phenotypes from the Biobank of Japan<sup>266</sup> and the GWAS summary statistics from the study of years of schooling GWAS by ref. 267. For family-based analysis, we analysed 102 GWAS for 34 traits compiled in ref. 54; for each trait, there were three GWAS: unrelated people, direct effect and indirect effect. Summary statistics for all our polygenic tests of selection for all of these 696 = 563 + 30 + 1 + 102 GWAS are presented in Supplementary Table 5.

## Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

## Data availability

Aligned sequences for the newly reported data from 10,016 ancient individuals are available through the European Nucleotide Archive under an accession number PRJEB106907. This includes complete genetic data for all newly reported individuals alongside point estimates of their dates and geographical categorization into five regions of West Eurasia, which is the full data used in the present study (Supplementary Tables 1–3). Our release of these data without restrictions to enable studies of natural selection has written approval from third-party sample custodians. Please contact the corresponding author (D.R.) for any questions regarding metadata not used in this study—such as skeletal codes, latitudes and longitudes, site names, site descriptions, physical anthropology and cultural context—which will be reported in future work that should be the references for studies of population history and archaeology (as a backup in case these studies are not published in reasonable time, we have deposited a version of the Supplementary Tables with the full archaeological metadata for all these individuals to Harvard Dataverse (<https://doi.org/10.7910/DVN/7RUV9N>) with a 15-year embargo so that it will go public in the year 2041). Imputed genomes for ancient individuals are also available at the Harvard Dataverse (<https://doi.org/10.7910/DVN/7RUV9N>). High-coverage (30×) phase 3 sequences from the 1000 Genomes Project were downloaded from `ftp.1000genomes.ebi.ac.uk/vol1/ftp/data_collections/1000G_2504_high_coverage/working/20201028_3202_phased`. Genome coordinates were converted to GRCh37/hg19 using `CrossMap` (v0.5.2). The processed 1000 Genomes data used as the imputation reference panel

and for downstream analyses are publicly available at the Harvard Dataverse (<https://doi.org/10.7910/DVN/7RVV9N>). Individual-level data from the UK Biobank used in this study were obtained from the UK Biobank Resource and are available to eligible researchers upon application through the UK Biobank Access Management System, in accordance with UK Biobank policies.

## Code availability

An interactive web application for this study is available (<http://reich-ages.rc.hms.harvard.edu>). The PQLseqPy software is available on GitHub (<https://github.com/mokar2001/PQLseqPy>). A permanently archived version of this package is available at Harvard Dataverse (<https://doi.org/10.7910/DVN/7RVV9N>). All code for the forward-in-time SLiM simulations are available on GitHub (<https://github.com/AnnabelPerry/slim-selection-simulations>).

61. Yang, J. et al. Genetic signatures of high-altitude adaptation in Tibetans. *Proc. Natl Acad. Sci. USA* **114**, 4189–4194 (2017).
62. Yang, J., Zaitlen, N. A., Goddard, M. E., Visscher, P. M. & Price, A. L. Advantages and pitfalls in the application of mixed-model association methods. *Nat. Genet.* **46**, 100–106 (2014).
63. Taus, T., Futschik, A. & Schlötterer, C. Quantifying selection with pool-seq time series data. *Mol. Biol. Evol.* **34**, 3023–3034 (2017).
64. Crow, J. F. *An Introduction to Population Genetics Theory* (Scientific Publishers, 2017).
65. Sun, S. et al. Heritability estimation and differential analysis of count data with generalized linear mixed models in genomic sequencing studies. *Bioinformatics* **35**, 487–496 (2019).
66. Loh, P.-R. et al. Efficient Bayesian mixed model analysis increases association power in large cohorts. *Nat. Genet.* **47**, 284–290 (2015).
67. Listgarten, J. et al. Improved linear mixed models for genome-wide association studies. *Nat. Methods* **9**, 525–526 (2012).
68. Zaidi, A. A. & Mathieson, I. Demographic history mediates the effect of stratification on polygenic scores. *eLife* **9**, e61548 (2020).
69. Zhou, X. & Stephens, M. Genome-wide efficient mixed-model analysis for association studies. *Nat. Genet.* **44**, 821–824 (2012).
70. Bulik-Sullivan, B. K. et al. LD score regression distinguishes confounding from polygenicity in genome-wide association studies. *Nat. Genet.* **47**, 291–295 (2015).
71. Shi, H. et al. Population-specific causal disease effect sizes in functionally important regions impacted by selection. *Nat. Commun.* **12**, 1098 (2021).
72. Gazal, S. et al. Linkage disequilibrium-dependent architecture of human complex traits shows action of negative selection. *Nat. Genet.* **49**, 1421–1427 (2017).
73. Gazal, S., Marquez-Luna, C., Finucane, H. K. & Price, A. L. Reconciling S-LDSC and LDKF functional enrichment estimates. *Nat. Genet.* **51**, 1202–1204 (2019).
74. Haller, B. C. & Messer, P. W. SLiM 4: multispecies eco-evolutionary modeling. *Am. Nat.* **201**, E127–E139 (2023).
75. Damgaard, P. D. B. et al. 137 Ancient human genomes from across the Eurasian steppes. *Nature* **557**, 369–374 (2018).
76. Jensen, T. Z. T. et al. A 5700 year-old human genome and oral microbiome from chewed birch pitch. *Nat. Commun.* **10**, 5520 (2019).
77. Ullinger, J. et al. A bioarchaeological investigation of fraternal stillborn twins from Tell el-Hesi. *East. Archaeol.* **85**, 228–237 (2022).
78. Olalde, I. et al. A common genetic origin for early farmers from Mediterranean Cardial and Central European LBK cultures. *Mol. Biol. Evol.* <https://doi.org/10.1093/molbev/msv181> (2015).
79. Cassidy, L. M. et al. A dynastic elite in monumental Neolithic society. *Nature* **582**, 384–388 (2020).
80. Moots, H. M. et al. A genetic history of continuity and mobility in the Iron Age central Mediterranean. *Nat. Ecol. Evol.* **7**, 1515–1524 (2023).
81. Olalde, I. et al. A genetic history of the Balkans from Roman frontier to Slavic migrations. *Cell* **186**, 5472–5485.e9 (2023).
82. Haber, M. et al. A genetic history of the Near East from an aDNA time course sampling eight points in the past 4,000 years. *Am. J. Hum. Genet.* **107**, 149–157 (2020).
83. Nikitin, A. G. et al. A genomic history of the North Pontic region from the Neolithic to the Bronze Age. *Nature* **639**, 124–131 (2025).
84. Fernandes, D. M. et al. A genomic Neolithic time transect of hunter-farmer admixture in central Poland. *Sci. Rep.* **8**, 14879 (2018).
85. Altınışık, N. E. et al. A genomic snapshot of demographic and cultural dynamism in Upper Mesopotamia during the Neolithic transition. *Sci. Adv.* **8**, eabo3609 (2022).
86. Fowler, C. et al. A high-resolution picture of kinship practices in an Early Neolithic tomb. *Nature* **601**, 584–587 (2022).
87. van den Brink, E. C. M. et al. A Late Bronze Age II clay coffin from Tel Shaddud in the Central Jezreel Valley, Israel: context and historical implications. *Levant* **49**, 105–135 (2017).
88. Harney, É et al. A minimally destructive protocol for DNA extraction from ancient teeth. *Genome Res.* **31**, 472–483 (2021).
89. Haber, M. et al. A transient pulse of genetic admixture from the Crusaders in the Near East identified from ancient genome sequences. *Am. J. Hum. Genet.* **104**, 977–984 (2019).
90. Wohns, A. W. et al. A unified genealogy of modern and ancient genomes. *Science* **375**, eabi8264 (2022).
91. González-Fortes, G. et al. A western route of prehistoric human migration from Africa into the Iberian Peninsula. *Proc. R. Soc. B* **286**, 20182288 (2019).
92. Unterländer, M. et al. Ancestry and demography and descendants of Iron Age nomads of the Eurasian Steppe. *Nat. Commun.* **8**, 14615 (2017).
93. Dulas, K. et al. Ancient DNA at the edge of the world: continental immigration and the persistence of Neolithic male lineages in Bronze Age Orkney. *Proc. Natl Acad. Sci. USA* **119**, e2108001119 (2022).
94. Harney, É et al. Ancient DNA from Chalcolithic Israel reveals the role of population mixture in cultural transformation. *Nat. Commun.* **9**, 3336 (2018).
95. Zalloua, P. et al. Ancient DNA of Phoenician remains indicates discontinuity in the settlement history of Ibiza. *Sci. Rep.* **8**, 17567 (2018).
96. Skourtanioti, E. et al. Ancient DNA reveals admixture history and endogamy in the prehistoric Aegean. *Nat. Ecol. Evol.* **7**, 290–303 (2023).
97. Szécsényi-Nagy, A. et al. Ancient DNA reveals diverse community organizations in the 5th millennium BCE Carpathian Basin. *Nat. Commun.* **16**, 5318 (2025).
98. Wang, K. et al. Ancient DNA reveals reproductive barrier despite shared Avar-period culture. *Nature* **638**, 1007–1014 (2025).
99. Zeng, T. C. et al. Ancient DNA reveals the prehistory of the Uralic and Yeniseian peoples. *Nature* **644**, 122–132 (2025).
100. Arzelier, A. et al. Ancient DNA sheds light on the funerary practices of Late Neolithic collective burial in southern France. *Proc. R. Soc. B* **291**, rspb.2024.1215 (2024).
101. Lamnidis, T. C. et al. Ancient Fennoscandian genomes reveal origin and spread of Siberian ancestry in Europe. *Nat. Commun.* **9**, 5018 (2018).
102. O’Sullivan, N. et al. Ancient genome-wide analyses infer kinship structure in an early medieval Alemannic graveyard. *Sci. Adv.* **4**, eaao1262 (2018).
103. Rivollat, M. et al. Ancient genome-wide DNA from France highlights the complexity of interactions between Mesolithic hunter-gatherers and Neolithic farmers. *Sci. Adv.* **6**, eaaz5344 (2020).
104. Ebenesersdóttir, S. S. et al. Ancient genomes from Iceland reveal the making of a human population. *Science* **360**, 1028–1032 (2018).
105. Fregel, R. et al. Ancient genomes from North Africa evidence prehistoric migrations to the Maghreb from both the Levant and Europe. *Proc. Natl Acad. Sci. USA* **115**, 6774–6779 (2018).
106. Brunel, S. et al. Ancient genomes from present-day France unveil 7,000 years of its demographic history. *Proc. Natl Acad. Sci. USA* **117**, 12791–12798 (2020).
107. Martiniano, R. et al. Ancient genomes illuminate eastern Arabian population history and adaptation against malaria. *Cell Genom.* **4**, 100507 (2024).
108. Brace, S. et al. Ancient genomes indicate population replacement in Early Neolithic Britain. *Nat. Ecol. Evol.* **3**, 765–771 (2019).
109. Günther, T. et al. Ancient genomes link early farmers from Atapuerca in Spain to modern-day Basques. *Proc. Natl Acad. Sci. USA* **112**, 11917–11922 (2015).
110. Žegarac, A. et al. Ancient genomes provide insights into family structure and the heredity of social status in the Early Bronze Age of southeastern Europe. *Sci. Rep.* **11**, 10072 (2021).
111. Gerber, D. et al. Ancient genomes reveal Avar-Hungarian transformations in the 9th–10th centuries CE Carpathian Basin. *Sci. Adv.* **10**, eadq5864 (2024).
112. Foody, M. G. B. et al. Ancient genomes reveal cosmopolitan ancestry and maternal kinship patterns at post-Roman Worth Matravers, Dorset. *Antiquity* **99**, 1356–1371 (2025).
113. Gneocchi-Ruscione, G. A. et al. Ancient genomes reveal origin and rapid trans-Eurasian migration of 7th century Avar elites. *Cell* **185**, 1402–1413.e21 (2022).
114. Saupe, T. et al. Ancient genomes reveal structural shifts after the arrival of steppe-related ancestry in the Italian Peninsula. *Curr. Biol.* **31**, 2576–2591.e12 (2021).
115. Krzewińska, M. et al. Ancient genomes suggest the eastern Pontic-Caspian steppe as the source of western Iron Age nomads. *Sci. Adv.* **4**, eaat4457 (2018).
116. Gneocchi-Ruscione, G. A. et al. Ancient genomic time transect from the Central Asian Steppe unravels the history of the Scythians. *Sci. Adv.* **7**, eabe4414 (2021).
117. Wang, C.-C. et al. Ancient human genome-wide data from a 3000-year interval in the Caucasus corresponds with eco-geographic regions. *Nat. Commun.* **10**, 590 (2019).
118. Lazaridis, I. et al. Ancient human genomes suggest three ancestral populations for present-day Europeans. *Nature* **513**, 409–413 (2014).
119. Ariano, B. et al. Ancient Maltese genomes and the genetic geography of Neolithic Europe. *Curr. Biol.* **32**, 2668–2680.e6 (2022).
120. Michel, M. et al. Ancient *Plasmodium* genomes shed light on the history of human malaria. *Nature* **631**, 125–133 (2024).
121. Antonio, M. L. et al. Ancient Rome: a genetic crossroads of Europe and the Mediterranean. *Science* **366**, 708–714 (2019).
122. Veselka, B. et al. Assembling ancestors: the manipulation of Neolithic and Gallo-Roman skeletal remains at Pommerœul, Belgium. *Antiquity* **98**, 1576–1591 (2024).
123. Scorrano, G. et al. Bioarchaeological and palaeogenomic portrait of two Pompeians that died during the eruption of Vesuvius in 79 AD. *Sci. Rep.* **12**, 6468 (2022).
124. Srigyan, M. et al. Bioarchaeological evidence of one of the earliest Islamic burials in the Levant. *Commun. Biol.* **5**, 554 (2022).
125. Zagorc, B. et al. Bioarchaeological perspectives on Late Antiquity in Dalmatia: paleogenetic, dietary, and population studies of the Hvar—Radošević burial site. *Archaeol. Anthropol. Sci.* **16**, 150 (2024).
126. Bagnasco, G. et al. Bioarchaeology aids the cultural understanding of six characters in search of their agency (Tarquinia, ninth–seventh century BC, central Italy). *Sci. Rep.* **14**, 11895 (2024).
127. Furtwängler, A. et al. Comparison of target enrichment strategies for ancient pathogen DNA. *Biotechniques* **69**, 455–459 (2020).
128. Cassidy, L. M. et al. Continental influx and pervasive matrilocality in Iron Age Britain. *Nature* **637**, 1136–1142 (2025).
129. Haber, M. et al. Continuity and admixture in the last five millennia of Levantine history from Ancient Canaanite and present-day Lebanese genome sequences. *Am. J. Hum. Genet.* **101**, 274–282 (2017).
130. Linderholm, A. et al. Corded Ware cultural complexity uncovered using genomic and isotopic analysis from south-eastern Poland. *Sci. Rep.* **10**, 6885 (2020).
131. Olalde, I. et al. Derived immune and ancestral pigmentation alleles in a 7,000-year-old Mesolithic European. *Nature* **507**, 225–228 (2014).
132. Blöcher, J. et al. Descent, marriage, and residence practices of a 3,800-year-old pastoral community in Central Eurasia. *Proc. Natl Acad. Sci. USA* **120**, e2303574120 (2023).

133. Gokhman, D. et al. Differential DNA methylation of vocal and facial anatomy genes in modern humans. *Nat. Commun.* **11**, 1189 (2020).
134. Papac, L. et al. Dynamic changes in genomic and social structures in third millennium BCE central Europe. *Sci. Adv.* **7**, eabi6941 (2021).
135. Penske, S. et al. Early contact between late farming and pastoralist societies in southeastern Europe. *Nature* **620**, 358–365 (2023).
136. Hofmanová, Z. et al. Early farmers from across Europe directly descended from Neolithic Aegeans. *Proc. Natl Acad. Sci. USA* **113**, 6886–6891 (2016).
137. Moreno-Mayar, J. V. et al. Early human dispersals within the Americas. *Science* **362**, eaav2621 (2018).
138. Broushaki, F. et al. Early Neolithic genomes from the eastern Fertile Crescent. *Science* **353**, 499–503 (2016).
139. Scheib, C. L. et al. East Anglian Early Neolithic monument burial linked to contemporary Megaliths. *Ann. Hum. Biol.* **46**, 145–149 (2019).
140. Gretzinger, J. et al. Evidence for dynastic succession among early Celtic elites in Central Europe. *Nat. Hum. Behav.* **8**, 1467–1480 (2024).
141. Saag, L. et al. Extensive farming in Estonia started through a sex-biased migration from the steppe. *Curr. Biol.* **27**, 2185–2193.e6 (2017).
142. Rivollat, M. et al. Extensive pedigrees reveal the social organization of a Neolithic community. *Nature* **620**, 600–606 (2023).
143. Nedoluzhko, A. et al. First ancient DNA analysis of mummies from the post-Scythian Oglakhty cemetery in South Siberia. Preprint at *Research Square* <https://doi.org/10.21203/rs.3.rs-1993191/v1> (2022).
144. Valdiosera, C. et al. Four millennia of Iberian biomolecular prehistory illustrate the impact of prehistoric migrations at the far end of Eurasia. *Proc. Natl Acad. Sci. USA* **115**, 3428–3433 (2018).
145. Immel, A. et al. Gene-flow from steppe individuals into Cucuteni-Trypillia associated populations indicates long-standing contacts and gradual admixture. *Sci. Rep.* **10**, 4253 (2020).
146. Peltola, S. et al. Genetic admixture and language shift in the medieval Volga-Oka interfluve. *Curr. Biol.* **33**, 174–182.e10 (2023).
147. Saag, L. et al. Genetic ancestry changes in Stone to Bronze Age transition in the East European plain. *Sci. Adv.* **7**, eabd6535 (2021).
148. Kumar, V. et al. Genetic continuity of Bronze Age ancestry with increased steppe-related ancestry in Late Iron Age Uzbekistan. *Mol. Biol. Evol.* **38**, 4908–4917 (2021).
149. Mattila, T. M. et al. Genetic continuity, isolation, and gene flow in Stone Age Central and Eastern Europe. *Commun. Biol.* **6**, 793 (2023).
150. Marcus, J. H. et al. Genetic history from the Middle Neolithic to present on the Mediterranean island of Sardinia. *Nat. Commun.* **11**, 939 (2020).
151. Hui, R. et al. Genetic history of Cambridgeshire before and after the Black Death. *Sci. Adv.* **10**, eadi5903 (2024).
152. Stolarek, I. et al. Genetic history of East-Central Europe in the first millennium CE. *Genome Biol.* **24**, 173 (2023).
153. Lazaridis, I. et al. Genetic origins of the Minoans and Mycenaeans. *Nature* **548**, 214–218 (2017).
154. Gamba, C. et al. Genome flux and stasis in a five millennium transect of European prehistory. *Nat. Commun.* **5**, 5257 (2014).
155. Guarino-Vignon, P. et al. Genome-wide analysis of a collective grave from Mentesh Tepe provides insight into the population structure of Early Neolithic population in the South Caucasus. *Commun. Biol.* **6**, 319 (2023).
156. Novak, M. et al. Genome-wide analysis of nearly all the victims of a 6200 year old massacre. *PLoS ONE* **16**, e0247332 (2021).
157. Waldman, S. et al. Genome-wide data from medieval German Jews show that the Ashkenazi founder event pre-dated the 14th century. *Cell* **185**, 4703–4716.e16 (2022).
158. Immel, A. et al. Genome-wide study of a Neolithic Wartberg grave community reveals distinct HLA variation and hunter-gatherer ancestry. *Commun. Biol.* **4**, 113 (2021).
159. Brace, S. et al. Genomes from a medieval mass burial show Ashkenazi-associated hereditary diseases pre-date the 12th century. *Curr. Biol.* **32**, 4350–4359.e6 (2022).
160. Gelabert, P. et al. Genomes from Verteba cave suggest diversity within the Trypillians in Ukraine. *Sci. Rep.* **12**, 7242 (2022).
161. Begg, T. J. A. et al. Genomic analyses of hair from Ludwig van Beethoven. *Curr. Biol.* **33**, 1431–1447.e22 (2023).
162. Rodríguez-Varela, R. et al. Genomic analyses of pre-European conquest human remains from the Canary Islands reveal close affinity to modern North Africans. *Curr. Biol.* **27**, 3396–3402.e5 (2017).
163. Simões, L. G. et al. Genomic ancestry and social dynamics of the last hunter-gatherers of Atlantic France. *Proc. Natl Acad. Sci. USA* **121**, e2310545121 (2024).
164. Scorrano, G. et al. Genomic ancestry, diet and microbiomes of Upper Palaeolithic hunter-gatherers from San Teodoro cave. *Commun. Biol.* **5**, 1262 (2022).
165. Yu, H. et al. Genomic and dietary discontinuities during the Mesolithic and Neolithic in Sicily. *iScience* **25**, 104244 (2022).
166. Krzewińska, M. et al. Genomic and strontium isotope variation reveal immigration patterns in a Viking age town. *Curr. Biol.* **28**, 2730–2738.e10 (2018).
167. Skoglund, P. et al. Genomic diversity and admixture differs for Stone-Age Scandinavian foragers and farmers. *Science* **344**, 747–750 (2014).
168. Skourtanioti, E. et al. Genomic history of Neolithic to Bronze Age Anatolia, northern Levant, and southern Caucasus. *Cell* **181**, 1158–1175.e28 (2020).
169. Lazaridis, I. et al. Genomic insights into the origin of farming in the ancient Near East. *Nature* **536**, 419–424 (2016).
170. Martiniano, R. et al. Genomic signals of migration and continuity in Britain before the Anglo-Saxons. *Nat. Commun.* **7**, 10326 (2016).
171. Villalba-Mouco, V. et al. Genomic transformation and social organization during the Copper Age–Bronze Age transition in southern Iberia. *Sci. Adv.* **7**, eabi7038 (2021).
172. Seguin-Orlando, A. et al. Heterogeneous hunter-gatherer and steppe-related ancestries in Late Neolithic and Bell Beaker genomes from present-day France. *Curr. Biol.* **31**, 1072–1083.e10 (2021).
173. Wang, K. et al. High-coverage genome of the Tyrolean Iceman reveals unusually high Anatolian farmer ancestry. *Cell Genom.* **3**, 100377 (2023).
174. Alves, I. et al. Human genetic structure in Northwest France provides new insights into West European historical demography. *Nat. Commun.* **15**, 6710 (2024).
175. Ingman, T. et al. Human mobility at Tell Atchana (Alalakh), Hatay, Turkey during the 2nd millennium BC: integration of isotopic and genomic evidence. *PLoS ONE* **16**, e0241883 (2021).
176. Morez, A. et al. Imputed genomes and haplotype-based analyses of the Picts of early medieval Scotland reveal fine-scale relatedness between Iron Age, early medieval and the modern people of the UK. *PLoS Genet.* **19**, e1010360 (2023).
177. Schiffels, S. et al. Iron Age and Anglo-Saxon genomes from East England reveal British migration history. *Nat. Commun.* **7**, 10408 (2016).
178. Wang, X. et al. Isotopic and DNA analyses reveal multiscale PPNB mobility and migration across Southeastern Anatolia and the Southern Levant. *Proc. Natl Acad. Sci. USA* **120**, e2210611120 (2023).
179. Penske, S. et al. Kinship practices at the Early Bronze Age site of Leubingen in Central Germany. *Sci. Rep.* **14**, 3871 (2024).
180. Armit, I. et al. Kinship practices in Early Iron Age South-east Europe: genetic and isotopic analysis of burials from the Dolge njive barrow cemetery, Dolenjska, Slovenia. *Antiquity* **97**, 403–418 (2023).
181. Mittnik, A. et al. Kinship-based social inequality in Bronze Age Europe. *Science* **366**, 731–734 (2019).
182. Patterson, N. et al. Large-scale migration into Britain during the Middle to Late Bronze Age. *Nature* **601**, 588–594 (2022).
183. Feldman, M. et al. Late Pleistocene human genome suggests a local origin for the first farmers of central Anatolia. *Nat. Commun.* **10**, 1218 (2019).
184. Higgins, O. A. et al. Life history and ancestry of the Late Upper Palaeolithic infant from Grotta delle Mura, Italy. *Nat. Commun.* **15**, 8248 (2024).
185. Gyuris, B. et al. Long shared haplotypes identify the Southern Urals as a primary source for the 10th century Hungarians. *Cell* <https://doi.org/10.1016/j.cell.2025.09.002> (2025).
186. Olalde, I. et al. Lasting Lower Rhine–Meuse forager ancestry shaped Bell Beaker expansion. *Nature* <https://doi.org/10.1038/s41586-026-10111-8> (2026).
187. Burger, J. et al. Low prevalence of lactase persistence in Bronze Age Europe indicates ongoing strong selection over the last 3,000 years. *Curr. Biol.* **30**, 4307–4315.e13 (2020).
188. Haak, W. et al. Massive migration from the steppe was a source for Indo-European languages in Europe. *Nature* **522**, 207–211 (2015).
189. Sánchez-Quinto, F. et al. Megalithic tombs in western and northern Neolithic Europe were linked to a kindred society. *Proc. Natl Acad. Sci. USA* **116**, 9469–9474 (2019).
190. Cassidy, L. M. et al. Neolithic and Bronze Age migration to Ireland and establishment of the insular Atlantic genome. *Proc. Natl Acad. Sci. USA* **113**, 368–373 (2016).
191. Arzelier, A. et al. Neolithic genomic data from southern France showcase intensified interactions with hunter-gatherer communities. *iScience* **25**, 105387 (2022).
192. Gnecci-Ruscione, G. A. et al. Network of large pedigrees reveals social practices of Avar communities. *Nature* **629**, 376–383 (2024).
193. Seguin-Orlando, A. et al. No particular genomic features underpin the dramatic economic consequences of 17th century plague epidemics in Italy. *iScience* **24**, 102383 (2021).
194. Saag, L. et al. North Pontic crossroads: mobility in Ukraine from the Bronze Age to the early modern period. *Sci. Adv.* **11**, eadr0695 (2025).
195. Simões, L. G. et al. Northwest African Neolithic initiated by migrants from Iberia and Levant. *Nature* **618**, 550–556 (2023).
196. Fischer, C.-E. et al. Origin and mobility of Iron Age Gaulish groups in present-day France revealed through archaeogenomics. *iScience* **25**, 104094 (2022).
197. Posth, C. et al. Palaeogenomics of Upper Palaeolithic to Neolithic European hunter-gatherers. *Nature* **615**, 117–126 (2023).
198. González-Fortes, G. et al. Paleogenomic evidence for multi-generational mixing between Neolithic farmers and Mesolithic hunter-gatherers in the lower Danube basin. *Curr. Biol.* **27**, 1801–1810.e10 (2017).
199. Lipson, M. et al. Parallel palaeogenomic transects reveal complex genetic history of early European farmers. *Nature* **551**, 368–372 (2017).
200. Chyleński, M. et al. Patrilocal and hunter-gatherer-related ancestry of populations in East-Central Europe during the Middle Bronze Age. *Nat. Commun.* **14**, 4395 (2023).
201. Childebayeva, A. et al. Population genetics and signatures of selection in Early Neolithic European farmers. *Mol. Biol. Evol.* **39**, msac108 (2022).
202. Veeramah, K. R. et al. Population genomic analysis of elongated skulls reveals extensive female-biased immigration in early medieval Bavaria. *Proc. Natl Acad. Sci. USA* **115**, 3494–3499 (2018).
203. Allentoft, M. E. et al. Population genomics of Bronze Age Eurasia. *Nature* **522**, 167–172 (2015).
204. Günther, T. et al. Population genomics of Mesolithic Scandinavia: investigating early postglacial migration routes and high-latitude adaptation. *PLoS Biol.* **16**, e2003703 (2018).
205. Allentoft, M. E. et al. Population genomics of post-glacial western Eurasia. *Nature* **625**, 301–311 (2024).
206. Margaryan, A. et al. Population genomics of the Viking world. *Nature* **585**, 390–396 (2020).
207. Ringbauer, H. et al. Punic people were genetically diverse with almost no Levantine ancestors. *Nature* **643**, 139–147 (2025).
208. Freilich, S. et al. Reconstructing genetic histories and social organisation in Neolithic and Bronze Age Croatia. *Sci. Rep.* **11**, 16729 (2021).
209. Järve, M. et al. Shifts in the genetic landscape of the western Eurasian steppe associated with the beginning and end of the Scythian dominance. *Curr. Biol.* **29**, 2430–2441.e10 (2019).
210. Gelabert, P. et al. Social and genetic diversity among the first farmers of Central Europe. Preprint at *bioRxiv* <https://doi.org/10.1101/2023.07.07.548126> (2023).
211. Gelabert, P. et al. Social and genetic diversity in first farmers of central Europe. *Nat. Hum. Behav.* **9**, 53–64 (2024).

212. Koptekin, D. et al. Spatial and temporal heterogeneity in human mobility patterns in Holocene Southwest Asia and the East Mediterranean. *Curr. Biol.* **33**, 41–57.e15 (2023).
213. Antonio, M. L. et al. Stable population structure in Europe since the Iron Age, despite high mobility. *eLife* **13**, e79714 (2024).
214. Villalba-Mouco, V. et al. Survival of Late Pleistocene hunter-gatherer ancestry in the Iberian Peninsula. *Curr. Biol.* **29**, 1169–1177.e7 (2019).
215. Gretzinger, J. et al. The Anglo-Saxon migration and the formation of the early English gene pool. *Nature* **610**, 112–119 (2022).
216. Saag, L. et al. The arrival of Siberian ancestry connecting the eastern Baltic to Uralic speakers further East. *Curr. Biol.* **29**, 1701–1711.e16 (2019).
217. Olalde, I. et al. The Beaker phenomenon and the genomic transformation of Northwest Europe. *Nature* **555**, 190–196 (2018).
218. Kilinc, G. M. et al. The demographic development of the first farmers in Anatolia. *Curr. Biol.* **26**, 2659–2666 (2016).
219. Reitsma, L. J. et al. The diverse genetic origins of a Classical period Greek army. *Proc. Natl Acad. Sci. USA* **119**, e2205272119 (2022).
220. De Barros Damgaard, P. et al. The first horse herders and the impact of Early Bronze Age steppe expansions into Asia. *Science* **360**, eaar7711 (2018).
221. Narasimhan, V. M. et al. The formation of human populations in South and Central Asia. *Science* **365**, eaat7487 (2019).
222. Fu, Q. et al. The genetic history of Ice Age Europe. *Nature* **534**, 200–205 (2016).
223. Rodríguez-Varela, R. et al. The genetic history of Scandinavia from the Roman Iron Age to the present. *Cell* **186**, 32–46.e19 (2023).
224. Lazaridis, I. et al. The genetic history of the Southern Arc: a bridge between West Asia and Europe. *Science* **377**, eabm4247 (2022).
225. Aneli, S. et al. The genetic origin of Daunians and the Pan-Mediterranean southern Italian Iron Age context. *Mol. Biol. Evol.* **39**, msac014 (2022).
226. Máróti, Z. et al. The genetic origin of Huns, Avars, and conquering Hungarians. *Curr. Biol.* **32**, 2858–2870.e7 (2022).
227. Lazaridis, I. et al. The genetic origin of the Indo-Europeans. *Nature* **639**, 132–142 (2025).
228. Mitnik, A. et al. The genetic prehistory of the Baltic Sea region. *Nat. Commun.* **9**, 442 (2018).
229. Malmström, H. et al. The genomic ancestry of the Scandinavian Battle Axe Culture people and their relation to the broader Corded Ware horizon. *Proc. R. Soc. B Biol. Sci.* **286**, 20191528 (2019).
230. Mathieson, I. et al. The genomic history of southeastern Europe. *Nature* **555**, 197–203 (2018).
231. Clemente, F. et al. The genomic history of the Aegean palatial civilizations. *Cell* **184**, 2565–2586.e21 (2021).
232. Agranat-Tamir, L. et al. The genomic history of the Bronze Age southern Levant. *Cell* **181**, 1146–1157.e11 (2020).
233. Olalde, I. et al. The genomic history of the Iberian Peninsula over the past 8000 years. *Science* **363**, 1230–1234 (2019).
234. Marchi, N. et al. The genomic origins of the world's first farmers. *Cell* **185**, 1842–1859.e18 (2022).
235. Coutinho, A. et al. The Neolithic Pitted Ware culture foragers were culturally but not genetically influenced by the Battle Axe culture herders. *Am. J. Phys. Anthropol.* **172**, 638–649 (2020).
236. Jones, E. R. et al. The Neolithic transition in the Baltic was not driven by admixture with early European farmers. *Curr. Biol.* **27**, 576–582 (2017).
237. Posth, C. et al. The origin and legacy of the Etruscans through a 2000-year archeogenomic time transect. *Sci. Adv.* **7**, eabi7673 (2021).
238. Martiniano, R. et al. The population genomics of archaeological transition in west Iberia: investigation of ancient substructure using imputation and haplotype-based methods. *PLoS Genet.* **13**, e1006852 (2017).
239. Ghalichi, A. et al. The rise and transformation of Bronze Age pastoralists in the Caucasus. *Nature* **635**, 917–925 (2024).
240. Spyrou, M. A. et al. The source of the Black Death in fourteenth-century central Eurasia. *Nature* **606**, 718–724 (2022).
241. Fernandes, D. M. et al. The spread of steppe and Iranian-related ancestry in the islands of the western Mediterranean. *Nat. Ecol. Evol.* **4**, 334–345 (2020).
242. Rohland, N. et al. Three assays for in-solution enrichment of ancient human DNA at more than a million SNPs. *Genome Res.* **32**, 2068–2078 (2022).
243. Amorim, C. E. G. et al. Understanding 6th-century barbarian social organization and migration through paleogenomics. *Nat. Commun.* **9**, 3547 (2018).
244. Schroeder, H. et al. Unraveling ancestry, kinship, and violence in a Late Neolithic mass grave. *Proc. Natl Acad. Sci. USA* **116**, 10705–10710 (2019).
245. Jones, E. R. et al. Upper Palaeolithic genomes reveal deep roots of modern Eurasians. *Nat. Commun.* **6**, 8912 (2015).
246. Beneker, O. et al. Urbanization and genetic homogenization in the medieval Low Countries revealed through a ten-century paleogenomic study of the city of Sint-Truiden. *Genome Biol.* **26**, 127 (2025).
247. Bycroft, C. et al. The UK Biobank resource with deep phenotyping and genomic data. *Nature* **562**, 203–209 (2018).
248. Byrská-Bishop, M. et al. High-coverage whole-genome sequencing of the expanded 1000 Genomes Project cohort including 602 trios. *Cell* **185**, 3426–3440.e19 (2022).
249. Dabney, J. et al. Complete mitochondrial genome sequence of a Middle Pleistocene cave bear reconstructed from ultrashort DNA fragments. *Proc. Natl Acad. Sci. USA* **110**, 15758–15763 (2013).
250. Korlević, P. et al. Reducing microbial and human contamination in DNA extractions from ancient bones and teeth. *Biotechniques* **59**, 87–93 (2015).
251. Rohland, N., Harney, E., Mallick, S., Nordenfelt, S. & Reich, D. Partial uracil-DNA-glycosylase treatment for screening of ancient DNA. *Phil. Trans. R. Soc. Lond. B. Biol. Sci.* **370**, 20130624 (2015).
252. Briggs, A. W. & Heyn, P. Preparation of next-generation sequencing libraries from damaged DNA. *Methods Mol. Biol.* **840**, 143–154 (2012).
253. Fu, Q. et al. DNA analysis of an early modern human from Tianyuan Cave, China. *Proc. Natl Acad. Sci. USA* **110**, 2223–2227 (2013).
254. Rohland, N., Glocke, I., Aximu-Petri, A. & Meyer, M. Extraction of highly degraded DNA from ancient bones, teeth and sediments for high-throughput sequencing. *Nat. Protoc.* **13**, 2447–2461 (2018).
255. Gansauge, M.-T., Aximu-Petri, A., Nagel, S. & Meyer, M. Manual and automated preparation of single-stranded DNA libraries for the sequencing of DNA from ancient biological remains and other sources of highly degraded DNA. *Nat. Protoc.* **15**, 2279–2300 (2020).
256. Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **25**, 1754–1760 (2009).
257. Broad Institute. Picard toolkit. *GitHub* <https://broadinstitute.github.io/picard/> (2019).
258. Behar, D. M. et al. A “Copernican” reassessment of the human mitochondrial DNA tree from its root. *Am. J. Hum. Genet.* **90**, 675–684 (2012).
259. 1000 Genomes Project Consortium et al. A global reference for human genetic variation. *Nature* **526**, 68–74 (2015).
260. Danecek, P. et al. Twelve years of SAMtools and BCFtools. *GigaScience* **10**, giab008 (2021).
261. Zhao, H. et al. CrossMap: a versatile tool for coordinate conversion between genome assemblies. *Bioinformatics* **30**, 1006–1007 (2014).
262. Karczewski, K. J. et al. The mutational constraint spectrum quantified from variation in 141,456 humans. *Nature* **581**, 434–443 (2020).
263. Karczewski, K. J. et al. Pan-UK Biobank genome-wide association analyses enhance discovery and resolution of ancestry-enriched effects. *Nat. Genet.* **57**, 2408–2417 (2025).
264. Kim, A. et al. Inferring causal cell types of human diseases and risk variants from candidate regulatory elements. Preprint at *medRxiv* <https://doi.org/10.1101/2024.05.17.24307556> (2024).
265. Gazal, S. S-LDSC reference files. *Zenodo* <https://doi.org/10.5281/zenodo.10515792> (2024).
266. Sakae, S. et al. A cross-population atlas of genetic associations for 220 human phenotypes. *Nat. Genet.* **53**, 1415–1424 (2021).
267. Chen, T.-T. et al. Shared genetic architectures of educational attainment in East Asian and European populations. *Nat. Hum. Behav.* **8**, 562–575 (2024).
268. Shelton, J. F. et al. Trans-ancestry analysis reveals genetic and nongenetic associations with COVID-19 susceptibility and severity. *Nat. Genet.* **53**, 801–808 (2021).
269. UK Biobank Whole-Genome Sequencing Consortium. Whole-genome sequencing of 490,640 UK Biobank participants. *Nature* **645**, 692–701 (2025).
270. Nakagawa, S. & Schielzeth, H. A general and simple method for obtaining  $R^2$  from generalized linear mixed-effects models. *Methods Ecol. Evol.* **4**, 133–142 (2013).

**Acknowledgements** We thank B. Browning, S. Carmi, E. Koch, M. Lipson, I. Mathieson, V. Narasimhan, B. Neale, S. Rubinacci, G. Sella and S. Sunyaev for discussions during the initial development of this project. Revisions of this manuscript benefited from online comments by S. Gusev and G. Coop; discussions with S. Gusev, A. Young, J. Lee and H. Zeberg following the release of our initial preprint; and early sharing of data in press by O. Beneker and T. Kivisild. We are grateful to I. Lazaridis, H. Li, A. Micco, M. Nawaz, Z. Zhang and M. Zhao for bioinformatic support; and to N. Adamski, R. Bernardos, N. Broomandkshobacht, K. Callan, A. Claxton, O. Cheronet, E. Curtis, M. Ferry, T. Frost, I. Greenslade, E. Harney, L. Iliev, A. Kearns, J. Kellogg, A. M. Lawson, M. Michel, J. Oppenheimer, I. Patterson, S. Nordenfelt, L. Qiu, K. Stewardson, A. Szécsényi-Nagy, J. N. Workman and F. Zalzalá for wet laboratory support. We thank the more than 200 archaeologists and anthropologists who provided explicit permission for release of raw data for individuals with never-before-reported ancient DNA data they shared with us for the purposes of studies of selection, decoupled from the contextual information, which will be presented in papers on which they are co-authors and is necessary for any studies of the population history of these individuals. We specifically acknowledge I. Armit, A. Coppa, W. Haak, C. Laluzea-Fox, B. Llamas and L. Vyazov for facilitating various aspects of this permissions process. This research relied on data from the UK Biobank Resource permitted through Application 16549, and also received funding from the European Research Council under the European Union's Horizon 2020 research and innovation programme under grant agreement 834087 (COMMIO). We acknowledge the Research Computing Group at Harvard Medical School for their support of the computational analyses in this paper and the AGES web application; support from US National Institutes of Health grant HG012287; from the Allen Discovery Center program, a Paul G. Allen Frontiers Group advised program of the Allen Family Philanthropies; from John Templeton Foundation grant 61220; from a private gift by Jean-François Clin; and from the Howard Hughes Medical Institute (HHMI). This article is subject to HHMI's Open Access to Publications policy. HHMI laboratory heads have previously granted a non-exclusive CC BY 4.0 license to the public and a sublicensable license to HHMI in their research articles. Pursuant to those licenses, the author-accepted manuscript of this article can be made freely available under a CC BY 4.0 license immediately upon publication.

**Author contributions** A.A. and D.R. conceived the study and wrote the manuscript, with input from all other authors. A.A. and A.P. developed the forward-in-time simulation pipeline. A.A. and M.K. developed the AGES web application. A.A., M.K. and Z.L. developed the PQLseqPy Python package. A.A., A.P., A.R.B., M.K., Y.Z. and A.M. analysed the genetic data. S.G., N.P., X.Z., A.L.P., E.S.L. and D.R. supervised various aspects of the study. D.R. curated and edited the archaeological information. A.A., M.M. and S.M. performed the bioinformatic analyses. R.P., N.R. and D.R. supervised the wet laboratory work.

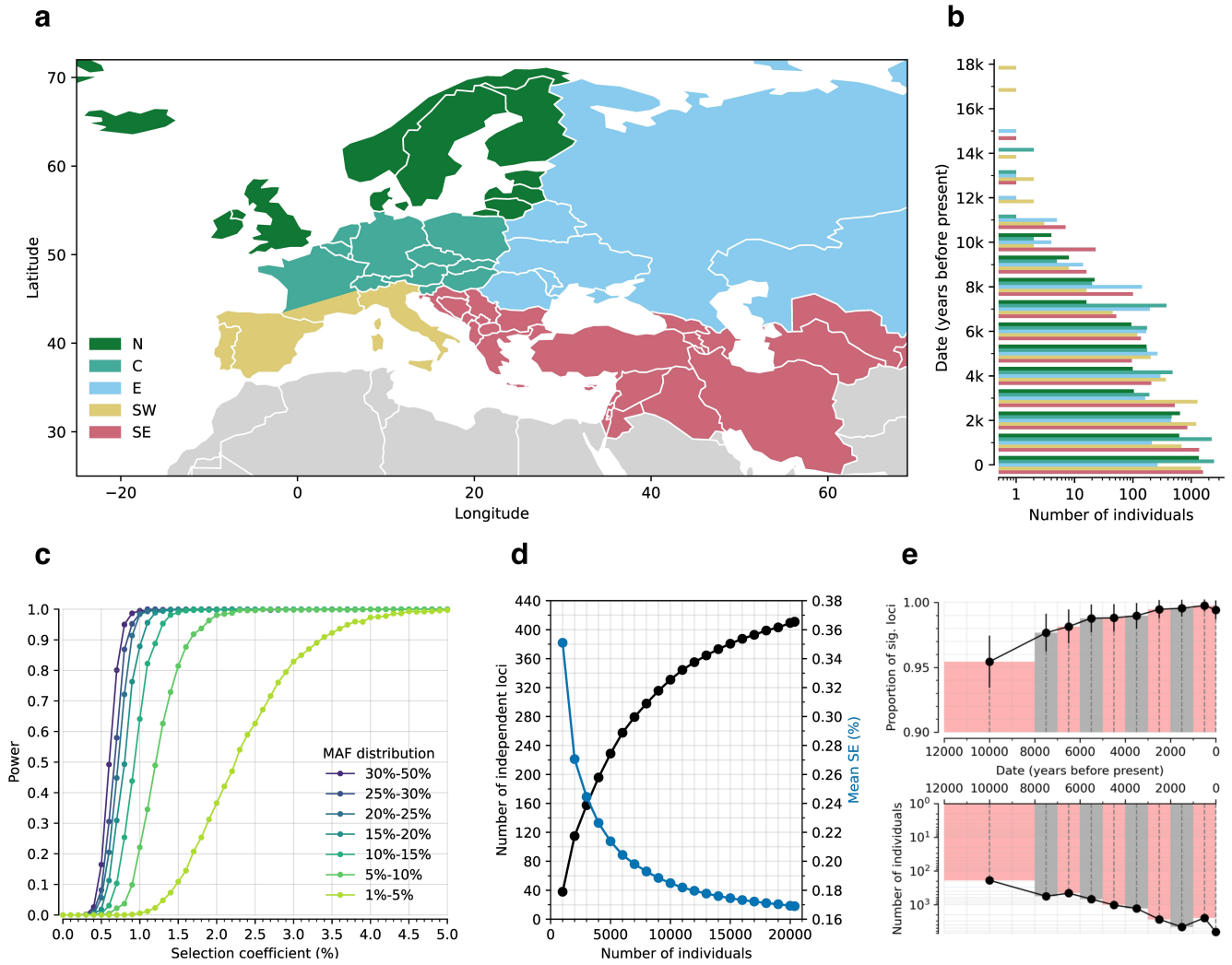
**Competing interests** The authors declare no competing interests.

**Additional information**

**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1038/s41586-026-10358-1>.

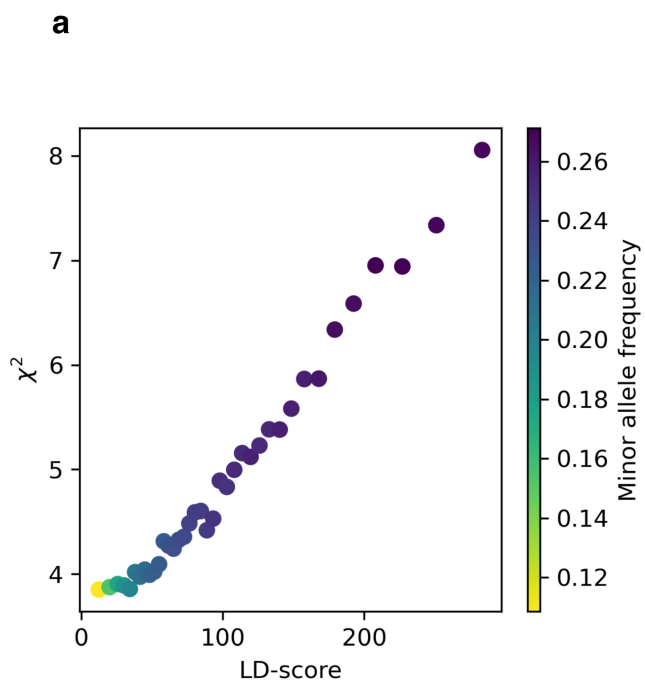
**Correspondence and requests for materials** should be addressed to Ali Akbari or David Reich. **Peer review information** *Nature* thanks the anonymous reviewer(s) for their contribution to the peer review of this work. Peer reviewer reports are available.

**Reprints and permissions information** is available at <http://www.nature.com/reprints>.



**Extended Data Fig. 1 | Spatiotemporal distribution of individuals and effect on power.** (a) Geographic origin: North (N), Central (C), East (E), Southwest (SW) and Southeast (SE). (b) Temporal distribution (x-axis on a logarithmic scale). (c) Power analysis based on simulations. Sample size, dates, and pattern of genetic relatedness are matched to real data. Power is defined as proportion of true positives expected at  $P < 5 \times 10^{-8}$ , based on two sided P values. We ran 20000 simulations for each selection coefficient, with minor allele frequency (MAF) at present (time=0) randomly drawn from the

MAF distribution in modern Europeans. (d) Number of independent and significant loci and standard error in selection coefficients as a function of sample size (from downsampling). (e) Effect of age on power. Data are divided into 10 non-overlapping periods; modern individuals are a separate bin. Top panel: proportion of 479 loci remaining significant after excluding 100 random individuals from each bin; error bars indicate 95% confidence intervals. Bottom panel: number of individuals per bin.

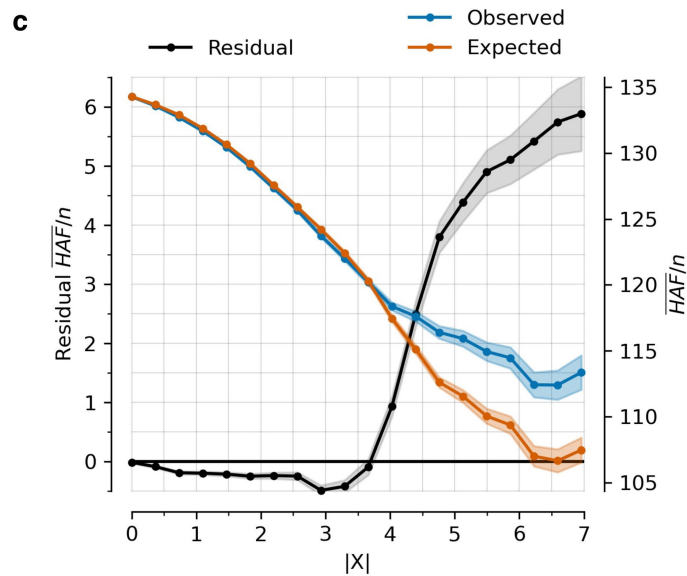
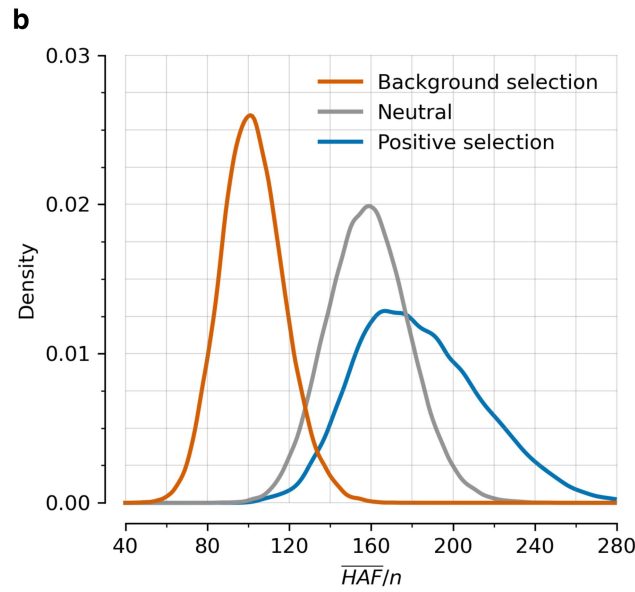
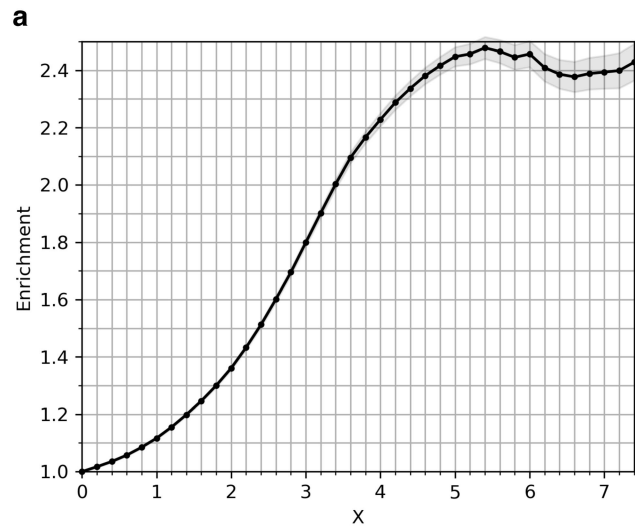


**Extended Data Fig. 2 | High proportion of genome affected by directional selection.** (a) LD score plot for nominal  $\chi^2$  statistics, with each point representing an LD score quantile. Values are averaged across each bin. (b) Mapping X-score to posterior probability ( $\pi$ ), False Discovery Rate (FDR),

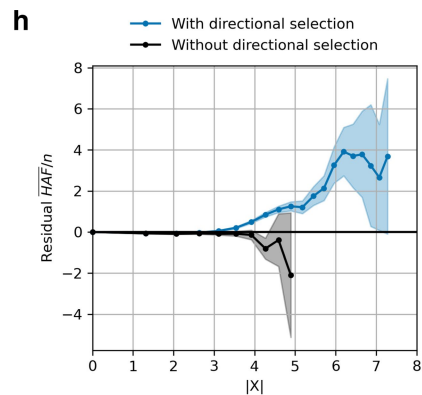
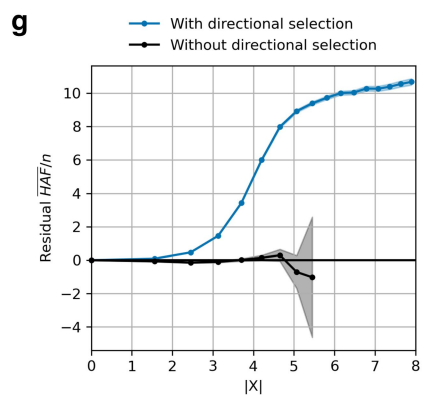
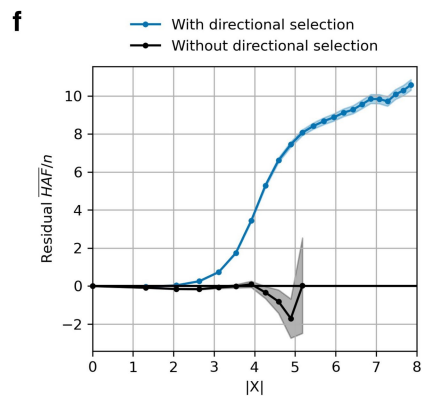
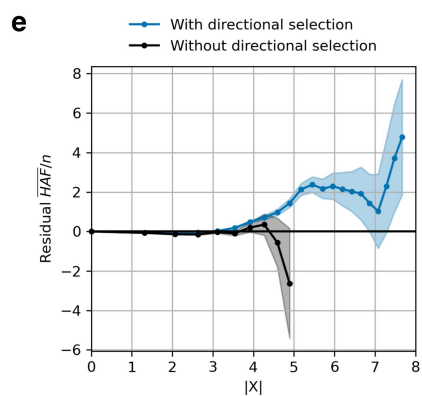
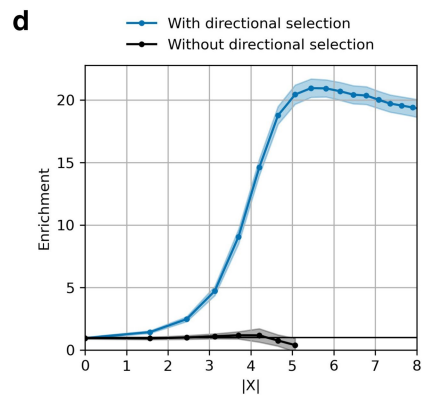
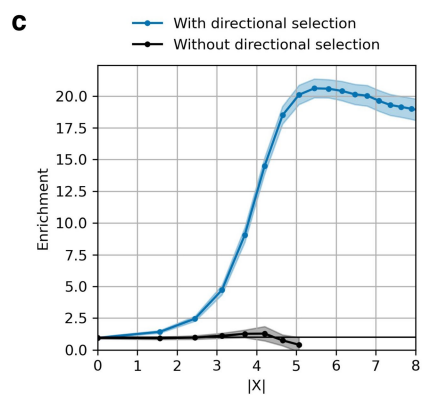
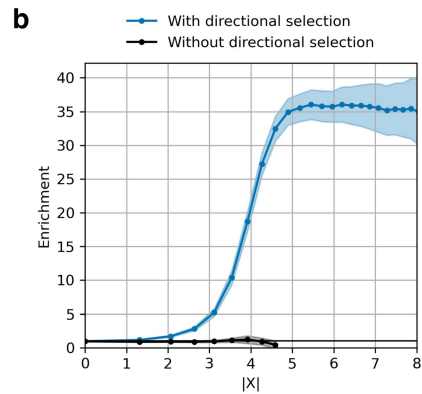
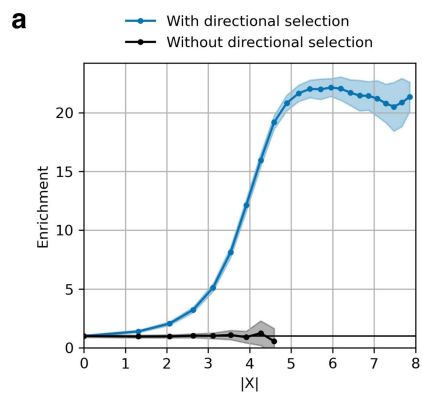
**b**

| $ X $  | $\pi$  | FDR    | N    | Percentage of genome in LD |
|--------|--------|--------|------|----------------------------|
| 5.4513 | 0.9900 | 0.0000 | 410  | 23.65±1.98                 |
| 5.4109 | 0.9800 | 0.0017 | 433  | 24.69±2.0                  |
| 5.3739 | 0.9710 | 0.0100 | 448  | 25.69±2.03                 |
| 5.3720 | 0.9700 | 0.0103 | 448  | 25.69±2.03                 |
| 5.3367 | 0.9600 | 0.0129 | 462  | 26.83±2.07                 |
| 5.3029 | 0.9500 | 0.0162 | 478  | 28.48±2.11                 |
| 5.2784 | 0.9424 | 0.0200 | 496  | 29.7±2.11                  |
| 5.2024 | 0.9161 | 0.0300 | 553  | 34.4±2.21                  |
| 5.1601 | 0.9000 | 0.0316 | 594  | 36.4±2.25                  |
| 5.0979 | 0.8758 | 0.0400 | 646  | 39.09±2.3                  |
| 5.0172 | 0.8423 | 0.0500 | 728  | 44.84±2.33                 |
| 4.9211 | 0.8000 | 0.0670 | 834  | 50.93±2.28                 |
| 4.7421 | 0.7199 | 0.1000 | 1108 | 65.42±2.13                 |
| 4.7000 | 0.7000 | 0.1124 | 1195 | 68.19±2.07                 |
| 4.4729 | 0.6000 | 0.1852 | 1742 | 84.84±1.56                 |
| 4.4351 | 0.5840 | 0.2000 | 1863 | 88.47±1.36                 |
| 4.2294 | 0.5000 | 0.2702 | 2650 | 95.32±0.78                 |
| 4.1587 | 0.4732 | 0.3000 | 3011 | 97.3±0.57                  |
| 3.8941 | 0.3814 | 0.4000 | 4727 | 99.46±0.18                 |
| 3.6057 | 0.2981 | 0.5000 | 7689 | 99.89±0.05                 |

number of independent loci excluding the HLA region (N), and the fraction of the genome (in base pairs) covered by stretches of LD ( $r^2 > 0.05$ ) around tag SNPs representing these loci. The stretches of LD are calculated in the modern genome using European populations from the 1000 Genomes Project.



**Extended Data Fig. 3 | Robustness of directional selection signals (related to Fig. 1a,c).** (a) Enrichment of SNPs significant in any of 452 UK Biobank GWAS studies for X-statistics with magnitudes larger than the threshold on the x-axis, adjusted for minor allele frequency and measures of background selection (McVicker-B, Murphy-phastCons, and Murphy-CADD). Background selection tends to be higher in functional genomic regions, so SNPs with higher  $|x|$  are more penalized than in Fig. 1a hence the lower plateau. (b) Simulating neutral, background, and positive selection for a 200 kb window around a focal SNP, with derived allele frequency drawn uniformly from [0,1]. Under positive selection, the focal SNP has a selection coefficient of  $s = 0.01$  for the favored allele. Under background selection, 20% of mutations are drawn from a gamma distribution with mean  $-0.03$  and shape parameter 0.206, and the remainder are neutral. The population size is constant at 20000 diploid individuals, mutation rate per base pair per generation is  $2 \times 10^{-8}$ , and recombination rate is 1 cM per 1 Mb. (c) Residual mean  $(HAF)/n$  for a haploid sample size  $n$  over 200 bp windows is observed minus expected value. Expected value is determined using a linear regression model with explanatory variables McVicker-B, Murphy-phastCons, Murphy-CADD, number of SNPs, and total heterozygosity, providing the expected mean  $(HAF)/n$  conditioned on them. Shaded areas show 95% confidence intervals.

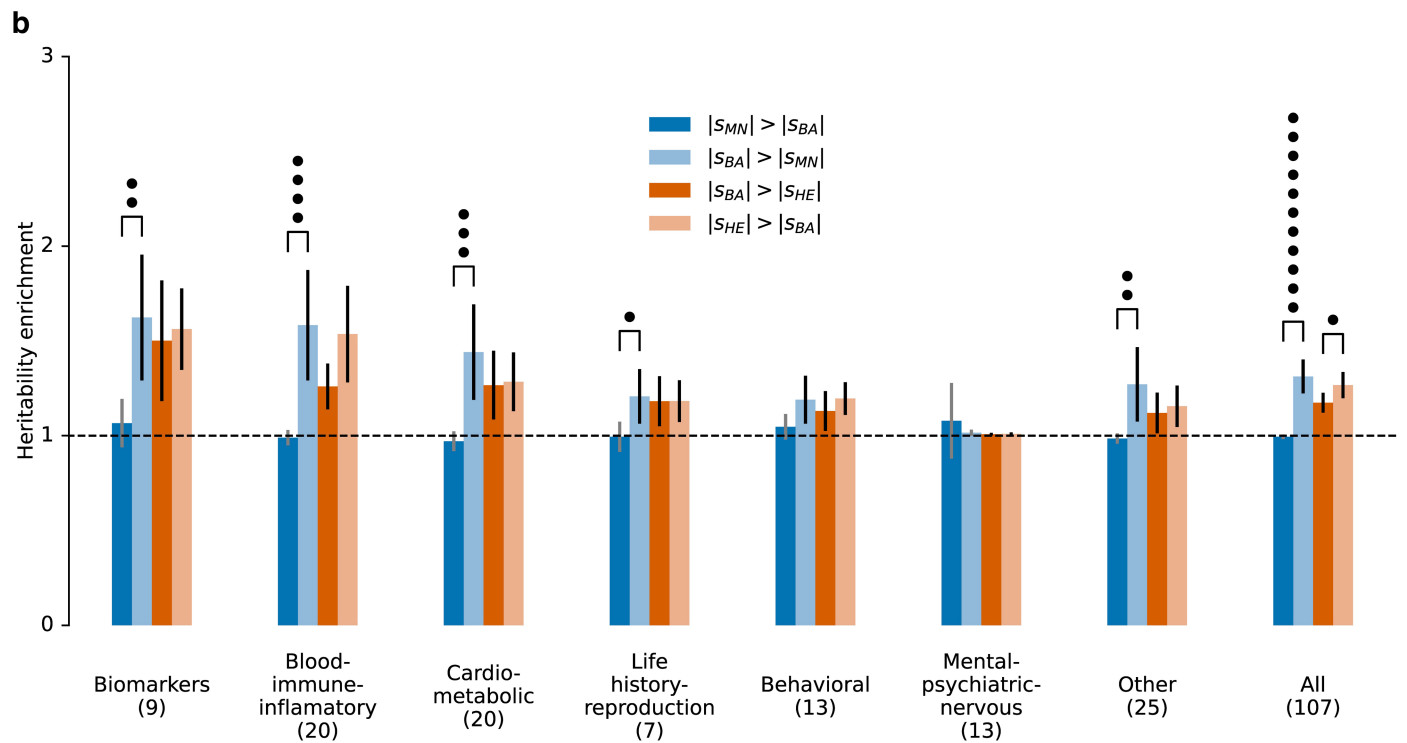
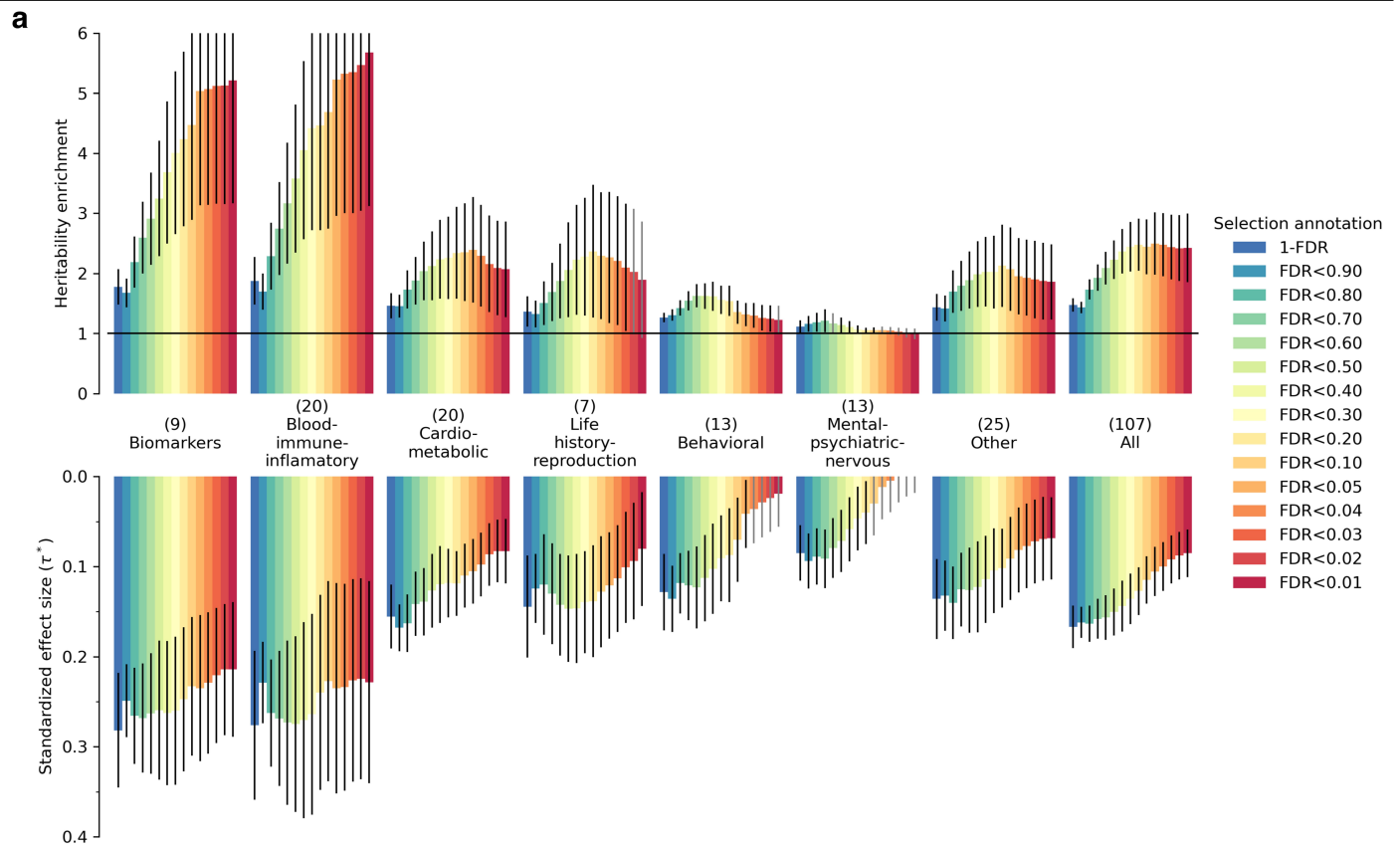


**Extended Data Fig. 4** | See next page for caption.

# Article

**Extended Data Fig. 4 | GWAS enrichment and residual HAF of selection signal under simulated models of selection. (a-d)** GWAS enrichment by X-score across four simulated models: **(a)** Model 1, polygenic directional selection with a stabilizing selection mechanism; **(b)** Model 2.1, soft sweep with a purifying selection mechanism; **(c)** Model 2.2, hard sweep with a purifying selection mechanism; **(d)** Model 3, polygenic directional selection with both stabilizing and purifying selection mechanisms. Shaded areas show 95% confidence intervals around the estimated enrichment. Each experiment includes 800 simulations (400 with directional selection and 400 without); blue denotes all simulations, and black denotes those without directional selection. **(e-h)** Residual mean (HAF)/n, for SNPs with  $|\chi|$  above the value on the x-axis, for a haploid sample size n over 200 bp windows across four simulated models: **(e)** Model 1, polygenic directional selection with a stabilizing selection

mechanism; **(f)** Model 2.1, soft sweep with a purifying selection mechanism; **(g)** Model 2.2, hard sweep with a purifying selection mechanism; **(h)** Model 3, polygenic directional selection with both stabilizing and purifying selection mechanisms. For each SNP, we compute mean (HAF)/n across haplotypes and residualize it as observed minus expected from a linear regression correcting for background selection using explanatory variables McVicker-B, coding region annotation, functional noncoding region annotation, number of SNPs, and total heterozygosity in a 200 kb window. Solid lines show the mean of this SNP-level residual statistic across SNPs meeting the  $|\chi|$  threshold; shaded areas show 95% confidence intervals. Each experiment includes 800 simulations (400 with directional selection and 400 without); blue denotes all simulations, and black denotes those without directional selection.

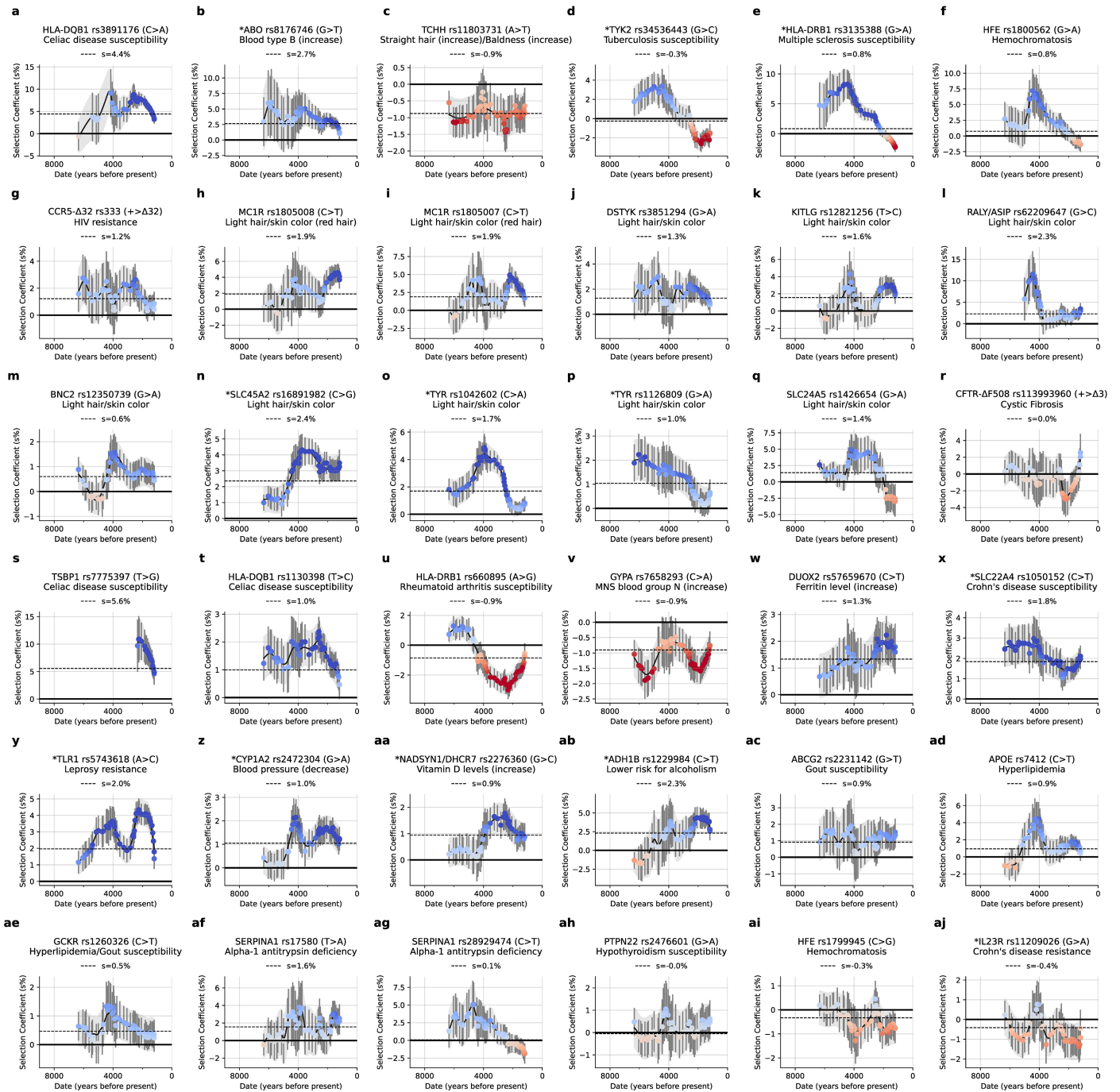


**Extended Data Fig. 5** | See next page for caption.

# Article

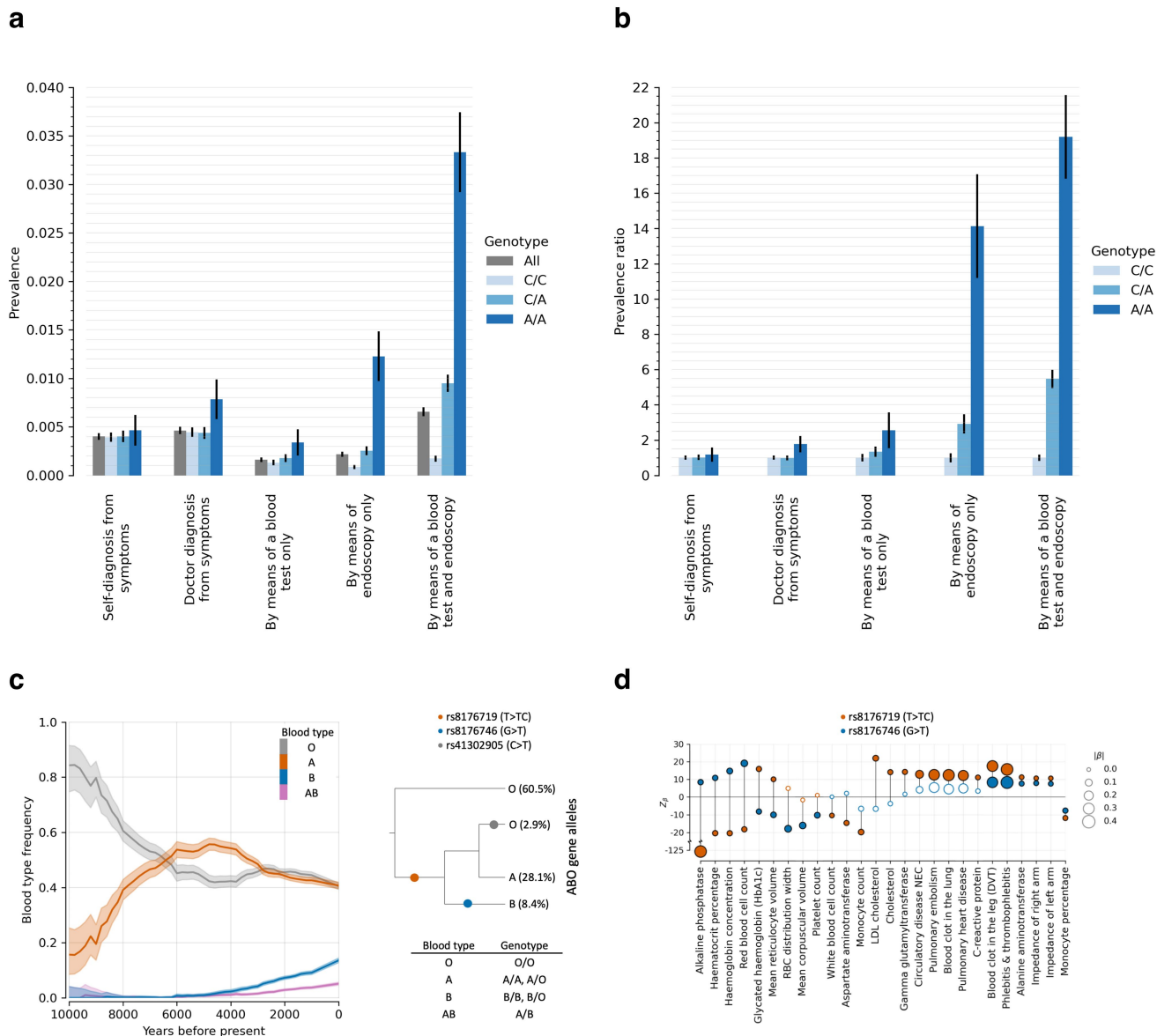
**Extended Data Fig. 5 | Stratified LD Score Regression shows that alleles affecting blood-immune-inflammatory and cardio-metabolic traits were unusually affected by selection, and that selection intensity increased in the Bronze Age (related to Fig. 1e).** (a) We annotated sites based on their inferred strength of selection—based on their FDR being above a specified threshold, or 1-FDR as a continuous annotation—and used Stratified LD Score Regression (S-LDSC) to study enrichment of GWAS signals and standardized effect sizes ( $r^*$ ) for traits in different functional categories. Our analysis adjusts for 97 annotations that are known to affect heritability and are part of the standard correction in S-LDSC (b) Tests for changes in selection intensity

during different cultural transitions: Mesolithic-Neolithic (MN) to Bronze Age (BA); and Bronze Age (BA) to Historical Era (HE). Each annotation is binary, identifying SNPs among the top 5% with the highest probability of experiencing stronger selection during one time period compared to another. This is determined using the estimated selection coefficient and standard error from models separately fit to each cultural period. Error bars show 95% confidence intervals of the estimated standardized effect sizes ( $r^*$ ) or heritability enrichment; dots represent significance of the pairwise comparisons of heritability enrichment.



**Extended Data Fig. 6 | How selection coefficients on single variants changed in intensity over time (for the gallery of 36 loci also highlighted in Fig. 3).** Time-variant selection coefficients are estimated by refitting our model in sliding windows of 2000 years, with a step size of 100 years, and a minimum of 2000 people per window. Color map represents the Z-score for the selection coefficient being non-zero in that window, ranging from -5 (dark red) to 5 (dark blue). The solid horizontal black line indicates 0, and the dashed horizontal black line represents the estimated selection coefficient using all data in this study (Fig. 3). Error bars show the 95% confidence interval, and

shaded areas represent smoothed point estimates with 95% confidence intervals, obtained using the GaussianProcessRegressor function from the Scikit-learn library in Python. We parameterize this function with  $\alpha = 1e-5$  and a  $1^*$ RationalQuadratic kernel, with `length_scale_bounds` set to `(1, 1e6)`. The x-values correspond to the median date of samples in each bin. Present-day samples are excluded because their disproportionate density shifts the median to the present, forcing the bin value to 0 and creating a -1000-year jump on the x-axis, causing computational and visual inconsistencies.

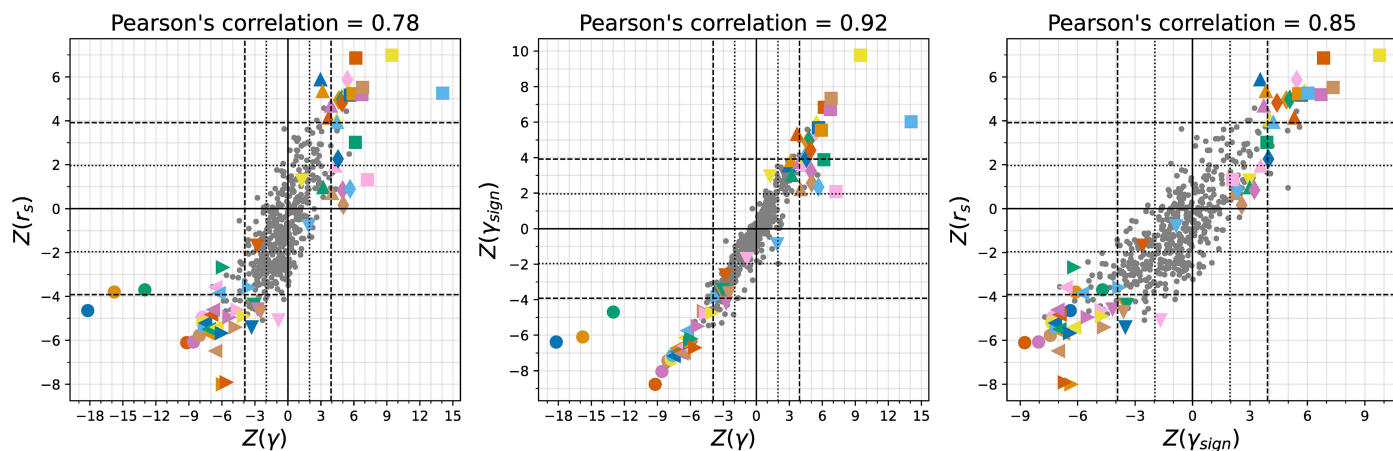


**Extended Data Fig. 7 | Genotype-phenotype correlations for the signals of selection for Coeliac disease at HLA and the ABO blood group locus.**

(a) Prevalence and (b) prevalence ratio of coeliac disease or gluten sensitivity (UK Biobank field 21068), conditioned on the genotype of rs3891176 (C > A) in 337,391 individuals of European ancestry from UK Biobank. The prevalence ratios are compared to the A/A genotype as a baseline. Bars show prevalence or prevalence ratio estimates; error bars are 95% confidence intervals. (c) Left:

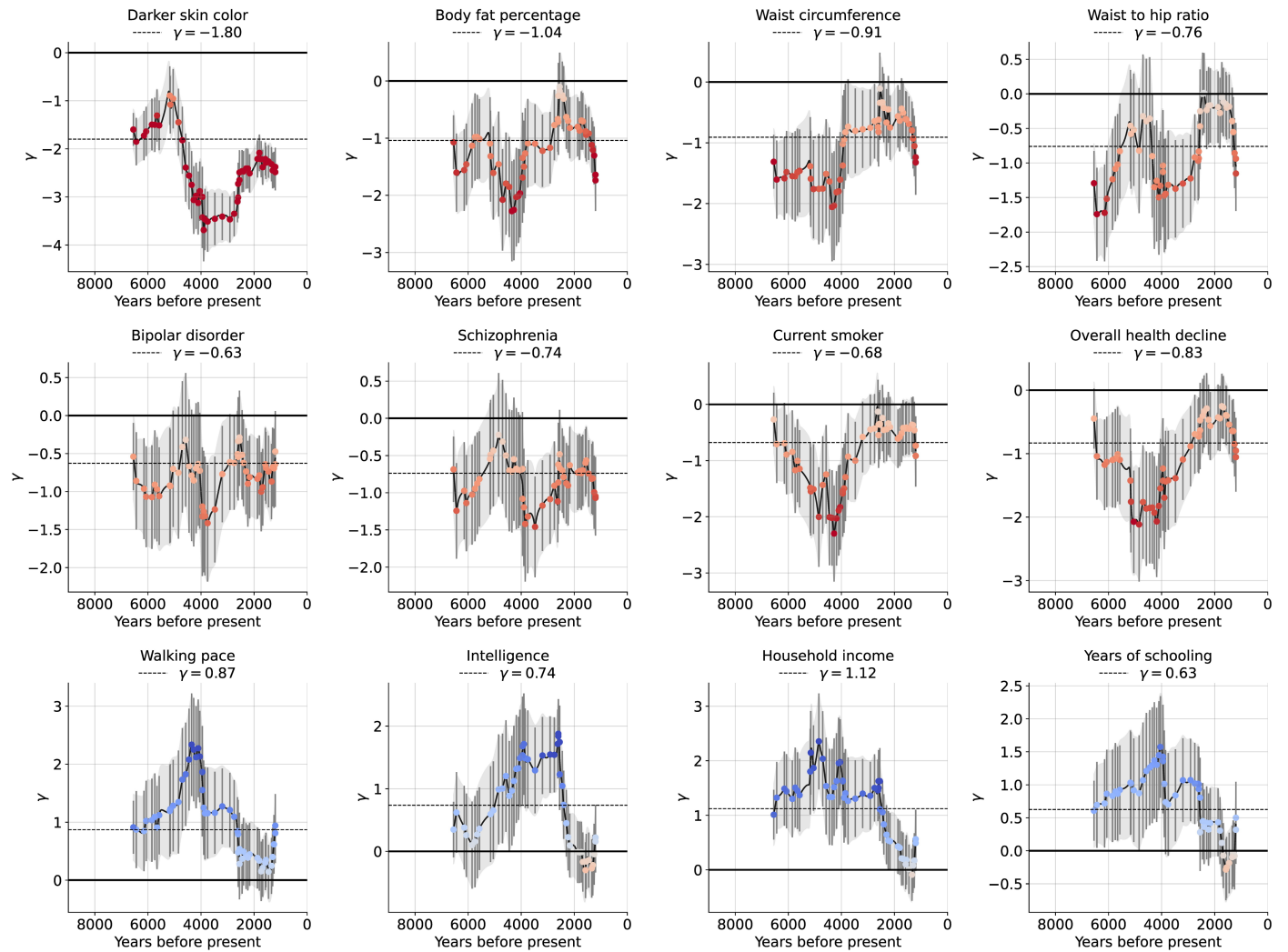
Blood type frequency trajectories for O, A, B, and AB estimated from our aDNA time series. Right: Genealogy of the ABO alleles approximated by Shelton et al.<sup>268</sup>. The allele frequencies are estimated from Europeans in the 1000 Genomes Project; shading gives 95% confidence interval around the estimated allele frequency trajectory. (d) Significant association to traits in UK Biobank for the two base pair insertion rs8176719 (T > TC) and SNP rs8176746 (G > T), approximating the alleles A and B.

- Darker skin color
- Black hair
- Dark brown hair
- Leg fat percentage (left)
- Leg fat percentage (right)
- Body fat percentage
- Leg fat mass (right)
- Arm fat percentage (right)
- Trunk fat percentage
- Arm fat percentage (left)
- Qualifications: None
- Waist circumference
- Heavy manual or physical job
- Time spend outdoors in summer
- Financial difficulties
- BMI (Yengo 2018)
- Overall health decline
- BMI
- Waist to hip ratio
- Schizophrenia (Trubetskoy 2022)
- Beer intake
- Rent from local authority
- Current tobacco smoking
- Bipolar disorder (Mullins 2021)
- Mouth/teeth dental problems
- Physical activity: None
- Type 2 diabetes
- Disability unemployment
- Hypertension
- Essential hypertension
- Reaction time (Davies 2018)
- Mental disorders related to tobacco
- Tobacco use disorder
- Heart failure (Shah 2020)
- Standing height
- Type 1 diabetes (Chiou 2021)
- Parental lifespan (Timmers 2020)
- Physical activity: Heavy DIY
- Calcium
- No dietary restriction
- Cheese intake
- IBD (DeLange 2017)
- Varicose vein surgery
- Smoking status: Never
- Noncognitive aspects of EA (Demange 2021)
- Reticulocyte count (Vuckovic 2020)
- Age first had sexual intercourse
- Education\_Years (Rietveld 2013)
- Years of schooling
- Hypothyroidism/myxoedema
- Hypothyroidism
- Qualifications: A levels
- Number of vehicles in household
- Malignant neoplasms of skin
- Fluid intelligence score
- Intelligence (Savage & Jansen 2018)
- Vitamin D
- Cereal Type: Muesli
- Cognitive aspects of EA (Demange 2021)
- Walking pace
- Use of sun/uv protection
- Household income
- Sunburn



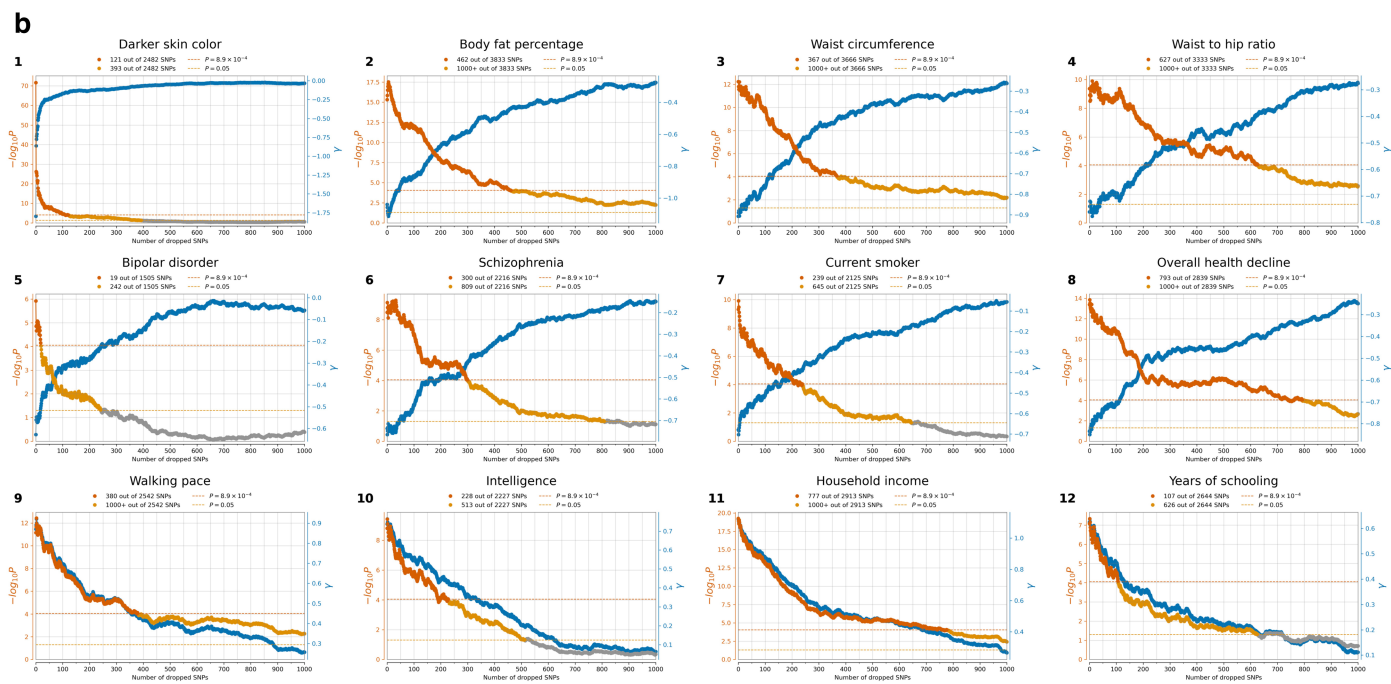
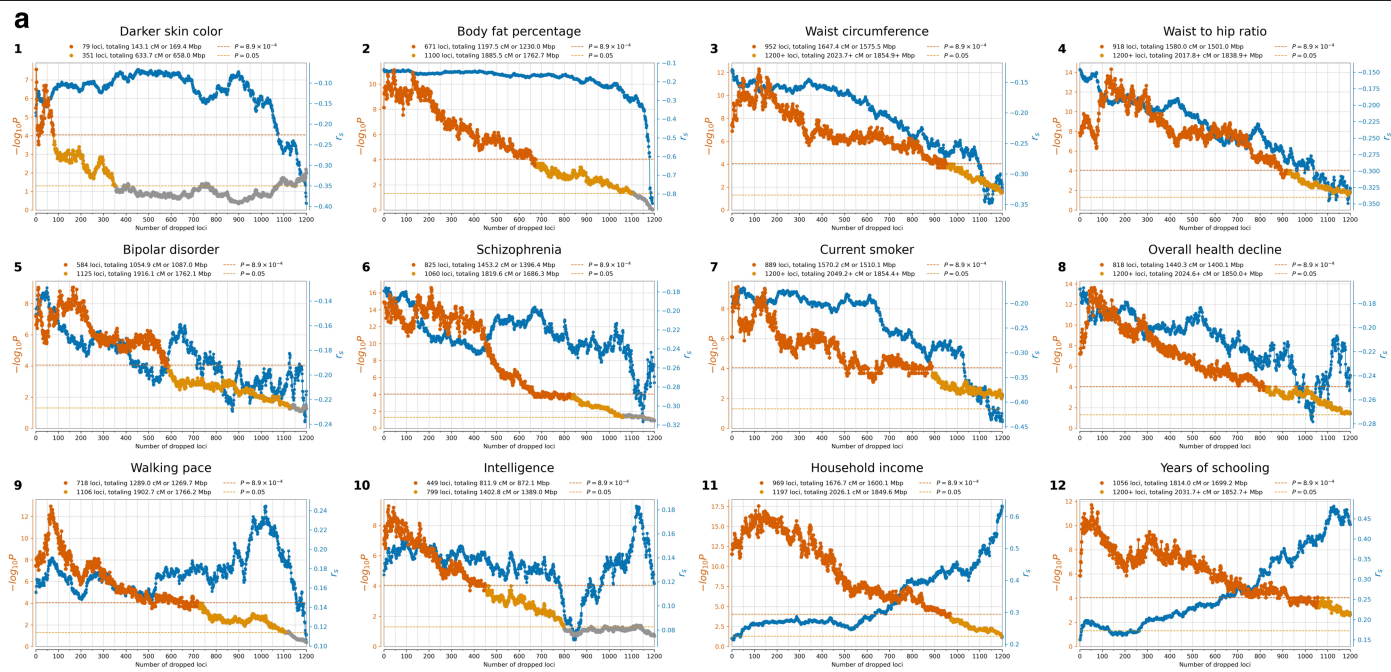
**Extended Data Fig. 8 | High correlation of 3 tests for polygenic selection** ( $\gamma$ ,  $\gamma_{\text{sign}}$ ,  $r_s$ ). Each dot represents a phenotype, some annotated by colors. Pearson's correlation for x and y axes at top. The dashed line indicates the

Bonferroni-corrected significance threshold ( $P = 8.9 \times 10^{-5}$ , correcting for 563 GWAS tested), and the dotted line indicates the uncorrected nominal significance threshold ( $P = 0.05$ ). All P values are two sided.



**Extended Data Fig. 9 | How coordinated selection on alleles affecting the same traits changed in intensity over time (gallery of 12 complex traits also highlighted in Fig. 4).** We estimate time-variant polygenic selection intensity  $\gamma$  by refitting our model in sliding windows of 2000 years, with a step size of 100 years, and a minimum of 2000 people per window. Color map represents the Z-score for the selection coefficient being non-zero in that window, ranging from -5 (dark red) to 5 (dark blue). The solid horizontal black line indicates 0, and the dashed horizontal black line represents the estimated  $\gamma$  using all data in this study (Fig. 4). Error bars show the 95% confidence interval

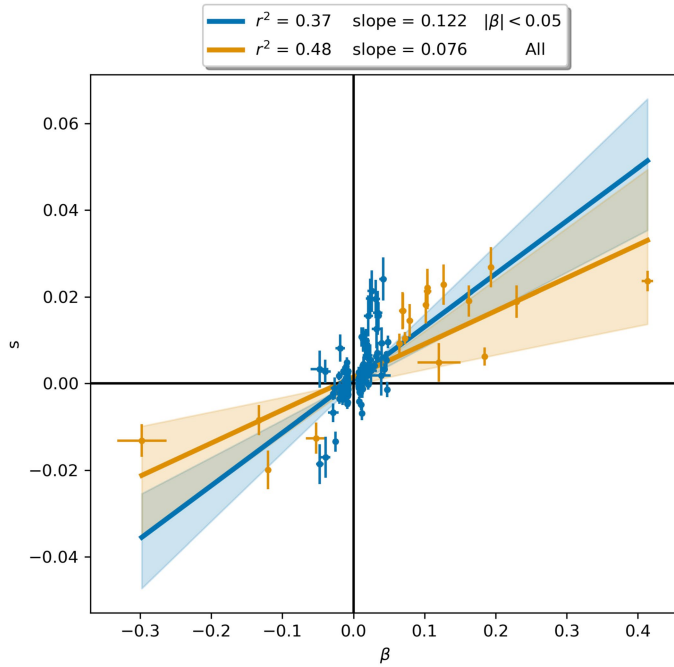
around the estimated  $\gamma$  in each window, and shaded areas represent smoothed point estimates with 95% confidence intervals, obtained using the GaussianProcessRegressor function from the Scikit-learn library in Python. We parameterize this function with  $\alpha = 1e-3$  and a  $1^*$ RationalQuadratic kernel, with length\_scale\_bounds set to (1, 1e6). The x-values correspond to the median date of samples in each bin. Present-day samples are excluded because their disproportionate density shifts the median to the present, forcing the bin value to 0 and creating a -1000-year jump on the x-axis, causing computational and visual inconsistencies.



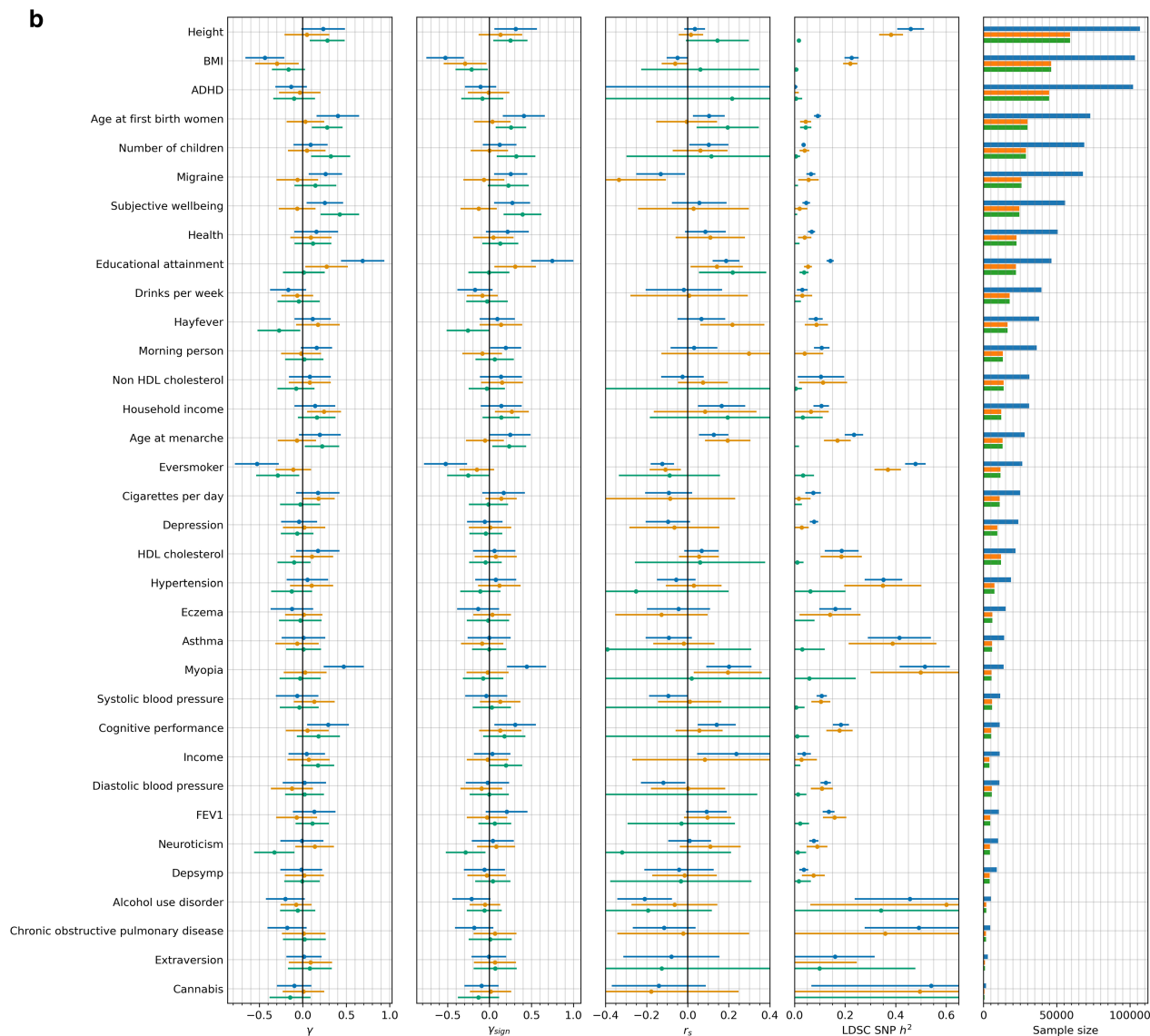
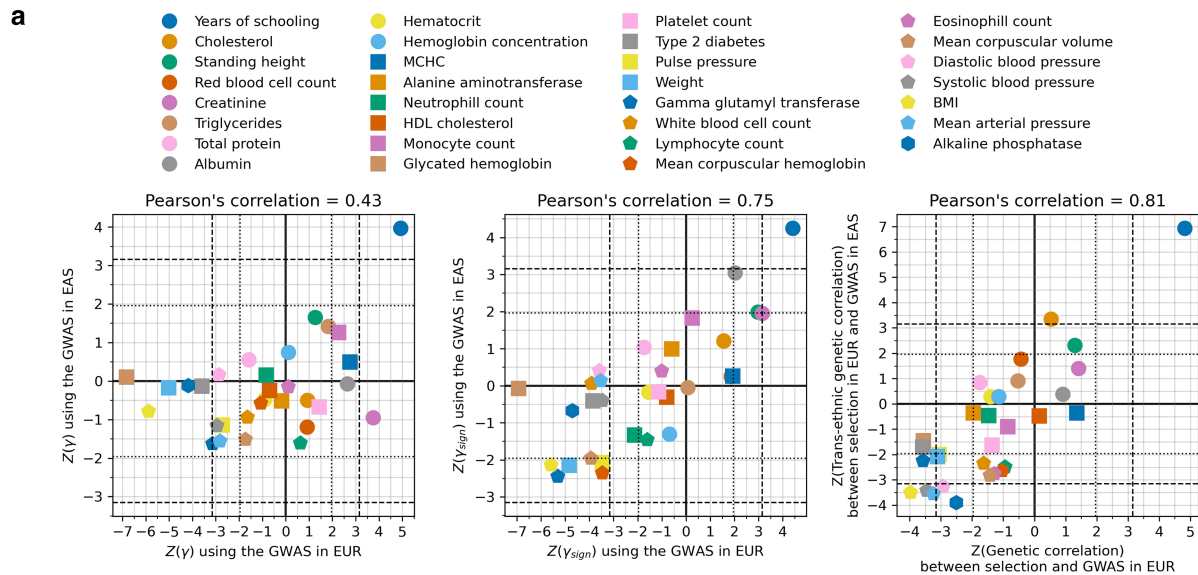
**Extended Data Fig. 10 | Estimating the minimum number of SNPs affected by selection for each trait (gallery of 12 traits also highlighted in Fig. 4).**

(a) Each panel shows the LDSC genetic correlation ( $r_s$ ) between the trait effect sizes and the selection coefficients ( $s$ ) as a function of number of dropped loci. The right axis displays  $r_s$  in blue; P-value on the left axis in orange. For each SNP, we define a priority score  $|\beta \times s \times f \times (1-f)|$ , where  $\beta$  is the GWAS effect size,  $s$  the selection coefficient, and  $f$  allele frequency. SNPs are sorted by priority score,

and in each iteration, a 2 cM region around the highest priority SNP is dropped,  $r_s$  is recalculated for the remaining genome, and this continues until no SNPs are left. (b) We similarly show  $\gamma$  estimates at right as a function of number of dropped SNPs (blue), and P-value for polygenic selection at left, with dark orange indicating  $P < 8.9 \times 10^{-5}$  (Bonferroni-corrected threshold for 563 GWAS tests), light orange  $P < 0.05$ , and gray otherwise. All P values are two sided.



**Extended Data Fig. 11 | Pigmentation is oligogenic but selection on it was polygenic.** Selection coefficient ( $s$ ) and effect size ( $\beta$ ) from the UK Biobank skin color phenotype for 150 independent SNPs passing the GWAS P-value threshold of  $p < 5 \times 10^{-8}$ . Following<sup>43</sup>, which applied ordinary least squares regression, we instead used orthogonal distance regression (ODR) to account for uncertainty in both variables. The orange line was fit to all SNPs (131 blue and 19 orange markers), whereas the blue line was fit only to SNPs with  $|\beta| < 0.05$  (131 blue markers). Neither the correlations (difference between Fisher Z-transformed Pearson's correlation,  $P = 0.22$ ) nor the slopes ( $P = 0.11$ ) differ significantly, consistent with a model in which selection for pigmentation had an equal impact on all variants in proportion to effect size. Shaded areas indicate 95% confidence bands. Points show  $\beta$  estimated in 415,018 UK Biobank individuals (x-axis) and  $s$  estimated from 20,374 unrelated individuals (y-axis); error bars show 95% confidence intervals. All P values are two sided.



Extended Data Fig. 12 | See next page for caption.

# Article

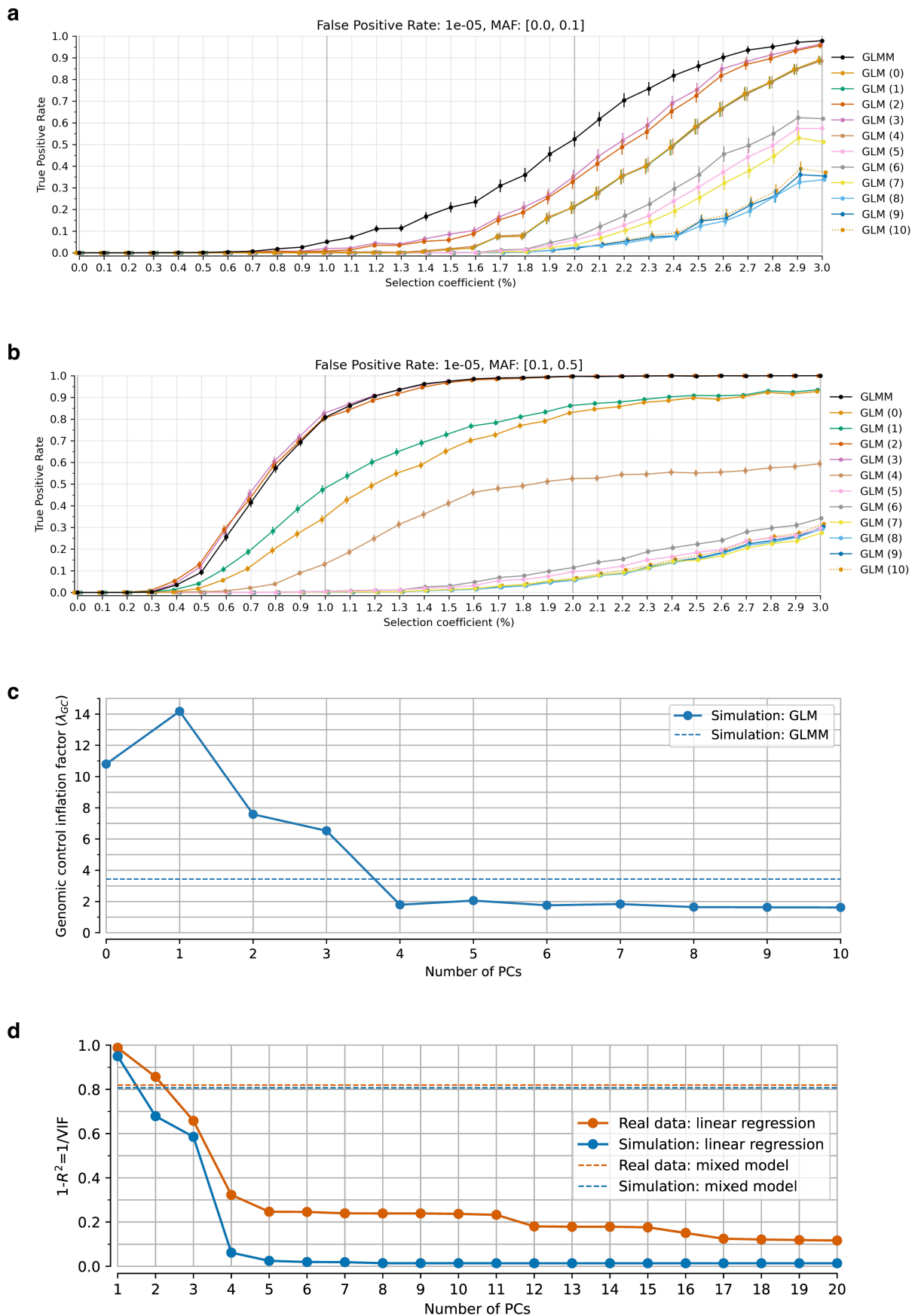
**Extended Data Fig. 12 | Replication of polygenic selection signals using effect sizes from East Asian GWAS and consistency between population-based and family-based GWAS.** (a) Replication of signals of polygenic selection using effect size estimates in East Asians whose population structure is uncorrelated to West Eurasians. We applied our polygenic selection test to 31 traits using pairs of GWAS studies for the trait, one from Europe and one from East Asia. We assessed  $\gamma$ ,  $\gamma_{\text{sign}}$ , and  $r_s$  in the two analyses and found them to be consistent. It is extremely difficult to believe that an estimate based on effect size estimates in East Asians could be an artifact of population structure, since structure within Europe and within East Asia are uncorrelated. This comes at the cost of reduced power relative to using European GWAS directly, driven by the smaller sample sizes of East Asian GWAS (median 30% of European GWAS), differences in LD structure between populations, and divergence in genetic architecture (allele frequencies and causal effect sizes)<sup>51-53</sup>. Despite reduced power, Pearson's correlations between Z-scores for European and East Asian GWAS are 0.43, 0.75, and 0.81 for  $\gamma$ ,  $\gamma_{\text{sign}}$ , and  $r_s$ , respectively, a positive correlation that must reflect real selection and would not be expected due to population structure. The dashed line indicates the Bonferroni-corrected significance threshold ( $P = 1.6 \times 10^{-3}$ , correcting for 31 traits), and the dotted line indicates the nominal significance threshold ( $P = 0.05$ ). Years of schooling is the only trait that reaches Bonferroni-corrected significance threshold for all three tests. (b) Consistency of polygenic selection signals using effect sizes estimated from GWAS of unrelated people and family-based GWAS. We analyze

the GWAS results reported in<sup>54</sup>, who compiled whole genotype-phenotype correlation data in families for 34 phenotypes, and report three GWAS for each phenotype: unrelated people ("population GWAS"), and direct and indirect effect GWAS obtained by differentiating between transmitted and untransmitted chromosomes within families. The first three columns show estimates for three polygenic tests of selection. The fourth column displays the estimated SNP heritability ( $h^2$ ) by LDSC. The fifth column shows the sample size, with the median sample size for family GWAS - 13,000 and for population GWAS - 31,000. Both are only a fraction of the UK Biobank cohort of -490,000 individuals of non-Finnish European ancestry<sup>269</sup>, explaining the reduced power and why confidence intervals are wide and often overlap zero. The Pearson's correlations between Z-scores from population GWAS and direct genetic effect GWAS are 0.51, 0.51, and 0.76 for  $\gamma$ ,  $\gamma_{\text{sign}}$ , and  $r_s$ , respectively. When restricted to population GWAS signals with nominal significance ( $P < 0.05$ ), the correlations increase to 0.84, 0.78, and 0.81, respectively. Among population GWAS in<sup>54</sup>, educational attainment, ever-smoker, and myopia reach the Bonferroni-corrected significance ( $P < 1.5 \times 10^{-3}$ , correcting for 34 traits) for all three tests. For direct genetic effects, educational attainment is the only trait with all three tests nominally significant ( $P < 0.05$ ). Blue indicates population GWAS, orange represents direct genetic effects from family GWAS, and green represents the average of paternal and maternal non-transmitted coefficients (NTCs) from family GWAS. Points show the estimated statistic; error bars indicate 95% confidence intervals. All P values are two sided.

|                        |                    |                     |                     |                        |                |                |                |                  |                    |                   |                  |
|------------------------|--------------------|---------------------|---------------------|------------------------|----------------|----------------|----------------|------------------|--------------------|-------------------|------------------|
| Body fat percentage    | 0.64<br>*****      |                     |                     |                        |                |                |                |                  |                    |                   |                  |
| Waist circumference    | 0.79<br>*****      | 0.88<br>*****       |                     |                        |                |                |                |                  |                    |                   |                  |
| Overall health decline | 0.51<br>*****      | 0.56<br>*****       | 0.57<br>*****       |                        |                |                |                |                  |                    |                   |                  |
| Walking pace           | -0.51<br>*****     | -0.66<br>*****      | -0.61<br>*****      | -0.7<br>*****          |                |                |                |                  |                    |                   |                  |
| Current smoker         | 0.36<br>*****      | 0.31<br>*****       | 0.34<br>*****       | 0.51<br>*****          | -0.42<br>***** |                |                |                  |                    |                   |                  |
| Intelligence           | -0.23<br>*****     | -0.21<br>*****      | -0.16<br>*****      | -0.39<br>*****         | 0.36<br>*****  | -0.3<br>*****  |                |                  |                    |                   |                  |
| Household income       | -0.35<br>*****     | -0.35<br>*****      | -0.31<br>*****      | -0.65<br>*****         | 0.51<br>*****  | -0.51<br>***** | 0.63<br>*****  |                  |                    |                   |                  |
| Years of schooling     | -0.39<br>*****     | -0.4<br>*****       | -0.35<br>*****      | -0.61<br>*****         | 0.53<br>*****  | -0.57<br>***** | 0.73<br>*****  | 0.82<br>*****    |                    |                   |                  |
| Darker skin color      | 0.03               | 0.06<br>****        | 0.04<br>*           | 0.02                   | -0.06<br>*     | 0.08<br>*      | -0.15<br>****  | -0.01            | -0.14<br>*****     |                   |                  |
| Bipolar disorder       | 0.01               | -0.05<br>*          | -0.04<br>*          | 0.11<br>*****          | -0.01          | 0.16<br>*****  | -0.08<br>***   | 0.02             | -0.0               | 0.05              |                  |
| Schizophrenia          | -0.04<br>*         | -0.09<br>*****      | -0.1<br>*****       | 0.15<br>*****          | 0.0            | 0.19<br>*****  | -0.22<br>***** | -0.15<br>*****   | -0.12<br>*****     | 0.0               | 0.69<br>*****    |
|                        | Waist to hip ratio | Body fat percentage | Waist circumference | Overall health decline | Walking pace   | Current smoker | Intelligence   | Household income | Years of schooling | Darker skin color | Bipolar disorder |

**Extended Data Fig. 13 | Correlations of polygenic scores for complex traits with strong evidence of coordinated selection (the same 12 traits highlighted in Fig. 4).** Genetic correlations of traits are computed using LDSC based on modern UK Biobank data (no ancient DNA are used in this analysis).

Asterisks indicate significance level (n asterisks represent a jackknife estimated  $P < 0.5 \times 10^{-n}$ ). All P values are two sided. A caution is that some of these genetic correlations may be inflated due to assortative mating.



Extended Data Fig. 14 | See next page for caption.

**Extended Data Fig. 14 | Comparison of GLMM and GLM performance and the impact of principal component covariates on inflation and power.**

**(a, b)** Comparing the performance of GLMM and a generalized linear model (GLM) with PC covariates. True positive rate at a false positive rate of  $10^{-5}$  versus selection coefficient for MAF ranges **(a)** [0, 0.1] and **(b)** [0.1, 0.5]. The black line shows the GLMM model, and the colored lines show the GLM model (with the number of PCs in parentheses). Selection coefficients range from 0.0 to 0.03 in 0.001 increments, with ~4,000 simulation replicates per selection coefficient. Error bars indicate 95% confidence intervals around estimated true positive rate. **(c, d)** Inclusion of additional principal components as covariates inflates variance and reduces power due to collinearity with time. **(c)** Genomic control inflation factor ( $\lambda_{GC}$ ) for the GLM and GLMM in simulation.  $\lambda_{GC}$  is defined as  $\text{median}(Z^2)/0.455$ , which quantifies the median inflation of the  $Z^2$  statistics relative to a chi-square distribution with one degree of freedom. Various factors, including frequency-dependent biases from imputation and the non-linear transformation of allele frequencies, which lead to a non-normal null distribution, as well as polygenic signal and confounding due to population structure, may contribute to this inflation (see Supplementary Section 2). **(d)** Loss of power due to collinearity between time and principal components.

The y-axis shows  $1 - R^2$  (equal to  $1/\text{VIF}$  - variance inflation factor), a proxy for statistical power—lower values indicate greater variance inflation and reduced power. Each dot represents a linear regression with time as the response and the number of PCs (x-axis) as covariates. Both simulated and real data reveal that including more PCs increases collinearity with time, diminishing power. In simulations, power drops sharply after three PCs, while in real data the decline is more gradual due to complex structure not captured in the simulation (so continued correction adds value).  $R^2$  is the coefficient of determination, indicating the proportion of variance in time explained by the PCs. VIF quantifies the inflation of the variance of the estimated coefficient due to collinearity with time. The dashed lines indicate  $1 - R_c^2$ , where  $R_c^2$  is the conditional  $R^2$  (see<sup>270</sup>).  $R_c^2$  is the proportion of variance in time explained by the fixed effects and the genetic random effect in a linear mixed model, where time is the response variable. The model includes an intercept as the fixed effect and a random effect with a GRM as its covariance structure. In this model, since the only fixed effect is the intercept, which contributes no variance,  $R_c^2$  reflects the variance explained by the genetic random effect. It measures the model's goodness of fit and is analogous to  $R^2$  in ordinary linear regression.

# Article

## Extended Data Table 1 | Re-evaluation of signals from five scans for selection in Holocene West Eurasia

|                          | Genome wide significant loci |         |              |              | Less stringent threshold |         |              |              |
|--------------------------|------------------------------|---------|--------------|--------------|--------------------------|---------|--------------|--------------|
|                          | Total                        | Pass QC | $\pi > 99\%$ | $\pi > 50\%$ | Total                    | Pass QC | $\pi > 99\%$ | $\pi > 50\%$ |
| Mathieson et al. 2015    | 12                           | 11      | 10           | 10           |                          |         |              |              |
| Field et al. 2016        | 3                            | 3       | 1            | 3            | 37                       | 36      | 3            | 11           |
| Le et al. 2022           | 24                           | 22      | 9            | 10           |                          |         |              |              |
| Kerner et al. 2023       | 3                            | 3       | 3            | 3            | 139                      | 125     | 14           | 23           |
| Irving-Pease et al. 2024 | 21                           | 21      | 14           | 17           |                          |         |              |              |

Significance according to our analysis for loci identified in five previous scans for selection in Holocene West Eurasia (all but Field et al. are ancient DNA-based scans). The less stringent P-value thresholds are  $10^{-5}$  for Field et al.<sup>13</sup> and 0.05 for Kerner et al. 2023. The cumulative number of non-HLA signals identified as genome-wide significant and confirmed in our re-analysis with a posterior probability of selection  $\pi > 99\%$  is 18 (4% of the 410 non-HLA loci with  $\pi > 99\%$ ). Of these, 8 were found in Mathieson et al., Field et al. added 0, Le et al. added 3, Kerner et al. added 0, and Irving-Pease et al. added 7.