A selection pressure landscape for 870 human polygenic traits

Weichen Song (song628196@gmail.com)
Shanghai Mental Health Center https://orcid.org/0000-0003-3197-6236
Weihao Pan
Shanghai Mental Health Center
Wei Qian
Shanghai Mental Health Center
Weidi Wang
Shanghai Mental Health Center
Shunying Yu
Shanghai Mental Health Center
Min Zhao
Shanghai Mental Health Center
Guan Ning Lin
Shanghai Mental Health Center

Article

Keywords: Natural Selection, Summary Statistics, European Ancestry, Polygenic Adaptation, Genetic Studies

Posted Date: December 1st, 2020

DOI: https://doi.org/10.21203/rs.3.rs-90125/v1

Version of Record: A version of this preprint was published at Nature Human Behaviour on November 15th, 2021. See the published version at https://doi.org/10.1038/s41562-021-01231-4.

A selection pressure landscape for 870 human polygenic traits

Abstract

Characterizing the natural selection of complex traits is essential for understanding human evolution and biological or pathological mechanisms. To fulfill this requirement, we leveraged Genome-wide summary statistics for 870 polygenic traits and quantified the selection pressure of different forms and time scales on them in European ancestry. We found that 88% of traits underwent polygenic adaptation in the past 2000 years. At the present time and Neolithic period, selection pressure showed profound alteration. Traits related to pigmentation, impedance, and nutrition intake exhibited strong selection signals across different time scales. Our result provided an overview of selection pressure on various human polygenic traits, which served as a foundation for further populational and medical genetic studies.

Main

The genetic architecture of present-day human is shaped by the profound selection pressures in the long history[1]. Understanding the patterns of natural selection can provide valuable insights into the mechanisms of biological process[2], the origin of human psychological characteristics[3], and the historical events of anthropology[4]. For public health and clinical medicine, the study of evolution promotes our knowledge of disease mechanisms and susceptibility[5,6], and aids precision medicine by highlighting the intolerant genetic variants[7]. The explosive growth of all branches of anthropology, biology and medicine demands a comprehensive understanding of natural selection,

both for heritable diseases and non-disease traits.

Quantifying the selection pressure, especially on human polygenic traits, is a complex task[1]. Unlike simple traits dominated by a single gene or variant, selection pressure on complex traits often results in polygenic adaptation[8], where subtle modification on a large number of variants segregates into phenotypic alteration. Polygenic adaptation could accomplish different forms of selection, such as purifying selection, balancing selection, and hard and soft sweeps[1,8]. Furthermore, the revolutions of culture and productivity in human history have profoundly distorted the existing selection pressure on human society[9], which led to distinct adaptation patterns at different time scales. Undoubtedly, a comprehensive understanding of natural selection should cover all these aspects. So far, a few studies have managed to generate a multi-aspect picture of selection pressure for a single polygenic disease, such as attention-deficit/hyperactivity disorder[10] and schizophrenia[11]. Yet, an intact overview covering all types of human traits is still lacking.

With the tremendous advancement of Genome-Wide Association Study (GWAS)[12] and various efficient analytical tools of population genetics[13], we're now able to study the selection pressure of human polygenic traits from a multi-dimensional perspective. Here, we leveraged GWAS summary statistics of 870 traits and applied various methods to quantify their selection pressure in different forms and at different time scales. We also performed cross-sectional and longitudinal comparisons to illustrate the essential characteristics of human adaptation. Together, our result provides a reference landscape for future genetic studies regarding human

evolution.

Result

By filtration in traitDB[12] database and literature research (Method), we collected the GWAS summary statistics of 870 polygenic traits with adequate power, 738 of which were carried out primarily in the UK Biobank. These traits were separated into 15 categories (Figure 1 and Table S1). To comprehensively evaluate the selection pressure on them, we adopted different methods to quantify the natural selection at four different time scale: present time, recent history (2,000 to 3,000 years), pan-Neolithic period and since human speciation (Figure 1). These metrics were then integrated to analyze the relations and discrepancy of selection pressures among different time scales and traits.

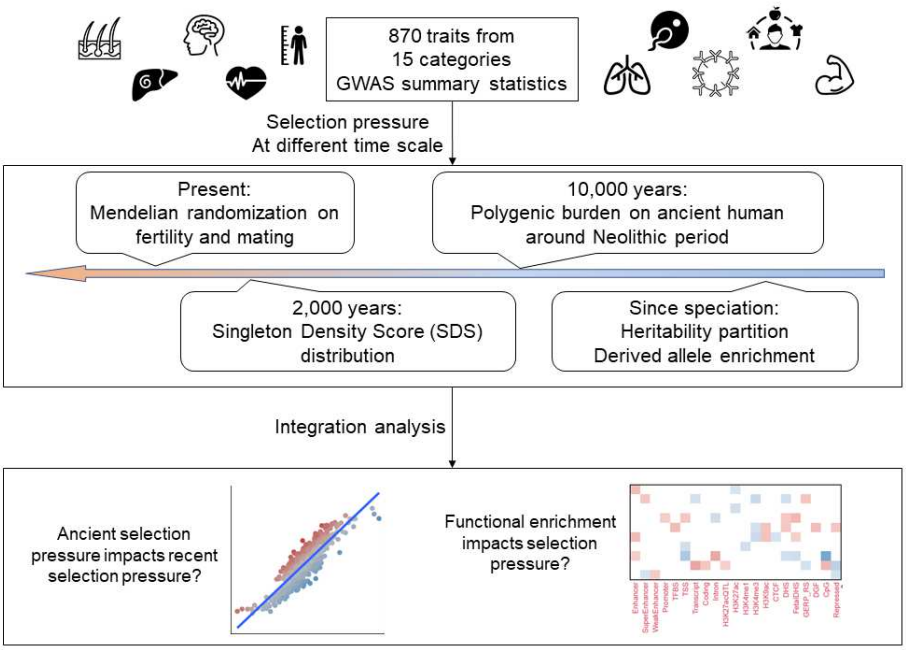

Figure 1 Flowchart of the study.

Widespread impacts of impedance traits on fertility in modern society

Our analysis started at the selection pressure at the present time. We applied Mendelian randomization (MR) to evaluate whether each trait could impact human fertility (i.e., number of children) and mating success (i.e., number of sexual partners). We found that among all 539 traits with valid MR results (Method and Figure S1), 40 (7.4%) of traits had a causal effect on male's number of children (ncm), whereas 32 (5.9%) impacted that of female's (ncf) (Table S2). Divided by category (Figure 2A and 2B), we found that 52% (23/44) of impedance traits (IMP) like height (Z_{ncm}=8.09, Z_{ncf}=4.91) and body mass index (BMI) (Z_{ncm}=7.11, Z_{ncf}=4.79) were causally related to male's number of children (for women, the proportion was 30% [14/47]). Effect of dermatology traits (DER) on fertility exhibited gender-specificity: 38% (5/13) of DER influenced ncm, but none affected ncf. The risk of any polygenic disease had no impact on fertility, providing no evidence for the previous hypothesis that some heritable diseases were not eliminated by natural selection because of their reproductive advantage[14].

As for mating success (Figure S2), IMP also had a profound impact (44% of IMP impacted male's number of sexual partners [nsm]; 12% affected that of female [nsf]). Interestingly, among all polygenic diseases, Schizophrenia (Z_{nsm}=7.37) and Attention deficit hyperactivity disorder (Z_{nsm}=4.62) increased nsm, in line with previous finding[14]. For males, the impact on fertility of a trait was profoundly positively correlated with its impact on the mating success (Figure S2; Pearson Correlation

Coefficient [PCC] =0.47, $p=9.30\times10^{-31}$); however, this was not true for female (Figure S2C; PCC=-0.10, p=0.02). This discrepancy supported the evolutionary psychology theory that male and female adopted distinct sexual strategies[15].

We then analyzed the difference in selection pressure between genders. In general, impacts on fertility were similar for males and females (Figure 2C; PCC= 0.38, $p=6.85\times10^{-31}$). A few exceptions included left leg impedance (Z_{ncm}=-4.30, Z_{ncf}=1.50) and Ease of skin tanning (Z_{ncm}=5.68, Z_{ncf}=-3.20). The impacts of mating success were even more similar between genders (Figure S2; PCC= 0.64, $p=9.18\times10^{-106}$). Notably, high intelligence could significantly reduce fertility, especially for females (Z_{ncm}=-5.13, Z_{ncf}=-7.55); however, it increased the expected number of sexual partners of females (Z_{nsf}=7.05) (Figure S1).

We also applied Causal Analysis Using Summary Effect Estimates (CAUSE)[16] to all MR results to analyze the role of genetic correlation. We found that most of the results were better explained by the causal effect (see supplementary methods and results for detail).

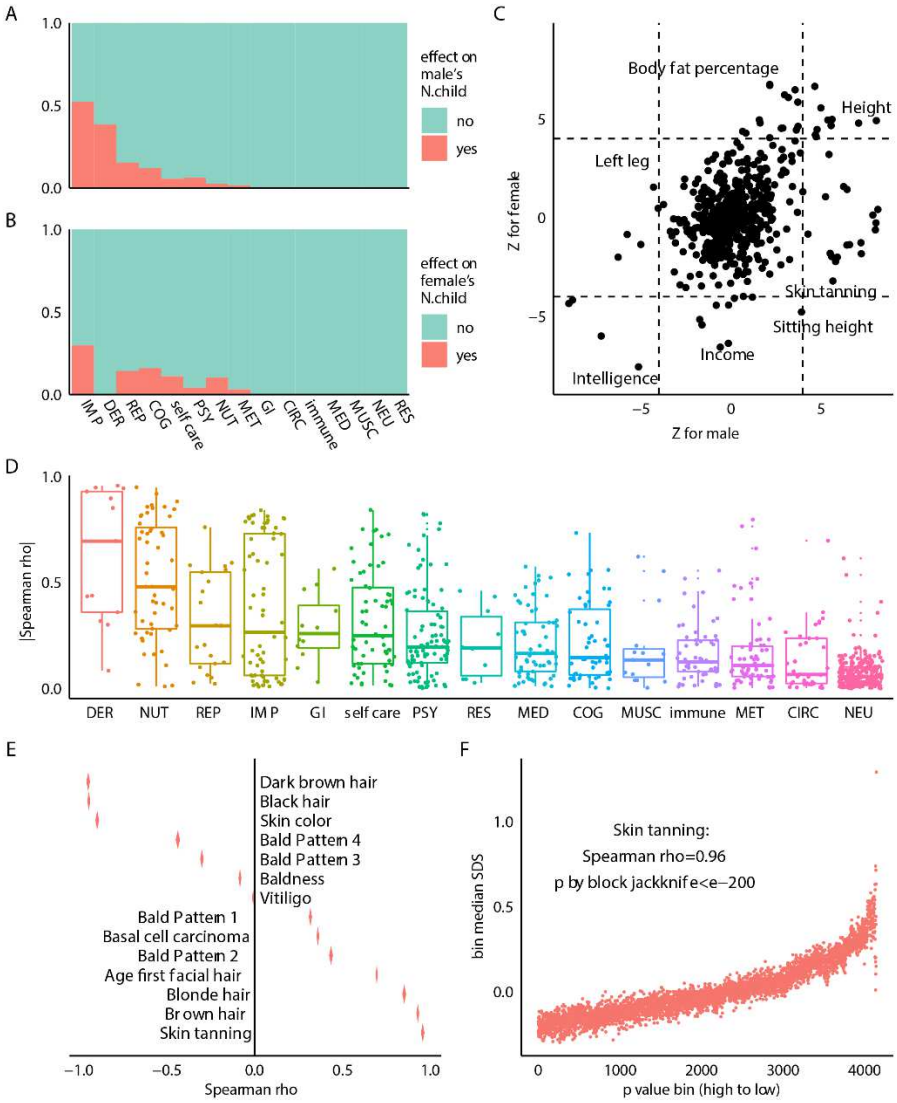

Figure 2 Selection pressure at present and recent history. A&B: Proportion of traits showing MR causal effects on number of children of male (A) and female (B) for each category. C: Comparison of MR z scores between male (x axis) and female (y axis). Dashed lines indicated significance threshold (|z|>4). Texts indicated selected traits with results of special interests. D: Distribution of absolute Spearman correlation ($|\rho_{SDS}|$) between tSDS and GWAS p value for each category. Upper and lower margins of box indicated first and third quartiles of ρ_{SDS}, and the thickened line indicated median ρ_{SDS}. E: ρ_{SDS} for all dermatology traits. The diagonal of the rhombus indicated ρ_{SDS}, and the width of rhombus indicated 95% confidence interval of ρ_{SDS}. F: Scatter plot showing the correlation between tSDS and GWAS p value for trait "Ease of skin tanning". Each point represented a bin of 1000 SNPs ranked by their GWAS p value. Y-axis indicated the bin median tSDS. DER: dermatology. NUT: nutrition. REP: reproduction. IMP: impedance. GI: Gastrointestinal. PSY: psychiatry. RES: respiratory. MED: medication. COG: social-cognition. MUSC: musculoskeletal. MET: metabolism. CIRC: circulation. NEU: neurology.

Most heritable traits underwent significant polygenic adaptation in the past 2000 years

Next, we extended our analysis to the recent history (past 2000-3000 years). Selection pressure at this time scale was measured by the Spearman correlation between SNP-trait association p-value and trait-increasing Singleton Density Score (tSDS), termed ρ_{SDS}, as applied by Field et al.[17] (Method). High tSDS for a SNP indicated that the trait-increasing allele of this SNP had an elevated frequency in the past 2,000~3,000 years. At the significance threshold of $p<0.05/870$, we found that 88% (761/870) of polygenic traits had a significant correlation between GWAS p-value and tSDS (ρ_{SDS}; Table S3). Previous analysis has found that population stratification between UKB and other GWAS might exaggerate the estimated polygenic adaptation[18]. In our study, GWAS from UKB showed a larger magnitude of ρ_{SDS} (p by permutation [pp]=0.001, Figure S3). This result was mainly driven by the excess number of DER and nutrition intake-related traits (NUT) in UKB (pp=0.08 after removing DER and NUT). We reasoned that the population stratification of GWAS outside UKB did not significantly inflate the observed adaptation. Thus, the high proportion (88%) of significant traits truly reflects recent adaptation prevalence.

To further verify this high prevalence of recent adaptation detected by ρ_{SDS}, we applied another method with a distinct statistic model, namely, Reconstructing the History of Polygenic Score[19] based on RELATE[20] (RHPS-RELATE, method and discussion). As shown in Table S3, the Polygenic Risk Score (PRS) alteration in the past 100 generations (roughly equivalent to 2,800 years[20]) was generally in

accordance with ρ_{SDS} (PCC=0.25, p=3.96×10^{-13}). Among 755 traits with significant non-zero ρ_{SDS}, 13.8% (104/755) showed consistent significant alteration of PRS (p by Tx test [pt]<0.05/870, method), and 26.1% (197/755) showed nominal significant alteration (pt<0.05). Notably, RHPS-RELATE also highlighted those traits with most profound ρ_{SDS}, such as Ease of skin tanning (p for $\rho_{SDS}<10^{-100}$; $pt<10^{-100}$) and raw vegetable intake (p for $\rho_{SDS}<10^{-100}$; pt=2.69×10^{-51}) (Table S3). In general, RHPS-RELATE supported the result of ρ_{SDS} analysis, albeit at a smaller statistical power. In the following section, we still considered ρ_{SDS} as the main result.

Among all traits, DER generally showed the most significant signals (median $|\rho_{SDS}|$=0.69, figure 1D-E), followed by NUT (median $|\rho_{SDS}|$=0.48; ρ_{SDS} =-0.95 for 'raw vegetable intake', Figure S3) and reproduction-related traits (REP; median $|\rho_{SDS}|$=0.30; ρ_{SDS} =-0.58 for 'Polycystic ovary syndrome', Figure S3). Ease of skin tanning was the trait with the most significant adaptation (ρ_{SDS} =-0.96; Figure 1F). Ever drinkers (ρ_{SDS} =-0.82) and sitting height (ρ_{SDS} =0.84) were also among traits with an extreme signal of adaptation ($|\rho_{SDS}|$>0.8), which made up 3.3% of all traits (Figure S3). Neurological traits like brain structures exhibited the least polygenic adaptation (median |ρ|=0.05).

Compared with non-disease traits, the adaptative pressure on polygenic disease was generally negative (median ρ_{SDS} =-0.08; permutation p=3.22×10^{-6}), especially for early-onset disease (median ρ_{SDS} =-0.12; Figure S4). However, the magnitude of adaptative pressure for diseases was similar to traits (median $|\rho_{SDS}|$=0.16, permutation p [disease vs. trait] =0.1). The most profound negative adaptation was found for high

cholesterol (ρ_{SDS} =-0.66, Figure S4). Exceptionally, we found positive adaptation on a few diseases like skin cancers and inflammatory bowel diseases (ρ_{SDS} >0.2; Figure S4), and even early-onset diseases like Attention deficit hyperactivity disorder (ρ_{SDS} =0.20) and Anorexia nervosa (ρ_{SDS} =0.16) (Table S3). This result suggested that some diseases might be by-products of other positive selection events.

Hunter-gather ancestry impacted natural selection around Neolithic

During the Neolithic (about 10,000 years ago[21]), human society underwent profound revolutions, and the selection pressure is also believed to be distinct from both ancient times and recent history[21]. To quantify the selection pressure at pan-Neolithic period, we downloaded three ancient human genome datasets (Neolithic[22], pro-Neolithic[23] and near east farmer[24], Table S4) and calculated the polygenic burden (measured by both allele counts and polygenic scores; Method and Figure S5) for each trait on all ancient human[10]. We applied linear regression to see whether the polygenic burden was altered along time and percent of hunter-gather ancestry (%HG). As shown in Figure 3A and Table S5, after controlling covariances (e.g., latitude, longitude, genotyping coverage, etc.) and multiple tests, the polygenic burden of 78 traits was significantly associated with %HG, whereas six traits were associated with time in one of three datasets. DER (7 out of 13) and NUT (18 out of 52) were most predominantly associated with %HG (Figure 3A). Ease of skin tanning was the most significant trait (t_{HG}=20.3, p=1.74×10^{-38} [Figure 3B]). In the near east dataset, we observed that the selection pressure on skin tanning was dependent on latitude (Figure 3C): it was favored by selection at low latitude region (latitude<50), but was

suppressed at high latitude region. After controlling the impact of latitude, we found a general ascending trend of it ($t_{neareast}$=5.81, p=2.29×10^{-8} [Figure 3C]). We also found a nominally significant increment in the pro-Neolithic period ($t_{proneolithic}$=4.25, p=0.0009), but a non-significant increment in Neolithic ($t_{neolithic}$=0.92, p=0.36), for Ease of skin tanning (Figure S6).

By analyzing the t statistics for all traits (Figure 3D), we found that t_{HG} was positively associated with $t_{neareast}$ (PCC=0.55, p=6.27×10^{-69}) and $t_{proneolithic}$ (PCC=0.61, p=4.59×10^{-89}), but was negatively correlated with $t_{neolithic}$ (PCC=-0.29, p=1.50×10^{--17}). This result suggested that traits related to hunter-gatherer ancestry were favored by natural selection in the pro-Neolithic period and near east farmer society but were suppressed by natural selection at the Neolithic period. This pattern was also true for polygenic diseases, albeit at smaller magnitude (PCC=0.18, 0.48, -0.23, respectively, green points and texts in Figure 3D). This pattern was partly because polygenic diseases had less association with %HG (median $|t_{HG}|$=0.85 for disease, permutation p [disease vs. trait] =7.96×10^{-12}). As shown in diagonal plots in Figure 3D, we also confirmed that polygenic diseases generally suffered more negative selection pressure than non-disease traits. This was most significant in near east farmer (median $t_{neareast}$=-0.66, permutation p =4.31×10^{-5}) and pro-Neolithic period (median $t_{proneolithic}$=-0.42, permutation p=0.004), but not significant in Neolithic period (median $t_{neolithic}$=-0.17, permutation p [disease vs trait] =0.34). Among the 13 exceptions (i.e. diseases with $t_{neareast}$, $t_{proneolithic}$ or $t_{neolithic}$>0 and p<0.05), we found immunological diseases like Crohn's disease ($t_{proneolithic}$=2.86, p=0.013), Atopic dermatitis ($t_{neolithic}$=2.61, p=0.01)

and periodontitis ($t_{proneolithic}$=2.48, p=0.029), as well as fracture ($t_{neareast}$=2.73, p=0.007) (Figure S6). These diseases were positively selected during the pan-Neolithic period.

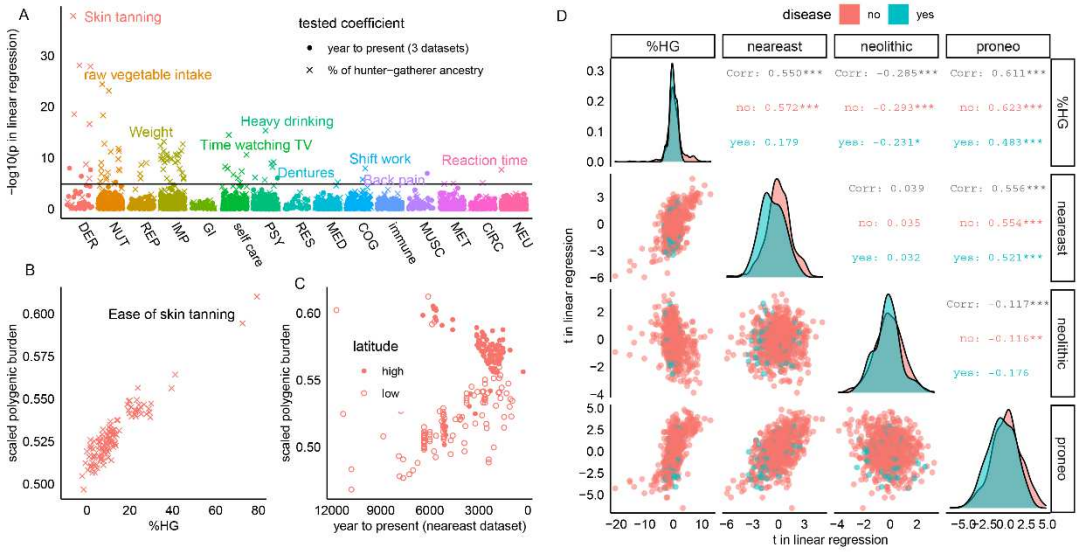

Figure 3 Selection pressure at pan-Neolithic period. A: Manhattan plot showing the p value of linear regression. The regression was between scaled genetic burden and either time to present (round dot) or percent of hunter-gatherer ancestry (%HG) (cross). B: Relation between genetic burden of "Ease of skin tanning" and %HG. Each dot denoted an ancient human in Neolithic dataset, its y-axis showed the genetic burden of "Ease of skin tanning" carried by this individual, and the y-axis showed the %HG of this individual. C: Similar to B, but for living times in neareast dataset. D: Relation among four selection metrices. Each dot represented a trait, and its x and y axis showed the t value of linear regression on two out of four selection metrices (Neolithic time, Neolithic %HG, neareast time and pro-Neolithic time). Red color corresponded to non-disease traits, and green color corresponded to polygenic diseases. Texts in upper triangle showed the Pearson Correlation Coefficients for symmetric plots in lower triangle. *: p<0.05; **: p<0.01; ***: p<0.001. Diagonal plots showed the distribution of t values.

SNPs at regions under background selection and constraint genes exhibited significant heritability enrichment of polygenic traits

To expand our analysis to a more ancient time scale, we collected several metrics that detected genomic regions undergoing different forms of selection[11,25–29], then applied linkage disequilibrium (LD) score regression (LDSC)[30] to see whether heritability of each trait enriched in these regions (Method and Figure S7). As shown in Figure 4A and Table S6, we detected widespread heritability enrichment in genomic regions with

low average coalescence times[26] ($ASMC_{avg}$, a metric that measures background selection in the past several hundred thousand years), around mutation-sensitive genes (indicated by high probability of LOF-intolerant, pLi[31]) and in regions with low LD or high conservation[30]. For $ASMC_{avg}$, the p-value by LDSC was significantly inflated compared to the null distribution (Figure 4B). Traits showing highest enrichment in low $ASMC_{avg}$ regions included body water mass (z=-7.32), intelligence (z=-5.65) and schizophrenia (z=-5.55) (Figure 4C). Similar enrichment was also observed for mutation-sensitive genes, especially for schizophrenia (z=6.23), intelligence (z=4.61) and neutrophil count (z=4.40) (Figure S8). Consistently, variants with high deleteriousness (measured by CADD[7]) were also significantly associated with polygenic traits (Figure S8). Heritability for traits like large artery stroke (z=8.84) and ever drinkers (z=7.68) were enriched in high-CADD variants whose alternative allele increased the traits (CADD+) (Figure S8). We also analyzed other forms of selection (Method and Table S6). We found that heritability of Large artery stroke was significantly enriched in regions undergoing soft sweep (z=4.08), whereas heritability of beer intake enriched in genomic regions suffering from balancing selection (z=3.83).

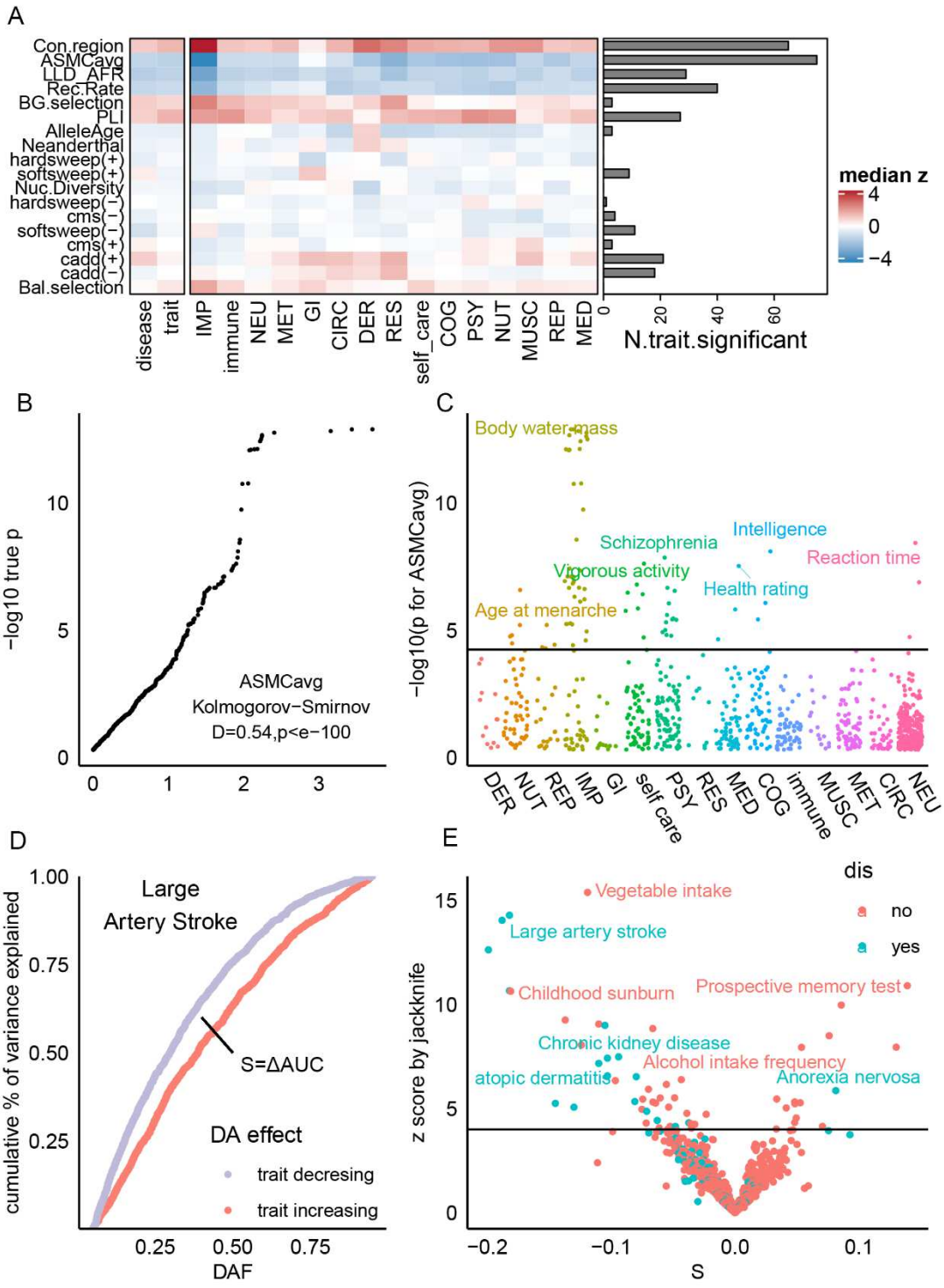

Figure 4 Selection pressure since human speciation. A: Heatmap showing the median LDSC enrichment Z score on each genomic annotation for each category. Bar plots denoted the total number of traits showing significant enrichment in the corresponding annotations. B: QQ plot for LDSC p value of ASMCavg enrichment. C: Manhattan plot for ASMCavg enrichment. D: Effect-frequency distribution for trait "Large artery stroke". Each dot denoted a SNP, x-axis showed its derived allele frequency (DAF), y-axis showed the cumulative proportion of variance explained by all SNPs with DAF smaller than this SNP. We separated SNPs according to their DA effect (increase or decrease the trait). The area between the curve of these two distributions, named S, was used to measure natural selection. E: Volcano plot for the statistical analysis of S.

By comparing the number of traits reaching the significance threshold for each annotation between diseases and non-disease traits (Figure S9), we found that CADD+ was predominantly associated with polygenic diseases (Odds Ratio [OR]=9.58, $p=5.69\times10^{-6}$). CADD+ z scores for diseases were also generally larger than non-disease traits (pp=0.0002, Figure S9). This result confirmed that highly deleterious variants make more contribution to diseases than non-disease traits. Interestingly, Atrial fibrillation (z=3.45), Anorexia nervosa (z=3.35) and Rheumatoid Arthritis (z=3.16) showed heritability enrichment in deleterious variants whose alternative alleles decreased disease risk (CADD(-), Table S6). This result suggested that these diseases' risk might be increased by the negative selection eliminating their protective alleles. Additionally, we found that Conserved regions (pp=0.003, Figure S9) and low ASMCavg regions (pp=0.07, Figure S9) tended to contribute to the non-disease traits.

Cerebrovascular disease suffered natural selection since human speciation

Lastly, we analyzed the overall selection pressure since human speciation by calculating the distribution of derived allele (DA) effects, similar to a previous tool, SbayesS[32]. Since SbayesS did not consider the direction of the DA effect, we borrowed their idea and calculated the difference of area under the effect-frequency distribution curve ($S=\Delta AUC$; Figure 4D and method) to measure the purifying selection since speciation (Methods and Table S7). At the significance threshold of $p<0.05/870$, 67 traits had S not equal to zero, which were mainly related to nutrition

(15), neurology (11), and psychiatric (9). For cerebrovascular diseases like large artery stroke (Figure 4D), the SNPs whose DA had a large risk-increasing effect generally had small DA frequency (S=-1.88, z=-14). This was also true for different kinds of stroke or intracerebral hemorrhage (Figure S10), suggesting an overall negative selection of cerebrovascular disease. On the other hand, memory was favored by natural selection (S=0.14; Figure 4E). Polygenic diseases were generally suppressed by selection (mean S=-0.03; permutation p [disease vs. trait] =8.56×10^{-6}; Figure S10). However, primary sclerosing cholangitis, anorexia nervosa and atrial fibrillation showed signals of being favored by natural selection (S>0 and p<0.05/870; Figure 4E and Figure S10).

Another interesting fact is that most traits had AUC>0.5 for both trait-increasing DA (AUC_{inc}>0.5: 868 out of 870 traits) and trait-decreasing DA (AUC_{dec}>0.5: 860 out of 870 traits) (Table S7), supporting the notion that purifying selection widely exists in human traits[33]. The highest purifying selection was observed for the blood level of Lipoprotein A (AUC_{inc}=0.70, AUC_{dec}=0.64, Table S7). However, we found exceptions on extreme height (AUC_{inc}=0.45, AUC_{dec}=0.45, Table S7) and extreme BMI (AUC_{inc}=0.46, AUC_{dec}=0.45, Table S7), suggesting a bi-directional positive selection on extreme impedance.

We additionally analyzed the distribution of trait-increasing alleles among derived and ancestral alleles (Table S7 and Figure S10). We found that alleles that promote ten traits, including anorexia nervosa, fresh fruit intake, hair and skin color, etc. were mainly ancestral (percent of trait-increasing DA [%DA] <50, p by binomial

test [pb]<0.05/870). In contrast, those promote ten traits like cerebrovascular diseases, alcohol intake and gout, etc. were mainly DA (%DA>50, pb<0.05/870) (Figure S10 and Table S7). Compared with non-disease traits, polygenic diseases were primarily promoted by DA (pp [disease vs. trait] =0.002, Figure S10).

Dramatic change of selection pressure at Neolithic and present time

So far, we have quantified the selection pressure at four different time scales, which allowed us to analyze the relation among these selection pressure. We reasoned that if environmental pressure were identical throughout the history, strength of selection at a later time would be dependent on that of ancient times, and the violation of this principle may reflect a modification of environmental pressure. Thus, we applied linear regression on the scaled selection metrics (Method and Table S8) to analyze whether ancient selection strength could predict recent selection. As shown in Figure 5A, ρ_{SDS} could be best predicted by ancient selection metrics (R^2=0.74, $p<10^{-100}$), especially by %HG ($p=3.60\times10^{-62}$) and $t_{neareast}$ ($p=3.28\times10^{-35}$). On the other hand, ncm (R^2=0.09, $p=3.45\times10^{-11}$) and $t_{neolithic}$ (R^2=0.13, $p=1.80\times10^{-23}$) were poorly explained by ancient selection metrics. For all seven tested metrics in Figure 5A, the prediction was more precise on non-disease traits than disease (Table S9), and the jackknife p-value for five of them was lower than the significance threshold of 0.05/7. For example, R^2 for ρ_{SDS} prediction was 0.76 on non-disease traits but dropped to 0.47 when applied on polygenic diseases (p by jackknife $=9.95\times10^{-5}$, Table S9 and Method). This result suggested that selection pressure at Neolithic and the present time was profoundly altered and deviated from ancient selection pressure, and the

deviation was more profound for polygenic diseases.

We further analyzed the traits that contribute to this selection deviation. For ncm (Figure 5B and Table S10), traits like arm impedance (Z score of residuals [Z_{resid}]=-5.54), educational attainment (Z_{resid} =-3.85), Loneliness (Z_{resid} =-3.48) and intelligence (Z_{resid} =-2.71), etc. had significantly lower Z_{ncm} than predicted by ancient selection metrics, whereas traits like Bald pattern 1 (Z_{resid} =2.75) tended to have higher Z_{ncm} than expected. For $t_{neolithic}$ (Figure 5C and Table S10), the volume of left vessel (Z_{resid} =3.30) and atopic dermatitis (Z_{resid} =3.20) had larger-than-expected $t_{neolithic}$, whereas trouble of concentration (Z_{resid} =-3.37) was lower than expected. We also summarized traits showing the largest deviation measured by other metrics (ncf, ρ_{SDS}, $t_{neareast}$, t_{proneo} and t_{HG}; Figure S11 and Table S10). We found that traits like intelligence, back pain and pigmentation-related traits showed deviation across many different time scales.

To gain a continuous view on the adaptation trajectories, we applied RHPS-RELATE to infer the population-mean PRS trajectory of each trait, then applied time series clustering to elucidate its pattern. As shown in Figure 5D and Table S11, the trajectory of 434 and 308 traits were grouped into clusters 1 and 2, respectively, which generally showed accelerating monotonic increasing or decreasing trends since about 500 generations ago. Typical representatives were Raw vegetable intake (Figure S12, pt between 496 and 96 generations $<10^{-100}$) and Atopic Dermatitis (Figure S12, pt between 496 and 96 generations $<10^{-100}$). On the other hand, 13 and 10 traits were grouped into clusters 3 and 4, characterized by a sharp turnover of adaptation directions around 133 generations ago (Figure 5D). Thes traits included intelligence

and Insomnia (Figure S12), etc.

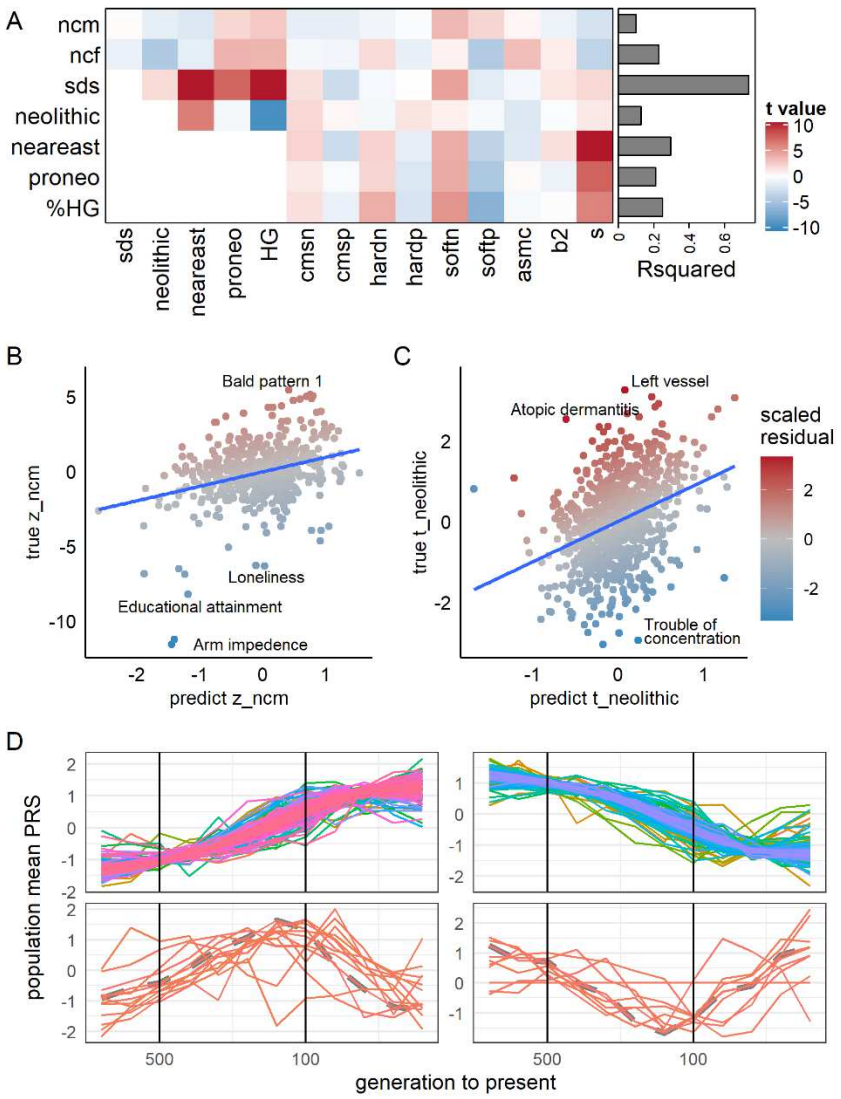

Figure 5 Relation among selection pressure at different time scales. A: Heatmap showing the t value of linear regression which used ancient selection pressure (columns) to predict recent selection pressure (rows). Bar plots denoted the R2 for corresponding linear regression. B: Each dot denoted a trait; x-axis showed the Zncm predicted by linear regression, y-axis showed the true Zncm, and color denoted the scaled residual in the linear model. C: Similar to B, but for tneolithic. D: Population-average polygenic risk score trajectory for 765 traits, grouped into four clusters according to their time series similarity. Y-axis showed z-score of PRS.

Functional genomic architectures of polygenic traits had a moderate impact on the selection pressure

Despite the relation among selection pressure at different time scales, we were also interested in their relationship with the traits' genetic architectures. Thus, we applied step-wise linear regression on each selection pressure metrics to explore whether the genetic characteristics (e.g., functional genomics enrichment, %DA, variant deleteriousness, etc.) could determine the selection pressure of the trait. We found that functional genomic patterns explained 7% (pro-Neolithic) ~ 18% (ASMCavg) of the variance in selection pressure (R^2=0.07~0.18; $p<8.22\times10^{-6}$; Figure 6A). Adding conservation annotations (annotations that are directly related to natural selection like LLD and allele age[30]) into the model increased the R^2 by 0.02 (Neolithic) ~ 0.49 (SDS). This increment was mainly driven by the inclusion of CADD(+) and CADD(-) (Figure S13); for the model of S, R^2 increased from 0.11 to 0.58 after the inclusion of CADD(+) and CADD(-).

We also analyzed the sign and significance of regression coefficients for each linear model (Table S12). As expected, the regression coefficient of CADD(+) was negative in seven out of ten linear models (i.e., traits promoted by high CADD variants were negatively selected; Figure 6A), especially for ρ_{SDS} ($t_{CADD(+)}$=-15) and S ($t_{CADD(+)}$=-14, Figure 6B). An unexpected exception was ASMCavg ($t_{CADD(+)}$=3.91), where high CADD(+) enrichment led to low enrichment in region under background selection (Figure S13). Another unexpected result was the positive relationship

between mutation-sensitive genes and Neolithic natural selection (t_{PLI}=3.47), which indicated that traits with heritability enriched in mutation-sensitive genes tended to be positively selected at Neolithic (Figure S13). We also found significant contribution of %DA to SDS (t_{DA}=16, Figure 6C) and %HG (t_{DA}=12). For all functional genomics annotation, the most significant relation was found between CpG and ASMCavg (t_{CpG}=-8.5, Figure 6D).

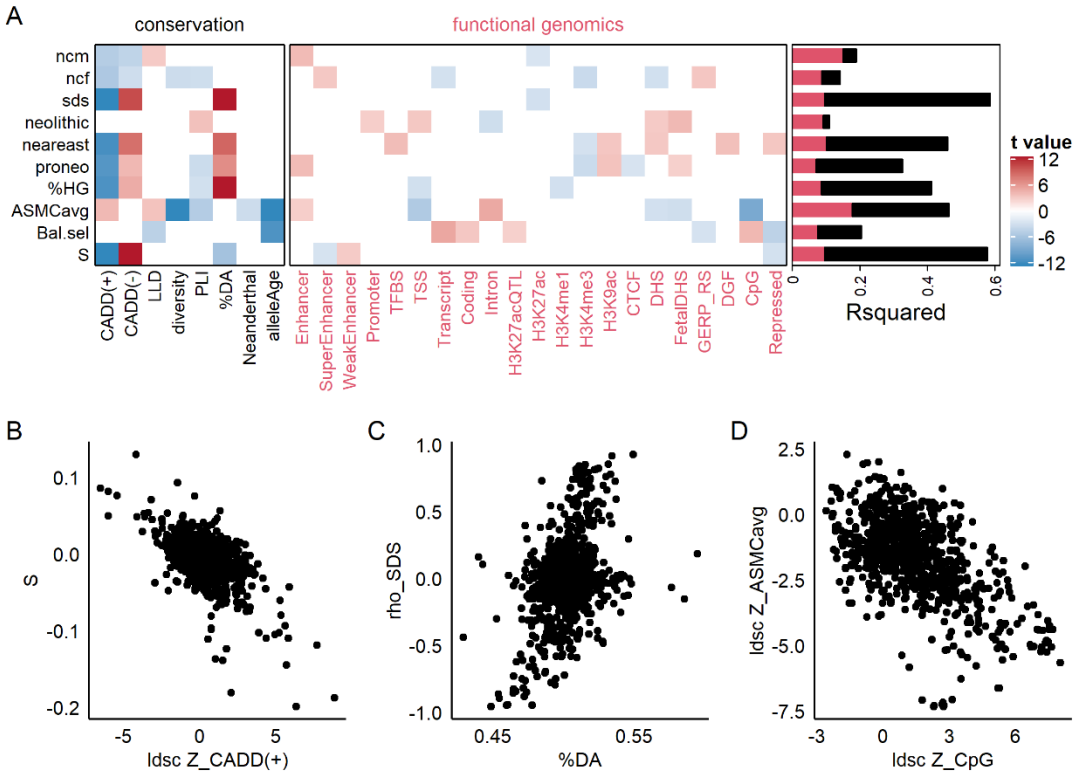

Figure 6 Genomic architectures impacted selection pressure. A: Heatmap showing the t value of linear regression which used functional genomic architecture and genetic conservation characteristic (columns) to predict selection pressure (rows). Red bars denoted the R^2 for corresponding linear regression using functional genomic architectures alone, and black bar denoted R^2 for linear regression using all predictors. B-D: Scatter plots showing the most significant contribution of genomic characteristics to selection pressures. Each dot represented one trait.

Discussion

In the current study, we quantified the selection pressure of human polygenic traits at four different time scales and in various forms. We analyzed the essential characteristics of selection pressure, such as its prevalence and strength, its uneven distribution among time points and trait categories, as well as its association with genetic architectures.

By analyzing the tSDS correlation and PRS trajectory, we found a widespread recent polygenic adaptation among different kinds of traits. The observation that polygenic adaptation was common among complex traits has long been questioned by researchers[8]. For one thing, the population stratification is known[18] to inflate the signal of ρ_{SDS}; for another, existing studies on polygenic adaptation usually focused on single trait[34,35]. In our study, the use of RHPS-RELATE could overcome the challenge of false-positive observation. First of all, false-positive ρ_{SDS} findings were mainly driven by a large number of SNPs with small effects[18,20], whereas RHPS-RELATE only analyzed top loci with large effect[19,20]. Secondly, we included various European populations from 1000G into RHPS-RELATE, which compensated the population stratification of GWAS. Lastly, RHPS-RELATE relied on a different statistical test (Tx test) to analyze the significance of adaptation, such that the potential inherited bias of ρ_{SDS} was avoided. Since ρ_{SDS} and RHPS-RELATE gave convergent results, we suggest that the widespread recent polygenic adaptation was plausible.

We found that pigmentation, impedance, and dietary traits were continuously under intense selection pressure across various time scales. Pigmentation is one of the most thoroughly studied examples of human evolution. The tremendous spatiotemporal variations of skin color reflected the complex balancing between UV damage, Vitamin D requirements and heat regulation[36]. With Ease of skin tanning as an example, our result also revealed a complex evolutionary history of pigmentation: dark skin was promoted before the Neolithic period and in recent history but had inconsistent adaptation during the Neolithic period. The body size and dietary habits, on the other hand, were mostly shaped by the trade-off among energy allocations on growth, maintenance, digestion and other functions[37,38]. Our result also suggested that among other factors that might impact energy allocations, such as ecology, climate and migration, genetic factors profoundly influenced the evolution of impedance and nutrition intake traits.

We also discovered a profound deviation of selection pressure at the Neolithic period and present time, in accordance with the radical change of culture and society at these periods. The tremendous agricultural revolution at Neolithic[39] and the industrial revolution in the past centuries[40] have thoroughly reshaped our society. This was accompanied by genetic reformation at loci related to diet[41], disease susceptibility[42] and reproductive behaviors[43]. Our result expanded these findings to a broader view of polygenic adaptation: genetic reformation at Neolithic and the present time was common and widespread, not restricted to a few traits or variants. This finding also supported the notion that the social revolution could significantly distort

the natural selection of human beings[9].

Another counterintuitive discovery is the positive adaptation of diseases like Anorexia nervosa and inflammatory bowel diseases (IBD). In the evolutionary perspective of Anorexia nervosa[44], foraging for food is typical behavior when facing the threat of starvation, thus will be favored by selection at periods of food supply shortage. As for IBD, researchers have suggested that the disease vulnerability may be associated with high defense against pathogen[45], which provided survival advantages in ancient societies with poor sanitary conditions. Since human beings have just solved starvation and sanitation threats recently and incompletely, natural selection does not have enough power to eliminate these diseases. Consistently, our findings further indicated that these diseases might be favored by natural selection at specific time points.

Our study has limitations. The currently available large-scale GWAS were dominated by European participants[46], especially UK Biobank[47], which significantly restricted the universality of GWAS-based genetic studies. Inevitably, our result had little power to dissect mainland Europe's subpopulation, not to mention the broad population in the rest of the world. The power to explore more ancient history (more than 100,000 years ago) is also limited since the available tools suitable for such a long time scale could only detect a few sweeps at a single loci[1]. In the future, the development of multi-ethnic GWAS, ancient human genome analysis, and analytical tools for more extended time scales will eventually achieve the intact landscape of human evolution.

In conclusion, we provided a global overview of natural selection on human polygenic traits and its essential characteristics, which could serve as a foundation for future studies regarding human genetics and evolution.

Method

GWAS filtration and preprocessing

We downloaded all GWAS summary statistics from traitDB[12] release 1 and retained those conducted solely in cohorts of European ancestry. Since traitDB was released at November, 2019, we additionally conducted literature research to search for all GWAS of European ancestry published between October, 2019 to April, 2020. We downloaded only GWAS summary statistics that were publicly available. All these GWAS were filtered according to the following criteria: sample size >10,000, SNP-based heritability (h2) calculated by LDSC >0.01 and z score of h2>4. For the duplicated phenotypes in the remaining GWAS, we chose those GWAS with larger sample size, with more participants outside UKB, with more professional definition of phenotypes (e.g. diagnosed disease rather than self-reported disease), and those conducted by professional consortium. For some meta-analysis that reported only z score of each variant, we used the result from its largest cohort with detailed summary statistics (i.e. effect size and standard error) as a substitute. If no detailed summary statistics were available, we did not include it in our study. We separated all included GWAS into 15 categories, which were slightly modified from the definition of traitDB (Table S1 and supplementary method). For all included polygenic diseases, we

additionally separated them according to the onset age: diseases that preliminary onset before reproductive age (18 years old) were labeled "early", diseases that preliminary onset after reproductive age (50 years old) were labeled "late", and the remaining diseases were labeled "lifetime".

We manually modified all summary statistics into a uniform format. Specifically, we log-transformed all Odds Ratio to ensure zero-centered effect size and modified the sign of the effect size to ensure that 1) A1 allele was the effect allele; 2) positive effect size was corresponded to literally "trait increasing" (see example in supplementary methods). We further removed all variants without rs ID, not recorded in 1000 Genome phase 3 (1000G)[48], had missing information, or had MAF<0.01 in European participants of 1000G. Since many of the GWAS did not provide allele frequency, we uniformly discarded all frequency information and used 1000G alternative allele frequency instead.

Mendelian randomization

To measure the human fertility and mating success, we downloaded the GWAS summary statistics of number of children (ncm and ncf) and number of sexual partners (nsm and nsf) for both sexes from Benjamin Neale Lab (http://www.nealelab.is/uk-biobank). For each of the 870 traits, we selected the SNP with $p<5\times10^{-8}$ as the instruments. We retained all instruments that were presented in outcome GWAS, then pruned at the threshold of LD>0.01 in 1000G. The data harmonization was applied separately for each exposure-outcome pair by TwoSampleMR R package[49].

For each exposure-outcome pair, we first calculated the per-instrument MR effects by Wald ratio, then meta-analyzed the results for all instruments with three methods: 1)Inverse variance-weighted (IVW), which was considered the primary results; 2) Weighted median (WM) method[50], which was relatively robust when some of the instruments were invalid; and 3) Egger regression[51], which allowed for non-zero directional pleiotropy.

Adjustment of pleiotropy and genetic correlation for Mendelian randomization

To get rid of the influence of outliers and pleiotropy in a uniform manner, we applied a step wise outlier removal test for each exposure-outcome pair. Specifically, we first applied three sensitivity tests (Cochran's Q test, Rucker's Q test and Egger intercept test)[52] on all instruments. If p values of any of these tests <0.05, we applied MR-PRESSO outlier test[53] to calculated the observed Residual Sum of Squares (RSS_{obs}) for all instruments, and ranked them in the descending order of RSS_{obs}. We removed top 1 instrument and repeated the three MR analysis and three sensitivity analysis on the remaining instruments. If p values of any sensitivity tests were still <0.05, we repeated this procedure by removing top 2, top 3, … top (n-3) instruments, until all sensitivity tests had p value >0.05 (leftmost black point in Figure S13). The MR results at this step were considered the final result. If p values of any sensitivity tests were <0.05 throughout the procedure, we denoted the MR results for this exposure-outcome pair as NA. For exposure-outcome pairs with less than three instruments, we provided the MR results (Wald ratio or IVW) in the Table S2, but did not consider them in the downstream analysis. After we got outlier-free MR results for

all pairs, we defined the significant causal effect as following: z score of IVW estimation >4 or <-4, estimation of IVW, WM and Egger regression were of the same direction. For those results not reaching significance criteria, we still included them in the correlation analysis (Figure 2C). We also applied CAUSE to distinguish causality from genetic correlation (see supplementary method and result).

tSDS analysis

Singleton density score (SDS)[17] is a metric that measures the density of singleton mutations on a haplotype tagging a tested SNP. Based on the assumption that positive selection distorts the ancestral genealogy of haplotype and leads to shorter terminal branches for the favored allele, SDS>0 indicate that derived allele (DA) of the tested SNP has an increased frequency in the recent history (about 100 generations)[17].

For each trait, we only included SNPs with DAF>0.05 and <0.95, and had a SDS calculated by Field et al.[17] using UK10K data[54]. Then, following the procedure of SDS authors[17], we modified the sign of SDS for each SNP according to the effect direction of its DA, which resulted in a tSDS metric. tSDS>0 indicated that the trait-increasing allele had an increased frequency. We re-normalized tSDS per DAF bin of 0.01 using z score method. We ranked all SNPs in the ascending order of their GWAS p value, and grouped them into consecutive bins of 1000 SNPs each. We calculated ρ_{SDS} as the Spearman correlation coefficients between median tSDS for each bin and the order of bins. To estimate the p value of ρ_{SDS}, we ordered all SNPs according to their physical positions (all chromosomes were treated as a concatenated meta-chromosome) and divided them into 1000 consecutive blocks. We applied jackknife

procedure by sequentially removed one block and repeated the ρ_{SDS} calculation on the remaining SNPs. We calculated the jackknife standard error to estimate the p value and 95% CI for ρ_{SDS} under uniform distribution.

Previous study[18] has found that combining GWAS outside UKB and SDS of UKB might cause false positive ρ_{SDS} due to population stratification. Since we found large proportion of significant ρ_{SDS}, we tested whether the significance was mainly contributed by GWAS outside UKB. Specifically, we separated all GWAS according to whether more than 70% of participants were from UKB, then compared the average $|\rho_{SDS}|$ between two groups by permutation test.

Polygenic burden of ancient human

To analyze the pan-Neolithic trajectory of polygenic burden for each trait, we downloaded ancient human genotype data for three studies, as Esteller-Cucala et al.[10] did: pro-Neolithic[23] (51 individuals), Neolithic[22] (180 individuals) and near east farmers[24] (281 individuals). The genotype data were transformed into ped and bim files using EIGENSOFT v6.1.4[55] and plink v1.07[56], with only SNPs that had a rs ID retained. For each trait and each ancient dataset, we retrieved all SNPs with GWAS $p<0.01$ and applied LD pruning in the ancient dataset using plink with parameter –indep 50 5 2 to get a list of independent trait-associated SNPs (tSNP). We excluded individuals with >90% missing information on the tSNP. Similar to the work of Esteller-Cucala et al.[10], we calculated the scaled polygenic burden f for each individual as following:

$$f = \frac{Number\ of\ copies\ of\ trait\ increasing\ alleles}{2 * Number\ of\ genotyped\ tSNP}$$

f is a metric between 0 and 1 that measured the percentage of polygenic burden this individual carried (f=1 indicated that the individual had homozygous trait-increasing allele on all non-missing tSNP). As a positive control, we also replaced the allele counts by the polygenic risk scores and repeated the entire analysis (see Supplementary methods and results).

In each dataset, we fitted a linear model to test for the relation between f and time to present, which reflected the polygenic adaptation on the traits. We included latitude, longitude, genotyping technique, sex, whether performed damage restrict, genotyping coverage, number of SNPs genotyped, fraction of library and inferred time of admixture as covariates (some of the covariates were not provided in particular datasets, see Supplementary method for detail). For Neolithic dataset, the percentage of hunter-gatherer ancestry (%HG) was also included as a predictive variable. From the regression result, we retrieved the t-statistics for time to present (t_{proneo}, $t_{neareast}$ and $t_{neolithic}$, respectively) and %HG (t_{HG}) as well as their p value as the measurements of polygenic adaptation. We also applied Pearson correlation analysis among these four metrices by GGally R package[57].

Heritability partition on genomic regions exhibiting different evidences of natural selection

We adopted a strategy similar to Pardiñas et al.[11] which first identified genomic regions under different selection pressure then partitioned the heritability of each trait to these regions by LDSC[30]. We obtained and reformatted following genomic annotations from literature (under hg19 position):

1) B2[29]. B2 is a metric of a set of composite likelihood ratio test statistics that are based on a mixture model. Regions with high B2 statistics were under balancing selection in about 10,000 generations. We downloaded the B2 statistics calculated by BalLerMix[29] on 1000 Genome CEU population[48], and annotated each SNP according to the B2 statistics of the region that covered it.

2) Combined Annotation Dependent Depletion (CADD)[7]. We downloaded CADD v1.3 for all 1000G SNV and indels and dichotomized all variants at the threshold of CADD-phred score >20. We further generated trait-specific annotations (CADD+ and CADD-) according to the effect of alternative alleles, i.e. variants whose alternative alleles increase the trait and had CADD-phred>20 were annotated "1" for CADD(+), whereas all other variants were annotated "0". Similarly, variants whose alternative alleles decrease the trait and had CADD-phred>20 were annotated "1" for CADD(-).

3) Composite of multiple signals (cms)[27]. cms integrated signals of several previous methods for detection of positive selection, and could detect positive selection in the last tens of thousands of years at a high resolution. We directly downloaded the genomic regions with significantly high cms from Grossman et al.[27] and generated trait-specific dichotomized annotations (cms+ and cms-).

4)Hard and soft sweep predicted by Trendsetter[28]. Trendsetter applied a penalized regression framework which took statistics at adjacent regions into account and predicted whether each genomic region has undergone hard sweep, soft sweep or neutral alteration. We downloaded the prediction results of CEU population and

labeled each genomic region by the label with highest probability, ang generated trait-specific annotations according to these labels (hard+, hard-, soft+ and soft-).

5)Neanderthal introgressions[25]. We downloaded the average posterior probability LA score from Sankararaman et al.[25] and dichotomized at the top 0.01 LA score.

6)pLi[31]. We curated the gene list of pLi>0.9 as mutation-intolerant genes and generated the annotation of physical position of these genes as well as the flanking regions of 100kb at both 3' and 5' end.

We generated these annotations for each trait and added them to the updated baseline annotations of LDSC[30] to apply heritability partition. We retrieved the z score for LDSC coefficient τ_c for each annotation as main results, which measured the heritability enrichment in each annotation on conditioned of all other annotations. Some of the annotations from the baseline model that also measured some aspects of natural selections were also retrieved (ASMCavg[26], B statistics[58], recombination rate[30], conserved regions[58], etc. see Supplementary method).

Analysis of derived allele distribution

The basic assumption of allele distribution analysis was that purifying selection tended to suppress the alleles with large effect on traits to a low frequency[33]. We first obtained the derived allele (DA) of about four million SNPs provided by 1000 Genome phase 1[48] and corresponding DA frequency (DAF) in UK10K[54] data curated by Field et al.[17]. For each trait, we included in our analysis SNPs meeting following criteria: has DA annotation, DAF between 0.05 and 0.95, GWAS z score>3 or <-3.

These SNPs were pruned by plink[56] with parameter –indep 50 5 2 and 1000G EUR as reference panels and then separated into two groups according to whether their DA increased the trait (i.e. z>0).

Similar to the work of Zeng et al.[33], we quantified the proportion of variance explained by each SNP ($\%variance$). However, we used a different estimation approach proposed by Shim et al.[59], which required less input variables:

$$\%variance = \frac{2 * DAF * (1 - DAF) * \beta^2}{2 * DAF * (1 - DAF) * \beta^2 + 2 * DAF * (1 - DAF) * N * se(\beta)^2}$$

Where β was the effect size, N was the sample size of GWAS, and se(β) was the standard error of effect size. Then, we ordered the two groups of SNPs in the ascending trend of DAF and calculated the cumulative $\%variance$ at each SNP, assuming an additive effect. The cumulative $\%variance$ was further scaled by total $\%variance$ by all SNPs in that group, resulting in two joint distribution of DAF and $\%variance$ from (0.05,0) to (0.95,1) (Figure 4D). We calculated the Area Under Curve (AUC) for each distribution; AUC>0.5 indicated that large effect (either trait-increasing of trait-decreasing effect) DA had a small DAF, which suggested the existence of purifying selection. To figure out the direction of selection, we calculated the difference of AUC:

$$S = \Delta AUC = AUC_{dec} - AUC_{inc}$$

S>0 indicated that the purifying selection pressure on trait-decreasing DA was larger than that on trait-increasing DA, thus the trait was favored by natural selection. To estimate the statistical significance of S, we randomly dropped half of the included SNPs and repeated the entire analysis for 100 times to estimate the standard error of S, then calculated the p value for the Null hypothesis of S=0 under normal

distribution.

Additionally, we calculated the proportion of included SNP whose DA had a trait-increasing effect (%DA), and calculated its corresponding p value by two-sided binomial test assuming the random proportion =50%.

Integrative analysis of all selection pressure

Before integrative analysis, we first applied linear regression to adjust all selection pressure metrices by three covariates calculated by LDSC[60]: λ_{GC}, mean X^2 and intercept. These covariates captured the confounders rather than genetic architecture of GWAS[60].

For each of the seven selection pressure metrices with definite time scale (Z_{ncm}, Z_{ncf}, ρ_{SDS}, $t_{neolithic}$, $t_{neareast}$, $t_{proneolithic}$, t_{HG}), we applied linear regression that included all metrices whose time scale were more ancient to it (Supplementary method). The regression was run on all traits with reliable results of the corresponding response variables (e.g. traits without homogenous MR results were not included in the regression of Z_{ncm} and Z_{ncf}). The obtained linear models were applied on the same data to generate predicted selection pressure metrices. We calculated the R^2 value on the predicted metrices using three set of traits: all traits, polygenic diseases and non-disease traits. Since the number of diseases (85) was much smaller than non-disease traits (785), we compared their R^2 with a subsampling method. Specifically, we randomly chose 85 non-disease traits and calculated R^2 on them for 1000 times, then calculated the standard error of them. We generated a normal distribution whose mean was equal to R^2 using all non-disease traits and standard error equal to the standard

error of 1000 subsampling tests. The R^2 for diseases was compared to this normal distribution to generate an empirical p value, as reported in Table S9.

To discover the traits showing deviation of selection pressure at each time scale, we calculated and z-scored the residuals of each linear model. Traits with scaled residual>3 or <-3 were highlighted.

Reconstructing the History of Polygenic Scores (RHPS) and Relate

To estimate the trajectory of population-mean PRS for each trait, we applied the Relate Reconstructing the History of Polygenic Scores (RHPS) method proposed by Edge et al.[19], which utilized local coalescent trees at GWAS locus to estimate the population-mean polygenic risk score (PRS) of a trait. As suggested by Edge et al.[19], for each GWAS we calculated the Bayesian Factors (BF) for each SNP:

$$BF = \frac{\sqrt{1-r}}{e^{(\frac{-z^2*r}{2})}}, r = \frac{0.1}{0.1+se^2}$$

Where z and se were the GWAS statistics for this SNP. Then, we partitioned all SNPs into 1702 consecutive and independent blocks provided by Pickrell et al.[61], and chose the SNP with largest BF from each block. To maximize the computational efficiency, we retained SNPs with BF>10000. The population mean PRS at ancient time t was estimated as

$$PRS(t) = 2\sum_{i \in G} \beta_i * p_i(t)$$

Where G denoted all SNPs retained, β_i denoted the GWAS effect size of SNP i, and $p_i(t)$ was the population frequency of SNP i and time t. To estimate $p_i(t)$, we applied Relate[20] to each of the retained SNP. Specifically, we retrieved all variants

within the ±100kb windows around the target SNP from non-Finnish European population of 1000G[48], and transferred into .haps format required by Relate. The ancestral genome sequence, genetic map of recombination rate and genetic distance, and genome mask of GRCh37 were downloaded from 1000G resource. We ran Relate on the variant data to estimate the focal tree in the 200kb windows with parameter –m 1.25×10^{-8} and -N 30000. The branch length and population size were re-normalized by EstimatePopulationSize function with three iterations and threshold of 0.7. The frequency of target SNP was estimated by DetectSelection function. We divided the output frequency by 808 (404 non-Finnish European individual×2) to obtain the population frequency per chromosome. To estimate the significance of polygenic adaptation during a time course, we applied the T_x test proposed by Edge et al.[19] to calculate the p value for PRS alteration in specific time window. For the analysis of recent history, we applied Tx test on the last five timepoints (96 generations ago to present).

Time course clustering of PRS trajectory

Time course clustering was conducted by dtwclust R package[62]. For all traits with PRS trajectory results from RHPS-RELATE (i.e. frequency trajectory available for at least three SNPs), we retained the result of last 12 timepoints (958 generations ago to present; detailed information of each timepoint was recorded in Table S11), since result at earlier time points was sparse. We z-scored the trajectory of each trait, calculated the similarity among traits by Dynamic time warping[63], and partitioned the traits by hierarchical clustering. We chose the number of clusters k=4 by comparing

the Silhouette coefficient for clustering results with k=2 to 10.

Impacts of genetic architectures on selection pressures

Despite the genomic annotations that directly measured selection pressure, we were also interested whether heritability enrichment in other annotations could impact natural selection. These annotations were roughly divided into two groups (columns of Figure 6A): those that are associated with selection, termed "conservative annotations" (black texts in Figure 6A), and those without evidence of direct association with selection, termed "functional genomic annotations" (red texts in Figure 6A). The annotations of MAF bin were discarded. Similar to the above section, all data were adjusted by λ_{GC}, mean X^2 and intercept prior to the analysis. For each of the ten selection pressure metrices (Figure 6A), we first fitted a full linear model including LDSC z scores of all conservative and functional genomic annotations, then applied a bi-directional step-wise regression aiming at maximization of Akaike Information Criterion (AIC) to obtain a simplified model. For the convenience of visualization, in Figure 6A we only showed the t value for annotations that reached $p<0.01$ in at least one simplified model. Full results for all simplified models can be found at Table S12.

To analyze the contribution of different groups of annotations, we subtracted three sub-models from each simplified model: (1) model containing only functional genomic annotations; (2) model containing functional genomic annotations and CADD annotations; and (3) model containing all annotations. We applied each of these models to generate predicted values for each selection metric and calculated

corresponding R^2 values for these precited values.

Statistical Analysis

For all comparison of metrices among groups, we applied two-sided Fisher-Pitman permutation test by oneway_test from the coin R package[64]. For all comparisons between two metrices, we applied Pearson correlation analysis. For comparisons between two distributions, we applied two-sided Kolmogorov-Smirnov test. The significance threshold was set at $p>0.05/870$ unless otherwise specified.

Reference

1. Mathieson, I. Human adaptation over the past 40,000 years. *Current Opinion in Genetics and Development* vol. 62 97–104 (2020).
2. Bersaglieri, T. *et al.* Genetic signatures of strong recent positive selection at the lactase gene. *Am. J. Hum. Genet.* **74**, 1111–1120 (2004).
3. Kelly, A. J., Dubbs, S. L. & Barlow, F. K. An evolutionary perspective on mate rejection. *Evol. Psychol.* **14**, 1–13 (2016).
4. Gibson, M. A. & Lawson, D. W. Applying evolutionary anthropology. *Evol. Anthropol.* **24**, 3–14 (2015).
5. Wells, J. C. K., Nesse, R. M., Sear, R., Johnstone, R. A. & Stearns, S. C. Evolutionary public health: introducing the concept. *The Lancet* vol. 390 500–509 (2017).
6. Currat, M. *et al.* Molecular analysis of the β-globin gene cluster in the niokholo mandenka population reveals a recent origin of the βs senegal mutation. *Am. J. Hum. Genet.* **70**, 207–223 (2002).
7. Kircher, M. *et al.* A general framework for estimating the relative pathogenicity of human genetic variants. *Nat. Genet.* **46**, 310–315 (2014).
8. Pritchard, J. K., Pickrell, J. K. & Coop, G. The Genetics of Human Adaptation: Hard Sweeps, Soft Sweeps, and Polygenic Adaptation. *Current Biology* vol. 20 R208–R215 (2010).
9. Laland, K. N., Odling-Smee, J. & Myles, S. How culture shaped the human genome: Bringing genetics and the human sciences together. *Nature Reviews*

Genetics vol. 11 137–148 (2010).

10. Esteller-Cucala, P. *et al.* Genomic analysis of the natural history of attention-deficit/hyperactivity disorder using Neanderthal and ancient Homo sapiens samples. *Sci. Rep.* **10**, (2020).
11. Pardiñas, A. F. *et al.* Common schizophrenia alleles are enriched in mutation-intolerant genes and in regions under strong background selection. *Nat. Genet.* **50**, 381–389 (2018).
12. Watanabe, K. *et al.* A global overview of pleiotropy and genetic architecture in complex traits. *Nat. Genet.* **51**, 1339–1348 (2019).
13. Hejase, H. A., Dukler, N. & Siepel, A. From Summary Statistics to Gene Trees: Methods for Inferring Positive Selection. *Trends in Genetics* vol. 36 243–258 (2020).
14. Lawn, R. B. *et al.* Schizophrenia risk and reproductive success: A Mendelian randomization study. *R. Soc. Open Sci.* **6**, 181049 (2019).
15. Bribiescas, R. G. *Men: Evolutionary and life history.* (Harvard University Press, 2009).
16. Morrison, J., Knoblauch, N., Marcus, J. H., Stephens, M. & He, X. Mendelian randomization accounting for correlated and uncorrelated pleiotropic effects using genome-wide summary statistics. *Nat. Genet.* **52**, 1–8 (2020).
17. Field, Y. *et al.* Detection of human adaptation during the past 2000 years. *Science (80-.).* **354**, 760–764 (2016).
18. Sohail, M. *et al.* Polygenic adaptation on height is overestimated due to

uncorrected stratification in genome-wide association studies. *Elife* **8**, (2019).

19. Edge, M. D. & Coop, G. Reconstructing the history of polygenic scores using coalescent trees. *Genetics* **211**, 235–262 (2019).

20. 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).

21. Janus, L. Two corner stones of the psychobiological development of mankind - The increase in frequency of pregnancies in the neolithic revolution and 'physiological prematurity'. in *Nutrition and Health* vol. 19 63–68 (SAGE Publications Ltd, 2007).

22. Lipson, M. *et al.* Parallel palaeogenomic transects reveal complex genetic history of early European farmers. *Nature* **551**, 368–372 (2017).

23. Fu, Q. *et al.* The genetic history of Ice Age Europe. *Nature* vol. 534 200–205 (2016).

24. Lazaridis, I. *et al.* Genomic insights into the origin of farming in the ancient Near East. *Nature* **536**, 419–424 (2016).

25. Sankararaman, S. *et al.* The genomic landscape of Neanderthal ancestry in present-day humans. *Nature* **507**, 354–357 (2014).

26. Palamara, P. F., Terhorst, J., Song, Y. S. & Price, A. L. High-throughput inference of pairwise coalescence times identifies signals of selection and enriched disease heritability. *Nat. Genet.* **50**, 1311–1317 (2018).

27. Grossman, S. R. *et al.* A composite of multiple signals distinguishes causal

variants in regions of positive selection. *Science (80-.).* **327**, 883–886 (2010).

28. Mughal, M. R. & DeGiorgio, M. Localizing and classifying adaptive targets with trend filtered regression. *Mol. Biol. Evol.* **36**, 252–270 (2019).

29. Cheng, X. & DeGiorgio, M. Flexible mixture model approaches that accommodate footprint size variability for robust detection of balancing selection. *Mol. Biol. Evol.* (2020) doi:10.1093/molbev/msaa134.

30. Gazal, S. *et al.* Linkage disequilibrium-dependent architecture of human complex traits shows action of negative selection. *Nat. Genet.* **49**, 1421–1427 (2017).

31. Lek, M. *et al.* Analysis of protein-coding genetic variation in 60,706 humans. *Nature* **536**, 285–291 (2016).

32. Zeng, J. *et al.* Bayesian analysis of GWAS summary data reveals differential signatures of natural selection across human complex traits and functional genomic categories. *bioRxiv* 752527 (2019) doi:10.1101/752527.

33. Zeng, J. *et al.* Bayesian analysis of GWAS summary data reveals differential signatures of natural selection across human complex traits and functional genomic categories. *bioRxiv* 752527 (2019) doi:10.1101/752527.

34. 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).

35. Daub, J. T. *et al.* Evidence for polygenic adaptation to pathogens in the human genome. *Mol. Biol. Evol.* **30**, 1544–1558 (2013).

36. Rocha, J. The Evolutionary History of Human Skin Pigmentation. *Journal of Molecular Evolution* vol. 88 77–87 (2020).

37. Luca, F., Perry, G. H. & Di Rienzo, A. Evolutionary adaptations to dietary changess. *Annual Review of Nutrition* vol. 30 291–314 (2010).

38. Little, M. A. Evolutionary Strategies for Body Size. *Frontiers in Endocrinology* vol. 11 107 (2020).

39. Bocquet-Appel, J. P. When the world's population took off: The springboard of the neolithic demographic transition. *Science* vol. 333 560–561 (2011).

40. Liao, Y., Loures, E. R., Deschamps, F., Brezinski, G. & Venâncio, A. The impact of the fourth industrial revolution: A cross-country/region comparison. *Producao* **28**, 20180061 (2018).

41. Cordain, L. *et al.* Origins and evolution of the Western diet: Health implications for the 21st century. *American Journal of Clinical Nutrition* vol. 81 341–354 (2005).

42. Comas, I. *et al.* Out-of-Africa migration and Neolithic coexpansion of Mycobacterium tuberculosis with modern humans. *Nat. Genet.* **45**, 1176–1182 (2013).

43. Papadimitriou, A. The Evolution of the Age at Menarche from Prehistorical to Modern Times. *Journal of Pediatric and Adolescent Gynecology* vol. 29 527–530 (2016).

44. Södersten, P., Brodin, U., Zandian, M. & Bergh, C. Eating Behavior and the Evolutionary Perspective on Anorexia Nervosa. *Front. Neurosci.* **13**, 596

(2019).

45. Ewald, P. W. & Swain Ewald, H. A. An evolutionary perspective on the causes and treatment of inflammatory bowel disease. *Curr. Opin. Gastroenterol.* **29**, 350–356 (2013).

46. Martin, A. R. *et al.* Clinical use of current polygenic risk scores may exacerbate health disparities. *Nat. Genet.* **51**, 584–591 (2019).

47. Bycroft, C. *et al.* The UK Biobank resource with deep phenotyping and genomic data. *Nature* **562**, 203–209 (2018).

48. 1000 Genomes Project Consortium. A global reference for human genetic variation. *Nature* vol. 526 68–74 (2015).

49. Hemani, G. *et al.* The MR-base platform supports systematic causal inference across the human phenome. *Elife* **7**, (2018).

50. Bowden, J., Davey Smith, G., Haycock, P. C. & Burgess, S. Consistent Estimation in Mendelian Randomization with Some Invalid Instruments Using a Weighted Median Estimator. *Genet. Epidemiol.* **40**, 304–314 (2016).

51. Bowden, J., Davey Smith, G. & Burgess, S. Mendelian randomization with invalid instruments: effect estimation and bias detection through Egger regression. *Int. J. Epidemiol.* **44**, 512–25 (2015).

52. Burgess, S., Bowden, J., Fall, T., Ingelsson, E. & Thompson, S. G. Sensitivity analyses for robust causal inference from mendelian randomization analyses with multiple genetic variants. *Epidemiology* vol. 28 30–42 (2017).

53. Verbanck, M., Chen, C. Y., Neale, B. & Do, R. Detection of widespread

horizontal pleiotropy in causal relationships inferred from Mendelian randomization between complex traits and diseases. *Nat. Genet.* **50**, 693–698 (2018).

54. Walter, K. *et al.* The UK10K project identifies rare variants in health and disease. *Nature* **526**, 82–89 (2015).

55. Galinsky, K. J. *et al.* Fast Principal-Component Analysis Reveals Convergent Evolution of ADH1B in Europe and East Asia. *Am. J. Hum. Genet.* **98**, 456–472 (2016).

56. Purcell, S. *et al.* PLINK: A tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.* **81**, 559–575 (2007).

57. Barret Schloerke *et al.* GGally.

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

59. Shim, H. *et al.* A Multivariate Genome-Wide Association Analysis of 10 LDL Subfractions, and Their Response to Statin Treatment, in 1868 Caucasians. *PLoS One* **10**, e0120758 (2015).

60. Bulik-Sullivan, B. K. *et al.* LD Score regression distinguishes confounding from polygenicity in genome-wide association studies. *Nat. Genet.* **47**, 291–295 (2015).

61. Berisa, T. & Pickrell, J. K. Approximately independent linkage disequilibrium blocks in human populations. *Bioinformatics* **32**, 283–285 (2016).

62. Sardá-Espinosa, A. Time-series clustering in R Using the dtwclust package. *R*

Journal vol. 11 22–43 (2019).

63. Berndt, D. J. & Clifford, J. Using dynamic time warping to find patterns in time series. *KDD Work.* **10**, 359–370 (1994).

64. Hothorn, T., Hornik, K., Wiel, M. A. van de & Zeileis, A. Implementing a Class of Permutation Tests: The **coin** Package. *J. Stat. Softw.* **28**, 1–23 (2008).

Figures

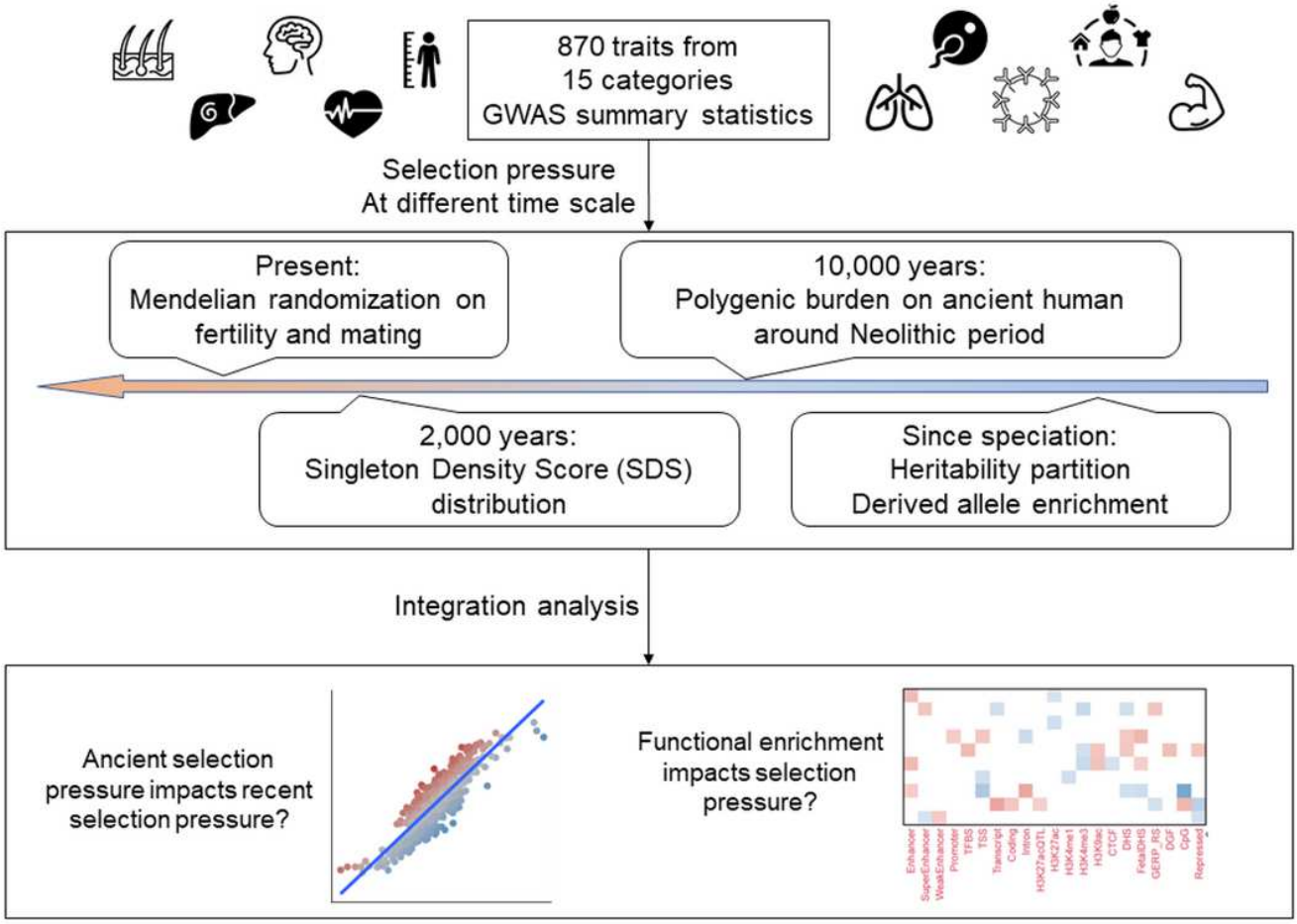

Figure 1

Flowchart of the study.

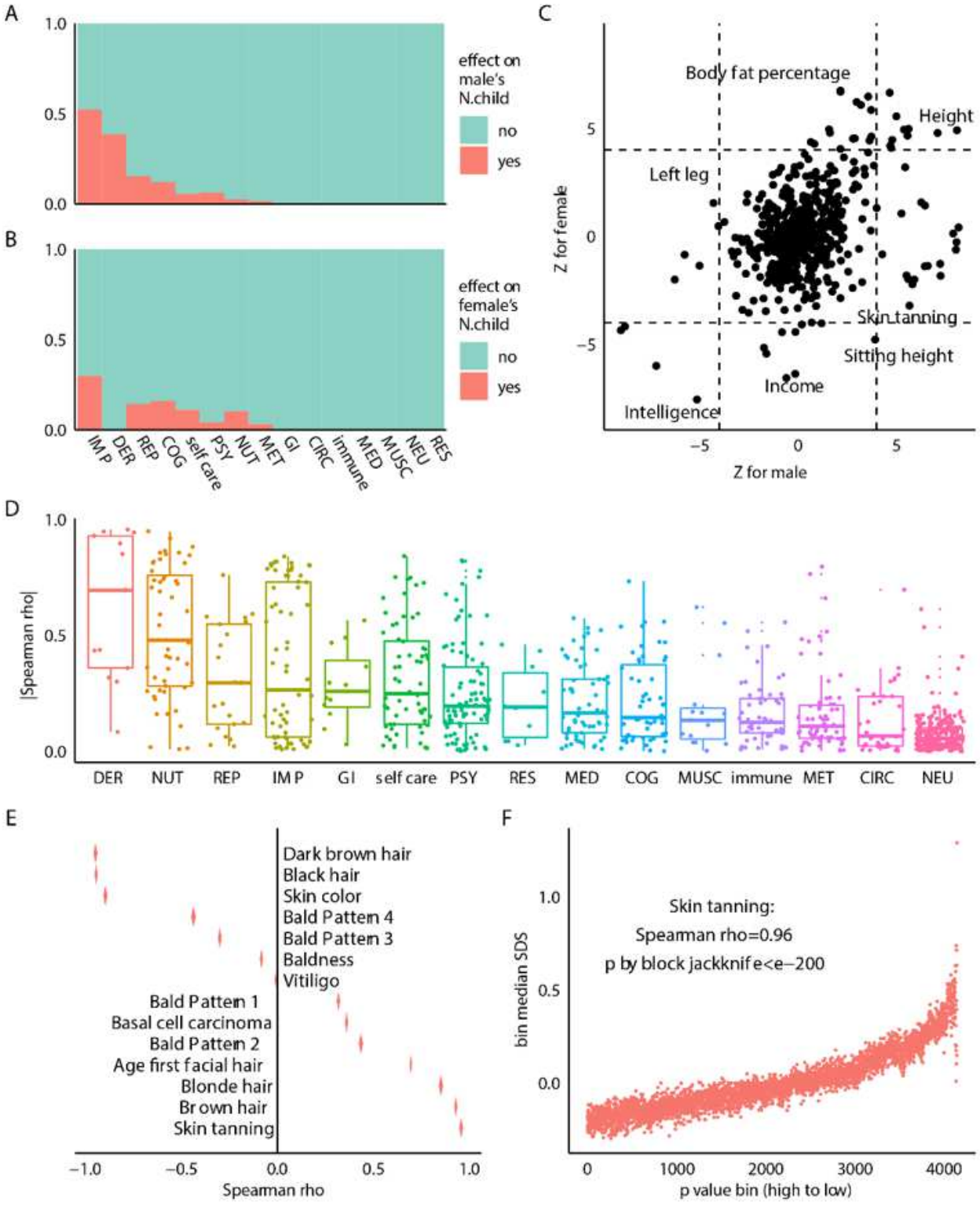

Figure 2

Selection pressure at present and recent history. A&B: Proportion of traits showing MR causal effects on number of children of male (A) and female (B) for each category. C: Comparison of MR z scores between male (x axis) and female (y axis). Dashed lines indicated significance threshold (|z|>4). Texts indicated selected traits with results of special interests. D: Distribution of absolute Spearman correlation (|ρSDS|) between tSDS and GWAS p value for each category. Upper and lower margins of box indicated first and

third quartiles of ρSDS, and the thickened line indicated median ρSDS. E: ρSDS for all dermatology traits. The diagonal of the rhombus indicated ρSDS, and the width of rhombus indicated 95% confidence interval of ρSDS. F: Scatter plot showing the correlation between tSDS and GWAS p value for trait “Ease of skin tanning”. Each point represented a bin of 1000 SNPs ranked by their GWAS p value. Y-axis indicated the bin median tSDS. DER: dermatology. NUT: nutrition. REP: reproduction. IMP: impedance. GI: Gastrointestinal. PSY: psychiatry. RES: respiratory. MED: medication. COG: social-cognition. MUSC: musculoskeletal. MET: metabolism. CIRC: circulation. NEU: neurology.

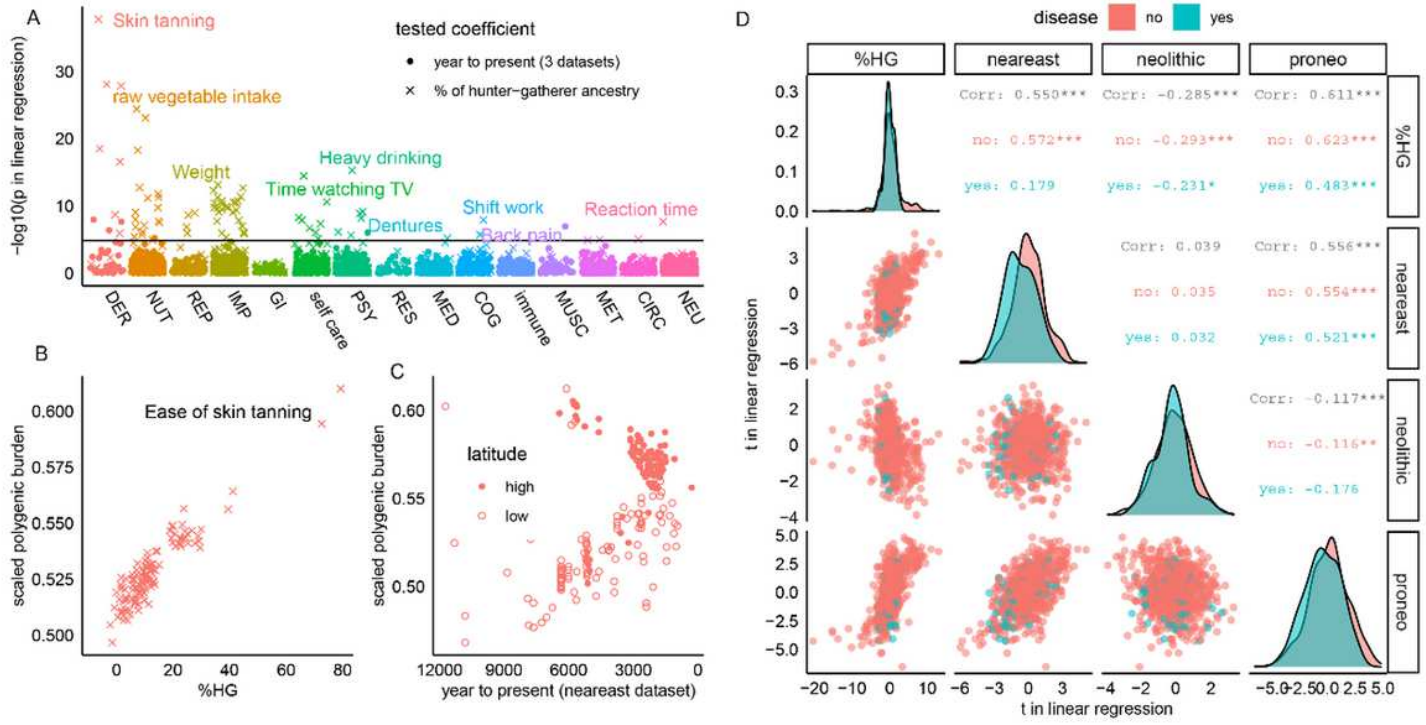

Figure 3

Selection pressure at pan-Neolithic period. A: Manhattan plot showing the p value of linear regression. The regression was between scaled genetic burden and either time to present (round dot) or percent of hunter-gatherer ancestry (%HG) (cross). B: Relation between genetic burden of “Ease of skin tanning” and %HG. Each dot denoted an ancient human in Neolithic dataset, its y-axis showed the genetic burden of “Ease of skin tanning” carried by this individual, and the y-axis showed the %HG of this individual. C: Similar to B, but for living times in neareast dataset. D: Relation among four selection metrices. Each dot represented a trait, and its x and y axis showed the t value of linear regression on two out of four selection metrices (Neolithic time, Neolithic %HG, neareast time and pro-Neolithic time). Red color corresponded to non-disease traits, and green color corresponded to polygenic diseases. Texts in upper triangle showed the Pearson Correlation Coefficients for symmetric plots in lower triangle. *: p<0.05; **: p<0.01; ***: p<0.001. Diagonal plots showed the distribution of t values.

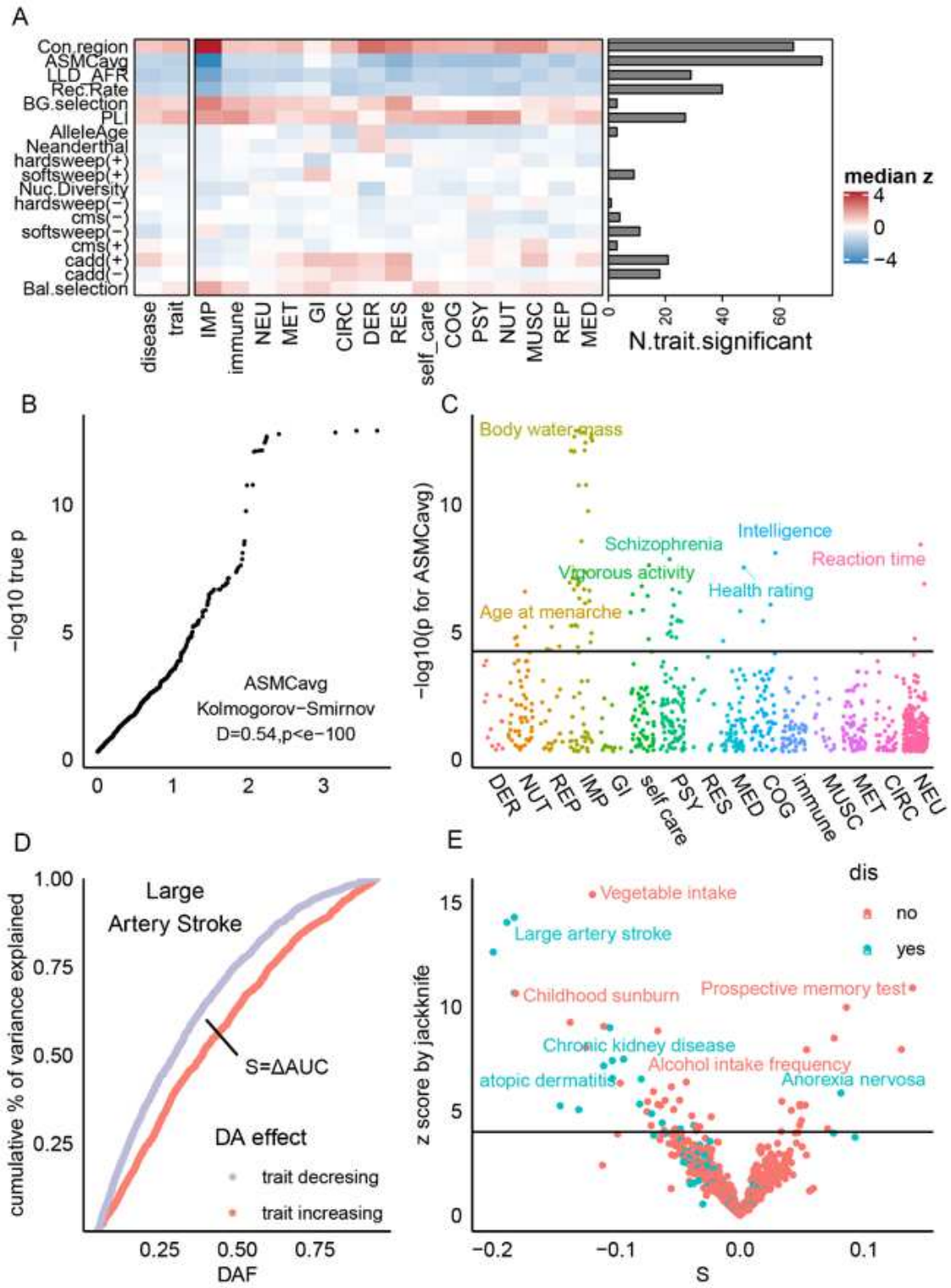

Figure 4

Selection pressure since human speciation. A: Heatmap showing the median LDSC enrichment Z score on each genomic annotation for each category. Bar plots denoted the total number of traits showing significant enrichment in the corresponding annotations. B: QQ plot for LDSC p value of ASMCavg enrichment. C: Manhattan plot for ASMCavg enrichment. D: Effect-frequency distribution for trait “Large artery stroke”. Each dot denoted a SNP, x-axis showed its derived allele frequency (DAF), y-axis showed

the cumulative proportion of variance explained by all SNPs with DAF smaller than this SNP. We separated SNPs according to their DA effect (increase or decrease the trait). The area between the curve of these two distributions, named S, was used to measure natural selection. E: Volcano plot for the statistical analysis of S.

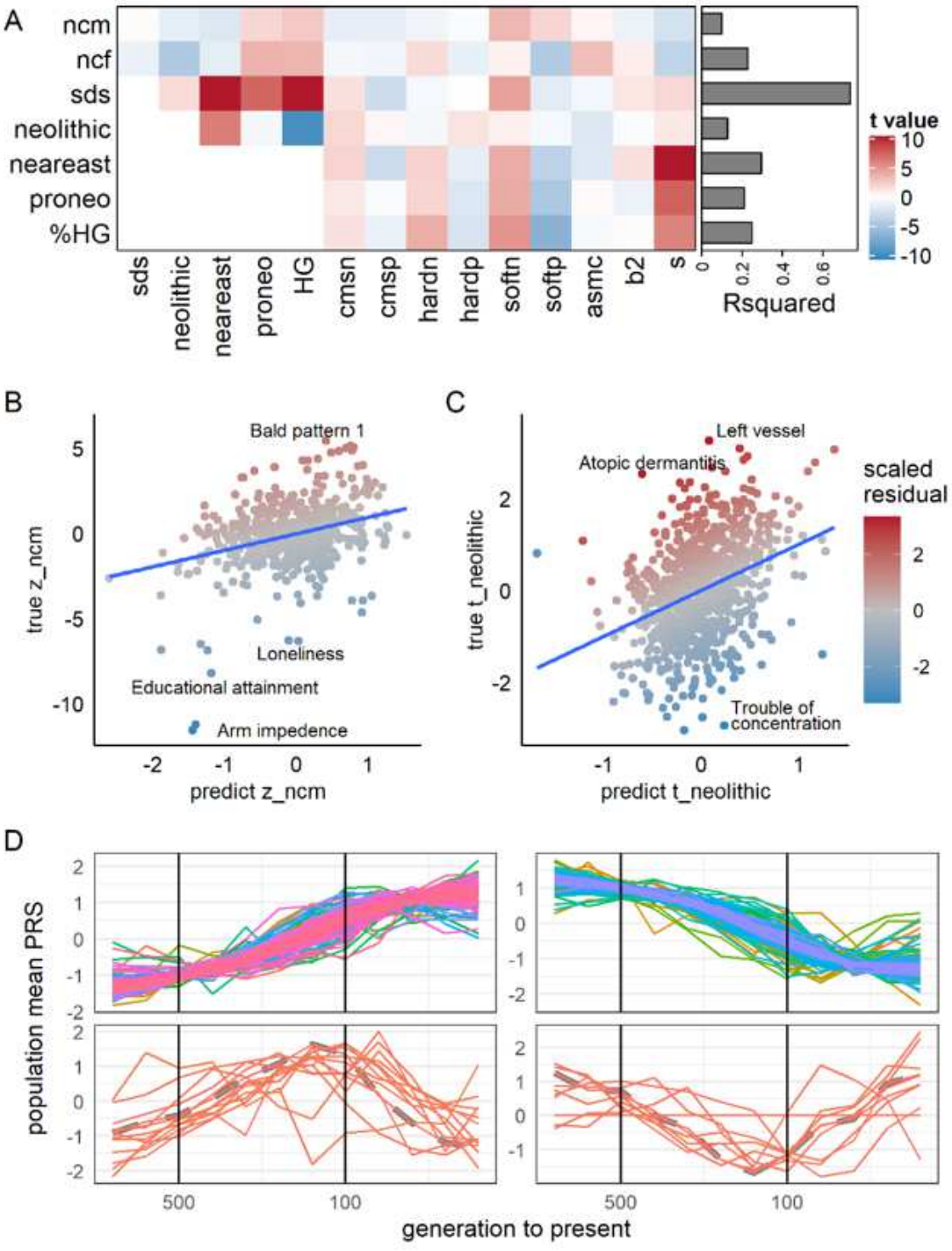

Figure 5

Relation among selection pressure at different time scales. A: Heatmap showing the t value of linear regression which used ancient selection pressure (columns) to predict recent selection pressure (rows). Bar plots denoted the R2 for corresponding linear regression. B: Each dot denoted a trait; x-axis showed the Zncm predicted by linear regression, y-axis showed the true Zncm, and color denoted the scaled residual in the linear model. C: Similar to B, but for tneolithic. D: Population-average polygenic risk score trajectory for 765 traits, grouped into four clusters according to their time series similarity. Y-axis showed z-score of PRS.

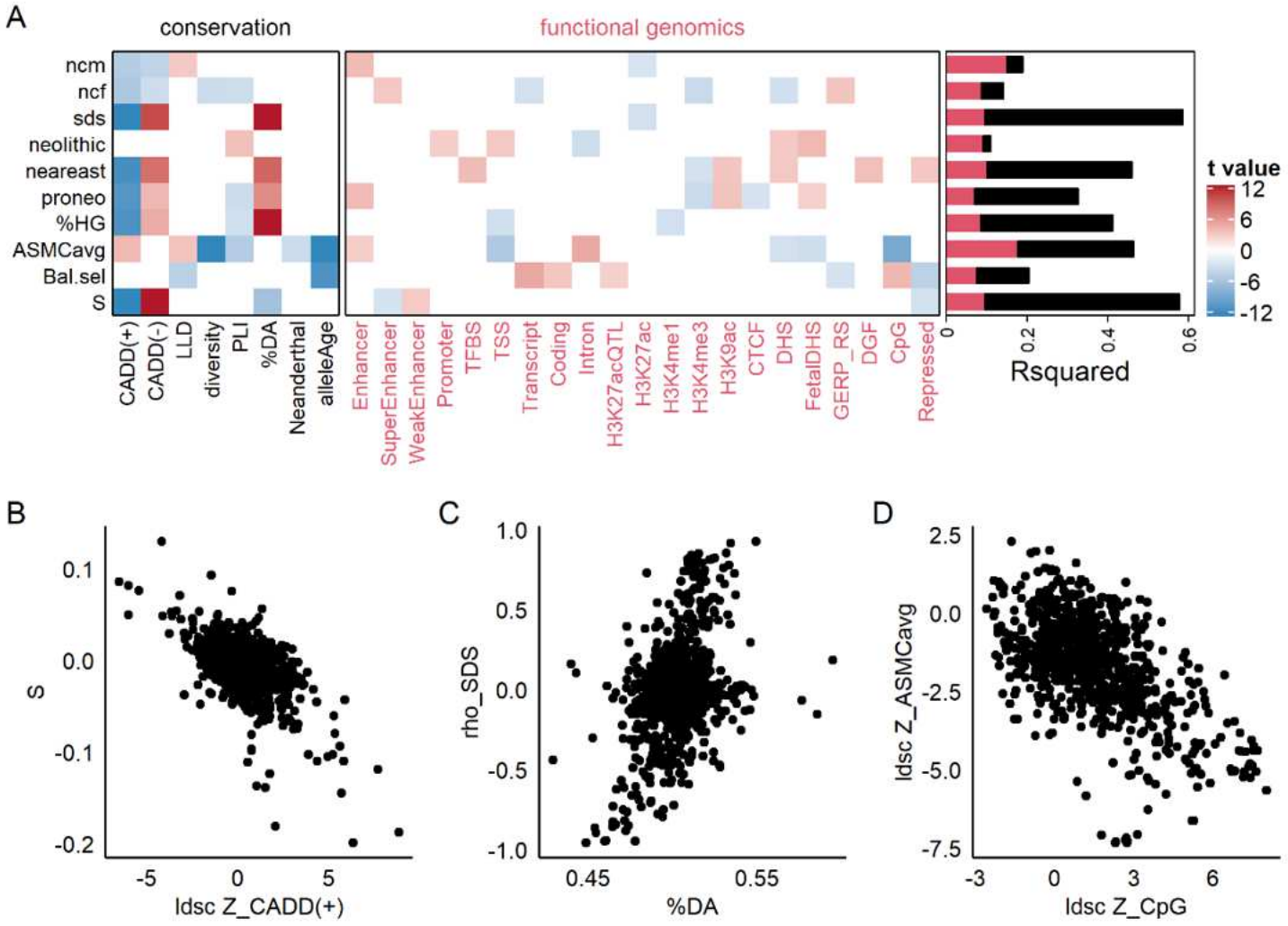

Figure 6

Genomic architectures impacted selection pressure. A: Heatmap showing the t value of linear regression which used functional genomic architecture and genetic conservation characteristic (columns) to predict selection pressure (rows). Red bars denoted the R2 for corresponding linear regression using functional genomic architectures alone, and black bar denoted R2 for linear regression using all predictors. B-D: Scatter plots showing the most significant contribution of genomic characteristics to selection pressures. Each dot represented one trait.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SupplementrayNotes.docx
- supp.rar