



OPEN

Genomic analysis of the natural history of attention-deficit/hyperactivity disorder using Neanderthal and ancient *Homo sapiens* samples

Paula Esteller-Cucala^{1,2,11}, Iago Maceda^{1,2}, Anders D. Børglum^{3,4,5}, Ditte Demontis^{3,4,5}, Stephen V. Faraone⁶, Bru Cormand^{7,8,9,10,12} & Oscar Lao^{1,2,12}

Attention-deficit/hyperactivity disorder (ADHD) is an impairing neurodevelopmental condition highly prevalent in current populations. Several hypotheses have been proposed to explain this paradox, mainly in the context of the Paleolithic versus Neolithic cultural shift but especially within the framework of the *mismatch theory*. This theory elaborates on how a particular trait once favoured in an ancient environment might become maladaptive upon environmental changes. However, given the lack of genomic data available for ADHD, these theories have not been empirically tested. We took advantage of the largest GWAS meta-analysis available for this disorder consisting of over 20,000 individuals diagnosed with ADHD and 35,000 controls, to assess the evolution of ADHD-associated alleles in European populations using archaic, ancient and modern human samples. We also included Approximate Bayesian computation coupled with deep learning analyses and singleton density scores to detect human adaptation. Our analyses indicate that ADHD-associated alleles are enriched in loss of function intolerant genes, supporting the role of selective pressures in this early-onset phenotype. Furthermore, we observed that the frequency of variants associated with ADHD has steadily decreased since Paleolithic times, particularly in Paleolithic European populations compared to samples from the Neolithic Fertile Crescent. We demonstrate this trend cannot be explained by African admixture nor Neanderthal introgression, since introgressed Neanderthal alleles are enriched in ADHD risk variants. All analyses performed support the presence of long-standing selective pressures acting against ADHD-associated alleles until recent times. Overall, our results are compatible with the mismatch theory for ADHD but suggest a much older time frame for the evolution of ADHD-associated alleles compared to previous hypotheses.

Attention-deficit/hyperactivity disorder (ADHD) is a common neurodevelopmental condition characterised by inattention, motor hyperactivity and impulsivity¹. Its worldwide prevalence is around 5% in children and adolescents and its symptoms often persist into adulthood². Twin studies have consistently shown high heritability rates

¹CNAG-CRG, Centre for Genomic Regulation (CRG), Barcelona Institute of Science and Technology (BIST), Barcelona, Spain. ²Universitat Pompeu Fabra (UPF), Barcelona, Spain. ³The Lundbeck Foundation Initiative for Integrative Psychiatric Research, iPSYCH, Aarhus, Denmark. ⁴Centre for Integrative Sequencing, iSEQ, and Aarhus Genome Centre, Aarhus, Denmark. ⁵Department of Biomedicine – Human Genetics, Aarhus University, Aarhus, Denmark. ⁶Departments of Psychiatry and of Neuroscience and Physiology, SUNY Upstate Medical University, Syracuse, NY, USA. ⁷Departament de Genètica, Microbiologia i Estadística, Facultat de Biologia, Universitat de Barcelona, Barcelona, Spain. ⁸Centro de Investigación Biomédica en Red de Enfermedades Raras (CIBERER), Instituto de Salud Carlos III, Madrid, Spain. ⁹Institut de Biomedicina de la Universitat de Barcelona (IBUB), Barcelona, Spain. ¹⁰Institut de Recerca Sant Joan de Déu (IR-SJD), Esplugues de Llobregat, Spain. ¹¹Present address: Institut de Biologia Evolutiva (UPF-CSIC), Barcelona, Spain. ¹²These authors contributed equally: Bru Cormand and Oscar Lao. ✉e-mail: bcormand@ub.edu; oscar.lao@cnag.crg.eu

of 70–80%³, a considerable fraction of which is explained by single nucleotide polymorphisms (SNPs)⁴. However, it has not been until recently that genome-wide association studies (GWAS) have reported significant variants associated with ADHD⁵. In practise, ADHD manifests as a notably impairing disorder with increased risk of psychiatric comorbidity, social dysfunction, injuries from traffic accidents, medical morbidity and premature death. These impairments translate into a reduced quality of life not only to ADHD patients themselves but also to their relatives^{2,6–8}.

The fact that ADHD poses such a burden on individual life quality⁹ and yet is so prevalent seems counterintuitive from an evolutionary perspective, as one might expect natural selection to purge the risk alleles from the population¹⁰. In this regard, several hypotheses have been put forward to explain the evolutionary persistence of this disorder, each of them emphasising on different aspects of the ADHD phenotype (Table 1). Such theories dispute that ADHD is the inevitable by-product of a continuous variable complex trait not necessarily under selective pressure¹¹. Instead, they resort to the evolutionary concept of *mismatch*¹², suggesting that traits which are specific to individuals diagnosed with ADHD increased the fitness of carriers in an ancestral environment and therefore spread in the population due to natural selection. However, as a result of the astonishingly fast cultural evolution of the human species and subsequent environmental changes, particularly in the last 10,000 years, ADHD traits have become maladaptive in present-day societies, and are therefore evolving to meet the demands of current environments^{13–15}. One of these theories, the hunter-farmer theory, states that ADHD traits specific to hunter-gatherers would have been environmentally beneficial well into the Neolithic Revolution, when non-ADHD traits characteristic of agriculturalism spread. Under this view, current humans diagnosed with ADHD should be enriched in alleles from ancient hunter-gatherers (Table 1)¹⁵.

Besides natural selection, a demographic factor that could also explain the observed prevalence of ADHD in current populations is archaic introgression. The sequencing of the Neanderthal¹⁶ and Denisovan¹⁷ genomes revealed an out-of-Africa introgression of genomic segments from archaic species into the genome of anatomically modern humans (AMH)^{18–22}. Moreover, even if introgressed sequences of archaic species were subjected to purifying selection shortly after admixture^{18,21,23,24}, the addition of particular introgressed regions could have helped AMH adapt to this new out-of-Africa environment²⁵. In fact, there is now mounting evidence of introgressed variants currently associated with human complex traits that have undergone positive selection in AMH^{21,22}, including neurological and psychiatric phenotypes²². These results support the mismatch theory by positing that some Neanderthal alleles could have been beneficial for AMH as they moved out of Africa but became detrimental in modern environments.

Overall, mismatch-related hypotheses in ADHD have been speculative⁹, but current genomic data allows to assess the evidence supporting each proposed model. Here we study the evolutionary nature of ADHD from a genomic perspective by analysing traces of selection in AMH and assessing the role that archaic interbreeding played in the genetic variation associated with this trait. For that purpose, we used ancient samples to analyse the evolution of ADHD-associated genetic load over time, using European present-day genomes as a baseline. Moreover, we also evaluated these theories in the context of the Neolithic transition, considering African admixture and archaic introgression.

Results

Evidence of selective pressures in ADHD variants and Loss of function (LoF) intolerant genes.

It has been hypothesised that genetic variants associated with early-onset traits under natural selection could be enriched in genes that are intolerant to loss of function mutations^{26,27}. In agreement with this hypothesis, a previous study²⁸ described that genetic variants associated with schizophrenia (SCZ) are enriched in mutation-intolerant genes and in regions under strong background selection. Given that ADHD is an early-onset condition, we assessed whether a similar trend could be observed in this phenotype. MAGMA analysis²⁹ statistically supported that LoF-intolerant genes (ExAC³⁰ pLI ≥ 0.9) expressed in brain are enriched in variants associated with ADHD ($p = 4.15e-3$; $p = 4.32e-3$ using genes expressed in brain²⁸), which is in line with previous results⁵, thus suggesting that ADHD could be under natural selection pressures. Consistently, a significant enrichment in the heritability from ADHD-associated SNPs located in conserved regions of the genome was also observed in the aforementioned study⁵.

Contribution of ancestral variants to ADHD. In order to characterise the nature of ADHD-associated variants we analysed the contribution of ancestral alleles to the ADHD-associated alleles controlling by the minor allele frequency (MAF) of each SNP. We observed that for statistically significant GWAS SNPs at $p \leq 1e-8$, the ancestral allele tended to be the risk allele as compared to SNPs that are not associated with ADHD in the GWAS analysis (i.e. those showing a GWAS $p \geq 0.9$) (odds ratio, OR = 1.34; one tail p based on 10,000 resamples controlling by MAF < 5e-5; Supplementary Fig. S1). This suggests that ADHD-associated variants tend to be enriched for the ancestral allele.

Time evolution of the frequency of current ADHD-risk alleles in ancient genomes. The evolution of ADHD load was assessed by computing Kendall's τ correlation between the average frequency of currently associated ADHD alleles (at GWAS significance < 0.01) (f_{ADHD} , see Methods) and date estimates of ancient European samples from three different datasets (Supplementary Fig. S2).

Alluding to the most relevant characteristics of each dataset, we refer to these as the Pre-Neolithic³¹, Near East³² and Neolithic³³ datasets. For the Near East and Neolithic datasets, which include more ancient samples, (151 and 84 genomes, respectively), we observed a significant negative τ between f_{ADHD} and sample age (Fig. 1): $\tau_{Near\ East} = -0.173$ ($p = 0.002$) and $\tau_{Neolithic} = -0.1574$ ($p = 0.034$). The Pre-Neolithic dataset (which after filtering consisted of 16 ancient samples) also showed a negative trend although it was not significant ($\tau_{Pre-Neolithic} = -0.317$, $p = 0.096$). These results support a decrease in the frequency of ADHD alleles over time.

Theory [Reference]	Prehistoric context	Description	Expected genetic outcome from ancient DNA analyses and tests on recent selective pressures
The wader theory ¹³	Human speciation (> 6 mya)	Based on the <i>Aquatic Ape Theory</i> ⁶⁷ , body hair loss in human females favoured them to enter water shores for protection from land predators. In this new environment, as infants could no longer cling to their mothers, ADHD traits (e.g., hypervocalisation) were able to attract maternal attention (e.g., breastfeeding, continuous monitoring) and thus more likely to survive.	1) ADHD-risk alleles are mostly the derived allele compared to chimpanzee. 2) Neanderthal-introgressed regions equally enriched in ADHD-risk and ADHD-protective alleles.
The fighter theory ¹³	Coexistence with Neanderthals (~60–30 kya)	ADHD-associated aggression in ancient humans favoured Neanderthal genocide. ADHD traits of impulsiveness, reduced inhibition control and inattention are the result of a battlefield mind set.	1) Neanderthal-introgressed regions in modern humans enriched in ADHD-protective alleles. 2) Decrease of ADHD-protective alleles during the time of coexistence with Neanderthals.
The response-readiness theory ¹⁴	Hunter-gatherers (~200–10 kya)	ADHD traits in response-ready individuals were advantageous over problem-solvers (without ADHD phenotype) in harsh and changing conditions. Main beneficial ADHD traits in the ancestral environment: (1) hyperactivity (exploratory behaviour), (2) impulsivity (ready-to-go behaviour) and (3) inattention (rapid-scanning of the surroundings)	1) Increase of ADHD-risk alleles over Pre-Neolithic times. 2) ADHD-protective alleles are recently positively selected.
The hunter-farmer theory ¹⁵	Hunter-gatherers (most intense hunting period, ~100 kya)	Modern individuals with ADHD are considered the remains of <i>hunters</i> , opposed to non-ADHD individuals (with <i>farmer</i> characteristics). Main beneficial ADHD traits in the ancestral environment: (1) hyperactivity (energetic and tireless behaviour), (2) impulsivity (ability to change the strategy quickly) and (3) inattention (detecting changes in the environment)	1) Increase of ADHD-risk alleles over Pre-Neolithic times. 2) Hunter-gatherer genetic component correlates with the frequency of ADHD-risk alleles in humans from the Neolithic. 3) ADHD-protective alleles are recently positively selected.

Table 1. Mismatch theories explaining the adaptive role of attention-deficit/hyperactivity disorder in ancestral environments and their expected genetic outcome in the conducted analyses. Abbreviations: kya, thousand (kilo) years ago; mya, million years ago; ADHD, attention-deficit/hyperactivity disorder.

Since GWAS measure associations between genotypes and phenotypes in a specific environment and population sample, some studies claim that no out-of-environment nor out-of-sample extrapolation can be performed^{34,35}. Hence, we inspected whether the decay trend of current ADHD-risk alleles is expected over time independently of the association of the alleles with the ADHD phenotype. For each of the three datasets, we built a Kendall's τ null distribution assuming independence of any association between the examined alleles and ADHD by randomising the risk alleles. This approximation is equivalent to examining sets of random SNPs matched to the allele frequencies of the ADHD-associated SNPs. Rejection of the null hypothesis was reported for all three datasets (p of 0.046, 0.003 and 0.016 for the Pre-Neolithic, Near East and Neolithic datasets, respectively, after 10,000 randomizations) (Supplementary Fig. S3). This supports the role of ADHD in the decay of current ADHD risk alleles over time in Europe. However, ADHD shows comorbidities with other psychiatric disorders, like autism spectrum disorder (ASD) or major depressive disorder (MDD)³⁶. In fact, the effect sizes of the ADHD-associated SNPs showed a relatively strong positive Kendall's τ correlation (Kendall's correlation of ~ 0.5 for all the ancient datasets) with those SNPs associated with ASD compared to other tested psychiatric disorders (Supplementary Table S1). The decay in the frequency of ADHD-risk alleles over time could be explained by demographic factors such as dilution of the ADHD-associated variants by constant admixture from a population enriched in ADHD-protective alleles. This process has been claimed as one of the reasons for the decay of Neanderthal ancestry in Europe, result of the constant migration from African populations³⁷. To test this, we studied the trend of f_{ADHD} derived from European-based ADHD-associated alleles in ancient and modern African populations. Since the temporal scale of the ancient African samples and the Near East dataset overlaps, we performed a multiple linear regression of f_{ADHD} against sample age considering the continent as a covariable. This analysis showed that ancient African samples have a higher f_{ADHD} of ADHD-associated alleles than the ancient European samples (slope of the continental location variable equals 0.0238, $p = 1.33e-12$). Similarly, when we compared African and European populations in the 1000 Genomes Phase 3 dataset, African samples also showed higher f_{ADHD} . This result is obtained independently of which set of SNPs defined for each of the ancient European datasets is considered (Wilcoxon test one tail $p = 2.31e-140$, one tail $p = 1.1e-76$ and one tail $p = 8.09e-32$ for Pre-Neolithic, Near East and Neolithic datasets, respectively). Additionally, the decay of ADHD-associated alleles over time could be explained by successive dilutions from European populations showing an excess of ADHD-protective alleles. Two main ancient migrations have been suggested to have shaped the genetic diversity of European populations: the steppe migration and the farmer Neolithic expansion³⁸. Steppe migration cannot explain the observed sample time trend, as it considers a more recent time period than the one our samples span³⁸. However, the admixture between hunter-gatherers (HG) and Neolithic farmers could explain the observed trend after the Neolithic expansion if HG had a lower burden of ADHD-associated alleles compared to Neolithic people. We observed a significant negative Kendall's τ correlation between HG and f_{ADHD} ($\tau = -0.16$, $p = 0.024$) in the Neolithic dataset; suggesting that Neolithic-ancestry individuals tend to carry more ADHD-risk alleles than HG. However, since the percentage of HG ancestry increases over the course of the Neolithic period³³, sample time could act as a confounder. Therefore, we repeated the analyses controlling for sample time and no statistically significant association between both variables was observed ($\tau = -0.1$, $p = 0.17$). Taken together, our analyses show that the observed trend between f_{ADHD} and sample time cannot be explained by ancestry dilution.

Quantification of polygenic adaptation. The previous correlation analyses between f_{ADHD} and sample age did not consider that the ancient populations might be related over time. We estimated the mean increase in allelic frequency of ADHD alleles per generation including the relationship between ancient populations in the Pre-Neolithic dataset in an Approximate Bayesian Computation coupled to Deep Learning (ABC-DL) using a simple model that considers the average selective pressure acting over all the loci and genetic drift (Fig. 2). After implementing the ABC-DL approach (see Methods), we first computed the *factor 2* statistic on 1,000 simulations to quantify the performance of the framework for properly estimating the slope. This statistic quantifies the number of times that the estimated mean from the posterior distribution is within the range 50% and 200% of the true value of the parameter used for the simulation³⁹. The mean of the posterior distribution computed by ABC-DL falls within the *factor 2* statistic 93.1% of the times, thus suggesting high confidence in using the mean of the posterior distribution as a proxy of the real value.

When applied to the observed data, the posterior distribution of the trend of the allelic frequency of each ADHD-risk allele per SNP and generation had a mean of $-7.72e-5$ (95% Credible Interval ranging from $-9.87e-5$ to $-5.9e-5$). This result corresponds to an average decay of 0.12 in the frequency of each ADHD-increasing risk allele since Palaeolithic times until present ($\sim 1,552$ generations assuming a generation time of 29 years⁴⁰) and supports the existence of a selective pressure towards current protective ADHD alleles from Paleolithic to Neolithic times.

To further evaluate whether selective forces acted on ADHD-associated alleles for the last 2,000–3,000 years, we used the trait Singleton Density Score (tSDS) computed on the UK10K data matched by the ADHD-increasing risk allele⁴¹. tSDS is suited for detecting polygenic adaptation acting less than 100 generations ago, and it is robust against putative confounder effects due to background selection⁴¹, which is expected to act on conserved genomic regions. Given that most complex-trait heritability is caused by SNPs that do not reach GWAS significance^{5,42,43}, we tested for genome-wide rank correlation between tSDS and GWAS p-values considering all SNPs⁴⁴. The Spearman correlation observed between tSDS and GWAS p-values was not expected under neutrality (Spearman $\rho = 0.197$, $p = 0.0217$). This positive correlation indicates that statistically significant p-values tend to be enriched by negative tSDS values, and non-significant p-values by tSDS close to 0. Positive values of tSDS imply that the frequency of the ADHD-risk allele increased in the past, whereas negative tSDS imply that the ADHD-protective allele is increasing in frequency⁴⁴. Our result supports that polygenic adaptation recently shaped the frequency of alleles that are protective for ADHD in one modern European population.

Current ADHD risk alleles in the context of archaic species. Genes involved in behaviour show differences between Neanderthal and modern humans⁴⁵ and current ADHD prevalence could potentially explain Neanderthal genocide (see the fighter theory in Table 1). To assess the role of archaic introgression in the loading of ADHD alleles into the *Homo sapiens* genome we estimated the mean effect size (β) of variants reported to be of Neanderthal origin. We used 1,151 tag SNPs from the Neanderthal introgressed haplotypes. We observed an increased average effect size compared to a null distribution estimated considering that Neanderthal introgressed alleles and ADHD effect are independent ($\beta = 0.013$, $p < 5e-5$, Supplementary Fig. S4). This suggests that variants of Neanderthal origin are enriched in ADHD-risk alleles.

Discussion

Our analyses support that natural selection has been acting against current ADHD-risk alleles for a long time, albeit the fraction of the causality explained by these SNPs remains relatively modest⁵ and the considered SNPs are most likely just linked to the actual causal variants.

First, we show that LoF-intolerant genes are enriched in signals of association with ADHD; given that ADHD is an early-onset condition, such result would support the hypothesis that it is under selective pressures^{26,27}. Second, the ancestral state of current ADHD-associated variants is enriched in risk alleles compared to the expected under neutrality. Finally, the relationship between sample age and f_{ADHD} computed in ancient samples consistently points towards a higher frequency of ADHD-associated alleles in older samples and a negative trend with decreasing f_{ADHD} in time in these samples (Fig. 1). We showed this decay cannot be explained by dilution due to admixture with hunter-gatherers or by continuous migration from African populations carrying lower levels of the considered ADHD-risk alleles, as African populations show higher frequencies of these alleles compared to European populations. These results do not imply that African populations have a higher prevalence of ADHD, as these markers have been identified in European-ancestry samples and it has been recently shown that GWAS-SNP portability among continents is limited^{34,35}. Moreover, by means of ABC-DL, we directly evidenced polygenic selective pressures against ADHD-risk alleles from 45,000 to 10,000 years. In fact, this pattern could also be traced back up to 2,000–3,000 years by using the tSDS estimates in a specific modern European population from the UK. We also proved that the decay of ADHD-risk alleles is not the demographic by-product of Neanderthal introgression, as we observed that introgressed Neanderthal alleles at SNPs that have been associated with ADHD are enriched for the risk allele instead. Once again, these results do not imply that ancient AMH or archaic species had an increased risk for ADHD, since the specific loci associated with the ADHD-like phenotype in the Neanderthal and ancient samples are unknown. However, we can make use of these results to contextualise the evolution of current ADHD variants in AMH.

Out of all the proposed hypotheses to explain the high prevalence of ADHD in current populations (Table 1), the fighter hypothesis, which assumes that ADHD-risk alleles are of anatomically modern human origin, does not predict the observed excess of Neanderthal introgressed ADHD alleles, nor the decay of f_{ADHD} in the period in which Neanderthals and AMH shared environment. The wader theory, which explains the rise in frequency of ADHD in the *Homo* lineage living in aquatic environments, cannot account for the enrichment of derived alleles associated with the protective ADHD phenotype compared to chimpanzees, nor for the differences in the proportions of introgressed ADHD-risk and ADHD-protective alleles either. This is because both Neanderthals and

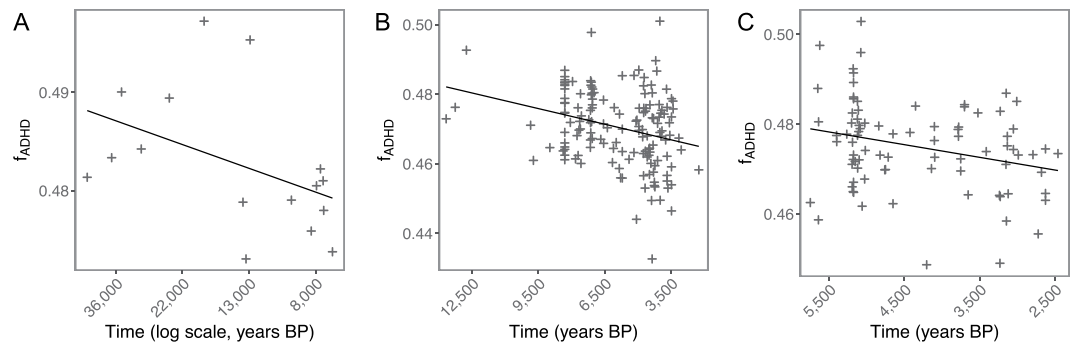


Figure 1. Temporal evolution of the f_{ADHD} in ancient samples using the Pre-Neolithic (A), Near East (B) and Neolithic (C) datasets. Each dot represents an individual defined by its f_{ADHD} and sample age. Date estimates are indicated in years BP (Before Present) and time scales are reversed to simplify interpretation. In the Neolithic dataset sample ages are shown in log scale.

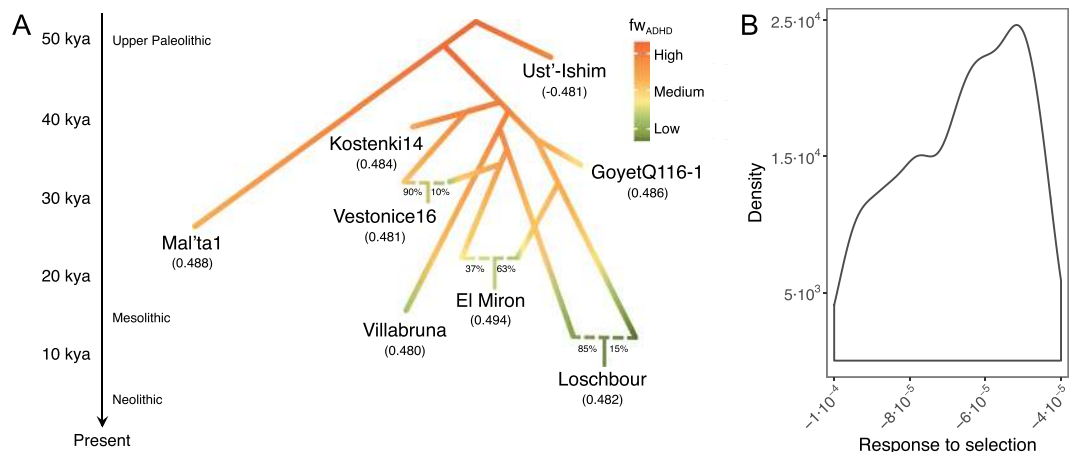


Figure 2. ABC-DL modelling using Pre-Neolithic samples. (A) Demographic model proposed by Fu *et al.* (see Figure 4 in³¹) in ancient samples from the Pre-Neolithic dataset with the mean f_{ADHD} contribution per SNP estimated using the effect size from a GWAS meta-analysis (see Fig. 1). (B) Posterior distribution of the parameter γ that quantifies the deviation of the frequency of the ADHD risk allele at locus j per generation due to selective pressures.

AMH share a common recent ancestor with respect to chimpanzees, yet with these analyses we see genomic differences in their ADHD load. Moreover, the response-readiness theory, which links ADHD to the hunter-gatherer environment, fails to explain why f_{ADHD} decreases over time during Pre-Neolithic times.

Our results on the decay of f_{ADHD} over time and the identification of recent signatures of positive selection are compatible with the mismatch hypothesis. Namely, the frequency of current ADHD-risk alleles was higher in the past, but these alleles became maladaptive due to recent environmental changes, to which Western European populations are still adapting. Within this context, the hunter-farmer hypothesis predicts a recent shift in the selective pressures due to agriculture during the Neolithic. However, this hypothesis cannot explain the observed negative association between the amount of hunter-gatherer ancestry and f_{ADHD} alleles in Neolithic samples. In fact, this result would be compatible with the hypothesis that selective pressures against current ADHD risk alleles were stronger in Paleolithic hunter-gatherers from Europe compared to Neolithic farmers from the Fertile Crescent. Nevertheless, the fact that the correlation is non-significant after controlling by sample time suggests that both populations were subjected to the same selective pressures. Furthermore, the hunter-farmer hypothesis cannot explain why current ADHD-risk alleles would have not been beneficial at least for the past 45,000 years, as this is the estimated age of the oldest sample included in our analyses. If selective pressures are acting on ADHD, our results call into question evolutionary hypotheses claiming the Neolithic Revolution to be the turning point for ADHD pressures. This conclusion falls in line with a study that suggested, after computing genetic risk scores on ancient human samples, a general improvement of the genomic health of ancient individuals over time. However, this trend was not seen for specific disease types such as cardiovascular disorders or neurological/psychological ones⁴⁶.

The exact biological and environmental factors that fuelled the selective pressures over current ADHD-risk alleles through time remain to be elucidated. ADHD is known to be comorbid with many other psychiatric phenotypes such as major depressive disorder or ADS⁴⁷. Notably, we observed a considerable correlation of the effect

sizes of ADHD with ASD. Furthermore, within the ADHD phenotype several characteristics are plausible candidates for claiming selective pressures in AMH¹⁴. Moreover, ADHD shows significant genetic risk correlations with phenotypes that are strong candidates for being under selective pressure (both negative and positive) such as weight-related traits, reproductive traits and parental longevity traits⁵.

Increasing knowledge on the biology of the disease and availability of high-quality ancient genomic data will pave the way to new studies on the evolutionary factors shaping the variation of ADHD-risk alleles, whose genetic history has been depicted for the first time in this study.

Methods

Human genome databases. *Ancient AMH genomes.* Genotype data from 51³¹, 281³² and 180³³ European ancient AMH were included in the analyses (Supplementary Table S2). Alluding to the most relevant characteristics of each dataset, we refer to these as the Pre-Neolithic³¹, Near East³² and Neolithic³³ datasets. Altogether, these three datasets include *Homo sapiens* samples spanning a time period from approximately 45,000 to 2,500 years ago (Supplementary Fig. S2).

Due to low-quality DNA, for most of the ancient samples only one DNA strand could be recovered and genotyped, resulting in hemizygous-like genomes. Additionally, given the heterogeneity of procedures for data production, each dataset was analysed independently. For the Near East dataset, three individuals (Ust'-Ishim, MA1, Kostenki14) were withdrawn from analyses as their sample dates were considered to strongly deviate from the remaining individuals in the dataset.

The number of tagSNPs obtained after the linkage disequilibrium (LD) pruning present in our ancient sequences was enriched by implementing a two-step filtering method. First, individuals with a number of genotype missing genotypes above 50% were removed. After such filtering, 16, 151 and 84 ancient genomes remained in the Pre-Neolithic, Near East and Neolithic datasets, respectively. Then, variants with genotype information below 90% of the remaining individuals were discarded. The resulting variants then underwent LD pruning using PLINK 1.9⁴⁸ with parameters *-indep 50 5 2*. ADHD-associated variants with an effect size p -value below 0.01 were retained (169,541 variants) and once filtered by LD, 3,276, 2,707 and 3,320 SNPs were obtained for the Pre-Neolithic, Near East and Neolithic datasets, respectively. The divergence in the number of SNPs is explained by the difference in the individuals included in the datasets.

Genotypes from 16 ancient African samples were retrieved⁴⁹: 6 east Africans (Tanzania, Kenya and Ethiopia), 7 south-central Africans (Malawi), and 3 southern Africans (South Africa). Prehistoric African sample dates range from ~8,100 to 400 years ago.

Modern AMH genomes. Modern human genomes of European and African origin were retrieved from 1000 Genomes Phase 3⁵⁰ (Supplementary Table S2). The database we used comprised information of 503 individuals from five European populations (CEU, Utah residents with Northern and Western European ancestry; TSI, Toscani in Italia; FIN, Finnish in Finland; GBR, British in England and Scotland; and IBS, Iberian population in Spain) and 661 individuals from 7 African populations (YRI, Yoruba in Nigeria; LWK, Luhya in Kenya; GWD, Gambian in Gambia; MSL, Mende in Sierra Leone; ESN, Esan in Nigeria; ASW, Americans of African Ancestry in SW USA; and ACB, African Caribbean in Barbados).

SNPs datasets. *Loci associated with ADHD and other psychiatric disorders.* The effect sizes ($\log(\text{Odds Ratio})$ or β) of ADHD association estimated in 8,047,421 SNPs were retrieved from a GWAS meta-analysis that included 11 cohort samples (19,099 cases and 34,194 controls) of European ancestry⁵ (Supplementary Table S2). Further analyses with this dataset included ADHD-associated SNP variants with a p -value of the association's effect size below 0.01, unless specified otherwise. The GWAS results for autism spectrum disorder (ASD)⁵¹, bipolar disorder (BD)⁵², major depressive disorder (MDD)⁵³ and schizophrenia (SCZ)⁵⁴ were retrieved from the Psychiatric Genomics Consortium (<https://www.med.unc.edu/pgc/>).

SDS from UK10K data. A list of trait singleton density score (tSDS) values calculated using samples from the United Kingdom^{41,44} was available through the Dryad Digital Repository⁴⁴ (Supplementary Table S2). Singletons were restricted to single nucleotide variants (i.e., excluding indels) thus resulting in a final dataset of 4,451,435 SNPs with valid SDS calls and a mean singleton distance of 1.4 Mb⁴⁴.

Dataset of Neanderthal-introgressed tagSNPs in modern humans. TagSNPs of Neanderthal-introgressed haplotypes present in current human populations were retrieved from²³ (Supplementary Table S2). Such variants were found after sequencing 1,523 geographically diverse non-African individuals. In aggregate, 1,340 Mb of the Neanderthal genome were recovered at a false discovery rate (FDR) of 5%^{23,55}.

Analyses. *LoF-intolerant gene enrichment of ADHD-risk variants.* In order to test whether ADHD-associated variants were enriched in LoF-intolerant genes, we used the list of genes defined as LoF ($pLI > 0.9$) from ExAC³⁰ (downloaded from <https://gnomad.broadinstitute.org/downloads>) and conducted a MAGMA analysis²⁹; we repeated the MAGMA analysis only using genes expressed in brain (i.e. the threshold to set if a gene was expressed in a given tissue is ≥ 1 FPKM).

Calculation of the mean frequency of ADHD-risk alleles (f_{ADHD}). For each ancient sample, a mean frequency of the number of ADHD-risk alleles (f_{ADHD}) was computed by combining information from genotype and p -value. This f_{ADHD} was calculated for each individual (i) as the average of the set of ADHD-associated SNPs ($p < 0.01$) with genotype call in that individual (G_i) according to:

$$f_{ADHD, i} = \frac{1}{I} \sum_{j=1}^I G_{i,j} \quad (1)$$

where $G_{i,j}$ corresponds to the number of copies of the risk allele at locus j in individual i (0, 1 or 2).

Contribution of ancestral alleles to ADHD. In order to define the ancestral allele of the SNPs used in⁵ we used the allele present in the chimpanzee (*Pan Troglodytes*) as a proxy. To obtain the chimpanzee allele we used the alignment between panTro4 and hg19 that is publicly available through the UCSC (<http://hgdownload.cse.ucsc.edu/goldenPath/hg19/vsPanTro4/reciprocalBest/>). We downloaded the alignments corresponding to the 22 autosomes present in humans in AXT format. Using a custom Python script, we obtained a FASTA file with the genome of the chimpanzee (panTro4) in the human genome coordinates (hg19). Then we created a BED file that contained the positions of the SNPs and the corresponding chimpanzee allele. We merged this file with the VCF that contains the SNPs being considered, adding the ancestral allele as an extra individual in the last position. In the cases where the result of the alignment was not conclusive, the genotype was labelled as *missing*. Moreover, if information of ancestral allele was not available, the SNP was not considered.

In order to quantify whether ADHD-risk alleles tend to be enriched for the derived or ancestral state, we inferred the ancestral state of 519,911 SNPs that do not contain A/T or C/G alleles and with a GWAS $p \leq 1e-8$ or with a GWAS $p \geq 0.9$, suggestive of strong GWAS association with ADHD or with no association, respectively. For each ADHD GWAS-association category, we computed the percentage of SNPs showing the ancestral state as the GWAS-risk ADHD allele, and calculated the odds ratio (OR). Since the power for detecting an association depends on the minor allele frequency (MAF) of each SNP, in order to estimate how likely is to observe such OR by chance controlling for MAF, we classified each SNP by MAF in bins of 0.02. Statistical significance was obtained out of 10,000 resamples. For each resample, we randomly distributed the SNPs of each MAF bin in the two GWAS-statistical significance categories while keeping the proportion of SNPs at each category. We then computed the OR with the resampled dataset and compared with the observed OR.

Evaluation of f_{ADHD} trend with time in ancient samples. In order to assess the relationship between the computed f_{ADHD} and sampling age, we computed the Kendall's τ correlation with the date estimate of the considered samples. We considered Kendall's τ correlation for these analyses as it allows for a more direct interpretation of the estimated value⁵⁶. To test whether the estimated correlation between sample age and f_{ADHD} is expected under the null hypothesis of independence of the risk allele and the ADHD phenotype (that is, the observed trend is due to demographic factors), for each SNP we randomly assigned the ADHD increasing risk allele, computed a new f_{ADHD} for each ancient individual and estimated the Kendall's correlation between the new f_{ADHD} and the age of the samples. This process was repeated 10,000 times and an empirical p-value was obtained out of this null distribution as the number of times the simulated correlation was greater than the observed one. In the case of the Near East dataset, an additional analysis was conducted to test whether the observed trend could be explained by constant migration from African populations with a putatively lower f_{ADHD} than European populations. Taking advantage of the ancient African dataset, which covers similar sampling ages to the one from Near East, we computed a multiple linear regression between f_{ADHD} , the continent of origin.

Evaluation of f_{ADHD} in modern African and European samples. We further computed the f_{ADHD} in each African and European samples from the 1000 Genomes Phase 3 using the SNPs from each ancient dataset. The difference between f_{ADHD} was tested by means of a Wilcoxon test as implemented in R.

Relationship between Hunter-gatherer ancestry and ADHD in ancient genomes. Aiming to validate some of the mismatch theories (Supplementary Table S2) which mainly ascribe ADHD-specific traits to foraging societies¹⁵, Kendall's correlation was computed between hunter-gatherer ancestry proportions (HG) reported for the Neolithic ancient genomes³³ and f_{ADHD} . Since both variables covariate with sampling age, we also estimated the partial Kendall's correlation between HG and f_{ADHD} controlling by sampling age using the *pcor.test* from the *pcor* package in R.

Correlation between the effect sizes of ADHD and other psychiatric disorders. The Kendall's τ correlation between the effect sizes of each disorder compared with ADHD was computed for the SNPs available in the Pre-Neolithic, Neolithic and Near East datasets weighted by the $-\log(P \text{ value})$ of the ADHD-associated SNP using the *test_indep* function from the *wdm* package in R.

Approximate Bayesian Computation coupled to a deep learning (ABC-DL) framework for estimating the ADHD selection coefficient. In order to estimate the temporal trend of the mean f_{ADHD} over all loci in ancient samples taking into account the fact that the considered populations are related, we implemented an ABC-DL framework. ABC is a statistical framework for computing the posterior distributions of parameters/models when the likelihood of the data given the parameters is not known but there is a way of generating simulated data using parameter values from prior distributions⁵⁷. Here, we constructed a frequency-based simulator using the framework proposed by Pickrell and Pritchard⁵⁸ to explain the change in allelic frequencies under neutrality, we define the evolution of the frequency of the ADHD-risk allele at each locus j ($f'_{ADHD,j}$) at t' generations given that it was known t generations ago as following a Brownian movement:

$$f'_{ADHD,j} \sim N\left(f_{ADHD,j} + \gamma\Delta t, d\Delta t\right) \quad (2)$$

$$\Delta t = t - t'$$

where γ is a parameter quantifying the deviation of the frequency of the ADHD risk allele at locus j per generation due to selective pressures and d is a parameter quantifying the amount of drift between time t and t' . Values of $\gamma > 0$ indicate positive selection for the risk ADHD allele, whereas negative values indicate positive selection for the protective ADHD allele. A value of $\gamma = 0$ indicates that the frequency of the considered locus evolves under neutrality. Given a known population tree topology, Eq. (2) provides an efficient way to simulate the $f_{ADHD,i,j}^t$ at each locus and sample at time t . The average $f_{ADHD,i}^t = \frac{1}{J_i} \sum_{j=1}^{J_i} f_{ADHD,i,j}^t$ for each individual i is computed assuming that each locus contributes equally to the phenotype, since we are interested in the average temporal trend, not in the specific value of each locus. We use the tree topology for the samples from Fu *et al.* (as depicted in Figure 4 in³¹) for samples *UstIshim*, *Malta1*, *Kostenski14*, *GoyetQ116-1*, *Villabruna*, *Loschbour*, *Vestonice16* and *ElMiron* in our simulations. We used Eq. (2) to generate simulations sampling the parameters from the prior distributions described in Supplementary Table S3; from all the considered parameters, we are only interested in γ . For this parameter, we set a uniform prior distribution ranging between $\gamma = -1e-4$ and $\gamma = 1e-4$ per generation, corresponding to an average change in frequency of each ADHD-risk allele between -0.15 to 0.15 in $\sim 1,552$ generations.

A key point of ABC is the definition of a proper set of non-redundant SS that recapitulates the parameters of interest. However, ascertaining informative non-redundant SS is traditionally difficult. Recently, Bai *et al.*⁵⁹ proposed coupling ABC to Deep Learning (DL) in order to ascertain the most informative summary statistics out of the data. A DL or a set of DLs are trained with a raw representation of the data (raw-SS) to predict as an output the parameter that produced the simulation. During the training phase, raw-SS is non-linearly transformed by the DL to predict the value of the parameter used for the simulation. In a further replication step, new simulations are generated and the DL is used to predict the value of the parameter. This predicted value is then used as summary statistic (SS-DL) in the classical ABC framework. This methodology has been recently applied in population genetics to distinguish among competing demographic models and estimating demographic parameters using whole-genome data^{60,61}. The $f_{ADHD,i}^t$ estimated at each ancient sample was used as the raw-SS input of the DL after standardizing each variable. In order to generate a consensus prediction of γ , we trained 10 independent DL using 10,000 simulations. Each DL was constructed in Encog 3 software⁶² using a feedforward architecture of four hidden layers of five neurons each layer. Neurons from the hidden layers used Elliott activation function⁶³ with a slope of 0.1 to prevent signal saturation. The DL was trained using resilient propagation with RPROP- and a Dropout rate of 0.1⁶⁴. Each DL was trained until the error was ≤ 0.01 or reached 10 000 iterations. The output was from an Elliott activation function and consisted on the value of γ used for the simulation scaled to range between 0 and 1. We used this prediction as SS in the ABC framework. ABC was conducted using the *abc* R package⁶⁵ with 1,000,000 new simulations, retaining the 1,000 closest and applying the local linear regression algorithm⁶⁶.

Signatures of recent (2,000–3,000 years) polygenic adaptation. To identify evidence of recent polygenic adaptation in ADHD during the last 2,000–3,000 years, the SDS statistic estimated for the UK10K data was used. Note this time-span does not correspond to a time-limited phenomenon but to the upper-limit resolution of the applied test⁴⁴. Following the approach in⁴⁴, trait-SDS (tSDS) estimates were obtained by matching the SDS value of the ADHD-risk allele. We clustered the tSDS values in bins of 1,000 SNPs according to their GWAS p-value. We computed the Spearman rank correlation between the tSDS values and the p-value of the association with ADHD. In order to assess whether the estimated Spearman rank correlation was expected under the null hypothesis of neutrality, we generated the null distribution out of 1,000 randomisations for each of which we split the genome in bins of 1 Mb, switched the risk allele of the SNPs of each bin with probability 0.5 and computed a new Spearman rank correlation. The resulting distribution was approached to a normal distribution and a one-tail p-value was computed.

Contribution of Neanderthal haplotypes to current ADHD-associated genomic diversity. A list of reported tagSNPs of Neanderthal introgressed haplotypes in modern humans²³ (Supplementary Table S2) was used to test for an enrichment of Neanderthal variants among ADHD risk alleles. Of these tagSNPs, only those variants that were described as homozygous in the Neanderthal Altai genome were retained and the effect size of the association at introgressed Neanderthal alleles were retrieved from the ADHD GWAS meta-analysis dataset. Aiming to inspect whether the Neanderthal-introgressed haplotypes had in aggregate a real effect on ADHD (either by being more disease-associated or protective variants than expected under the hypothesis of no association with the phenotype), the allele effect size of each of the previously analysed tagSNPs was permuted at random, and the mean β of the permuted SNPs was estimated. This process was repeated 10,000 times and simulated mean β values were used to build a null distribution.

Received: 30 September 2019; Accepted: 24 April 2020;

Published online: 25 May 2020

References

1. American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders*. (Arlington, VA: American Psychiatric Publishing, 2013).
2. Faraone, S. V. *et al.* Attention-deficit/hyperactivity disorder. *Nat. Rev. Dis. Prim* **1**, 15020 (2015).
3. Faraone, S. V. *et al.* Molecular genetics of attention-deficit/hyperactivity disorder. *Biological Psychiatry* **57**, 1313–1323 (2005).
4. Hawi, Z. *et al.* The molecular genetic architecture of attention deficit hyperactivity disorder. *Mol. Psychiatry* **20**, 289–297 (2015).

5. Demontis, D. *et al.* Discovery of the first genome-wide significant risk loci for attention deficit/hyperactivity disorder. *Nat. Genet.* **51**, 63–75 (2019).
6. Danckaerts, M. *et al.* The quality of life of children with attention deficit/hyperactivity disorder: A systematic review. *Eur. Child Adolesc. Psychiatry* **19**, 83–105 (2010).
7. Harpin, V. A. The effect of ADHD on the life of an individual, their family, and community from preschool to adult life. *Arch. Dis. Child.* **90**, i2–7 (2005).
8. Østergaard, S. D., Dalsgaard, S., Faraone, S. V., Munk-Olsen, T. & Laursen, T. M. Teenage Parenthood and Birth Rates for Individuals With and Without Attention-Deficit/Hyperactivity Disorder: A Nationwide Cohort Study. *J. Am. Acad. Child Adolesc. Psychiatry* **56**, 578–584.e3 (2017).
9. Thagaard, M. S., Faraone, S. V., Sonuga-Barke, E. J. & Ostergaard, S. D. Empirical tests of natural selection-based evolutionary accounts of ADHD: A systematic review. *Acta Neuropsychiatrica* **28**, 249–256 (2016).
10. Keller, M. C. & Miller, G. Resolving the paradox of common, harmful, heritable mental disorders: Which evolutionary genetic models work best? *Behav. Brain Sci.* **29**, 385–404 (2006).
11. Stearns, S. C. & Medzhitov, R. *Evolutionary medicine*. (Sinauer Associates, Inc, 2015).
12. Crawford, C. & Salmon, C. Psychopathology or adaptation? Genetic and evolutionary perspectives on individual differences and psychopathology. *Neuro endocrinology letters* **23 Suppl 4**, 39–45 (2002).
13. Shelley-Tremblay, J. F. & Rosén, L. A. Attention Deficit Hyperactivity Disorder: An Evolutionary Perspective. *J. Genet. Psychol.* **157**, 443–453 (1996).
14. Jensen, P. S. *et al.* Evolution and Revolution in Child Psychiatry: ADHD as a Disorder of Adaptation. *J. Am. Acad. Child Adolesc. Psychiatry* **36**, 1672–1681 (1997).
15. Hartmann, T. & Michael, P. *Attention deficit disorder: a different perception*. (Underwood Books, 1997).
16. Prüfer, K. *et al.* The complete genome sequence of a Neanderthal from the Altai Mountains. *Nature* **505**, 43–49 (2014).
17. Meyer, M. *et al.* A high-coverage genome sequence from an archaic Denisovan individual. *Science* (80) **338**, 222–226 (2012).
18. Patterson, N. *et al.* The genomic landscape of Neanderthal ancestry in present-day humans. *Nature* **507**, 354–7 (2014).
19. Vernot, B. & Akey, J. M. Complex history of admixture between modern humans and neandertals. *American Journal of Human Genetics* **96**, 448–453 (2015).
20. Vattathil, S. & Akey, J. M. Small Amounts of Archaic Admixture Provide Big Insights into Human History. *Cell* **163**, 281–284 (2015).
21. Sankararaman, S. *et al.* The Combined Landscape of Denisovan and Neanderthal Ancestry in Present-Day Humans. *Curr. Biol.* **26**, 1241–1247 (2016).
22. Simonti, C. N. *et al.* The phenotypic legacy of admixture between modern humans and Neandertals. *Science* (80) **351**, 737–741 (2016).
23. Vernot, B. *et al.* Excavating Neandertal and Denisovan DNA from the genomes of Melanesian individuals. *Science* (80) **352**, 235–9 (2016).
24. Currat, M. & Excoffier, L. Strong reproductive isolation between humans and Neanderthals inferred from observed patterns of introgression. *Proc. Natl. Acad. Sci. USA* **108**, 15129–34 (2011).
25. Nielsen, R. *et al.* Tracing the peopling of the world through genomics. *Nature* **541**, 302–310 (2017).
26. Smith, N. G. C. & Eyre-Walker, A. Human disease genes: Patterns and predictions. *Gene* **318**, 169–75 (2003).
27. Blekhnman, R. *et al.* Natural Selection on Genes that Underlie Human Disease Susceptibility. *Curr. Biol.* **18**, 883–9 (2008).
28. Pardiñas, A. F. *et al.* Common schizophrenia alleles are enriched in mutation-intolerant genes and in regions under strong background selection. *Nat. Genet.* **50**, 381–389 (2018).
29. de Leeuw, C. A., Mooij, J. M., Heskes, T. & Posthuma, D. MAGMA: Generalized Gene-Set Analysis of GWAS Data. *PLoS Comput. Biol.* **11**, e1004219 (2015).
30. Lek, M. *et al.* Analysis of protein-coding genetic variation in 60,706 humans. *Nature* **536**, 285–91 (2016).
31. Fu, Q. *et al.* The genetic history of Ice Age Europe. *Nature* **534**, 200–205 (2016).
32. Lazaridis, I. *et al.* Genomic insights into the origin of farming in the ancient Near East. *Nature* **536**, 419–24 (2016).
33. Lipson, M. *et al.* Parallel palaeogenomic transects reveal complex genetic history of early European farmers. *Nature* **551**, 368–372 (2017).
34. Martin, A. R. *et al.* Human Demographic History Impacts Genetic Risk Prediction across Diverse Populations. *Am. J. Hum. Genet.* **100**, 635–649 (2017).
35. Curtis, D. Polygenic risk score for schizophrenia is more strongly associated with ancestry than with schizophrenia. *Psychiatr. Genet.* **28**, 85–89 (2018).
36. Franke, B. *et al.* Live fast, die young? A review on the developmental trajectories of ADHD across the lifespan. *Eur. Neuropsychopharmacol.* **28**, 1059–1088 (2018).
37. Petr, M., Pääbo, S., Kelso, J. & Vernot, B. Limits of long-term selection against Neandertal introgression. *Proc. Natl. Acad. Sci. USA* **116**, 1639–1644 (2019).
38. Haak, W. *et al.* Massive migration from the steppe was a source for Indo-European languages in Europe. *Nature* **522**, 207–211 (2015).
39. Excoffier, L., Estoup, A. & Cornuet, J. M. Bayesian analysis of an admixture model with mutations and arbitrarily linked markers. *Genetics* **169**, 1727–1738 (2005).
40. Fenner, J. N. Cross-cultural estimation of the human generation interval for use in genetics-based population divergence studies. *Am. J. Phys. Anthropol.* **128**, 415–423 (2005).
41. The UK10K Project Consortium. The UK10K project identifies rare variants in health and disease. *Nature* **526**, 82–90 (2015).
42. Belsky, D. W. & Israel, S. Integrating Genetics and Social Science: Genetic Risk Scores. *Biodemography Soc. Biol.* **60**, 137–155 (2014).
43. Lee, S. H. *et al.* Genetic relationship between five psychiatric disorders estimated from genome-wide SNPs. *Nat. Genet.* **45**, 984–994 (2013).
44. Field, Y. *et al.* Detection of human adaptation during the past 2000 years. *Science* (80). **354**, 760–764 (2016).
45. Castellano, S. *et al.* Patterns of coding variation in the complete exomes of three Neandertals. *Proc. Natl. Acad. Sci.* **111**, 6666–71 (2014).
46. Berens, A. J., Cooper, T. L. & Lachance, J. The Genomic Health of Ancient Hominins. *Hum. Biol.* **89**, 7–19 (2017).
47. Antshel, K. M., Zhang-James, Y. & Faraone, S. V. The comorbidity of ADHD and autism spectrum disorder. *Expert Rev. Neurother* **13**, 1117–1128 (2013).
48. Chang, C. C. *et al.* Second-generation PLINK: rising to the challenge of larger and richer datasets. *Gigascience* **4**, 7 (2015).
49. Skoglund, P. *et al.* Reconstructing Prehistoric African Population Structure. *Cell* **171**(59–71), e21 (2017).
50. Auton, A. *et al.* A global reference for human genetic variation. *Nature* **526**, 68–74 (2015).
51. Grove, J. *et al.* Identification of common genetic risk variants for autism spectrum disorder. *Nat. Genet.* **51**, 431–444 (2019).
52. Stahl, E. A. *et al.* Genome-wide association study identifies 30 loci associated with bipolar disorder. *Nat. Genet.* **51**, 793–803 (2019).
53. Howard, D. M. *et al.* Genome-wide association study of depression phenotypes in UK Biobank identifies variants in excitatory synaptic pathways. *Nat. Commun.* **9**, 1470 (2018).
54. Lam, M. *et al.* Comparative genetic architectures of schizophrenia in East Asian and European populations. *Nat. Genet.* **51**, 1670–1678 (2019).
55. Vernot, B. & Akey, J. M. Resurrecting surviving Neandertal lineages from modern human genomes. *Science* (80) **343**, 1017–1021 (2014).

56. Howell, D. C. *Statistical methods for psychology*. (Wadsworth, Cengage Learning, 2010).
57. Sunnåker, M. *et al.* Approximate Bayesian Computation. *PLoS Comput. Biol.* **9**, e1002803 (2013).
58. Pickrell, J. K. & Pritchard, J. K. Inference of Population Splits and Mixtures from Genome-Wide Allele Frequency Data. *PLoS Genet.* **8**, e1002967 (2012).
59. Jiang, B., Wu, T.-Y., Zheng, C. & Wong, W. H. Learning summary statistic for approximate bayesian computation via deep neural network. *Stat. Sin.* **27**, 1595–1618 (2017).
60. Mondal, M., Bertranpetit, J. & Lao, O. Approximate Bayesian computation with deep learning supports a third archaic introgression in Asia and Oceania. *Nat. Commun.* **10**, 246 (2019).
61. Lorente-Galdos, B. *et al.* Whole-genome sequence analysis of a Pan African set of samples reveals archaic gene flow from an extinct basal population of modern humans into sub-Saharan populations. *Genome Biol.* **20**, 77 (2019).
62. Heaton, J. Programming Neural Networks with Encog 2 in Java. *J. Chem. Inf. Model.* **53**, 1689–1699 (2013).
63. Elliott, D. L., Elliott, D. L. & Elliott, D. L. *A better Activation Function for Artificial Neural Networks*. (1993).
64. Goodfellow, I., Bengio, Y. & Courville, A. *Deep learning*. (MIT Press, 2017).
65. Csilléry, K., François, O. & Blum, M. G. B. Abc: An R package for approximate Bayesian computation (ABC). *Methods Ecol. Evol.* **3**, 475–479 (2012).
66. Beaumont, M. A., Zhang, W. & Balding, D. J. Approximate Bayesian computation in population genetics. *Genetics* **162**, 2025–2035 (2002).
67. Morgan, E. *The descent of woman*. (Souvenir, 1985).

Acknowledgements

We thank U. M. Marigorta and J.A. Rodríguez for helpful comments and suggestions on the manuscript. We acknowledge the support of the Spanish Ministry of Science and Innovation through the Instituto de Salud Carlos III, the Generalitat de Catalunya through Departament de Salut and Departament d'Empresa i Coneixement, the co-funding by the Spanish Ministry of Science and Innovation with funds from the European Regional Development Fund (ERDF) corresponding to the 2014-2020 Smart Growth Operating Program. O.L. gratefully acknowledges the financial support from Ministerio de Economía y Competitividad (Ministry of Economy and Competitiveness) - RYC-2013-14797, BFU2015-68759-P and PGC2018-098574-B-I00, Generalitat de Catalunya (Government of Catalonia) - GRC 2017 SGR 937, EC | EU Horizon 2020 Research and Innovation Programme (Personalising Health and Care) - 667302 and EC | EU Horizon 2020 Research and Innovation Programme (Marie Skłodowska-Curie Action: Innovative Training Networks, ITN) - 643051. P.E.-C. gratefully acknowledges the financial support from the Government of Catalonia | Agència de Gestió d'Ajuts Universitaris i de Recerca (Agency for Management of University and Research Grants) - FI_B1_00113. I.M. gratefully acknowledges the financial support from the Government of Catalonia | Agència de Gestió d'Ajuts Universitaris i de Recerca (Agency for Management of University and Research Grants) - GRC 2014 SGR 615. A.D.B. and D.D. were supported by the Lundbeck Foundation (grant numbers R102-A9118 and R155-2014-1724). High-performance computer capacity for handling and statistical analysis of iPSYCH and PGC data on the GenomeDK HPC facility was provided by the Center for Genomics and Personalized Medicine, Aarhus University and Central Region Denmark, and Centre for Integrative Sequencing, iSEQ, Aarhus University (grant to A.D.B.). S.V.F. was supported by the European Union's Seventh Framework Programme for Research, Technological Development and Demonstration under grant agreement - 602805, the European Union's Horizon 2020 Research and Innovation Programme under grant agreements 667302 & 728018 and NIMH grants 5R01MH101519 and U01 MH109536-01. B.C. gratefully acknowledges the financial support from Ministerio de Economía y Competitividad (Ministry of Economy and Competitiveness) - SAF2015-68341-R and RTI2018-100968-B-I00, Generalitat de Catalunya | Agaur (Government of Catalonia) - GRC 2017 SGR 738, EC | EU Horizon 2020 Research and Innovation Programme (Personalising Health and Care) - 667302, EC | EU Horizon 2020 Research and Innovation Programme (Marie Skłodowska-Curie Action: Innovative Training Networks, ITN) - 643051 and EC | EU Horizon 2020 Research and Innovation Programme (Sustainable Food Security - Resilient and Resource - Efficient Value Chains) - 728018. This publication reflects only the authors' view and the European Commission is not responsible for any use that may be made of the information it contains.

Author contributions

P.E.-C., B.C. and O.L. conceived the study. A.D.B., D.D. and S.V.F. contributed to the study design and GWAS data analyses. P.E.-C., I.M. and O.L. performed analyses. P.E.-C., B.C. and O.L. wrote the manuscript with input from all co-authors.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at <https://doi.org/10.1038/s41598-020-65322-4>.

Correspondence and requests for materials should be addressed to B.C. or O.L.

Reprints and permissions information is available at www.nature.com/reprints.

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



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2020