Title

Mapping the genetic evolutionary timeline of human neural and cognitive traits

Author

Ilan Libedinsky[1], Yongbin Wei[1,2], Christiaan de Leeuw[1], James K. 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. Emory 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, PhD, De Boelelaan 1085, Amsterdam, 1081 HV, the Netherlands.

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

Abstract

Human evolution is characterised by extensive changes of body and brain, with perhaps one of the core developments being the fast increase in cranial capacity and brain volume. Paleontological records are the most direct method to study such changes, but they can unfortunately provide a limited view of how 'soft traits' such as brain function and cognitive abilities have evolved in humans. A potential complementary approach is to identify when particular genetic variants associated with human phenotypes (such as height, body mass index, intelligence, and also disease) have emerged in the 6-7 million years since we diverged from chimpanzees. In this study, we combine data from genome-wide association studies on human brain and cognitive traits with estimates of human genome dating. We systematically analyse the temporal emergence of genetic variants associated with modern-day human brain and cognitive phenotypes over the last five million years. Our analysis provides evidence that genetic variants related to neocortex structure (e.g., area, thickness; median evolutionary age = 400,170 years old), cognition (e.g., fluid intelligence; median age = 459,465), education (median age = 637,646), and psychiatric disorders (median age = 412,639) have emerged more recently in human evolution than expected by chance. In contrast, variants related to other physical traits, such as height (median age = 811,305) and body mass index (median age = 794,265), emerged relatively later. We further show that genes containing recent evolutionary modifications (from around 54,000 to 4,000 years ago) are linked to intelligence ($P = 2 \times 10^{-6}$) and neocortical surface area ($P = 6.7 \times 10^{-4}$), and that these genes tend to be highly expressed in cortical areas involved in language and speech (pars triangularis, $P = 6.2 \times 10^{-4}$). Elucidating the temporal dynamics of genetic variants associated with brain and cognition is another source of evidence to advance our understanding of human evolution.

Main

The human genome contains a footprint of our evolutionary history. Mapping the timeline of when specific genetic variants associated with modern-day phenotypes have emerged in our genome provides a unique window to examine the evolution of human features. This approach is particularly valuable for studying the evolutionary timeline of neural and cognitive traits[1], which leave no direct physical traces in the fossil record.

Human evolution involved dramatic changes to the brain and human cognitive abilities. Cranial capacity tripled over the past two million years of human evolution, a pace of encephalization that is unparalleled in other mammals[2], with the neocortex being among the most expanded brain regions[3]. This expansion is widely believed to be one of the strongest catalysers for the development of complex behaviour and advanced cognition in the human lineage[3–5]. Human encephalization is thought to be the outcome of a wide range of natural (e.g. climate), nutritional (e.g. diet), and social (e.g. group size, parental care)[6] selection pressures. These pressures have synergistically interacted with, and led to, modifications in the human genome at different periods across history[7]. In recent years, a rapidly growing number of Genome-Wide Association Studies (GWAS) have begun to unravel the genetic basis of modern human phenotypes[8], and in this study we combined dating estimates for the emergence of variants in the human genome[9] with these GWAS findings[8] to map the genetic evolutionary timeline of human brain and cognitive abilities. We show that the emergence of genetic variants associated with core human attributes, such as brain morphology, cognition, and neuropsychiatric conditions, follows a distinctive temporal pattern, revealing recent genetic modifications in human evolution.

Results

Genetic timelines of human phenotypes run parallel to evolutionary milestones

We began by examining when single nucleotide polymorphisms (SNPs) associated with a wide range of human phenotypes appeared in the human genome. Data from the Human Dating Genome project (HDG)[9] were used for this purpose. The HDG is 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 genetic variants across 22 chromosomes of the human genome covering ~200,000 generations. This database provides dating information from ~5.1M to ~100 years ago (Methods, Supplementary Methods). We first examined the genetic timeline of a total of 67,119 variants (36,506 unique lead SNPs; referred to as human-phenotypic SNPs), combined from the summary statistics of 2,538 GWAS[8] associated with a broad range of human phenotypes (e.g. eye shape, cancer, height, brain, cognition, behaviour). Using the HDG we computed the dates when human-phenotypic SNPs appeared and compared them to the dates for all other SNPs in the human genome. The combined genetic timeline of the 36,506 human-phenotypic SNPs (ranging from ~4.5M to ~2,000 years ago, median evolutionary age = 673,520 years old) displayed a bimodal distribution with two distinct peaks (Figure 1A): an 'old peak' of genetic variants arising from ~2.4M until ~280,000 years ago reaching a maximum at ~1.1M years ago, comprising ~85% out of all variants with high minor allele frequency (MAF >= 0.4; Supplementary Methods); and a second 'young peak' of genetic variants arising at ~280,000 until ~2,000 years ago reaching a maximum at ~55,000 years ago, a period including ~62% of variants with low MAF (MAF <= 0.1; Figure 1B).

Permutation testing comparing the timeline of human-phenotypic SNPs to that of all SNPs allowed to identify periods of time when there were more (or less) genetic variants than expected by chance (MAF-controlled, 10,000 permutations; see Supplementary Methods). This analysis revealed a significant increase in human-phenotypic SNPs at the beginning of the old peak (from ~1,920,000 to 1,900,000 years ago, $P = 2.5 \times 10^{-8}$). A high concentration of human-phenotypic variants was also found at the highest point of the young peak, ranging from ~61,000 to ~2,000 years ago (three consecutive significant bins, $P = 4.2$

x 10^{-16}, $P = 6.2 \times 10^{-27}$, and $P = 2.3 \times 10^{-40}$; Figure 1C). After the highest point of the old peak, three separate time periods show a slightly less human-phenotypic SNPs than expected (~900,000 to ~880,000 years ago, $P = 1.7 \times 10^{-4}$; ~800,000 to ~780,000 years ago, $P = 3.1 \times 10^{-4}$; ~500,000 to ~480,000 years ago, $P = 5.0 \times 10^{-5}$), and once after the beginning of the young peak (~160,000 to ~140,000 years ago, $P = 4.2 \times 10^{-4}$; Figure 1C). We observed this bimodal distribution again by selecting human-phenotypic SNPs from the EBI Catalog[10] as an alternative resource (including unique non-overlapping variants with GWAS Atlas, *n* = 88,362), resulting in a highly similar distribution (*rho* = 0.96, $P = 5.5 \times 10^{-59}$; Figure 1D, Supplementary Results). Figure S1 further shows the genetic timeline of human-phenotypic SNPs grouped by their functional genetic consequences (e.g. intergenic, missense).

Plotting the bimodal genetic timeline side-by-side with the emergence of evolutionary milestones in the human lineage identified several remarkable points of overlap: the old peak coincides with (among others, see also Discussion) the origin of *Homo erectus* (~1.8M years ago)[11] and the divergence of *Homo sapiens* and Neanderthals (~800,000 years ago)[12]. The beginning of the young peak co-occurs with the origin of modern humans in Africa (~200,000 years ago)[13,14] 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[15].

Psychiatric phenotypes influenced by evolutionarily recent genetic modifications

We utilised data from the GWAS Atlas[8] to gain insight into the genetic timeline of specific classes of human characteristics. This procedure involved grouping identified genetic variants according to a nested hierarchical organisation, ranging from global categories to specific phenotypes according to domain, chapter, subchapter, and trait levels. Using permutation testing, we assessed the median evolutionary age for each phenotype to determine whether it was younger or older than expected by random chance (null-distribution of equally sized random sets out of all human-phenotypic SNPs, details in

Methods). At the domain level (n = 28 domains, Bonferroni P value threshold < 1.7 x 10^{-3}; Figure 2, Figure S2), genetic variants related to 'Psychiatric' phenotypes presented an evolutionary age younger than expected by chance (median evolutionary age = 412,639 years old), a class of traits suggested to be linked to human brain evolution[5,16–20]. 'Ophthalmological' and 'Nutritional' phenotypes as well displayed a significantly younger evolutionary age (median age = 30,484 and 429,993, respectively). In contrast, genetic variants associated with 'Metabolic', 'Skeletal' and 'Immunological' phenotypes overall displayed the relatively oldest age estimates (median age = 780,323, 759,056, and 747,780, respectively; all P values listed in Table S1).

Sensitivity analyses

We performed a series of sensitivity analyses to test the robustness of the presented findings. We first confirmed our results by examining also the more fine-grained chapter (Figure S3), subchapter (Figure S4), and trait level (Figure S5), which all showed similar results with the findings reported above (for all results see Table S1). We further examined whether our findings were robust when controlled for differences in the number of SNPs per phenotype, the sample size of the included GWAS, and when taking into account variants with the most accurate date estimates (Supplementary Results). We also considered only independent SNPs in Linkage Disequilibrium (LD) blocks[21] (R^2 < 0.1) and outside the MHC (Supplementary Results). All these analyses confirmed our main findings.

We also specifically examined the effect of an expected positive association between the MAF and the age of emergence of a genetic variant (r = 0.38, P < 1 x 10^{-100}). We investigated whether the median evolutionary age of variants related to a phenotype deviated from the expected values when taking into account both the number of variants and their corresponding MAF (Supplementary Methods; 33,621 unique SNPs age and MAF estimates were available; 26 domains, Bonferroni P value threshold < 1.9 x 10^{-3}). The results of this sensitivity analysis suggest that selection pressure may have influenced the variants

linked to significant phenotypes[22]. When controlling both polygenicity and MAF, SNPs related to 'Psychiatric' phenotypes presented the most recent evolutionary age (median evolutionary age = 475,833 years old), followed by 'Activities' (e.g. looking after one's health, use of communication devices; median age = 644,450), and 'Environment' phenotypes (e.g. education, work; median age = 626,965). In contrast, SNPs associated with 'Neoplasm' and 'Metabolic' phenotypes exhibited the relatively oldest evolutionary age (median age = 591,690 and 785,572, respectively; all *P* values listed in Table S2). The effects observed in the 'Psychiatric', 'Activities', and 'Neoplasm' domains further remained statistically significant after also controlling for LD (by including SNPs with an R^2 below 0.1 and excluding SNPs within the MHC; Supplementary Results).

Neocortical variation linked to evolutionarily recent genetic modifications

Further examination was conducted on the genetic timeline of several brain structures by analysing the date estimates of 6,284 genetic variants (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 S3)[23]. The genetic timeline of BRAIN-SNPs ranged from ~3.6M to ~4,000 years ago (median evolutionary age = 716,318 years old; Figure S7). Neocortex SNPs (a subset of 130 SNPs out of the 2,476 BRAIN-SNPs; Supplementary Methods) exhibited a median evolutionary age significantly younger than expected by chance when controlling for the number of variants (median age = 400,170 years old, $P = 4.4 \times 10^{-4}$; Figure 3A; complete results in Table S4) and a trend towards significance when controlling for MAF as well ($P = 0.06$). The neocortex is believed to be the most expanded brain structure in the *Homo* lineage[24], now vastly exceeding the volume and size found in chimpanzees and bonobos[5,25], and substrate of complex cognitive abilities including multisensory processing and sociability[26,27]. No significant results were found 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 (Bonferroni P value threshold was applied < 5.5 x 10^{-3} to all tests performed; Supplementary Results).

Genes containing recent mutations are enriched for genes related to cortical surface area, intelligence, and neuropsychiatric disorders

An estimate of the evolutionary age of each gene was derived by computation of the median age of the variants 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[28]; Methods). Computed gene evolutionary age varied between ~3M to ~4,000 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; Supplementary Methods)[29]. 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), cognition (e.g. intelligence, sociability), and neuropsychiatric disorders (Supplementary Methods; included studies in Table S6). Gene-set analysis[30] (Methods) was used to examine the top 2,000 genes with the youngest evolutionary age (~10% from all genes; ranging from ~54,000 to ~4,000 years ago) and their association with genes-sets related to variation in these ten phenotypes (using Bonferroni P value threshold < 2.5 x 10^{-3}; Table S7). Genes containing recent mutations were found to be significantly enriched for genes involved in intelligence (b = 0.13, P = 2 x 10^{-6}; using a lower or higher number of genes revealed similar findings, Supplementary Results) and cortical area (b = 0.08, P = 6.7 x 10^{-4}). 'Young genes' were possibly also 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}; Figure 4C),

but these effects did not survive our stringent Bonferroni correction. As a control condition, we also examined the group of genes with oldest evolutionary age estimates (ranging from ~3M to ~350,000 years ago) and no significant enrichment could be found ($P > 0.05$; Figure 4C, Table S7). A direct comparison of enrichment effect sizes confirmed that 'young genes' have a greater enrichment than 'old genes' for these phenotypes (Supplementary Results).

It was further tested the median evolutionary age of the set of genes associated with the same brain, cognitive, and neuropsychiatric phenotypes ($q > 0.05$, FDR; Supplementary Methods) compared to a null-model of randomly selected genes (ten phenotypes, Bonferroni P value threshold $< 5 \times 10^{-3}$; 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; $P = 2.6 \times 10^{-3}$), suggesting that genes linked to these phenotypes contain relatively more recent mutations in the human genome than compared to all genes (Figure 4D, Figure S8; sensitivity analysis in Supplementary Results). These findings corroborate theories that link neuropsychiatric conditions to recent human evolution[16–20].

Genes containing recent mutations are highly expressed in language areas

We further used gene expression data from the Allen Human Brain Atlas[31] to investigate the transcriptomic pattern of the youngest genes. 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; analyses with other thresholds in Supplementary Results) were found to be highly expressed in the gyrus pars triangularis ($P = 6.2 \times 10^{-4}$; Figure 5, Table S10; sensitivity analyses in Supplementary Results), a region recognised as Broca's area and well known to be involved in speech and language, a human unique trait[32]. These genes were found to be more expressed in multimodal regions[33] 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 x 10^{-2}; Bonferroni P value threshold 2.5 x 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 x 10^{-3}).

Discussion

Mapping the evolutionary genetic timeline during human evolution reveals distinct timelines for traits related to human brain organisation and cognition. Traits related to neocortex, intelligence and psychiatric disorders are found to be relatively evolutionarily more recent, with particularly enrichment of variants related to these traits around a 'young peak' starting from ~280,000 and continuing until ~2,000 years ago (reaching the maximum point at ~55,000 years ago). It is important to acknowledge that the analyses conducted in this study made use of all available GWAS data, of which the majority (91%) focused on European populations. Consequently, using these analyses to draw conclusions about the genetic timelines of phenotypes among different populations is not feasible.

The two observed 'old' and 'young' peaks line up with important paleontological observations. The old peak highlights the origin of *Homo erectus*[11] and the divergence of *Homo sapiens* and Neanderthals[12], a period displaying an enrichment for genetic variants 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[13,14], 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)[34], followed by the Upper Paleolithic period (~45,000 years ago)[15], and the beginning of agriculture (~10,000 years ago)[35]. Comparative genetic analyses of archaic human ancestry revealed that a large number of adaptive changes in the human genome occurred around ~600,000 and ~200,000 years ago[36],

coinciding with the observed genetic timeline of the old and young peak from the human-phenotypic SNPs.

The Upper Paleolithic is considered to be a period in which important revolutions in cognition occurred[15], characterised by the first consistent evidence of symbolic behaviour (i.e. art carvings, ritual artefacts, musical instruments), technological advancement (i.e. microlithic stone tools and bone artefacts, upgraded hunting tools), and ecological expansion (i.e. long-distance trading of materials)[37]. Our findings suggest an enrichment around this period for genetic variants associated with brain and neuropsychiatric phenotypes of modern-day humans. Consistent with these results, a previous study reported an enrichment for biological processes related to postsynaptic membrane in genes associated with modern-specific human variants arising within the last ~300,000 years ago[38], which also aligns with our observation of an enrichment of neuronal processes in genes exhibiting evolutionary recent age (Table S8).

The genetic timeline of neocortical organisation of modern humans presents an enrichment of variants that emerged recently in evolution (Figure 3A). Genetic modifications are widely believed to have played a major role in the growth and remodelling of the neocortex in humans[7]. The human brain has undergone major changes in size and structure since the last common ancestor shared by modern humans and other primates[39–41], with the neocortex being the most expanded region[3]. These genetic and neocortical modifications potentially set humans apart from other primates[3–5], enabling higher cognitive abilities such as complex language, social cognition, and higher-order thinking[40,42]. 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 expanded continuously in history, with evidence it has even been decreasing in the last ~3,000 years ago[43].

One of the arguably most distinctive features between humans and other apes is our linguistic capacity[32], 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[32,44]. Our findings suggest that genes containing evolutionary recent variants are expressed in frontal cortical areas (Figure 5), most prominently in the gyrus pars triangularis - a key region within Broca's area and central to language processing and speech production[32]. These findings align with comparative transcriptomic studies indicating that multimodal and language areas of the human brain are enriched for genes that have undergone accelerated divergence during the human lineage (known as HAR genes)[5]. Our findings also converge with evidence of selection pressures acting on alleles involved in variation of surface area of the inferior frontal cortex[45].

Genetic modifications have driven the development of advanced cognitive skills[4,46] through adaptations in brain areas and circuits[47]. These genetic and neuronal changes may have also rendered the human brain vulnerable to dysfunction[19,20,48], potentially shaping certain brain disorders to be unique or specific to humans[16,18,20]. For instance, we observe that genetic variants related to the nervous system emerged before than variants underlying cognition (two-sided t-test, $t = 3.95$, $P = 7.9 \times 10^{-5}$) and psychiatric disorders ($t = 8.91$, $P = 6.9 \times 10^{-19}$; Figure S6). Previous studies have found that genes in human accelerated areas of the genome have been noted to be enriched for neurodevelopmental disorders[19]; 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) have contributed to vulnerability to the disorder[16,48]. Our findings provide several lines of supporting evidence for such theories. First, recent genetic modifications in the modern-day human genome reveal a particular enrichment for psychiatric traits (e.g. depression, alcohol intake; see Figure 2 and Table S1), in line with recent evidence showing that introgressed

variants from Neanderthals (variants relatively recent in human evolution, ~50,000 years ago) are linked to smoking, alcohol consumption, and mood-related traits[49]; and with findings of an overrepresentation of alleles conferring risk for mood-related traits specifically in the genomes of ancient farmers (~11,000 years old), but not in the genomes of evolutionarily earlier hunter-gatherer[50]. Second, genetic variants associated with intelligence, educational attainment, hippocampal volume, and psychiatric disorders appear to have a more recent evolutionary age than expected considering their polygenicity and MAF, suggesting the potential action of selection pressures on these phenotypes (Supplementary Results, Table S2). Third, genes intolerant to LoF variants tend to encompass more recently evolved mutations than genes tolerant to LoF (Figure 4B). Fourth, genes containing relatively recent modifications are particularly associated with genes involved in intelligence and neocortical surface area (Figure 4C), aspects exhibiting recent genetic selection[45] and often reported to be affected in the context of psychiatric and neurological conditions[51]. Fifth, genes related to neuropsychiatric conditions (SCZ, BD, and AD) contain more recent genetic modifications compared to other genes (Figure 4D).

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, constraining the replication of the presented results with alternative resources. In this study, we utilised all available SNPs date estimates. Dating precision varies among genetic variants, with certain SNPs demonstrating higher precision than others (e.g., lower confidence interval). We carried out sensitivity analyses focusing on genetic variants with the highest dating precision, which supported our primary findings (see Supplementary Results). One of our strategies involved matching SNP dating estimates with GWAS results. It is important to point out that GWAS results largely target common variants. This introduces a bias into our evolutionary age estimation of human phenotypes because it neglects rare

variants, which generally appear more recently in evolution, and entirely fixed variants, which tend to be older. Consequently, the median evolutionary age of traits should be approached with caution, and also take into account the lower and upper bounds of the evolutionary age (see Table S1 and Table S2). Our null models are designed by permuting human-phenotypic SNPs, ensuring that our statistical analyses account for this potential bias. In light of this, the current study focuses on human phenotypes that are still evolving, as opposed to traits that are fully evolved and have become fixed in the human genome. This is particularly relevant as human-phenotypic SNPs are associated with modern human trait variation, and it remains an open question whether these same SNPs were associated with similar traits in ancient hominins or other animals. Furthermore, the human evolutionary milestones were set side-by-side with the time distribution of genetic variants, which does not constitute causal evidence of forces shaping the genome.

Methods

Human Dating Genome Data. Human Dating Genome[9] (HDG) 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[9]. The HDG consisted of 13,689,983 SNPs of the human genome mapped across 22 chromosomes covering ~200,000 generations with dates ranging from 5,140,625 to 87.5 years ago (one generation was defined as 25 years), 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[8], 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; Supplementary Methods). 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[23] GWAS ($P < 5 \times 10^{-8}$, discovery sample; https://open.win.ox.ac.uk/ukbiobank/big40/ SNPs were grouped according to brain structure (see Supplementary Methods and Table S3).

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 Supplementary Methods). Sensitivity analysis was conducted controlling for both the number of SNPs and their corresponding MAF (Supplementary Methods).

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. The length

of genes exhibited a weak correlation with the median evolutionary age of genes (r = -0.058, P = 2.3 x 10^{-15}), suggesting that the gene length was not a confounder in the estimation of the evolutionary age.

Gene-set analysis. MAGMA competitive gene-set analysis[30] 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 Supplementary Methods) 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 a 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.

Figures

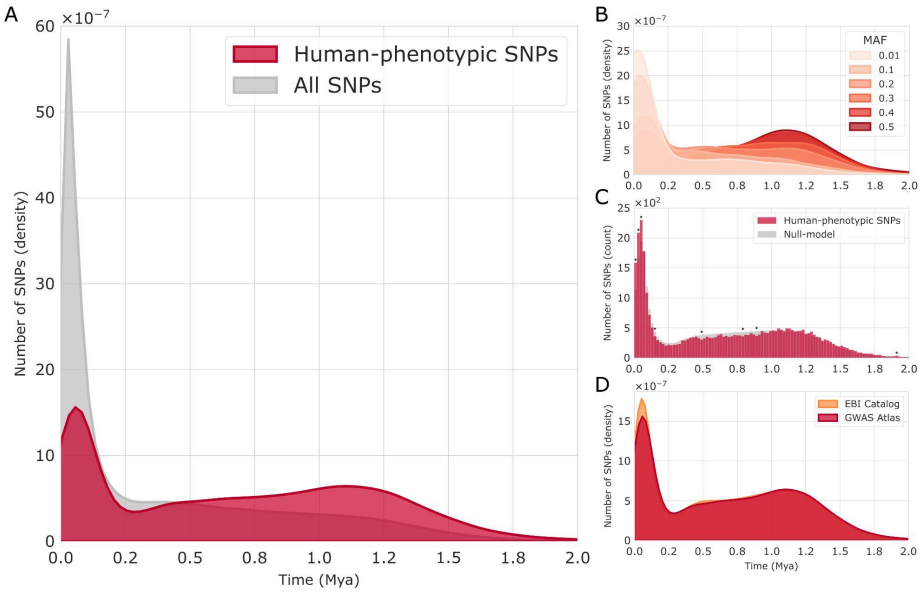

Figure 1. Genetic timeline of human-phenotypic SNPs. *(a)* Timeline of the density (normalised count; y-axis) of all significant human-phenotypic SNPs (red) associated with human traits extracted from the GWAS Atlas repository (2,538 GWAS studies, 36,506 unique SNPs) and the total set of SNPs (grey) extracted from the Human Dating Genome (HDG; ~13.5 million SNPs) across the last 2M years (x-axis). *(b)* Timeline of the density of human-phenotypic SNPs per minor allele frequency category (MAF; six bins ranges: 0 - 0.01, 0.01 - 0.1, 0.1 - 0.2, 0.2 - 0.3, 0.3 - 0.4, 0.4 - 0.5). *(c)* Absolute count of number of SNPs per time-bin (100 bins of ~20,000 years old) of human-phenotypic SNPs (red). Asterisks (*) denote bins where the number of human-phenotypic SNPs per age bin was significantly higher or lower than the null-model of random equally sized sets of SNPs selected from all SNPs from the HDG (MAF-controlled, 100 tests, Bonferroni *P* value threshold < 5×10^{-4}). *(d)* 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 (*rho* = 0.96, *P* = 5.5×10^{-59}). Mya, million years ago; MAF, Minor Allele Frequency.

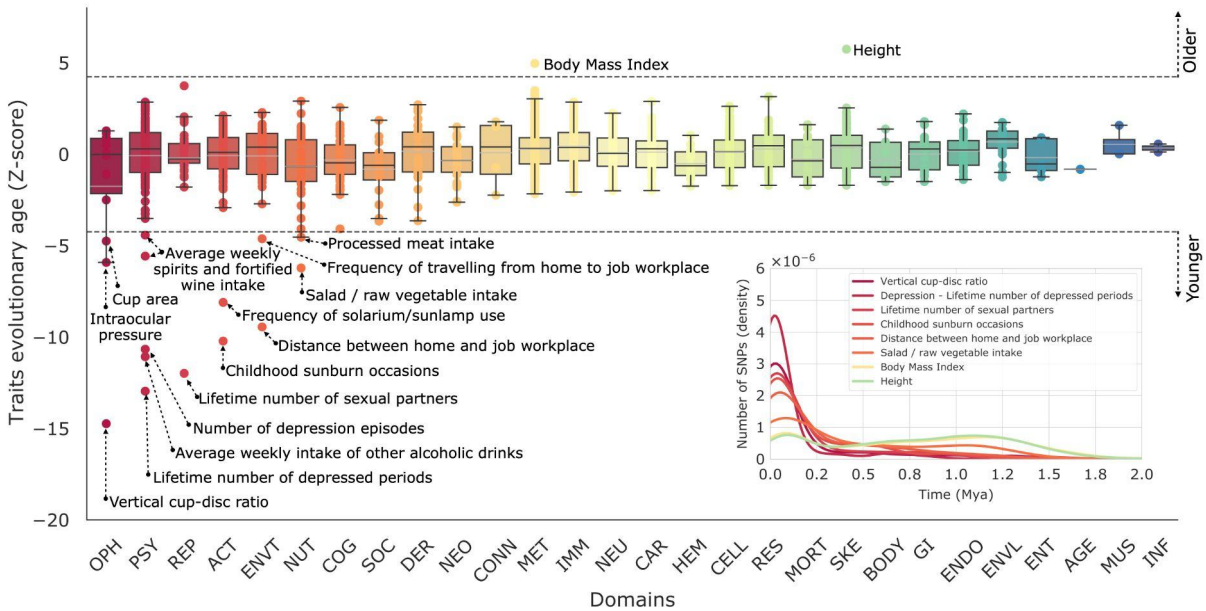

Figure 2. Genetic timeline of human traits. Expected age of traits (dots) under a null-model (see *Statistical inference* in Methods). Z-scores beyond the dotted line indicate traits (*n* = 2,251) with a median age of SNPs significantly younger (negative z-score values) or older (positive z-score values) than expected by chance (Bonferroni *P* value threshold < 2.2 x 10^{-5}). Black and grey lines within boxes denote domain median and mean z-scores, respectively. Exemplary genetic timelines of SNPs from significant traits are shown in the insert. Colours 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, haematological; CELL, cell; RES, respiratory; MORT, mortality; SKE, skeletal; BODY, body structures; GI, gastrointestinal; ENDO, endocrine; ENVL, environmental; ENT, ear, nose, throat; AGE, ageing; MUS, muscular; INF, infection.

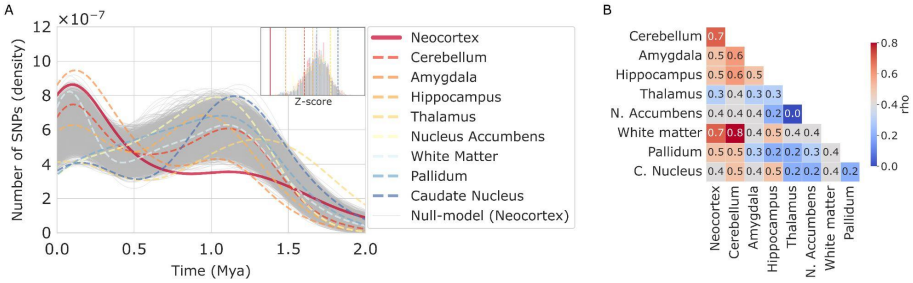

Figure 3. Genetic timeline of brain-imaging phenotypes. *(a)* Histogram showing the density (y-axis) of the number of SNPs per age bin (x-axis). Solid lines denote phenotypes (neocortex, red) significantly younger than expected by chance (see insert for the null-model of each brain phenotype). *(b)* 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.

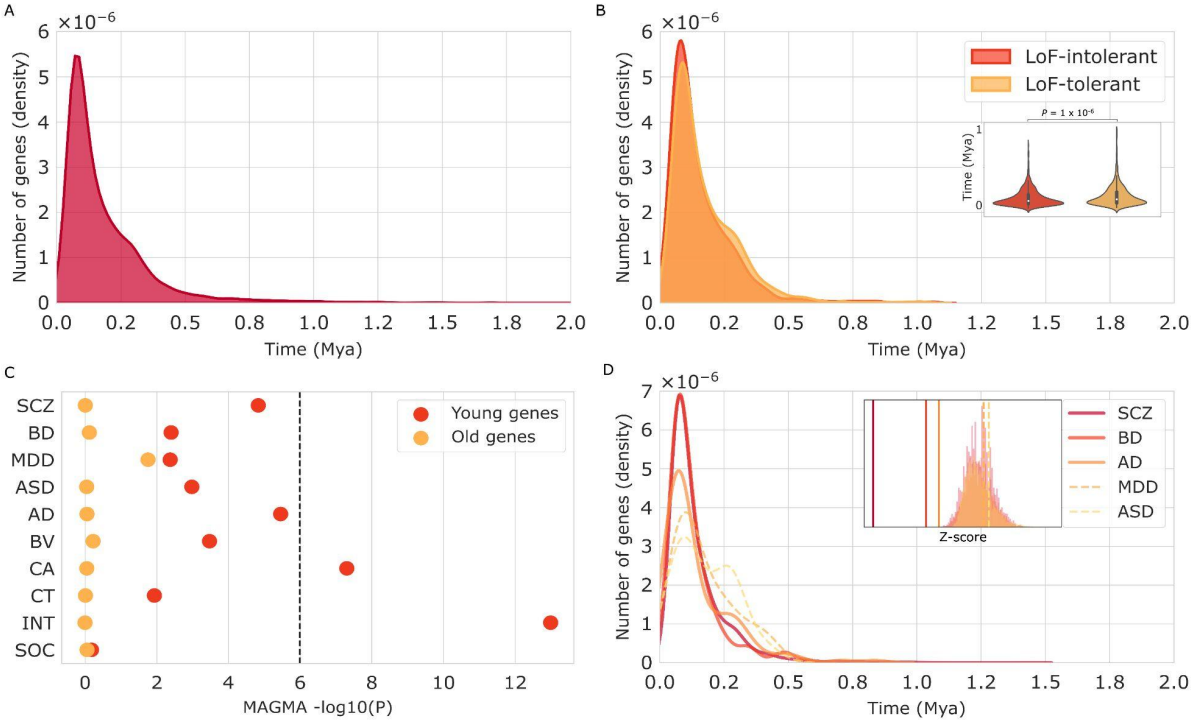

Figure 4. Genetic timeline of genes. *(a)* Timeline of genes (*n* = 18,328) with date estimates ranging from 2,965,600 to 3,803 years ago. Histogram depicts the density (y-axis) of the number of genes per age bin (x-axis; until 2M years ago is shown). *(b)* Timeline of the set of genes most intolerant 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)* 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. *(d)* 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). 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.

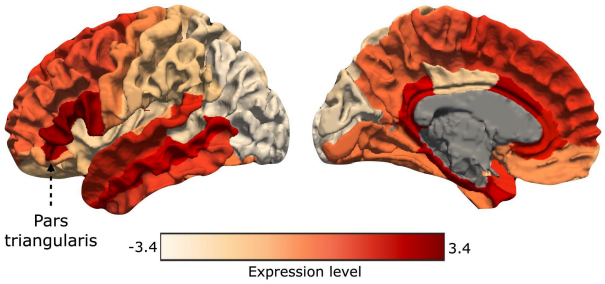

Figure 5. Transcriptomic brain map of evolutionarily recent genes. Brain plot of the normative expression levels of the top 1% genes with the youngest evolutionary age across brain areas (*n* = 164, ranging from 26,552 to 6,717 years ago), ranging from relatively low (white) to high (red) expression 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).

References

1. Irving-Pease, E.K., Muktupavela, R., Dannemann, M., and Racimo, F. (2021). Quantitative Human Paleogenetics: What can Ancient DNA Tell us About Complex Trait Evolution? Front. Genet. *12*, 703541. 10.3389/fgene.2021.703541.
2. Jerison, H. (2012). Evolution of the brain and intelligence (Elsevier).
3. Barton, R.A. (2007). Evolutionary specialization in mammalian cortical structure. J. Evol. Biol. *20*, 1504–1511. 10.1111/j.1420-9101.2007.01330.x.
4. Shultz, S., Nelson, E., and Dunbar, R.I.M. (2012). Hominin cognitive evolution: identifying patterns and processes in the fossil and archaeological record. Philos. Trans. R. Soc. B Biol. Sci. *367*, 2130–2140. 10.1098/rstb.2012.0115.
5. Wei, Y., De Lange, S.C., Scholtens, L.H., Watanabe, K., Ardesch, D.J., Jansen, P.R., Savage, J.E., Li, L., Preuss, T.M., Rilling, J.K., et al. (2019). Genetic mapping and evolutionary analysis of human-expanded cognitive networks. Nat. Commun. *10*, 4839. 10.1038/s41467-019-12764-8.
6. Sherwood, C.C., Subiaul, F., and Zawidzki, T.W. (2008). A natural history of the human mind: tracing evolutionary changes in brain and cognition. J. Anat. *212*, 426–454. 10.1111/j.1469-7580.2008.00868.x.
7. Preuss, T.M., Cáceres, M., Oldham, M.C., and Geschwind, D.H. (2004). Human brain evolution: insights from microarrays. Nat. Rev. Genet. *5*, 850–860. 10.1038/nrg1469.
8. Watanabe, K., Stringer, S., Frei, O., Umićević Mirkov, M., De Leeuw, C., Polderman, T.J.C., Van Der Sluis, S., Andreassen, O.A., Neale, B.M., and Posthuma, D. (2019). A global overview of pleiotropy and genetic architecture in complex traits. Nat. Genet. *51*, 1339–1348. 10.1038/s41588-019-0481-0.
9. Albers, P.K., and McVean, G. (2020). Dating genomic variants and shared ancestry in population-scale sequencing data. PLOS Biol. *18*, e3000586. 10.1371/journal.pbio.3000586.
10. Buniello, A., MacArthur, J.A.L., Cerezo, M., Harris, L.W., Hayhurst, J., Malangone, C., McMahon, A., Morales, J., Mountjoy, E., Sollis, E., et al. (2019). The NHGRI-EBI GWAS Catalog of published genome-wide association studies, targeted arrays and summary statistics 2019. Nucleic Acids Res. *47*, D1005–D1012. 10.1093/nar/gky1120.
11. Herries, A.I.R., Martin, J.M., Leece, A.B., Adams, J.W., Boschian, G., Joannes-Boyau, R., Edwards, T.R., Mallett, T., Massey, J., Murszewski, A., et al. (2020). Contemporaneity of Australopithecus, Paranthropus, and early Homo erectus in South Africa. Science *368*, eaaw7293. 10.1126/science.aaw7293.
12. Gómez-Robles, A. (2019). Dental evolutionary rates and its implications for the Neanderthal–modern human divergence. Sci. Adv. *5*, eaaw1268. 10.1126/sciadv.aaw1268.
13. Vigilant, L., Stoneking, M., Harpending, H., Hawkes, K., and Wilson, A.C. (1991). African Populations and the Evolution of Human Mitochondrial DNA. Science *253*, 1503–1507. 10.1126/science.1840702.
14. Hublin, J.-J., Ben-Ncer, A., Bailey, S.E., Freidline, S.E., Neubauer, S., Skinner, M.M., Bergmann, I., Le Cabec, A., Benazzi, S., Harvati, K., et al. (2017). New fossils from Jebel Irhoud, Morocco and the pan-African origin of Homo sapiens. Nature *546*, 289–292. 10.1038/nature22336.
15. Bar-Yosef, O. (2002). The Upper Paleolithic Revolution. Annu. Rev. Anthropol. *31*, 363–393. 10.1146/annurev.anthro.31.040402.085416.
16. Crow, T. (2000). Schizophrenia as the price that Homo sapiens pays for language: a resolution of the central paradox in the origin of the species. Brain Res. Rev. *31*, 118–129. 10.1016/S0165-0173(99)00029-6.
17. Rantala, M.J., Luoto, S., Borráz-León, J.I., and Krams, I. (2021). Bipolar disorder: An evolutionary psychoneuroimmunological approach. Neurosci. Biobehav. Rev. *122*, 28–37. 10.1016/j.neubiorev.2020.12.031.
18. Finch, C.E., and Austad, S.N. (2015). Commentary: is Alzheimer's disease uniquely

human? Neurobiol. Aging *36*, 553–555. 10.1016/j.neurobiolaging.2014.10.025.
19. Doan, R.N., Bae, B.-I., Cubelos, B., Chang, C., Hossain, A.A., Al-Saad, S., Mukaddes, N.M., Oner, O., Al-Saffar, M., Balkhy, S., et al. (2016). Mutations in Human Accelerated Regions Disrupt Cognition and Social Behavior. Cell *167*, 341-354.e12. 10.1016/j.cell.2016.08.071.
20. Pattabiraman, K., Muchnik, S.K., and Sestan, N. (2020). The evolution of the human brain and disease susceptibility. Curr. Opin. Genet. Dev. *65*, 91–97. 10.1016/j.gde.2020.05.004.
21. Slatkin, M. (2008). Linkage disequilibrium — understanding the evolutionary past and mapping the medical future. Nat. Rev. Genet. *9*, 477–485. 10.1038/nrg2361.
22. Pritchard, J.K., Pickrell, J.K., and Coop, G. (2010). The Genetics of Human Adaptation: Hard Sweeps, Soft Sweeps, and Polygenic Adaptation. Curr. Biol. *20*, R208–R215. 10.1016/j.cub.2009.11.055.
23. Smith, S.M., Douaud, G., Chen, W., Hanayik, T., Alfaro-Almagro, F., Sharp, K., and Elliott, L.T. (2021). An expanded set of genome-wide association studies of brain imaging phenotypes in UK Biobank. Nat. Neurosci. *24*, 737–745. 10.1038/s41593-021-00826-4.
24. Rilling, J.K. (2014). Comparative primate neuroimaging: insights into human brain evolution. Trends Cogn. Sci. *18*, 46–55. 10.1016/j.tics.2013.09.013.
25. Hopkins, W.D., and Rilling, J.K. (2000). 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*, 739–748. 10.1037/0735-7044.114.4.739.
26. Ghazanfar, A., and Schroeder, C. (2006). Is neocortex essentially multisensory? Trends Cogn. Sci. *10*, 278–285. 10.1016/j.tics.2006.04.008.
27. Barton, R.A. (1996). Neocortex size and behavioural ecology in primates. Proc. R. Soc. Lond. B Biol. Sci. *263*, 173–177.
28. Yates, A., Akanni, W., Amode, M.R., Barrell, D., Billis, K., Carvalho-Silva, D., Cummins, C., Clapham, P., Fitzgerald, S., Gil, L., et al. (2016). Ensembl 2016. Nucleic Acids Res. *44*, D710–D716. 10.1093/nar/gkv1157.
29. Balasubramanian, S., Fu, Y., Pawashe, M., McGillivray, P., Jin, M., Liu, J., Karczewski, K.J., MacArthur, D.G., and Gerstein, M. (2017). Using ALoFT to determine the impact of putative loss-of-function variants in protein-coding genes. Nat. Commun. *8*, 382. 10.1038/s41467-017-00443-5.
30. De Leeuw, C.A., Mooij, J.M., Heskes, T., and Posthuma, D. (2015). MAGMA: Generalized Gene-Set Analysis of GWAS Data. PLOS Comput. Biol. *11*, e1004219. 10.1371/journal.pcbi.1004219.
31. Hawrylycz, M.J., Lein, E.S., Guillozet-Bongaarts, A.L., Shen, E.H., Ng, L., Miller, J.A., Van De Lagemaat, L.N., Smith, K.A., Ebbert, A., Riley, Z.L., et al. (2012). An anatomically comprehensive atlas of the adult human brain transcriptome. Nature *489*, 391–399. 10.1038/nature11405.
32. Friederici, A.D. (2017). Evolution of the neural language network. Psychon. Bull. Rev. *24*, 41–47. 10.3758/s13423-016-1090-x.
33. Mesulam, M. (1998). From sensation to cognition. Brain *121*, 1013–1052. 10.1093/brain/121.6.1013.
34. Posth, C., Renaud, G., Mittnik, A., Drucker, D.G., Rougier, H., Cupillard, C., Valentin, F., Thevenet, C., Furtwängler, A., Wißing, C., et al. (2016). Pleistocene Mitochondrial Genomes Suggest a Single Major Dispersal of Non-Africans and a Late Glacial Population Turnover in Europe. Curr. Biol. *26*, 827–833. 10.1016/j.cub.2016.01.037.
35. Moore, A.M. (2014). The transition from foraging to farming in Southwest Asia: Present problems and future directions. In Foraging and Farming (Routledge), pp. 620–631.
36. Schaefer, N.K., Shapiro, B., and Green, R.E. (2021). An ancestral recombination graph of human, Neanderthal, and Denisovan genomes. Sci. Adv. *7*, eabc0776. 10.1126/sciadv.abc0776.

37. d'Errico, F., Henshilwood, C., Lawson, G., Vanhaeren, M., Tillier, A.-M., Soressi, M., Bresson, F., Maureille, B., Nowell, A., and Lakarra, J. (2003). Archaeological evidence for the emergence of language, symbolism, and music–an alternative multidisciplinary perspective. J. World Prehistory *17*, 1–70.
38. Andirkó, A., Moriano, J., Vitriolo, A., Kuhlwilm, M., Testa, G., and Boeckx, C. (2022). Temporal mapping of derived high-frequency gene variants supports the mosaic nature of the evolution of Homo sapiens. Sci. Rep. *12*, 9937. 10.1038/s41598-022-13589-0.
39. Van Essen, D.C., Donahue, C.J., and Glasser, M.F. (2018). Development and Evolution of Cerebral and Cerebellar Cortex. Brain. Behav. Evol. *91*, 158–169. 10.1159/000489943.
40. Ardesch, D.J., Scholtens, L.H., Li, L., Preuss, T.M., Rilling, J.K., and Van Den Heuvel, M.P. (2019). Evolutionary expansion of connectivity between multimodal association areas in the human brain compared with chimpanzees. Proc. Natl. Acad. Sci. *116*, 7101–7106. 10.1073/pnas.1818512116.
41. Finlay, B.L., and Darlington, R.B. (1995). Linked Regularities in the Development and Evolution of Mammalian Brains. Science *268*, 1578–1584. 10.1126/science.7777856.
42. van den Heuvel, M.P., Ardesch, D.J., Scholtens, L.H., de Lange, S.C., van Haren, N.E.M., Sommer, I.E.C., Dannlowski, U., Repple, J., Preuss, T.M., Hopkins, W.D., et al. (2023). Human and chimpanzee shared and divergent neurobiological systems for general and specific cognitive brain functions. Proc. Natl. Acad. Sci. *120*, e2218565120. 10.1073/pnas.2218565120.
43. DeSilva, J.M., Traniello, J.F., Claxton, A.G., and Fannin, L.D. (2021). 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.
44. Rilling, J.K., Glasser, M.F., Preuss, T.M., Ma, X., Zhao, T., Hu, X., and Behrens, T.E.J. (2008). The evolution of the arcuate fasciculus revealed with comparative DTI. Nat. Neurosci. *11*, 426–428. 10.1038/nn2072.
45. Tilot, A.K., Khramtsova, E.A., Liang, D., Grasby, K.L., Jahanshad, N., Painter, J., Colodro-Conde, L., Bralten, J., Hibar, D.P., Lind, P.A., et al. (2021). The Evolutionary History of Common Genetic Variants Influencing Human Cortical Surface Area. Cereb. Cortex *31*, 1873–1887. 10.1093/cercor/bhaa327.
46. Bailey, D.H., and Geary, D.C. (2009). Hominid brain evolution. Hum. Nat. *20*, 67–79.
47. Ardesch, D.J., Scholtens, L.H., De Lange, S.C., Roumazeilles, L., Khrapitchev, A.A., Preuss, T.M., Rilling, J.K., Mars, R.B., and Van Den Heuvel, M.P. (2022). Scaling Principles of White Matter Connectivity in the Human and Nonhuman Primate Brain. Cereb. Cortex *32*, 2831–2842. 10.1093/cercor/bhab384.
48. Van Den Heuvel, M.P., Scholtens, L.H., De Lange, S.C., Pijnenburg, R., Cahn, W., Van Haren, N.E.M., Sommer, I.E., Bozzali, M., Koch, K., Boks, M.P., et al. (2019). Evolutionary modifications in human brain connectivity associated with schizophrenia. Brain *142*, 3991–4002. 10.1093/brain/awz330.
49. Dannemann, M., Milaneschi, Y., Yermakovich, D., Stiglbauer, V., Kariis, H.M., Krebs, K., Friese, M.A., Otte, C., Estonian Biobank Research Team, Esko, T., et al. (2022). Neandertal introgression partitions the genetic landscape of neuropsychiatric disorders and associated behavioral phenotypes. Transl. Psychiatry *12*, 433. 10.1038/s41398-022-02196-2.
50. Irving-Pease, E.K., Refoyo-Martínez, A., Ingason, A., Pearson, A., Fischer, A., Barrie, W., Sjögren, K.-G., Halgren, A.S., Macleod, R., and Demeter, F. (2022). The Selection Landscape and Genetic Legacy of Ancient Eurasians. bioRxiv.
51. Thompson, P.M., Jahanshad, N., Ching, C.R.K., Salminen, L.E., Thomopoulos, S.I., Bright, J., Baune, B.T., Bertolín, S., Bralten, J., Bruin, W.B., et al. (2020). 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. 10.1038/s41398-020-0705-1.