Title

Genetic timeline of human brain and cognitive traits

Authors

Ilan Libedinsky[1], Yongbin Wei[1,2], Christiaan de Leeuw[1], James Rilling[3,4,5,6,7], Danielle Posthuma[1,8], Martijn P. van den Heuvel[1,8,*]

Affiliations

1. Department of Complex Trait Genetics, Center for Neurogenomics and Cognitive Research, Amsterdam Neuroscience, Vrije Universiteit Amsterdam, Amsterdam, 1081HV, the Netherlands.

2. School of Artificial Intelligence, Beijing University of Posts and Telecommunications, Beijing, 100876, China.

3. Yerkes National Primate Research Center, Emory University, Atlanta, GA 30329, USA.

4. Center for Translational Social Neuroscience, Emory University, Atlanta, GA 30329, USA.

5. Department of Anthropology, Emory University, Atlanta, GA 30322, USA.

6. Silvio O. Conte Center for Oxytocin and Social Cognition, Emory University, Atlanta, GA 30322, USA.

7. Department of Psychiatry and Behavioral Sciences, Emory University, Atlanta, GA 30322, USA.

8. Department of Child and Adolescent Psychiatry and Psychology, Section Complex Trait Genetics, Amsterdam Neuroscience, Vrije Universiteit Medical Center, Amsterdam UMC, Amsterdam 1081HV, the Netherlands.

Corresponding author: Martijn P. van den Heuvel

E-mail: martijn.vanden.heuvel@vu.nl

Abstract

Humans have undergone several anatomical adaptations throughout evolution. Paleontological records are a prime method of studying these adaptations, but they can unfortunately provide only a limited view of how modifications of 'soft traits' such as brain and cognition have contributed to the emergence of *Homo sapiens*. An additional approach includes the examination of when genetic variations associated with human phenotypes emerged in our history. Combining data from genome-wide association studies (GWAS) with dating data on the human genome, we systematically analysed the temporal emergence of single-nucleotide polymorphisms (SNPs) associated with modern-day human phenotypes over the last five million years. We show the genetic timeline of human-characteristic phenotypes to follow a distinct pattern with two bursts of genetic variation that co-emerge with milestones in the human lineage. Our findings suggest that SNPs associated with neocortical, neuropsychiatric, and ophthalmological traits appeared relatively recently in hominin evolution, with genes containing recently emerged SNPs linked to intelligence and neocortical area.

Main Text

Introduction

The genome contains a footprint of human history. Mapping the genetic timeline of when specific single-nucleotide polymorphisms (SNPs) associated with modern-day phenotypes first emerged in our genome provides a window to examine the evolutionary tempo of characteristic features of humans during hominin evolution, in particular traits that do not fossilize and thus are difficult to directly assess using common paleontological methods (1).

One such aspect of hominin evolution is the remarkable evolution of human brain organization and function. Cranial capacity of modern-day humans has expanded more than 1.5 times since *Homo erectus* (2) [who lived around ~2M years ago (3)], with the neocortex

among the most expanded brain regions (4, 5). Expansion of cortical areas is widely believed to be one of the strongest catalyzers for the development of complex behavior and advanced cognition in hominini (4, 6, 7). This expansion is thought to be the result of a wide range of natural (e.g. climate), nutritional (e.g. diet), and social forces (e.g. group size, parental care) (8) that synergistically interacted with modifications and adaptations in the human genome (9). A growing number of Genome-Wide Association Studies (GWAS) have started to unravel the genetic background of a wide range of modern human phenotypes (10). In this study, we integrate these GWAS findings with dating estimates of the emergence of variants in the human genome (11) to map the genetic evolutionary timeline of brain and cognitive phenotypes. We show that the genetic timeline of SNPs associated with general human phenotypes follows a distinctive time pattern with two characteristic peaks that co-emerge with important events in the homo lineage. The evolutionary oldest peak is enriched for genetic variants linked to skeletal and metabolic traits, and the most recent peak is enriched for variants related to human-characteristic traits associated with brain, cognition, and neuropsychiatric conditions.

Results

Genetic timelines of human phenotypes run parallel to evolutionary milestones

We started by reconstructing the genetic timeline of when SNPs associated with human phenotypes appeared in the human genome. We used data from the Human Dating Genome project (HDG) (11), a database that, by tracing back alleles to a mutation event in the genome of a common ancestor, describes estimates of the emergence date of ~13.5 million SNPs across 22 chromosomes of the human genome covering ~200,000 generations from 5,140,625 to 87 years ago (*Methods*, *SI Appendix*). We combined this data with a total of 67,119 SNPs (36,506 unique SNPs; referred to as human-phenotypic SNPs) from the summary statistics of 2,538 GWAS (10), associated with a broad range of human phenotypes (e.g. hair color, eye shape, cancer, height, brain, behavior). Combining these

datasets enabled us to compute the dates of when human-phenotypic SNPs appeared in the human genome [for sensitivity analyses were performed controlling for SNPs in Linkage Disequilibrium (LD) dependence see *SI Appendix*]. The combined genetic timeline of the 36,506 human-phenotypic SNPs (ranging from ~4.5M to ~1,700 years ago, median evolutionary age = 673,520 years old) revealed a bimodal distribution with two distinct peaks (Figure 1A): an 'old peak' of SNPs arising from ~2.4M until ~280,000 years ago reaching a maximum at ~1.1M years ago, comprising ~85% out of all high minor allele frequency SNPs (MAF >= 0.4; *SI Appendix*); and a second 'young peak' of SNPs arising at ~280,000 until ~1,700 years ago reaching a maximum at ~55,000 years ago, a period including ~62% of low MAF SNPs (MAF <= 0.1; Figure 1D). This bimodal distribution was absent from the genetic timeline of all SNPs, suggesting that it is a specific characteristic of human-phenotypic SNPs (Figure 1B shows the genetic timeline of all SNPs).

Statistical testing revealed a significant large increment of human-phenotypic SNPs at the beginning of the old peak (from 1,920,075 to 1,900,075 years ago, $P = 2.5 \times 10^{-8}$, permutation test, 10,000 permutations, *SI Appendix*), and at the highest point of the young peak from 61,625 to 1,675 years ago (three consecutive significant bins, $P = 4.2 \times 10^{-16}$, $P = 6.2 \times 10^{-27}$, and $P = 2.3 \times 10^{-40}$; Figure 1A). Figure S1 further shows the genetic timeline of human-phenotypic SNPs grouped by their functional genetic consequences (e.g. intergenic, missense). We found this bimodal distribution again by selecting human-phenotypic SNPs from the EBI Catalog (12) as an alternative resource (including unique non-overlapping SNPs with GWAS Atlas, $n = 88{,}362$), resulting in a highly similar distribution (*rho* = 0.96, $P = 5.5 \times 10^{-59}$; Figure 1C, *SI Appendix*).

Plotting the bimodal genetic timeline side-by-side with reports on evolutionary milestones in the human lineage identified several remarkable points of overlap (Figure 1A): the old peak coincides with (among others, see also Discussion) the origin of *Homo erectus* (~1.8M years ago) (3) and the divergence of *Homo sapiens* and Neanderthals (~800,000 years ago) (13). The beginning of the young peak co-occurs with the origin of modern

humans in Africa (~200,000 years ago) (14, 15) and the highest point of the young peak (~55,000 years ago) matches with the transition to the Upper Paleolithic (~45,000 years ago), a period considered as the origin of modern human behavior (16).

Ophthalmological and psychiatric phenotypes influenced by evolutionary recent mutations

We used data on the nested hierarchical organization from the GWAS Atlas (10) to gain insight into the genetic timeline of specific classes of human phenotypes (domain, chapter, subchapter, and trait level were examined). At the domain level (n = 28 domains, Bonferroni P value threshold < 1.7×10^{-3}; Figure 2), SNPs related to 'Ophthalmological' phenotypes presented the most recent evolutionary age (n = 823 SNPs, median evolutionary age = 30,484 years old), younger than expected by random chance ($P = 2.1 \times 10^{-75}$; null-distribution of equally sized random sets out of all human-phenotypic SNPs, details in *Methods*; sensitivity analysis controlling for LD in *SI Appendix*). This domain of traits is widely noted to have been under strong adaptive changes in human evolution (17) and argued to include several distinctive features between *Homo sapiens* and Neanderthals (18). 'Psychiatric' phenotypes (n = 4,022, median age = 412,639, $P = 7.8 \times 10^{-65}$), a class of traits suggested to be linked to hominin evolution (7, 19–23), and 'Nutritional' phenotypes (n = 929, median age = 429,993, $P = 4.2 \times 10^{-14}$) followed in second and third place. In contrast, SNPs associated with variation of 'Metabolic' (n = 10,287, median age = 780,323, $P = 5.9 \times 10^{-36}$), 'Skeletal' (n = 4,452, median age = 759,056, $P = 1.5 \times 10^{-9}$) and 'Immunological' (n = 4,197, median age = 747,780, $P = 4.3 \times 10^{-7}$; Figure S2) phenotypes displayed the relatively oldest evolutionary age, in line with important dietary and skeletal changes in *Homo erectus* (24). Results on the more fine-grained chapter (Figure S3), subchapter (Figure S4), and trait level (Figure S5) confirmed these results (*SI Appendix*; complete results in Table S1). Sensitivity analyses showed that these findings were robust against differences in the

number of SNPs per phenotype, the sample size of included GWAS, the quality of SNP dates accuracy, and correcting for LD dependence among SNPs (25) (*SI Appendix*).

Neocortex phenotype variation linked to evolutionary recent mutations

Further examination was conducted on the genetic timeline of several brain structures by analysing the date estimates of 6,284 SNPs (2,476 unique BRAIN-SNPs) collated from extensive GWAS on 1,744 neuroimaging-derived phenotypes from the UK Biobank (e.g. cortical surface, myelin, volume of brain structures; Table S2) (26). The genetic timeline of BRAIN-SNPs ranged from 3,621,625 to 4,450 years ago (median evolutionary age = 28,652 years old; Figure S6).

Neocortex SNPs (a subset of n = 130 SNPs out of the 2,476 BRAIN-SNPs) exhibit a median evolutionary age significantly younger than expected by chance ($P = 4.4 \times 10^{-4}$; Figure 3A; all P values in Table S3; *SI Appendix*). The neocortex is indeed believed to be among the most expanded brain structures in the *Homo* lineage (27), displaying substantial divergence in terms of volume and size compared to chimpanzees and bonobos (7, 28). We did not find such significant results for date estimates of SNPs related to white matter (n = 676), cerebellum (n = 144), amygdala (n = 39), hippocampus (n = 30), or thalamus (n = 19) among others (all nine brain structures listed in Table S2; Bonferroni P value threshold was applied $< 5.5 \times 10^{-3}$ to all tests performed; *SI Appendix*).

Genes containing recent mutations are enriched for genes related to brain structure, cognition and disorders, and expressed in language areas

An estimate of the evolutionary age of each gene was derived by computation of the median age of the SNPs located within the start and stop position of each gene, resulting in date estimates for 18,328 human genes [out of 26,836 genes, GRCh37 human assembly (29); *Methods*]. Computed gene evolutionary age varied between 2,965,600 to 3,803 years ago with a peak density ~70,000 years ago (Figure 4A). We evaluated whether the resilience of

genes against mutations is related to their evolutionary age, by ranking genes according to intolerance to Loss-of-Function (LoF; *SI Appendix*) (30). The top 2,000 most intolerant LoF genes were found to be on average younger than the top tolerant LoF genes (t = -4.89, P = 1.0 x 10^{-6}; Figure 4B), indicating that genes intolerant to LoF contain genetic variants relatively recent in evolution compared to tolerant LoF genes.

We next examined whether genes containing recent evolutionary modifications of the human genome are enriched for genes related to variation in brain (e.g. cortical area, thickness, and volume), cognitive (e.g. intelligence, sociability), and neuropsychiatric phenotypes (*SI Appendix*; included studies in Table S5). We used gene-set analysis (*Methods*) (31) to examine the top 2,000 genes with the youngest evolutionary age (~10% from all genes) and their association with genes-sets related to variation in these ten phenotypes (using Bonferroni P value threshold < 2.5 x 10^{-3}). Genes containing recent mutations were found to be significantly enriched for genes involved in intelligence (b = 0.13, P = 2 x 10^{-6}) and cortical area (b = 0.08, P = 6.7 x 10^{-4}). We also observed a trend level enrichment of these 'young genes' to be associated with Alzheimer's disease (AD; b = 0.06, P = 4.2 x 10^{-3}) and schizophrenia (SCZ; b = 0.06, P = 7.9 x 10^{-3}), effects not significant after Bonferroni correction. The set of 'young genes' was also consistently enriched for intelligence when using a lower (n = 1,000) or higher (n = 3,000) number of genes (*SI Appendix*). We also examined the genes with the oldest evolutionary age and no significant enrichment was found (P > 0.05; Figure 4D; Table S6). A direct comparison of the enrichment effect sizes confirms that 'young genes' had a greater enrichment than 'old genes' for these phenotypes (*SI Appendix*).

It was also examined this relationship the other way around, testing the median evolutionary age of the total set of genes associated with the same brain, cognitive, and neuropsychiatric phenotypes (q > 0.05, FDR; *SI Appendix*) compared to a null-model of randomly selected genes (ten phenotypes, Bonferroni P value threshold < 5 x 10^{-3}; *Methods*). Genes related to intelligence were indeed found to display a younger evolutionary

age than expected by chance (n = 2,102 genes; $P = 6.4 \times 10^{-17}$), as well genes related to SCZ (n = 2,275; $P = 8.1 \times 10^{-10}$), bipolar disorder (BD; n = 357; $P = 5.5 \times 10^{-5}$), sociability (n = 250; $P = 1.4 \times 10^{-4}$), brain volume (n = 170; $P = 1.6 \times 10^{-4}$), and AD (n = 143; $P = 2.6 \times 10^{-3}$; sensitivity analysis in *SI Appendix*), suggesting that genes linked to these phenotypes contain relatively more recent mutations in the human genome than compared to all genes (Figure 4C, Figure S7). These findings corroborate theories that link neuropsychiatric conditions to recent human evolution (19–21).

Gene expression data from the Allen Human Brain Atlas (32) was used to investigate the transcriptomic pattern of the top 1% genes containing the evolutionary most recent mutations (we opted for a more confined set of genes than the previous analysis to generate a null-model that evenly samples across the whole distribution; sensitivity analysis with other thresholds in *SI Appendix*). We examined their joint pattern of expression across cortical areas [Bonferroni P value threshold < 1.5×10^{-3}; null-random-gene model (33); *SI Appendix*]. This set of genes (n = 164 out of 16,344 genes with both expression and dating estimates, mapped from 26,552 to 6,717 years ago) was found to be highly expressed in the gyrus pars triangularis ($P = 6.2 \times 10^{-4}$; Figure 4E, Table S9; sensitivity analyses in *SI Appendix*), a region recognized as Broca's area and well known to be involved in speech and language, a human unique trait (34). These genes were found to be more expressed in multimodal regions (35) than primary regions (two-sided t-test, t = 2.5, $P = 1.2 \times 10^{-2}$), and a trend towards higher expression compared to unimodal regions (t = 2.1, $P = 3.1 \times 10^{-2}$; Bonferroni P value threshold 2.5×10^{-2}). Expression levels in cortical regions of the genes containing the oldest variants were found to not be different from chance level (P > 0.05; Bonferroni P value threshold < 1.5×10^{-3}).

Discussion

Mapping the evolutionary genetic timeline during hominin evolution reveals two epochs of increased emergence of SNPs; an 'old peak' between ~2.4M and ~280,000 years ago, and a

'young peak' starting from ~280,000 to ~1,700 years ago (reaching the maximum point at ~55,000 years ago). The evolutionary most recent peak is enriched for genetic variants related to human-characteristic traits associated with brain, cognition, and neuropsychiatric conditions.

The two peaks line up with important paleontological observations. The old peak highlights the origin of *Homo erectus* (3), the first signs of fire domestication (36), and the divergence of *Homo sapiens* and Neanderthals (13), a period displaying an enrichment for SNPs associated with skeletal and metabolic phenotypes. The start of the young peak is consistent with studies pinpointing the African origin of modern humans ~300,000 to ~200,000 years ago (14, 15), marking a critical period of increased genetic diversification in the human genome reaching the maximum at ~55,000 years ago, overlapping with the dispersion of *Homo sapiens* out of Africa (~60,000 years ago) (37), followed by the Upper Paleolithic period (~45,000 years ago) (16), and the beginning of agriculture (~10,000 years ago) (38). The Upper Paleolithic is considered to be a period in which important revolutions in humankind occurred (16), characterized by the first consistent evidence of symbolic behavior (i.e. art carvings, ritual artifacts, musical instruments), technological advancement (i.e. microlithic stone tools and bone artifacts, upgraded hunting tools), and ecological expansion (i.e. long-distance trading of materials) (39). Our findings suggest an enrichment around this period for genetic variants associated with ophthalmological, brain, and neuropsychiatric phenotypes of modern-day humans. These results further support recent genetic comparisons of archaic hominin ancestry revealing that a large number of adaptive changes in the human genome occurred around ~600,000 and ~200,000 years ago (40), coinciding with the observed genetic timeline of the old and young peak from the human-phenotypic SNPs.

The genetic timeline of SNPs associated with variation in neocortical organization of modern humans presents an overrepresentation of variants that emerged recently in evolution. Evidence indicates that this structure appeared in evolution more recently than other subcortical structures (41) and genetic modifications are widely believed to have played a major role in the growth and remodeling of the neocortex in humans (9). The human brain has undergone major changes in brain size and structure since the last common ancestor shared by modern humans and other primates (5, 41–43), with the neocortex among the most expanded regions (4). Important to mention in this context is that the studied SNPs are linked to the variation of brain size in modern-day humans and thus not necessarily to an expansion of the brain per se. Indeed, human brain size has not grown constantly in history, with evidence proposing that it had been decreasing in the last ~3,000 years ago (44).

One of the arguably most distinctive features between humans and apes is our linguistic capacity (34), with the human neocortex particularly expanded around language-related areas (such as Brodmann areas 44 and 45) as well as an expanded arcuate fasciculus pathway interconnecting language-related areas (34, 45). Our findings suggest that genes containing evolutionary recent variants are expressed in frontal cortical areas (Figure 4E), most prominently in the gyrus pars triangularis - a key region within Broca's area and central to language processing and speech production (34). These findings are in line with comparative transcriptomic studies that indicate multimodal and language areas of the human brain are enriched for genes that have undergone accelerated divergence during the human lineage (HAR genes) (7).

Genetic modifications have driven the development of advanced cognitive skills (6, 46) through adaptations in brain areas and circuits (43). These genetic and neuronal modifications may have also rendered the human brain vulnerable to dysfunction (22, 23, 47), potentially explaining why many brain disorders are uniquely found within modern

humans (19, 21–23, 47). For example, genes in human accelerated areas of the genome have been noted to be enriched for neurodevelopmental disorders (22); and the strong genetic basis and human-specific character of schizophrenia have led to evolutionary theories that genetic adaptations of the human brain [related to the emergence of language (19, 47)] have contributed to vulnerability to the disorder (19). Our findings provide several lines of supporting evidence for such theories. First, genes intolerant to LoF tend to encompass more recently evolved SNPs than tolerant LoF genes. Second, recent genetic modifications in the modern-day human genome reveal a particular enrichment for psychiatric traits (e.g. depression, alcohol intake; *SI Appendix*), in line with recent evidence showing that introgressed variants from Neanderthals (variants relatively recent in human evolution) are linked to smoking, alcohol consumption, and mood-related traits (48); and with findings of an overrepresentation of alleles conferring risk for mood-related traits specifically in ancient farmers genome (~11,000 years old) but not in evolutionary previous hunter-gatherer genome (49). Third, genes related to neuropsychiatric conditions (SCZ, BD, and AD) were found to contain more recent genetic modifications compared to other genes. Fourth, genes containing relatively recent modifications were found to be particularly associated with genes involved in intelligence and size of the neocortex, aspects often reported to be affected in the context of psychiatric and neurological conditions (50).

Mapping the evolutionary genetic timeline of human-characteristic traits has inherent methodological limitations. To our knowledge, there is scarce available data on comprehensive dating estimates of the emergence of SNPs in the human genome, limiting the replication of the presented results with alternative resources. SNP dating estimates were matched to GWAS results which usually only include common variants, thus excluding rare variants linked to human phenotypes from our analyses. Gene-set analysis provided evidence that genes containing recent modifications are statistically enriched for genes associated with brain and cognitive phenotypes. This does not necessarily imply that

evolutionary old variants have never been implicated in evolution with brain or cognitive phenotypes. Furthermore, we set side-by-side the evolutionary milestones with the time distribution of SNPs, but this does not include causal evidence of forces shaping the genome. We note that human-phenotypic SNPs are associated with modern human trait variation, leaving it unknown whether they are also associated with the same trait in ancient hominins or other animals.

Methods

Human Dating Genome Data. Human Dating Genome (HDG) (11) data was downloaded from https://human.genome.dating/. HDG was developed by inferring the time of the most recent common ancestor between individual genomes, using two nonparametric approaches to estimate the date of origin of genetic variants by means of a recombination clock and mutation clock (11). The HDG consisted of 13,689,983 SNPs of the human genome mapped across 22 chromosomes with dates ranging from 5,140,625 to 87.5 years ago, with no assumptions about the demographic or selective processes that shaped the underlying genealogy. The present study used the median age estimates from both clocks combined on the two large-scale sequencing datasets.

Human-phenotypic SNPs. Significant lead SNPs were selected from a wide range of human GWAS [GWAS Atlas (10), https://atlas.ctglab.nl, access date June 2021]. In total, date estimates of unique 36,506 SNPs across 2,538 GWAS from human phenotypes were obtained (ranging from 4,556,425 to 1,681 years ago; *SI Appendix*). The genetic timeline was computed by counting the number of human-phenotypic SNPs per time-bin; SNP density per bin was computed by dividing SNP count per bin by the total number across all time-bins.

BRAIN-SNPs. A set of 6,284 SNPs (2,476 unique SNPs; BRAIN-SNPs) significantly associated across 1,744 brain phenotypes was derived based on the extensive UK Biobank BIG40 (26) GWAS ($P < 5 \times 10^{-8}$, discovery sample; https://open.win.ox.ac.uk/ukbiobank/big40/ SNPs were grouped according to brain structure (*SI Appendix*, Table S2).

Statistical inference. Permutation testing was used to statistically test the median evolutionary age of a set of SNPs or genes associated with a phenotype. A null-model was built by computing the median evolutionary age of equally sized random SNPs (or genes) at 10,000 iterations, assigning a z-score across iterations and a matching *P* value to the effect of interest (description of other permutation tests in *SI Appendix*).

Genes evolutionary age estimation. The genetic timeline of genes was computed by taking the median evolutionary age of the SNPs located within a gene, defined by the SNPs coordinate that is within the start and stop position of the gene (GRCh37; sensitivity analysis of 1-kb window both sides resulted in similar results). SNPs located outside of genes or assigned to two or more genes were excluded. Median evolutionary age estimates were obtained for 18,328 genes (out of 26,836 genes) with at least one SNP with a date estimate (median of ~100 SNPs within genes) ranging from 2,965,600 to 3,800 years ago.

Gene-set analysis. MAGMA competitive gene-set analysis (31) was performed (one-sided positive direction) on the set of the top 2,000 genes (~10% of all genes) with youngest/oldest median evolutionary age, testing for the enrichment of effects on the gene-based statistics (details of gene analysis in *SI Appendix*) for five brain disorders [SCZ, BD, AD, major depressive disorder (MDD), and autism spectrum disorder (ASD)], three brain (brain volume, thickness, and area) and two cognitive phenotypes (intelligence and sociability; ten phenotypes tested twice for each young/old gene-set, Bonferroni *P* value threshold < 2.5 x

10^{-3}). Gene-set analysis was conditioned for gene size, LD levels between SNPs in genes, mean minor allele count in genes, and log values of these three variables.

Data availability. HDG is available at https://human.genome.dating. GWAS Atlas is available at https://atlas.ctglab.nl. EBI Catalog is available at https://www.ebi.ac.uk/gwas. UK Biobank BIG40 is available at https://open.win.ox.ac.uk/ukbiobank/big40. Summary statistics of GWAS from brain disorders, intelligence, and brain volume are available at https://www.med.unc.edu/pgc/download-results; from cortical area and thickness are available at https://enigma.ini.usc.edu/research/download-enigma-gwas-results; and from sociability is available at https://www.repository.cam.ac.uk/handle/1810/277812. Cortical gene microarray transcriptome expression data from the Allen Human Brain Atlas is available at http://human.brain-map.org/static/download.

Acknowledgments. M.P.v.d.H. is supported by a VIDI (452-16-015) grant from the Netherlands Organization for Scientific Research (NWO) and an European Research Council consolidator grant (ID 101001062). D.P. is supported by an NWO Gravitation project BRAINSCAPES: A Roadmap from Neurogenetics to Neurobiology (024.004.012) and a European Research Council advanced grant (ID 834057).

Author Contributions. M.P.v.d.H. and I.L. conceived the project. I.L. analysed the data. M.P.v.d.H, and I.L. wrote the manuscript. Y.W., C.d.L., J.R., and D.P. provided expertise and feedback on the text and analyses.

Competing interests

The authors declare no competing interests.

Figures

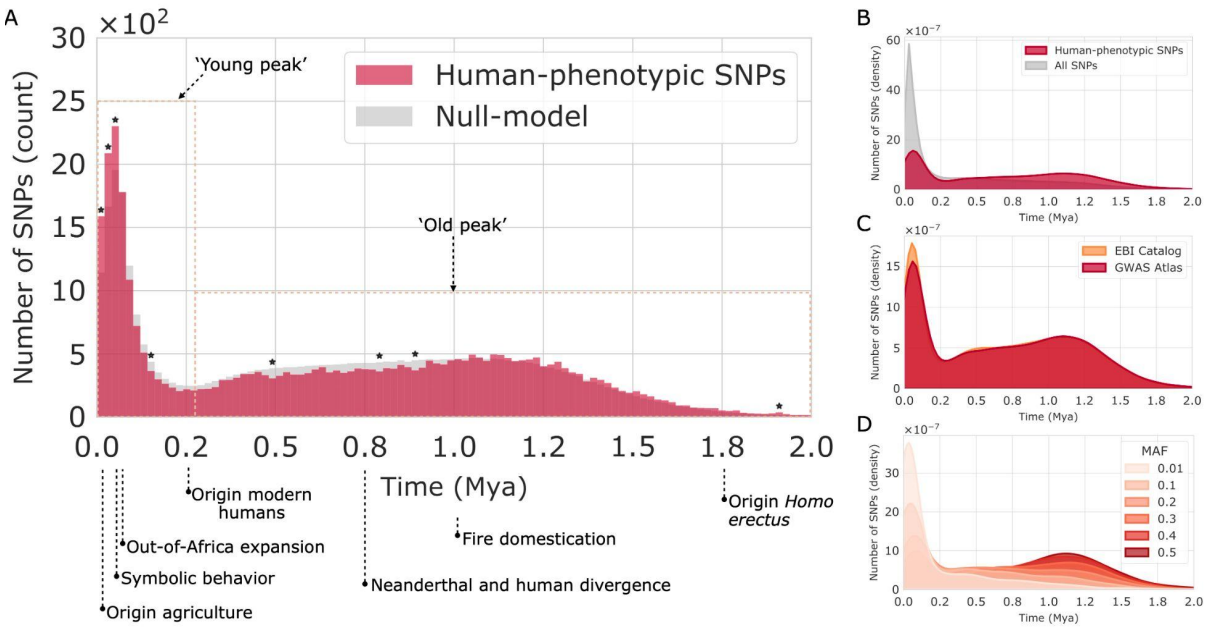

Figure 1. Genetic timeline of human-phenotypic SNPs.
(a) Absolute count (y-axis) of number of SNPs (x-axis) per time-bin (100 bins of ~20,000 years old) of all significant human-phenotypic SNPs (red) associated with human traits extracted from the GWAS Atlas repository (2,538 GWAS studies, 36,506 unique SNPs). Human-phenotypic SNPs date estimates extracted from the Human Dating Genome (HDG) database range from 4,556,425 - 1,681 years ago. Asterisks (*) denote bins where the number of human-phenotypic SNPs per age bin significantly exceeds the null-model of random equally sized sets of SNPs selected from all SNPs from the HDG (100 tests, Bonferroni *P* value threshold < 5 x 10^{-4}). X-axis flags important human evolutionary milestones in the last 2M years: the origin of *Homo erectus* (~1.8M years ago), first signs of fire domestication (~1M years ago), Neanderthal and *Homo sapiens* divergence (~800,000 years ago), the origin of modern humans (~200,000 years ago), the most recent out of Africa expansion (~60,000 years ago), consistent evidence of symbolic behavior (cave art carvings and decoration, ritual artifacts, musical instruments; ~45,000 years ago), and the origin of agriculture (~10,000 years ago). *(b)* Timeline of the density (normalized count) of human-phenotypic SNPs (red) and the total set of SNPs included in the HDG (gray) across the last 2M years. *(c)* Validation: timeline of the density of human-phenotypic SNPs (GWAS Atlas) compared to the alternative EBI Catalog of GWAS results of human traits showing high consistency between both distributions of genetic variation. *(d)* Timeline of the density of human-phenotypic SNPs per MAF category (six bins with MAF ranges: 0 - 0.01, 0.01 - 0.1, 0.1 - 0.2, 0.2 - 0.3, 0.3 - 0.4, 0.4 - 0.5). Mya, million years ago; MAF, Minor Allele Frequency.

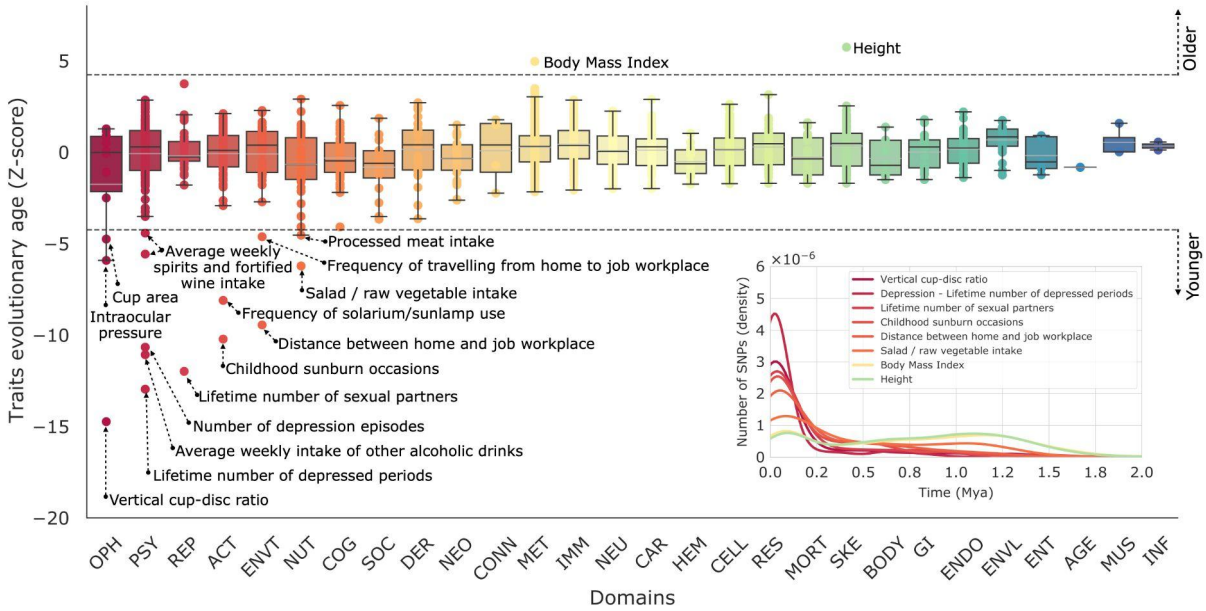

Figure 2. Genetic timeline of human traits. Expected age of traits (dots) under a null-model (*Methods*). Z-scores beyond the dotted line indicate traits (n = 2,251) with a median age of SNPs significantly younger (negative z-score values) or older (positive z-score values) than expected by chance (Bonferroni P value threshold < 2.2 x 10^{-5}). Black and gray lines within boxes denote domain median and mean z-scores, respectively. Exemplary genetic timelines of SNPs from significant traits are shown in the insert. Colors denote the different domains in GWAS Atlas. OPH, ophthalmological; PSY, psychiatric; REP, reproduction; ACT, activities; ENVT, environment; NUT, nutritional; COG, cognitive; SOC, social interactions; DER, dermatological; NEO, neoplasms; CONN, connective tissue; MET, metabolic; IMM, immunological; NEU, neurological; CAR, cardiovascular; HEM, hematological; CELL, cell; RES, respiratory; MORT, mortality; SKE, skeletal; BODY, body structures; GI, gastrointestinal; ENDO, endocrine; ENVL, environmental; ENT, ear, nose, throat; AGE, aging; MUS, muscular; INF, infection.

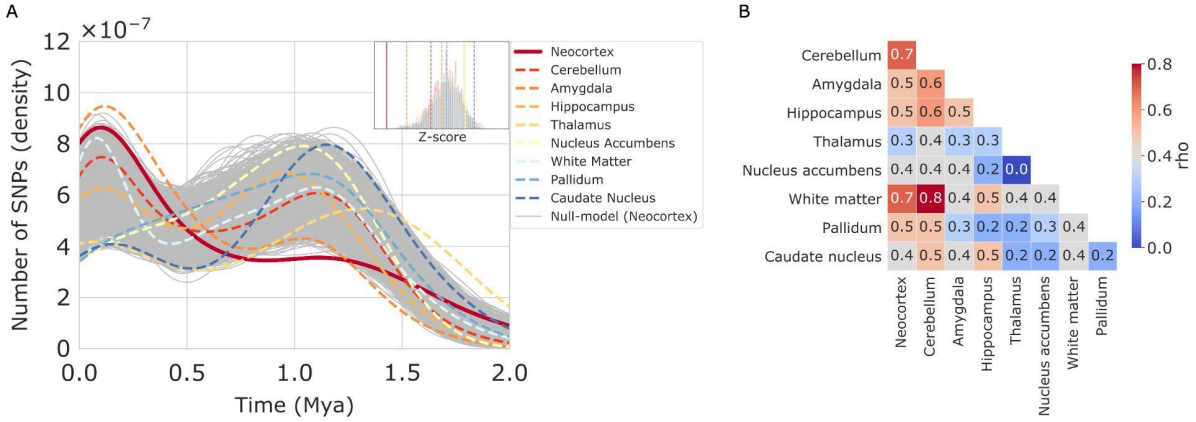

Figure 3. Genetic timeline of brain-imaging SNPs. *(a)* Histogram showing the density (y-axis) of the number of SNPs per age bin (x-axis). Solid lines denote phenotypes (cortex, red) significantly younger than expected by chance (see insert for the null-model of each brain phenotype). *(b)* Heatmap of the co-fluctuation between the timelines of SNPs related to brain structures (Spearman's correlation). All correlations are significant (Bonferroni *P* value threshold < 1.4 x 10^{-3}) except for the thalamus which only presents a significant correlation with white matter and cerebellum; and for pallidum which has a significant correlation with neocortex, cerebellum, nucleus accumbens, white matter, and caudate nucleus. Mya, million years ago.

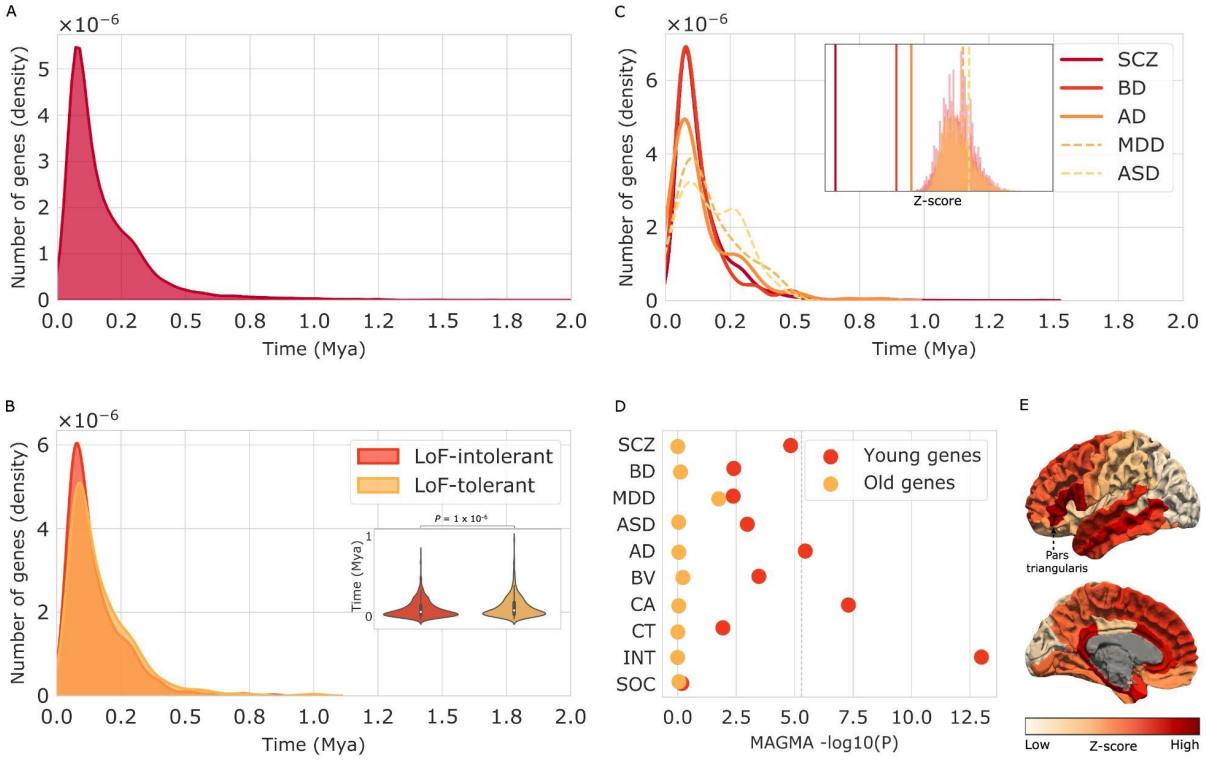

Figure 4. Genetic timeline of genes. *(a)* Timeline of genes (n = 18,328) with date estimates ranging from 2,965,600 to 3,803 years ago. Histogram depicts the density (y-axis) of the number of genes per age bin (x-axis; until 2M years ago is shown). *(b)* Timeline of the set of genes most intolerant to loss-of-function (LoF; orange) and the most tolerant to LoF (yellow); insertion (y-axis genes-age) shows LoF-intolerant genes (red) to be significantly younger (pair-sample t-test, $P = 1 \times 10^{-6}$) than LoF-tolerant genes (yellow). *(c)* Timeline of genes associated with five major brain disorders, including SCZ, BD, AD, MDD, and ASD; solid lines denote phenotypes significantly different from the null model (Bonferroni P value threshold $< 5 \times 10^{-3}$, SCZ, BD, AD; see insert for the null model distribution of each condition). (*d*) Plots show P values (x-axis) of MAGMA gene-set analysis of the top oldest (yellow dots) and youngest (red dots) genes for enrichment of genes related to neuropsychiatric and brain phenotypes (y-axis). Dotted line indicates Bonferroni P value threshold $< 2.5 \times 10^{-3}$. Figure shows strong enrichment of the young genes for intelligence, and cortical area (also, AD and SCZ nominally significant). In contrast, the old genes are not found to be significantly enriched in disorders and/or brain aspects. *(e)* Brain plot of the normative expression levels of the genes with the youngest evolutionary age across brain areas, ranging from relatively low (yellow) to high (red) expression levels (z-scores). Figure shows high expression in Broca's area (pars triangularis, a region part of Brodmann Area 45) a key brain area for language ($P = 6.2 \times 10^{-4}$, showing expression levels significantly exceeding the null-model of random genes expression, see *Methods*). Mya, million years ago; LoF, loss-of-function; SCZ, schizophrenia; BD, bipolar disorder; AD, Alzheimer's disease; MDD, major depressive disorder; ASD, autism spectrum disorder; BV, brain volume; CA, cortical area; CT, cortical thickness; INT, intelligence; SOC, sociability.

References

1. E. K. Irving-Pease, R. Muktupavela, M. Dannemann, F. Racimo, Quantitative Human Paleogenetics: What can Ancient DNA Tell us About Complex Trait Evolution? *Front. Genet.* **12**, 703541 (2021).
2. P. D. Gingerich, Pattern and rate in the Plio-Pleistocene evolution of modern human brain size. *Sci Rep* **12**, 11216 (2022).
3. A. I. R. Herries, *et al.*, Contemporaneity of Australopithecus, Paranthropus, and early Homo erectus in South Africa. *Science* **368** (2020).
4. R. A. Barton, Evolutionary specialization in mammalian cortical structure. *J. Evol. Biol.* **20**, 1504–1511 (2007).
5. D. C. Van Essen, C. J. Donahue, M. F. Glasser, Development and Evolution of Cerebral and Cerebellar Cortex. *Brain Behav Evol* **91**, 158–169 (2018).
6. S. Shultz, E. Nelson, R. I. Dunbar, Hominin cognitive evolution: identifying patterns and processes in the fossil and archaeological record. *Philos Trans R Soc Lond B Biol Sci* **367**, 2130–40 (2012).
7. Y. Wei, *et al.*, Genetic mapping and evolutionary analysis of human-expanded cognitive networks. *Nat Commun* **10**, 4839 (2019).
8. C. C. Sherwood, F. Subiaul, T. W. Zawidzki, A natural history of the human mind: tracing evolutionary changes in brain and cognition. *J Anat* **212**, 426–54 (2008).
9. T. M. Preuss, M. Caceres, M. C. Oldham, D. H. Geschwind, Human brain evolution: insights from microarrays. *Nat Rev Genet* **5**, 850–60 (2004).
10. K. Watanabe, *et al.*, A global overview of pleiotropy and genetic architecture in complex traits. *Nat Genet* **51**, 1339–1348 (2019).
11. P. K. Albers, G. McVean, Dating genomic variants and shared ancestry in population-scale sequencing data. *PLoS Biol* **18**, e3000586 (2020).
12. A. Buniello, *et al.*, The NHGRI-EBI GWAS Catalog of published genome-wide association studies, targeted arrays and summary statistics 2019. *Nucleic Acids Res* **47**, D1005–D1012 (2019).
13. A. Gomez-Robles, Dental evolutionary rates and its implications for the Neanderthal-modern human divergence. *Sci Adv* **5**, eaaw1268 (2019).
14. L. Vigilant, M. Stoneking, H. Harpending, K. Hawkes, A. C. Wilson, African populations and the evolution of human mitochondrial DNA. *Science* **253**, 1503–7 (1991).
15. J. J. Hublin, *et al.*, New fossils from Jebel Irhoud, Morocco and the pan-African origin of Homo sapiens. *Nature* **546**, 289–292 (2017).
16. O. Bar-Yosef, The upper paleolithic revolution. *Annu. Rev. Anthropol.*, 363–393 (2002).
17. H. Kobayashi, S. Kohshima, Unique morphology of the human eye and its adaptive meaning: comparative studies on external morphology of the primate eye. *J Hum Evol* **40**, 419–35 (2001).
18. E. Pearce, C. Stringer, R. I. Dunbar, New insights into differences in brain organization between Neanderthals and anatomically modern humans. *Proc. R. Soc. B Biol. Sci.* **280**, 20130168 (2013).
19. T. J. Crow, Schizophrenia as the price that homo sapiens pays for language: a resolution of the central paradox in the origin of the species. *Brain Res Brain Res Rev* **31**, 118–29 (2000).
20. M. J. Rantala, S. Luoto, J. I. Borraz-Leon, I. Krams, Bipolar disorder: An evolutionary psychoneuroimmunological approach. *Neurosci Biobehav Rev* **122**, 28–37 (2021).
21. C. E. Finch, S. N. Austad, Commentary: is Alzheimer's disease uniquely human? *Neurobiol Aging* **36**, 553–5 (2015).
22. R. N. Doan, *et al.*, Mutations in Human Accelerated Regions Disrupt Cognition and Social Behavior. *Cell* **167**, 341-354 e12 (2016).
23. K. Pattabiraman, S. K. Muchnik, N. Sestan, The evolution of the human brain and disease susceptibility. *Curr Opin Genet Dev* **65**, 91–97 (2020).

24. S. C. Antón, Natural history of Homo erectus. *Am. J. Phys. Anthropol. Off. Publ. Am. Assoc. Phys. Anthropol.* **122**, 126–170 (2003).
25. M. Slatkin, Linkage disequilibrium--understanding the evolutionary past and mapping the medical future. *Nat Rev Genet* **9**, 477–85 (2008).
26. S. M. Smith, *et al.*, An expanded set of genome-wide association studies of brain imaging phenotypes in UK Biobank. *Nat Neurosci* **24**, 737–745 (2021).
27. J. K. Rilling, Comparative primate neuroimaging: insights into human brain evolution. *Trends Cogn Sci* **18**, 46–55 (2014).
28. W. D. Hopkins, J. K. Rilling, A comparative MRI study of the relationship between neuroanatomical asymmetry and interhemispheric connectivity in primates: implication for the evolution of functional asymmetries. *Behav Neurosci* **114**, 739–48 (2000).
29. A. Yates, *et al.*, Ensembl 2016. *Nucleic Acids Res.* **44**, D710–D716 (2016).
30. S. Balasubramanian, *et al.*, Using ALoFT to determine the impact of putative loss-of-function variants in protein-coding genes. *Nat Commun* **8**, 382 (2017).
31. C. A. de Leeuw, J. M. Mooij, T. Heskes, D. Posthuma, MAGMA: generalized gene-set analysis of GWAS data. *PLoS Comput Biol* **11**, e1004219 (2015).
32. M. J. Hawrylycz, *et al.*, An anatomically comprehensive atlas of the adult human brain transcriptome. *Nature* **489**, 391–399 (2012).
33. Y. Wei, *et al.*, Statistical testing in transcriptomic-neuroimaging studies: A how-to and evaluation of methods assessing spatial and gene specificity. *Hum Brain Mapp* **43**, 885–901 (2022).
34. A. D. Friederici, Evolution of the neural language network. *Psychon Bull Rev* **24**, 41–47 (2017).
35. M. M. Mesulam, From sensation to cognition. *Brain* **121 (Pt 6)**, 1013–52 (1998).
36. F. Berna, *et al.*, Microstratigraphic evidence of in situ fire in the Acheulean strata of Wonderwerk Cave, Northern Cape province, South Africa. *Proc Natl Acad Sci U A* **109**, E1215-20 (2012).
37. C. Posth, *et al.*, Pleistocene Mitochondrial Genomes Suggest a Single Major Dispersal of Non-Africans and a Late Glacial Population Turnover in Europe. *Curr Biol* **26**, 827–33 (2016).
38. A. M. Moore, “The transition from foraging to farming in Southwest Asia: Present problems and future directions” in *Foraging and Farming*, (Routledge, 2014), pp. 620–631.
39. F. d’Errico, *et al.*, Archaeological evidence for the emergence of language, symbolism, and music–an alternative multidisciplinary perspective. *J. World Prehistory* **17**, 1–70 (2003).
40. N. K. Schaefer, B. Shapiro, R. E. Green, An ancestral recombination graph of human, Neanderthal, and Denisovan genomes. *Sci Adv* **7** (2021).
41. B. L. Finlay, R. B. Darlington, Linked Regularities in the Development and Evolution of Mammalian Brains. *Science* **268**, 1578–1584 (1995).
42. D. J. Ardesch, *et al.*, Evolutionary expansion of connectivity between multimodal association areas in the human brain compared with chimpanzees. *Proc. Natl. Acad. Sci.* **116**, 7101–7106 (2019).
43. D. J. Ardesch, *et al.*, Scaling Principles of White Matter Connectivity in the Human and Nonhuman Primate Brain. *Cereb Cortex* **32**, 2831–2842 (2022).
44. J. M. DeSilva, J. F. Traniello, A. G. Claxton, L. D. Fannin, When and why did human brains decrease in size? A new change-point analysis and insights from brain evolution in ants. *Front. Ecol. Evol.*, 712 (2021).
45. J. K. Rilling, *et al.*, The evolution of the arcuate fasciculus revealed with comparative DTI. *Nat Neurosci* **11**, 426–8 (2008).
46. D. H. Bailey, D. C. Geary, Hominid brain evolution. *Hum. Nat.* **20**, 67–79 (2009).
47. M. P. van den Heuvel, *et al.*, Evolutionary modifications in human brain connectivity associated with schizophrenia. *Brain* **142**, 3991–4002 (2019).
48. M. Dannemann, *et al.*, Neandertal introgression partitions the genetic landscape of

neuropsychiatric disorders and associated behavioral phenotypes. *Transl Psychiatry* **12**, 433 (2022).
49. E. K. Irving-Pease, *et al.*, “The Selection Landscape and Genetic Legacy of Ancient Eurasians” (Evolutionary Biology, 2022) https:/doi.org/10.1101/2022.09.22.509027 (October 17, 2022).
50. P. M. Thompson, *et al.*, ENIGMA and global neuroscience: A decade of large-scale studies of the brain in health and disease across more than 40 countries. *Transl Psychiatry* **10**, 100 (2020).