Evolution of mental traits linked to cultural development in modern humans and Neanderthals

Yuka Moriya[1], Naoko T. Fujito[2], Masahiro Terahara[1], Naoyuki Takahata[2], Yoko Satta[2], Kosuke M. Teshima[3], Toshiyuki Hayakawa[4*]

[1]Graduate School of Systems Life Sciences, Kyushu University, Fukuoka, Japan
[2]School of Advanced Sciences, SOKENDAI (The Graduate University for Advanced Studies), Hayama, Kanagawa, Japan
[3]Department of Biology, Kyushu University, Fukuoka, Japan
[4]Faculty of Arts and Science, Kyushu University, Fukuoka, Japan

* Corresponding author:
Toshiyuki Hayakawa
Kyushu University, 744 Motooka, Nishi-ku, Fukuoka 819-0395, Japan
Phone: +81-92-802-5827; E-mail: thayakawa@artsci.kyushu-u.ac.jp

Abstract

A cumulative process of cultural evolution occurred globally alongside the spread of anatomically modern humans (AMHs). This process of evolution was likely accompanied by global changes in mental and behavioral phenotypes, resulting in cultural differences between AMHs and Neanderthals. Globally, selective increases of frequencies were detected in alleles, which contributed to the suppression of neuroticism and advanced learning. This finding indicates that neuroticism was suppressed and learning behaviors developed through the global spread of AMHs. It suggests that global adaptation to psychosocial stress occurred, which contributed to cumulative cultural evolution by increasing cooperative learning. A comparison of phenotypic trends at the population level indicates that such globally occurring adaptive evolution, entailing neuroticism suppression and increased learning behaviors has been more significant in AMHs than in Neanderthals. However, adaptive evolution toward greater intelligence was more significant in Neanderthals than in AMHs. These findings suggest that mental traits involved in cultural development evolved differently in AMHs and Neanderthals, and that learning behaviors such as cooperative learning, rather than intelligence, played an important role in AMH evolution.

Introduction

Anatomically modern humans (AMHs; *Homo sapiens*) emerged in Africa and expanded their habitat to almost every corner of the globe (1). While spreading out, AMHs developed increasing technological sophistication and adaptive knowledge, which promoted efficiency, enabling their survival in new environments (2). Presumably, the adaptation of mental and behavioral phenotypes would have occurred during this cumulative process of cultural evolution, prompting a global evolutionary trend based on the distribution of associated alleles. One such trend would have been adaptive evolution of the *ST8SIA2* gene (3, 4), associated with brain/neural function (5). Selective increases of its promoter types (haplotypes of three promoter single nucleotide polymorphisms [SNPs]), annotated as schizophrenia-associated non-risk types and showing lower promoter activity than the risk types, induced a trend toward lower promoter activity at the population level (population promoter activity: PPA) through out-migration from Africa of AMHs. This phenomenon can be interpreted as an adaptation to psychosocial stress induced during the process of adapting to surrounding social environments (3, 4) because psychosocial stress is the sole schizophrenia environmental risk factor dealt with by brain functions (6, 7) (Fig. S1).

Considering the above discussion, we applied evolutionary trends of schizophrenia-associated genetic loci showing selective increases of non-risk allelic types as yardsticks through the establishment of a new population functional score (PFS). Consequently, we detected global adaptive evolution of mental phenotypes involved in cumulative cultural evolution by examining the dominance of similar trends in their evolution (Fig. S2). To elucidate the significance of such adaptive evolution in AMHs, we further determined differences between phenotypic trends in AMHs and Neanderthals using a new population polygenic score (PPS). We subsequently examined positive selection on genetic loci associated with mental phenotypes showing such trend differences in Neanderthals.

Results

A yardstick PFS-26 for non-AFR

As a first step, we developed a PFS reflecting allelic function at the population level by

focusing on the functional difference (*D*) between alleles. We could therefore apply the functional trend of the *ST8SIA2* gene as a yardstick in the analysis using alleles for which no functional information was available. Next, we examined the correlation in the PFS distributions of 26 populations (PFS-26s) with a global spread, encompassing Africa (AFR), Europe (EUR), South Asia (SAS), East/Southeast Asia (EAS), and the Americas (AMR), derived from the 1000 Genomes Project (8). Therefore, experimental estimation of *D* was unnecessary (see Materials and Methods; Figs. S2 and S3). The PFS-26 of the *ST8SIA2* gene ($\text{PFS-26}_{\text{ST8SIA2}}$) was obtained by determining the difference in promoter activity referring to the ancestral risk type, which showed the highest activity (Fig. 1A). PPA and PFS were negatively correlated, and $\text{PFS-26}_{\text{ST8SIA2}}$ showed a correlation coefficient (ρ) of -0.99 with the PPAs of the 26 populations, indicating the precise translation of PPA into PFS. Because $\text{PFS-26}_{\text{ST8SIA2}}$ mainly reflects the selective increase of schizophrenia-associated non-risk types in non-AFR (3, 4) (see Fig. 1A), we used it as a yardstick for examining selection in non-AFR.

A yardstick PFS-26 for AFR

To apply PFS-26 as a yardstick reflecting the selective increase of alleles in AFR, we next examined positive selection on schizophrenia-associated SNPs (9) (SCZ1 in Table 1). We detected signals of positive selection on non-risk alleles of three SNPs of brain/neural function-related genes implicated by fine-mapping (10): rs11693094 ($F_c = 0.024$, $P = 0.003$) in the *ZNF804A* gene, rs9922678 ($F_c = 0.109$, $P = 4.3 \times 10^{-4}$) in the *GRIN2A* gene, and rs1501357 ($F_c = 0.028$, $P = 0.003$) in the *HCN1* gene. The operational durations of this positive selection were estimated to be 17,100, 25,200 and 33,600 years for rs11693094, rs9922678, and rs1501357, respectively. These findings indicate that the detected positive selection could be an adaptation to psychosocial stress during the spread of AMHs. As for the non-risk alleles of rs11693094 and rs9922678, signals were detected in EUR and EAS, respectively. A signal on the non-risk allele of rs1501357 was also detected in AFR. We therefore selected its PFS-26 ($\text{PFS-26}_{\text{rs1501357}}$) as a yardstick for examining positive selection in AFR (Fig. 1B).

Criterion in a correlation analysis

SNPs associated with phenotypes were obtained from genome-wide association studies (GWASs) (see Table 1 for abbreviations of the GWAS SNP sets), and index SNPs were therefore used instead of causal SNPs to obtain PFS-26s. Nearly complete linkages between index and causal SNPs can only be assumed in populations targeted in the GWAS analyses. Accordingly, the PFS-26s would not have been completely identical to those from causal SNPs. To assess the degree of discrepancy, we focused on concordant SNP pairs showing a tight linkage between the same allelic type of a particular phenotype in EUR because we obtained index SNPs from GWASs that mainly targeted the European population or meta-population containing Europeans (Fig. S4). The size of the associated region was typically <200 kb in length (11). In these regions, 743 concordant SNP pairs identified independently in different GWASs for the same phenotype (neuroticism, intelligence, schizophrenia, BMI, height, and type 2 diabetes; see Materials and Methods) showed a tight linkage ($r^2 > 0.81$). Lastly, assuming that these tight-linked SNPs were index and causal SNPs, we examined correlations between PFS-26s in SNP pairs. The ρ-values were mainly distributed within a range from 0.9 to 1 (71% [527/743]) (Fig. S5A). Along a sliding window of $\rho = 0.05$, the PFS-26 proportion was consistently 6% (one-tenth of the highest proportion) or less after a window shift covering PFS-26s showing a ρ-value <0.9 (Fig. S5B). This indicated that when PFS-26 from a causal SNP was completely identical to the yardstick PFS-26s (i.e., $\rho = 1$), the PFS-26 from the index SNP generally showed a value of $\rho > 0.9$ in the correlation with the yardstick PFS-26. Thus, a ρ-value >0.9 was considered a reasonable criterion for assessing a high correlation with the yardstick PFS-26s.

The PFS-26s of non-risk alleles of rs11693094 and rs9922678, which showed positive selection in non-AFR (see above), exhibited a ρ-value >0.9 with PFS-$26_{ST8SIA2}$ (ρ = 0.92 [rs11693094] and 0.92 [rs9922678]; see Fig. S6). The ρ-value of PFS-$26_{rs1501357}$ with PFS-$26_{ST8SIA2}$ was 0.68 (below the threshold of 0.9). These findings indicate that PFS-$26_{ST8SIA2}$ reflects the PFS distribution resulting from selective allelic increase in non-AFR and are consistent with the assertion that PFS-$26_{ST8SIA2}$ does not adequately document the selective increase of alleles in AFR. The criterion of $\rho > 0.9$ with the yardstick PFS-26s is therefore appropriate for detecting genetic loci showing a PFS distribution matching that of

the yardstick PFS-26s. Furthermore, even if index SNPs are not guaranteed to be causal SNPs, this finding also suggests that index SNPs could be used instead of causal SNPs in our analysis.

Convergent evolution of PFS-26s

The observed rate of PFS-26s exhibiting a ρ-value >0.9 with the yardstick PFS-26s (yardstick-like PFS-26s) was examined for each allelic type (Table 1). The rate for non-risk alleles/negative-effect alleles was compared with that for risk alleles/positive-effect alleles in each GWAS SNP set. Using PFS-26$_{ST8SIA2}$, Bonferroni-corrected significant differences ($\alpha = 0.01/13$; $P < 7.7 \times 10^{-4}$) were found in SNP sets associated with neuroticism (NEU1 and NEU2), educational attainment (EDU), and body mass index (BMI) (see Table 1), and the rates for negative-effect alleles of NEU1, NEU2, and BMI and positive-effect alleles of EDU were higher than those for the opposite-effect alleles ($P = 3.7 \times 10^{-5}$ [NEU1], $P = 7.2 \times 10^{-5}$ [NEU2], $P = 1.1 \times 10^{-6}$ [BMI], and $P < 7.7 \times 10^{-4}$ [EDU]; binomial test) (Table 1). The use of PFS-26$_{rs1501357}$ showed the occurrence of Bonferroni-corrected significance in NEU1, EDU, SCZ1, and the depression-associated SNP set (DEP). The rates for negative-effect alleles of NEU1, positive-effect alleles of EDU, and non-risk alleles of SCZ1/DEP were higher than those for the opposite-effect alleles ($P < 7.7 \times 10^{-4}$ [NEU1, EDU and SCZ1] and $P = 6.5 \times 10^{-5}$ [DEP]; binomial test) (Table 1). The significant occurrence rate of the yardstick-like PFS-26s represents convergent evolution of PFS-26s through the global spread of AMHs (Fig. S2).

The PFS-26s characteristic of phenotypes

To examine the involvement of the yardstick PFS-26s in PFS-26s unique to each phenotype, we obtained the PFS-26s characteristic of each allelic type showing a significant occurrence rate of the yardstick-like PFS-26s (see Fig. S7). We identified characteristic PFS-26s using the median ρ-value as a threshold to identify a similarity trend in a pairwise comparison of PFS-26s. These characteristic PFS-26s showed that ρ-values above the median value significantly occurred with the same allelic types rather than with the opposite-effect types ($\alpha = 0.05$/number of alleles). They could be considered to reflect the unique allele distribution

in a certain phenotype and were obtained in NEU1, NEU2, EDU, and BMI (Fig. 2 and Fig. S8). Strikingly, the distribution of ρ-values with the yardstick PFS-26s was restricted to a positive correlation in characteristic PFS-26s of negative-effect alleles of NEU1 (100% [8/8] in PFS-$26_{ST8SIA2}$ and PFS-$26_{rs1501357}$), NEU2 (100% [3/3] in PFS-$26_{ST8SIA2}$), and BMI (100% [7/7] in PFS-$26_{ST8SIA2}$), and positive-effect alleles of EDU (96% [2136/2223] in PFS-$26_{ST8SIA2}$; 98% [2180/2223] in PFS-$26_{rs1501357}$) (Fig. 2A–C and Fig. S8).

Domination of the yardstick PFS-26s

The degree to which positive correlations with the yardstick PFS-26s dominated in the characteristic PFS-26s was elucidated by determining the correlation in characteristic PFS-26s between ρ-values with the yardstick PFS-26s and the average ρ-values with other characteristic PFS-26s (Fig. 2D–F; Fig. S9). As for PFS-$26_{ST8SIA2}$, high R^2 values with Bonferroni-corrected significant positive correlations ($\alpha = 0.01/3$; $P < 3.3 \times 10^{-3}$) were found in characteristic PFS-26s of NEU1 and EDU (Fig. 2D and E). Higher ρ-values with PFS-$26_{ST8SIA2}$ were associated with higher average ρ-values with other characteristic PFS-26s in negative-effect alleles of NEU1 ($R^2 = 0.84$, $P = 1.5 \times 10^{-3}$) and positive-effect alleles of EDU ($R^2 = 0.43$, $P = 2.6 \times 10^{-275}$) (Fig. 2D and E). A high R^2 value with Bonferroni-corrected significant positive correlation was found for PFS-$26_{rs1501357}$ in EDU (Fig. 2F). Higher ρ-values with PFS-$26_{rs1501357}$ were associated with higher average ρ-values with other characteristic PFS-26s in positive-effect alleles of EDU ($R^2 = 0.53$, $P < 3.3 \times 10^{-3}$) (Fig. 2F). Taken together with the significant occurrence rate of the yardstick-like PFS-26s (Table 1) and an exclusive restriction toward positive ρ values with the yardstick PFS-26s in the characteristic PFS-26s (Fig. 2A–C), the observed association shown in Fig. 2D–F indicated that a trend of becoming similar to the yardstick PFS-26s dominated for the PFS-26s unique to the negative-effect alleles of NEU1 and the positive-effect alleles of EDU.

To confirm the uniqueness of the domination of the yardstick PFS-26s in NEU1 and EDU, we applied the PFS-26s of alleles already established as selective targets (rs4988235 in the *MCM6* gene, rs1800414 in the *OCA2* gene, and rs3827760 in the *EDAR* gene) (12–14) as yardsticks (control-yardstick PFS-26s). The control-yardstick PFS-26s showed no dominant trend of becoming similar to them in the characteristic PFS-26s of NEU1 and EDU

(Table S1 and Fig. S10).

Global adaptation in AMHs

The dominant trend of becoming similar to the yardstick PFS-26s was detected only in the PFS-26s of neuroticism negative-effect alleles and educational attainment positive-effect alleles (Table 1 and Fig. 2). Considered together with the genetic correlation between lower neuroticism and increased educational attainment (15), this dominant trend indicates that singular global adaptive evolution has driven the suppression of neuroticism while increasing behaviors with positive effects on educational attainment. Neuroticism is a personality trait associated with sensitivity to psychosocial stress (16), whereas educational attainment is a behavioral phenotype primarily determined by environmental factors. Because the progress of cooperative learning via smooth communication that is achieved by overcoming psychosocial stress has a positive effect on educational attainment, psychosocial stress is considered an influential environmental factor. Tolerance to psychosocial stress is thus a shared function between neuroticism negative-effect alleles, educational attainment positive-effect alleles, and schizophrenia non-risk alleles (see also Fig. S1). We therefore posit that the dominant trend of becoming similar to the yardstick PFS-26s in the PFS-26s of neuroticism negative-effect alleles and educational attainment positive-effect alleles indicates a polygenic trend toward higher frequencies of allelic types conferring tolerance to psychosocial stress through the global spread of AMHs. Simply put, it indicates global polygenic adaptation to psychosocial stress.

Assessment of Neanderthal phenotypes

We further examined the significance of detected global adaptive evolution in AMHs by comparing phenotypic trends for AMHs and Neanderthals (Altai, Vindija, and Chagyrskaya). Accordingly, we developed a new PPS by summarizing a polygenic phenotypic trend at the population level. This was based conceptually on the polygenic risk score/polygenic score (PRS/PS) in the GWAS analyses (see Materials and Methods). As previously noted, SNPs used in this study were index SNPs, which are not necessarily causal SNPs. We therefore examined the conservation of linkages between AMHs and Neanderthals using the 743

concordant SNP pairs showing a strong linkage ($r^2 > 0.81$) in EUR. Because Neanderthal genome sequences are unphased, we selected 726 SNP pairs in which both SNPs are homozygous in Altai whose genome coverage was higher than Vindija and Chagyrskaya. Of these pairs, 589 (81%) showed complete linkages in Altai, indicating that using index SNPs is appropriate even in Neanderthals.

In every GWAS, over 99% of genetic loci were found in Neanderthals (see Table S2). This finding indicates that genetic loci associated with phenotypes targeted in this study have been well-conserved between AMHs and Neanderthals, and the estimation of PPS was performed using almost all of the genetic loci identified in the GWASs. Effect sizes showed a biased distribution toward smaller values in every GWAS (Fig. S11). These results indicate that phenotype expression depends on the orchestration of genetic loci having small effect sizes. Considering them together with almost complete conservation of genetic loci, we concluded that exclusion of genetic loci not obtained in Neanderthals was not so significant in our PPS comparison of AMHs and Neanderthals. Furthermore, even if Neanderthals had their own effect size in each genetic loci, differences from those of AMHs would be negligible.

Significant phenotypic trends in AMHs

Given the conservation of gene repertories between AMHs and Neanderthals and the biased distribution toward smaller effect sizes (Fig. S11), we assumed that gene repertories and effect sizes were similar for AMHs and Neanderthals. We compared phenotypic trends in AMHs and Neanderthals using the PPS of positive-effect alleles (PPSp) (Fig. 3A; Fig. S12). In this comparative analysis, we focused on non-disease phenotypes because global adaptive evolution was only detected in these phenotypes (see Table 1 and Fig. 2). As for Neanderthals, the PPSp of each GWAS was estimated as a possible distribution because only three individuals were available (Fig. 3A; see Materials and Methods). When comparing PPSp in AMHs and Neanderthals, we detected Bonferroni-corrected significance ($\alpha = 0.01/7$; $P < 1.4 \times 10^{-3}$) in NEU2, EDU, BMI, and the intelligence-associated SNPs (INT) but not in NEU1, the height-associated SNPs (HEI), and the chronotype-associated SNPs (CHR) (Fig. 3A; Fig. S12).

For BMI, PPSp was higher in Neanderthals than in AMHs ($P = 1.0 \times 10^{-4}$; Fig. 3A). This finding and the absence of a significant difference found in HEI (Fig. 3A; Fig. S12) were consistent with differences in body proportions between AMHs and Neanderthals, with Neanderthals being slightly shorter and significantly heavier than AMHs (17). These findings indicate that PPSp can be used to compare phenotypic trends of AMHs and Neanderthals.

For NEU2, the PPSp in AMHs was significantly lower than that in Neanderthals ($P = 1.0 \times 10^{-4}$; Fig. 3A). However, the PPSp for EDU was significantly higher in AMHs than in Neanderthals ($P < 1.0 \times 10^{-4}$; Fig. 3A). These findings are consistent with a view of global adaptive evolution toward the suppression of neuroticism and increased learning behaviors in AMHs. Given these consistencies in the two phenotypes, such global adaptive evolution is likely to have been more significant in AMHs than in Neanderthals. Thus, global adaptation to psychosocial stress was more significant in AMHs than in Neanderthals.

Phenotypic trends in Neanderthals

Surprisingly, for INT, the PPSp of AMHs was significantly lower than that of Neanderthals ($P < 1.0 \times 10^{-4}$; Fig. 3A). Therefore, adaptive evolution toward greater intelligence was possibly more significant in Neanderthals than in AMHs. To assess this possibility, we also examined positive selection on genetic loci of NEU2, EDU, INT, and BMI in Neanderthals. Three Neanderthal individuals, Altai, Chagyrskaya, and Vindija, showed similar heterozygosity: 0.021%, 0.020%, and 0.019% (18). Because the ages of these individuals differed (Altai: 122,000 years; Chagyrskaya: 80,000 years; Vindija: 52,000 years) (18, 19), it can be assumed that heterozygosity was maintained constantly in the Neanderthal lineage. Considering the age difference of about 70,000 years between the oldest individual (Altai) and the youngest (Vindija), we examined the change in heterozygosity between them to look for positive selection (see Fig. S13). Our PFS analysis (see Table 1 and Fig. 2) revealed that a clear signal of positive selection was not necessarily detectable in all genetic loci identified by the GWAS. According to the difference in heterozygosity between Vindija and Altai (see Materials and Methods), the SNP sites, which are homozygous in both individuals, were categorized into three groups: Vindija<Altai (lower in Vindija), Vindija≈Altai (no difference) and Vindija>Altai (higher in Vindija) (Figs. S13 and S14). Positive selection at a particular

site would reduce the heterozygosity of the region surrounding the site, and differences in heterozygosity between positive-effect alleles and negative-effect alleles would thus be observable, especially in the Vindija<Altai group and/or the Vindija>Altai group (Fig. S15). A significant difference ($P < 0.05$; Kolmogorov-Smirnov test) was detected in the Vindija<Altai group for INT and in the Vindija>Altai group for BMI (Fig. 3B and C).

The Vindija>Altai group for BMI showed that a trend toward lower heterozygosity in Vindija than in Altai was stronger in the positive-effect alleles than in the negative-effect alleles ($P = 0.0117$; Fig. 3C; Fig. S16). The higher PPSp for BMI in Neanderthals compared with that in AMHs remained unchanged even when Vindija was excluded (Fig. 3A; Fig. S17A), indicating that the detected difference in PPSp did not depend only on Vindija. These findings support the view that adaptive evolution toward higher BMI was more significant in Neanderthals than in AMHs. This interpretation is consistent with adaptive acquisition of body proportion (i.e., high BMI) in the Neanderthal lineage (17).

For INT, the Vindija<Altai group revealed that a trend toward lower heterozygosity in Vindija than in Altai was stronger for the positive-effect alleles than for the negative-effect alleles ($P = 0.0075$; Fig. 3B; Fig. S16). This means that intelligence-associated genetic loci were selective targets in Neanderthal populations. Furthermore, the higher PPSp for INT in Neanderthals compared with that in AMHs remained unchanged even when Vindija was excluded (Fig. 3A; Fig. S17B). These findings support the view that adaptive evolution toward greater intelligence was more significant in Neanderthals than in AMHs. We considered this finding together with the greater significance of global adaptive evolution toward the suppression of neuroticism and increased learning behaviors in AMHs than in Neanderthals. Accordingly, we posited that the mental traits associated with cultural development evolved differently in AMHs and Neanderthals (Fig. S18).

Discussion

The yardstick PFS-26s ($\text{PFS-26}_{\text{ST8SIA2}}$ and $\text{PFS-26}_{\text{rs1501357}}$; Fig. 1) are just two PFS distribution patterns resulting from global adaptation to psychosocial stress. Given that other patterns would have been present but are currently unknown, the yardstick PFS-26s should

not be misinterpreted as reflecting a difference in adaptation to psychosocial stress between populations.

NEU1 but not NEU2 showed a dominant trend of becoming similar to PFS-$26_{ST8SIA2}$ (Table 1; Fig. 2A and D; Figs. S8A and S19A). However, NEU2 showed lower PPSp in AMHs than in Neanderthals, whereas NEU1 showed no significant difference in PPSp between AMHs and Neanderthals (Fig. 3A). In addition, for NEU1, the dominant trend of becoming similar to the yardstick PFS-26s was detected with the PFS-$26_{ST8SIA2}$ but not with the PFS-$26_{rs1501357}$ (Table 1; Fig. 2A and D; Figs. S8C and S19B). These incompatibilities relating to NEU1 and NEU2 could suggest a minim necessary locus size (~100 SNPs) for our PFS and PPS analyses.

Although psychosocial stress is one of the environmental risk factors associated with the onset of schizophrenia and depression (6, 7, 20, 21), we did not detect any dominant trend toward becoming similar to the yardstick PFS-26s for SCZ1, SCZ2, or DEP. This could be attributed to the inclusion of their genetic loci associated with other environmental risk factors. The detection power of our analysis would have been reduced by the limited number of genetic loci associated with sensitivity to psychosocial stress. Importantly, this finding indicates that the yardstick-like PFS-26s did not result from negative selection against schizophrenia and depression.

Simply meeting the criterion of $\rho > 0.9$ with the yardstick PFS-26s is insufficient to warrant the conclusion that certain PFS-26s are involved in adaptation to psychosocial stress. To elucidate such involvement in each yardstick-like PFS-26 of NEU1, NEU2, and EDU, detection of positive selection would be needed, which is beyond the scope of this study.

The negative-effect alleles of BMI showed a significant occurrence rate of the yardstick-like PFS-26s (Table 1). This raises the possibility that adaptation to psychosocial stress also influenced physical phenotypes.

Henrich (2) argued that once individuals have evolved so that they can learn from each other with sufficient accuracy, social groups develop what could be termed ‘collective brains’, which drive cumulative cultural evolution. It is probable that overcoming psychosocial stress in the spread of AMH (i.e., global adaptation to psychosocial stress)

expanded these collective brains by increasing the degree of social interaction and population size, thereby accelerating cumulative cultural evolution (4). We propose that the current genetic basis of neuroticism and educational attainment was established through such an evolutionary process (Fig. S20).

It appears that global adaptation to psychosocial stress was more significant in AMHs than in Neanderthals. This adaptation and its involvement in the enlargement of collective brains, leading to the evolution of collective brains, would have played a more important role in AMHs than in Neanderthals. This finding is consistent with larger population size and faster cultural evolution in AMHs than in Neanderthals (2, 18).

As noted above, the trend toward greater intelligence was plausibly more significant in Neanderthals than in AMHs. This trend and the greater significance of the collective brain in AMHs than in Neanderthals suggests that Neanderthals may have depended on functional improvements of individual brains rather than on the enlargement of collective brains for their survival. This could be related to larger average brain volume in Neanderthals (1,450 cc) than in AMHs (1,328 cc) (22). As Henrich (2) has argued, this difference could be related to the replacement of Neanderthals by AMHs.

In conclusion, the findings in this study suggest that mental traits (neuroticism, learning behaviors, and intelligence) associated with cultural development evolved differently in AMHs and Neanderthals, and that learning behaviors such as cooperative learning, rather than intelligence, played an important role in AMH evolution. We named ourselves *Homo sapiens* (wise man). The findings in this study may prompt a reconsideration of our self-recognition.

Materials and Methods

SNPs associated with phenotypes

Single-nucleotide polymorphisms (SNPs) associated with different phenotypes were obtained from genome-wide association studies (GWASs) reported to date (9, 10, 23–33) (see Table 1). SNPs identified according to genome-wide significance ($P < 5.0 \times 10^{-8}$) were

used in this study.

Psychosocial stress is an environmental risk factor for the onset of schizophrenia and depressive disorders (6, 7, 20, 21). Neuroticism is also an environmental risk factor for the onset of conditions such as schizophrenia and depression (23), and individuals demonstrating higher neuroticism levels are vulnerable to social stress. We previously pointed to the possible role of adaptation to psychosocial stress in learning via communication (3, 4). Therefore, we selected GWASs of neuroticism (NEU1 and NEU2) (23, 24), educational attainment (EDU) (25), intelligence (INT) (26), schizophrenia (SCZ1 and SCZ2) (9, 10), and depressive disorder (DEP) (27) for inclusion in this study. In addition to educational attainment, we chose chronotype as a behavioral phenotype, and included a GWAS that examined it (CHR) (28). We also focused on GWASs conducted on physiological phenotypes (29–33): body mass index (BMI), height, (HEI), vitamin D level (VID), type 2 diabetes (T2D), and migraine (MIG) to examine whether any physiological factors contributed to adaptation to psychosocial stress.

Because the GWASs included those that did not apply conditional and joint multiple-SNP (COJO) analysis (34), we used SNP sets identified without COJO analysis, with the exceptions of HEI for which only SNP sets identified by COJO analysis were available. Referring to the database of the 1000 Genomes Project (hereinafter abbreviated as 1KGP) (8), we eliminated SNPs for which allelic frequencies were unavailable as well as SNPs located on the X chromosome and indels. For SNPs located on the X chromosome, the difference in scores (see below) among populations would have been inaccurate if the sex ratios of the populations differed. For example, the sex ratio for FIN in the 1KGP differed significantly from the expected 1:1 ratio of males to females ($P = 0.02$). In the GWASs used in this study, the SNPs on the X chromosome were as follows: 3 in SCZ1, 5 in SCZ2, 57 in EDU, 3 in CHR, 1 in VID, and 2 in MIG. Thus, they constituted a very small proportion of the SNPs in each GWAS. In this study, we mainly used GWASs that identified close to or more than 100 associated genetic loci to maintain statistical power for our analyses (see Tables S3–5). However, we also included GWASs that identified considerably fewer than 100 associated genetic loci (for neuroticism, intelligence, schizophrenia and type 2 diabetes) (35–43) in the case of linkage analysis using concordant SNP pairs (see the main text).

DNA sequence data

The frequencies of each allele of the SNPs were retrieved from phase 3 of the 1KGP database (8), covering 2,504 individuals (5,008 sequences) from 26 different global populations. Referring to this database, we classified these populations into five meta-populations: Africa (AFR), Europe (EUR), South Asia (SAS), East and Southeast Asia (EAS), and the Americas (AMR).

DNA sequences of archaic humans, namely three Neanderthals (Vindija, Altai, and Chagyrskaya), were obtained from public databases: http://cdna.eva.mpg.de/neandertal/Vindija/; http://cdna.eva.mpg.de/neandertal/altai/; and http://ftp.eva.mpg.de/neandertal/Chagyrskaya/VCF/).

Population functional score

Three SNPs (rs3759916, rs3759915, and rs3759914) in the promoter region of the *ST8SIA2* gene, which encodes a sialyltransferase synthesizing a linear homopolymer of sialic acid (polysialic acid) in the brain (5), are associated with schizophrenia risk (44–46). Selective increases of non-risk haplotypes of these three promoter SNPs (TCT and CGC types) occurred during the spread of AMHs (3, 4). These two promoter types show lower promoter activities than risk types, such as TGT and CGT types (3), resulting in a trend toward lower population promoter activity (PPA) in the spread of AMHs (4). This trend in the *ST8SIA2* gene can be interpreted as an adaptation to psychosocial stress and has been established in populations worldwide through the spread of AMHs (4). Thus, an examination of local adaptation is not sufficient on its own to identify similarity with this global functional trend in SNPs associated with phenotypes. Furthermore, because PPAs are estimated using both frequencies and the promoter activities of promoter types, they cannot be estimated for GWAS SNPs, whose functions are usually unknown. By focusing on differences in allelic function, we established a score that enabled us to examine the similarity of any SNPs to the global trend of population function of the *ST8SIA2* gene, even if the allelic functions of those SNPs were unknown. SNPs identified from GWAS analysis by their association with phenotypes are biallelic. By setting D as the difference in allelic function (e.g., promoter

activity) in allele A1 from allele A2 at a biallelic SNP, and assuming additive allelic effects, the functional difference was expressed as 2*D*, *D*, and 0 for the A1/A1 homozygote, A1/A2 heterozygote, and A2/A2 homozygote, respectively. The functional difference at the population level (population function score: PFS) was obtained as follows:

$$PFS = \frac{2Dn_{11} + Dn_{12}}{2(n_{11} + n_{12} + n_{22})} = \frac{2n_{11} + n_{12}}{2(n_{11} + n_{12} + n_{22})} D,$$

where n_{11}, n_{12}, and n_{22} are respectively the numbers of A1/A1 homozygotes, A1/A2 heterozygotes, and A2/A2 homozygotes in the population. Therefore, the PFS could also be defined simply as:

$$PFS = f_{A1} D,$$

where f_{A1} is A1 allele frequency $(2n_{11} + n_{12})/2(n_{11} + n_{12} + n_{22})$.

The PFS of the *ST8SIA2* gene was obtained using the actual functional difference (i.e., promoter activity). The TGT type is the ancestral type and shows the highest promoter activity (3). Thus, the PFS of the *ST8SIA2* gene ($PFS_{ST8SIA2}$) was calculated as follows:

$$PFS_{ST8SIA2} = f_{CGT}D_{CGT} + f_{TCT}D_{TCT} + f_{CGC}D_{CGC},$$

where D_{CGT}, D_{TCT}, and D_{CGC} are differences in promoter activities from the TGT type in the CGT, TCT, and CGC types, and f_{CGT}, $f_{TCT,}$ and f_{CGC} are frequencies of the CGT, TCT, and CGC types, respectively.

Importantly, *D* of GWAS SNP was cancelled out in the estimation of the correlation coefficient (ρ) (i.e., similarity) between the distributions of PFSs from 26 populations (PFS-26; global trend of population function). Thus, correlation analysis using PFS-26 was performed to examine the similarity of the global trend of population function without requiring experimental estimation of allelic function (see Figs. S2 and S3).

In addition to eliminating the previously mentioned SNPs associated with phenotypes, we excluded those for which multiple derived alleles were indicated and those for which only derived alleles are found in AMHs. Table S3 shows numbers of SNPs used in the PFS analysis.

Correlation with yardstick PFS-26s

The observed rate of high correlation with certain PFS-26s used as a yardstick was obtained

for each allele type group (e.g., risk alleles of SCZ1). A binomial test was applied to examine whether this rate increased for certain allelic types compared with the opposite allelic types. The standard observed rate of PFS-26s showing a rate of high correlation with the yardstick PFS-26 is unknown in SNPs associated with phenotypes. We therefore compared the rate for negative-effect/non-risk alleles with that for positive-effect/risk alleles for each GWAS. Because populations show genetic kinship with one another, the assumption of independence among populations (see Fig. S3) does not hold when examining the correlation between PFS-26s. To enable strict assessment of the high correlation between PFS-26s, we needed a criterion that accounted for this genetic kinship (see the main text).

Identification of characteristic PFS-26s

Similar PFS-26s uniquely found in a certain allelic type (e.g., schizophrenia risk alleles) were identified in a pairwise comparison of PFS-26s (Fig. S7). For their identification, we focused on the number of PFS-26s ($N_{>\mathrm{MED}}$) showing ρ-values exceeding the median value of all pairwise comparisons of PFS-26s. A particular type, of which $N_{>\mathrm{MED}}$ is significantly larger in comparing with the same allelic types than in that with the opposite-effect types (α = 0.05/number of alleles; χ^2 test), is categorized as a characteristic PFS-26. In the χ^2 test, results were not considered if $N_{>\mathrm{MED}}$ was <10.

Detecting positive selection

We applied the F_c statistic to detect positive selection on SNPs in three meta-populations: AFR, EUR, and EAS (3, 47, 48). To detect all possible genetic loci, we modified a previously reported simulation method (3, 47) by including alleles showing an S-value (number of segregating sites) of at least 100. The Benjamini–Hochberg correction was performed for multiple testing. Significant discovery was then examined at a false discovery rate of < 0.1.

Timing estimation on positive selection

We estimated the operational duration of positive selection following Fujito et al (3).

Population polygenic score

Following the identification of SNPs associated with certain phenotypes in GWASs, the polygenic risk score (PRS)/polygenic score (PS) is generally applied to assess the degree of the phenotypic trend in individual (49). This score was obtained as follows:

$$PS = \sum_{i=1}^{n} X_i \beta_{i,}$$

where n is the total number of associated SNPs, X is the number of alleles at associated SNPs in individuals, and β is the effect size estimated in GWAS. Accordingly, the population polygenic score (PPS) was developed to assess the phenotypic trend at the population level as follows:

$$PPS = \sum_{i=1}^{n} F_i \, \beta_i,$$

where n is total number of associated SNPs, F is the frequency of alleles at associated SNPs in a certain population, and β is the effect size estimated in GWASs.

Estimation of PPSs in Neanderthals

Using Bayesian estimation based on allele frequencies in three Neanderthal individuals, we first obtained a possible distribution of the allele frequency at each genetic locus and subsequently applied this distribution to calculate the PPS. Lastly, we estimated the Neanderthal PPS in each GWAS as a distribution from 10,000 replications of the calculation. In addition to eliminating previously mentioned SNPs associated with phenotypes, we excluded SNPs not available in any Neanderthal individuals. The numbers of SNPs used in the PPS estimations are shown in Table S4.

Classification of SNPs in Neanderthals

Homozygous SNP sites (i.e., positive-effect allele/positive-effect allele or negative-effect allele/negative-effect allele) in both Altai and Vindija were first selected as targets, and heterozygosity was calculated for the region available in both individuals by focusing on a 200-kb region (100-kb upstream and downstream) surrounding the target SNPs in AMHs. Target SNPs showing a 0 value for heterozygosity and those for which a 200-kb region was unavailable in AMHs were eliminated. As a final step, using the calculated heterozygosity

value, we categorized target SNPs into three groups: Vindija<Altai, Vindija≈Altai, and Vindija>Altai according to the binomial distribution of heterozygosity for each SNP in Altai (see Figs. S13–S15). The Vindija<Altai group and the Vindija>Altai group contained genetic loci for which heterozygosity for Vindija was lower and higher than a 99% confidence interval (CI) of binomial distribution of heterozygosity for Altai, respectively. Genetic loci showing heterozygosity for Vindija within a 99% CI of heterozygosity for Altai were categorized in the Vindija≈Altai group. Table S5 shows the final number of SNPs used in the analysis.

Data availability

The genomes used are available from the 1000 Genomes Project (phase 3 release, https://www.internationalgenome.org/).

Acknowledgements

This research was supported by the Japanese Society for the Promotion of Science (grant nos. JP16K07535, JP19K06866, and JP24K09628 to T.H.). We thank Edanz (https://jp.edanz.com/ac) for editing a draft of this manuscript.

Author contributions

T.H. conceived and supervised the study. Y.M., N.T.F., and T.H. collected sequence data. N.T.F., M.T., N.T., Y.S., and T.H. performed analysis detecting signals of positive selection. T.H. and Y.M. performed correlation analyses. Y.M. and T.H. performed analyses of phenotypic trends. Y.M., N.T.F., N.T., Y.S., and K.M.T. read and commented on the pre-final version of the manuscript. T.H. wrote the paper.

Competing interest declaration

The authors declare no competing interest.

References

1. Nielsen, R., Akey, J.M., Jakobsson, M., Pritchard, J.K., Tishkoff, S., Willerslev, E. Tracing the peopling of the world through genomics. Nature. 2017; 541(7637):302-310. https://doi.org/10.1038/nature21347 PMID: 28102248.
2. Henrich, J. The secret of our success: how culture is driving human evolution, domesticating our species, and making us smarter. Princeton, NJ: Princeton University Press; 2016.
3. Fujito, N.T., Satta, Y., Hane, M., Matsui, A., Yashima, K., Kitajima, K., Sato, C., Takahata, N., Hayakawa, T. Positive selection on schizophrenia-associated ST8SIA2 gene in post-glacial Asia. PLoS One. 2018; 13(7):e0200278. https://doi.org/10.1371/journal.pone.0200278 PMID: 30044798.
4. Hayakawa, T., Terahara, M., Fujito, N.T., Matsunaga, T., Teshima, K.M., Hane, M., Kitajima, K., Sato, C., Takahata, N., Satta, Y. Lower promoter activity of the ST8SIA2 gene has been favored in evolving human collective brains. PLoS ONE. 2021; 16(12): e0259897. https://doi.org/10.1371/journal.pone.0259897 PMID: 34914745.
5. Sato, C., & Hane, M. Mental disorders and an acidic glycan-from the perspective of polysialic acid (PSA/polySia) and the synthesizing enzyme, ST8SIA2. Glycoconj J. 2018; 35(4): 353-373. https://doi.org/10.1007/s10719-018-9832-9 PMID: 30058042.
6. van Os, J., Kenis, G., Rutten, B.P. The environment and schizophrenia. Nature. 2010; 468(7321):203-212. https://doi.org/10.1038/nature09563 PMID: 21068828.
7. Howes, O.D., McCutcheon, R., Owen, M.J., Murray, R.M. The Role of Genes, Stress, and Dopamine in the Development of Schizophrenia. Biol Psychiatry. 2017; 81(1):9-20. https://doi.org/10.1016/j.biopsych.2016.07.014 PMID: 27720198.
8. The 1000 Genomes Project Consortium. A global reference for human genetic variation.

Nature. 2015; 526(7571): 68-74. https://doi.org/10.1038/nature15393. PMID: 26432245.

9. Schizophrenia Working Group of the Psychiatric Genomics Consortium. Biological insights from 108 schizophrenia-associated genetic loci. Nature. 2014; 511(7510): 421-7. https://doi.org/10.1038/nature13595 PMID: 25056061.
10. Trubetskoy, V., Pardiñas, A.F., Qi, T., Panagiotaropoulou, G., Awasthi, S., Bigdeli, T.B., Bryois, J., Chen, C.Y., Dennison, C.A., Hall, L.S., Lam, M., et al. Mapping genomic loci implicates genes and synaptic biology in schizophrenia. Nature. 2022; 604(7906): 502-508. https://doi.org/10.1038/s41586-022-04434-5. PMID: 35396580.
11. The 1000 Genomes Project Consortium. An integrated map of genetic variation from 1,092 human genomes. Nature. 2012; 491(7422): 56-65. https://doi,org/10.1038/nature11632. PMID: 23128226.
12. Tishkoff, S.A., Reed, F.A., Ranciaro, A., Voight, B.F., Babbitt, C.C., Silverman, J.S., Powell, K., Mortensen, H.M., Hirbo, J.B., Osman, M., Ibrahim, M., et al. Convergent adaptation of human lactase persistence in Africa and Europe. Nat Genet. 2007; 39(1): 31-40. https://doi.org/10.1038/ng1946. PMID: 17159977.
13. Fujimoto, A., Kimura, R., Ohashi, J., Omi, K., Yuliwulandari, R., Batubara, L., Mustofa, M.S., Samakkarn, U., Settheetham-Ishida, W., Ishida, T., Morishita, Y., et al. A scan for genetic determinants of human hair morphology: EDAR is associated with Asian hair thickness. Hum Mol Genet. 2008; 17(6): 835-843. https://doi.org/10.1093/hmg/ddm355. PMID: 18065779.
14. Sturm., R.A., and Duffy, D.L. Human pigmentation genes under environmental selection. Genome Biol. 2012; 13(9): 248. https://doi.org/10.1186/gb-2012-13-9-248. PMID: 23110848.
15. Hill, W.D., Marioni, R.E., Maghzian, O., Ritchie, S.J., Hagenaars, S.P., McIntosh, A.M., Gale, C.R., Davies, G., Deary, I.J. A combined analysis of genetically correlated traits identified 187 loci and a role for neurogenesis and myelination in intelligence. Mol Psychiatry. 2019; 24(2): 169-181. https://doi.org/10.1038/s41380-017-0001-5. PMID: 29326435.
16. Goldberg, L.R. The structure of phenotypic personality traits. Am Psychol. 1993;48(1): 26-34. https://doi.org/10.1037//0003-066x.48.1.26. PMID: 8427480.

17. Pomeroy, E. Review: The different adaptive trajectories in Neanderthals and Homo sapiens and their implications for contemporary human physiological variation. Comp Biochem Physiol A Mol Integr Physiol. 2023; 280: 111420. https://doi.org/10.1016/J.CBPA.2023.111420. PMID: 37001690.
18. Mafessoni, F., Grote, S., de Filippo, C., Slon, V., Kolobova, K.A., Viola, B., Markin, S.V., Chintalapati, M., Peyrégne, S., Skov, L., Skoglund, P., Krivoshapkin, A.I., Derevianko, A.P., Meyer, M., Kelso, J., Peter, B., Prüfer, K., Pääbo, S. A high-coverage Neandertal genome from Chagyrskaya Cave. Proc Natl Acad Sci U S A. 2020; 117(26): 15132–15136. https://doi.org/10.1073/pnas.2004944117. PMID: 32546518.
19. Prüfer, K., de Filippo, C., Grote, S., Mafessoni, F., Korlević, P., Hajdinjak, M., Vernot, B., Skov, L., Hsieh, P., Peyrégne, S., Reher, D., Hopfe, C., Nagel, S., Maricic, T., Fu, Q., Theunert, C., Rogers, R., Skoglund, P., Chintalapati, M., Dannemann, M., Nelson, B.J., Key, F.M., Rudan, P., Kućan, Ž., Gušić, I., Golovanova, L.V., Doronichev, V.B., Patterson, N., Reich, D., Eichler, E.E., Slatkin, M., Schierup, M.H., Andrés, A.M., Kelso, J., Meyer, M., Pääbo, S. A high-coverage Neandertal genome from Vindija Cave in Croatia. Science. 2017; 358(6363): 655–658. https://doi.org/10.1126/science.aao1887. PMID: 28982794.
20. Rodriguez, V., Alameda, L., Trotta, G., Spinazzola, E., Marino, P., Matheson, S.L., Laurens, K.R., Murray, R.M., Vassos, E. Environmental risk factors in bipolar disorder and psychotic depression: a systematic review and meta-analysis of prospective studies. Schizophr Bull. 2021; 47(4): 959-974. https://doi.org/10.1093/schbul/sbaa197. PMID: 33479726.
21. Slavich, G.M., & Irwin, M.R. From stress to inflammation and major depressive disorder: a social signal transduction theory of depression. Psychol Bull. 2014; 140(3): 774-815. https://doi.org/10.1037/a0035302. PMID: 24417575.
22. Neubauer, S., Hublin, J.J., Gunz, P. The evolution of modern human brain shape. Sci Adv. 2018; 4(1): eaao5961. https://www.science.org/doi/10.1126/sciadv.aao5961. PMID: 29376123.
23. Nagel, M., Jansen, P.R., Stringer, S., Watanabe, K., de Leeuw, C.A., Bryois, J., Savage, J.E., Hammerschlag, A.R., Skene, N.G., Nuñoz-Manchado, A.B.; 23andMe Research Team; White, T., et al. Meta-analysis of genome-wide association studies for neuroticism

in 449,484 individuals identifies novel genetic loci and pathways. Nat Genet. 2018; 50(7): 920-927. https://doi.org/10.1038/s41588-018-0151-7. PMID: 29942085.

24. Luciano, M., Hagenaars, S.P., Davies, G., Hill, W.D., Clarke, T.K., Shirali, M., Harris, S.E., Marioni, R.E., Liewald, D.C., Fawns-Ritchie, C. Adams, M.J., et al. Association analysis in over 329,000 individuals identifies 116 independent variants influencing neuroticism. Nat Genet. 2018; 50(1): 6-11. https://doi.org/10.1038/s41588-017-0013-8. PMID: 29255261.
25. Okbay, A., Wu, Y., Wang, N., Jayashankar, H., Bennett, M., Nehzati, S.M., Sidorenko, J., Kweon, H., Goldman, G., Gjorgjieva, T., Jiang, Y., et al. Polygenic prediction of educational attainment within and between families from genome-wide association analyses in 3 million individuals. Nat Genet. 2022; 54(4): 437-449. https://doi.org/10.1038/s41588-022-01016-z. PMID: 35361970.
26. Savage, J.E., Jansen, P.R., Stringer, S., Watanabe, K., Bryois, J., de Leeuw, C.A., Nagel, M., Awasthi, S., Barr, P.B., Coleman, J.R.I., Grasby, K.L., et al. Genome-wide association meta-analysis in 269,867 individuals identifies new genetic and functional links to intelligence. Nat Genet. 2018; 50(7): 912-919. https://doi.org/10.1038/s41588-018-0152-6. PMID: 29942086.
27. Als, T.D., Kurki, M.I., Grove, J., Voloudakis, G., Therrien, K., Tasanko, E., Nielsen, T.T., Naamanka, J., Veerapen, K., Levey, D.F., Bendl, J., et al. Depression pathophysiology, risk prediction of recurrence and comorbid psychiatric disorders using genome-wide analyses. Nat Med. 2023; 29(7): 1832-1844. https://doi.org/10.1038/s41591-023-02352-1. PMID: 37464041.
28. Jones, S.E., Lane, J.M., Wood, A.R., van Hees, V.T., Tyrrell, J., Beaumont, R.N., Jeffries, A.R., Dashti, H.S., Hillsdon, M., Ruth, K.S., Tuke, M.A., et al. Genome-wide association analyses of chronotype in 697,828 individuals provides insights into circadian rhythms. Nat Commun. 2019; 10(1): 343. https://doi.org/10.1038/s41467-018-08259-7. PMID: 30696823.
29. Huang, J., Huffman, J., Huang, Y., Do Valle, Í., Assimes, T., Raghavan, S., Voight, B., Liu, C., Barabási, A., Huang, R., Hui, Q., Nguyen, X., Ho, Y., Djousse, L., Lynch, J., Vujkovic, M., Tcheandjieu, C., Tang, H., Damrauer, S., Reaven, P., Miller, D., Phillips, L.,

Ng, M., Graff, M., Haiman, C., Loos, R., North, K., Yengo, L., Smith, G., Saleheen, D., Gaziano, J., Rader, D., Tsao, P., Cho, K., Chang, K., Wilson, P., Sun, Y., O'Donnell, C. Genomics and phenomics of body mass index reveals a complex disease network. Nat. Commun. 2022; 13(1): 1–10. https://doi.org/10.1038/s41467-022-35553-2. PMID: 36581621.

30. Yengo, L., Vedantam, S., Marouli, E., Sidorenko, J., Bartell, E., Sakaue, S., Graff, M., Eliasen, A.U., Jiang, Y., Raghavan, S., Miao, J., et al. A saturated map of common genetic variants associated with human height. Nature. 2022; 610(7933): 704-712. https://doi.org/10.1038/s41586-022-05275-y. PMID: 36224396.
31. Revez, J.A., Lin, T., Qiano, Z., Xue, A., Holtz, Y., Zhu, Z., Zeng, J., Wang, H., Sidorenko, J., Kemper, K.E., Vinkhuyzen, A.A.E., et al. Genome-wide association study identifies 143 loci associated with 25 hydroxyvitamin D concentration. Nat Commun. 2020; 11(1): 1647. https://doi.org/10.1038/s41467-020-15421-7. PMID: 32242144.
32. Xue, A., Wu, Y., Zhu, Z., Zhang, F., Kemper, K.E., Zheng, Z., Yengo, L., Lloyd-Jones, L.R., Sidorenko, J., Wu. Y.; eQTLGen Consortium; McRae, A.F., et al. Genome-wide association analyses identify 143 risk variants and putative regulatory mechanisms for type 2 diabetes. Nat Commun. 2018; 9(1): 2941. https://doi.org/10.1038/s41467-018-04951-w. PMID: 30054458.
33. Hautakangas, H., Winsvold, B.S., Ruotsalainen, S.E., Bjornsdottir, G., Harder, A.V.E., Kogelman, L.J.A., Thomas, L.F., Noordam, R., Benner, C., Gormley, P., Artto, V., et al. Genome-wide analysis of 102,084 migraine cases identified 123 risk loci and subtype-specific risk alleles. Nat Genet. 2022; 54(2): 152-160. https://doi.org/10.1038/s41588-021-00990-0. PMID: 35115687.
34. Yang, J., Ferreira, T., Morris, A.P., Medland, S.E., Genetic Investigation of ANthropometric Traits (GIANT) Consortium, DIAbetes Genetics Replication And Meta-analysis (DIAGRAM) Consortium, Madden, P.A.F., Heath, A.C., Martin, N.G., Montgomery, G.W., Weedon, M.N., et al. Conditional and joint multiple-SNP analysis of GWAS summary statistics identifies additional variants influencing complex traits. Nat Genet. 2012; 44(4): 369-375. https://doi.org/10.1038/ng2213. PMID: 22426310.
35. Genetics of Personality Consortium; de Moor, M.H., et al. Meta-analysis of genome-wide

association studies for neuroticism, and the polygenic association with major depressive disorder. JAMA Psychiatry. 2015; 72(7): 642-650. https://doi.org/10.1001/jamapsychiatry.2015.0554. PMID: 25993607.

36. Okbay, A., Baselmans, B.M., De Neve, J.E., Turley, P., Nivard, M.G., Fontana, M.A., Meddens, S.F., Linnér, R.K., Rietveld, C.A., Derringer, J., Gratten, J., et al. Genetic variants associated with subjective well-being, depressive symptoms, and neuroticism identified through genome-wide analyses. Nat Genet. 2016; 48(6): 624-633. https://doi.org/10.1038/ng.3552. PMID: 27089181.
37. Smith, D.J., Escott-Price, V., Davies, G., Bailey, M.E., Colodro-Conde, L., Ward, J., Vedernikov, A., Marioni, R., Cullen, B. Lyall, D., et al. Genome-wide analysis of over 106000 individuals identifies 9 neuroticism-associated loci. Mol Psychiatry. 2016; 21(6): 749-757. https://doi.org/10.1038/mp.2016.49. PMID: 27067015.
38. Lo, M.T., Hinds, D.A., Tung, J.Y., Franz, C., Fan, C.C., Wang, Y., Smeland, O.B., Schork, A., Holland, D., Kauppi, K., Sanyal, N., et al. Genome-wide analyses for personality traits identify six genomic loci and show correlation with psychiatric disorders. Nat Genet. 2017; 49(1): 152-156. https://doi,org/10.1038/ng.3736. PMID: 27918536.
39. Sniekers, S., Stringer, S., Watanabe, K., Jansen, P.R., Coleman, J.R.I., Krapohl, E., Taskesen, E., Hammerschlag, A.R., Okbay, A, Zabaneh, D., Amin, N., et al. Genome-wide association meta-analysis of 78,308 individuals identifies new loci and genes influencing human intelligence. Nat Genet. 2017; 49(7): 1107-1112. https://doi.org/10.1038/ng.3869. PMID: 28530673.
40. Li, Z., Chen, J., Yu, H., He, L., Xu, Y., Zhang, D., Yi, Q., Li, C., Li, X., Shen, J., Song, Z., et al. Genome-wide association analysis identifies 30 new susceptibility loci for schizophrenia. Nat Genet. 2017; 49(11): 1576-1583. https://doi.org/10.1038/ng.3973. PMID: 28991256.
41. Lam, M., Chen, C.Y., Li, Z., Martin, A.R., Bryois, J., Ma, X., Gaspar, H., Ikeda, M., Benyamin, B., Brown, B.C., Liu, R., et al. Comparative genetic architectures of schizophrenia in East Asian and European populations. Nat Genet. 2019; 51(12): 1670-1678. https://doi.org/10.1038/s41588-019-0512-x. PMID: 31740837.
42. Yengo, L., Sidorenko, J., Kemper, K.E., Zheng, Z., Wood, A.R., Weedon, M.N., Frayling,

T.M., Hirschhorn, J., Yang, J., Visscher, P.M., GIANT Consortium. Meta-analysis of genome-wide association studies for height and body mass index in ~700000 individuals of European ancestry. Hum Mol Genet. 2018; 27(20): 3641-3649. https://doi.org/10.1093/hmg/ddy271. PMID: 30124842.

43. Kooner, J.S., Saleheen, D., Sim, X., Sehmi, J., Zhang, W., Frossard, P., Been, L.F., Chia, K.S., Dimas, A.S., Hassanali, N., Jafar, T., et al. Genome-wide association study in individuals of South Asian ancestry identifies six new type 2 diabetes susceptibility loci. Nat Genet. 2011; 43(10): 984-989. https://doi,org/10.1038/ng.921. PMID: 21874001.
44. Arai, M., Yamada, K., Toyota, T., Obata, N., Haga, S., Yoshida, Y., Nakamura, K., Minabe, Y., Ujike, H., Sora, I. et al. Association between polymorphisms in the promoter region of the sialyltransferase 8B (SIAT8B) gene and schizophrenia. Biol Psychiatry. 2006; 59(7):652-659. https://doi.org/10.1016/j.biopsych.2005.08.016 PMID: 16229822.
45. Tao, R., Li, C., Zheng, Y., Qin, W., Zhang, J., Li, X., Xu, Y., Shi, Y.Y., Feng, G., He, L. Positive association between SIAT8B and schizophrenia in the Chinese Han population. Schizophr Res. 2007; 90(1-3):108-114. https://doi.org/10.1016/j.schres.2006.09.029 PMID: 17126533.
46. Gilabert-Juan, J., Nacher, J., Sanjuán, J., Molto, M.D. Sex-specific association of the ST8SIAII gene with schizophrenia in a Spanish population. Psychiatry Res. 2013; 210(3):1293-1295. https://doi.org/10.1016/j.psychres.2013.09.001. PMID: 24070986.
47. Fujito, N.T., Satta, Y., Hayakawa, T., Takahata, N. A new inference method for detecting an ongoing selective sweep. Genes Genet Syst. 2018; 93(4): 149-161. https://doi.org/10.1266/ggs.18-00008. PMID: 30270233.
48. Satta, Y., Zheng, W., Nishiyama, K.V., Iwasaki, R.L., Hayakawa, T., Fujito, N.T., Takahata, N. Two-dimensional site frequency spectrum for detecting, classifying and dating incomplete selective sweeps. Genes Genet Syst. 2020; 94(6): 283-300. https://doi.org/10.1266/ggs.19-00012. PMID: 31827009.
49. Lambert, S., Abraham, G., Inouye, M. Towards clinical utility of polygenic risk scores, Hum Mol Genet. 2019; 28(R2): R133–R142. https://doi.org/10.1093/hmg/ddz187. PMID: 31363735.

Figures

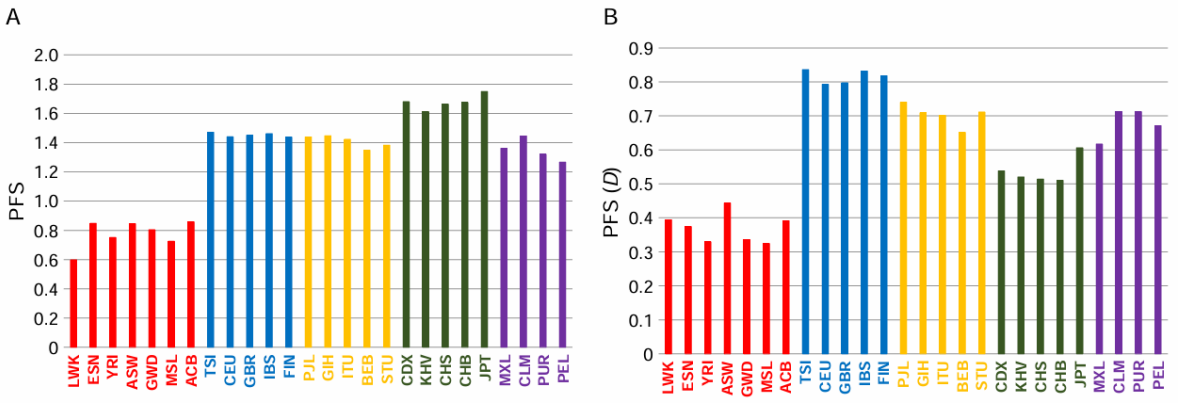

Fig. 1. (A) PFSs of the promoter types of the *ST8SIA2* gene. (B) PFSs of the schizophrenia non-risk allele of rs1501357. Populations are represented by abbreviations used in the 1000 Genomes Project (8).

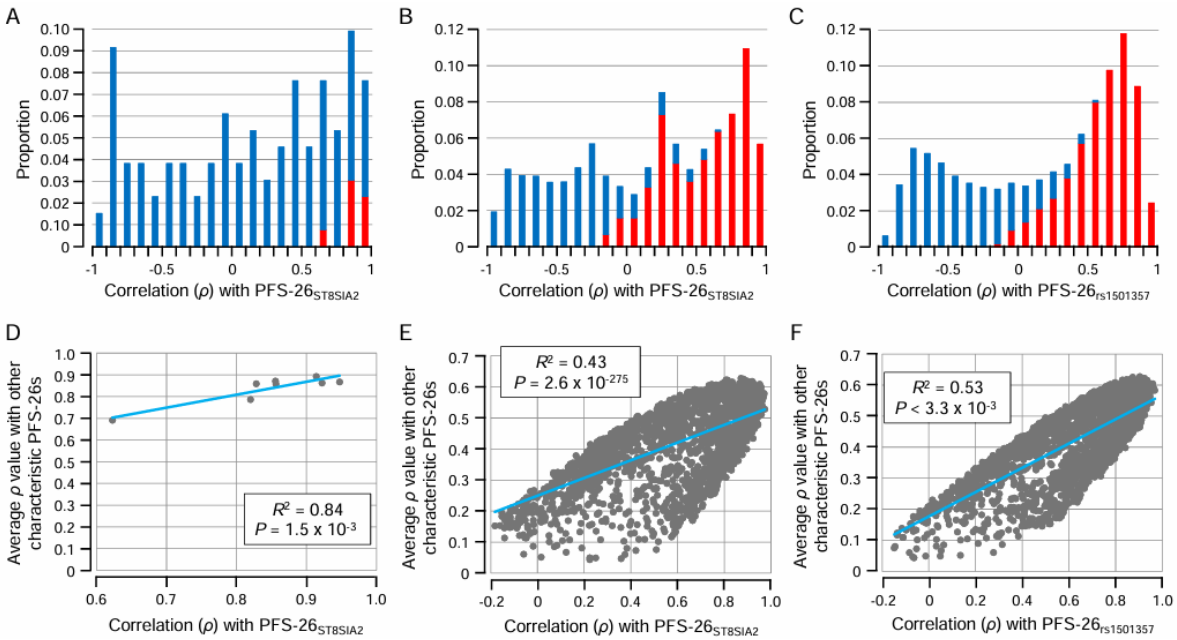

Fig. 2. Involvement of the yardstick PFS-26s in the evolution of alleles associated with neuroticism and educational attainment. (A) Distribution of correlation coefficients with PFS-26$_{ST8SIA2}$ in PFS-26s of negative-effect alleles of NEU1. (B) Distribution of correlation coefficients with PFS-26$_{ST8SIA2}$ in PFS-26s of positive-effect alleles of EDU. (C) Distribution of correlation coefficients with PFS-26$_{rs1501357}$ in PFS-26s of positive-effect alleles of EDU. The red part of each bar corresponds to the characteristic PFS-26s. (D) Correlation trend with PFS-26$_{ST8SIA2}$ in the characteristic PFS-26s of negative-effect alleles of NEU1. (E) Correlation trend with PFS-26$_{ST8SIA2}$ in the characteristic PFS-26s of positive-effect alleles of EDU. (F) Correlation trend with PFS-26$_{rs1501357}$ in the characteristic PFS-26s of positive-effect alleles of EDU.

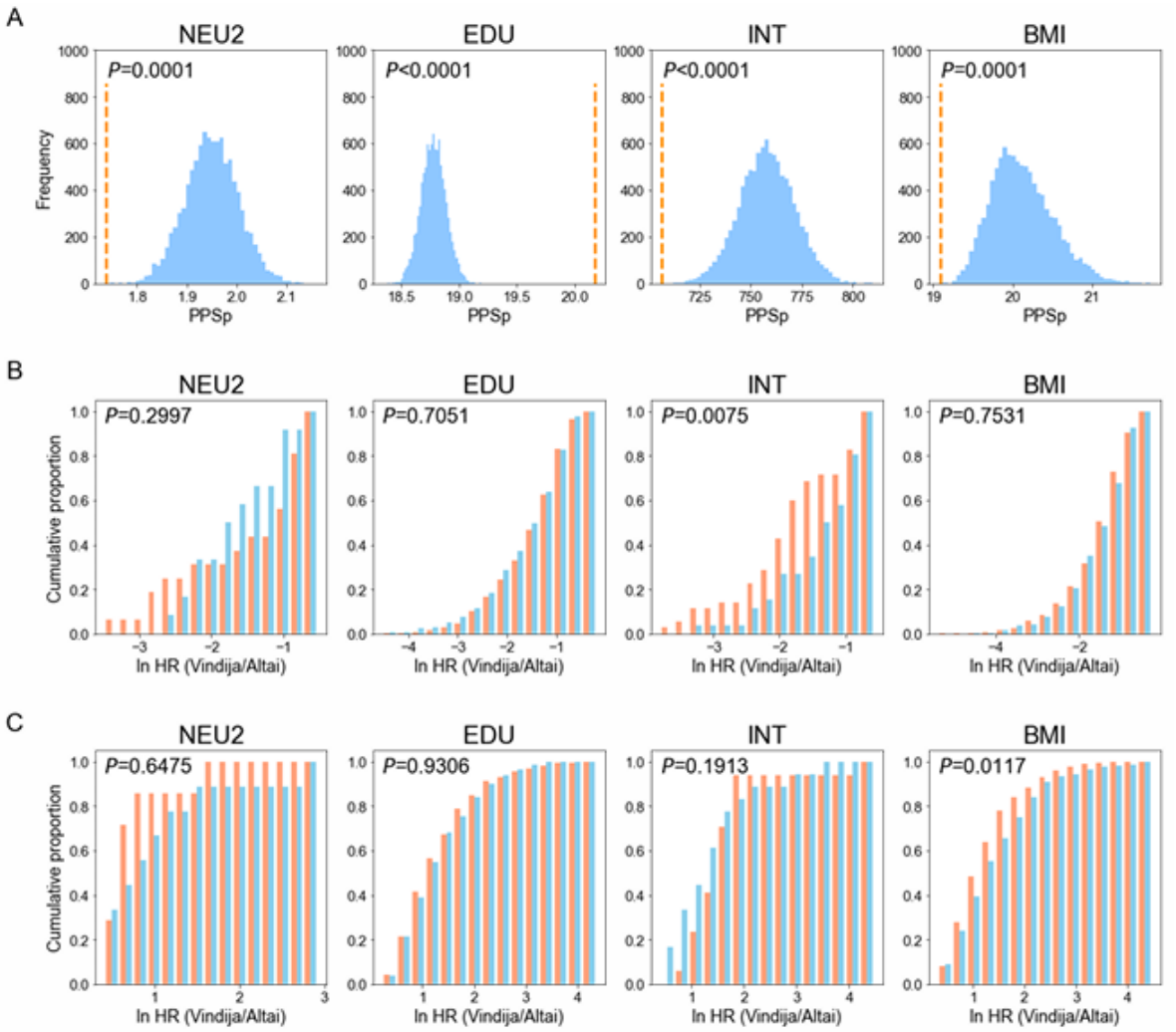

Fig. 3. Phenotypic trends in Neanderthals. (A) A comparison of PPSp in AMHs and Neanderthals. The vertical dotted line shows a PPSp of AMHs. PPSp of Neanderthals is represented as its possible distribution. (B) Distribution of the heterozygosity ratio (HR) in the Vindija<Altai group. The red and blue bars represent positive-effect alleles and negative-effect alleles, respectively. (C) The distribution of HR in the Vindija>Altai group. Red and blue bars respectively represent positive-effect alleles and negative-effect alleles.

Table

Table 1. Observed rate of yardstick-like PFS-26s.

	Phenotype	GWAS	$PFS\text{-}26_{ST8SIA2}$		$PFS\text{-}26_{rs1501357}$ (HCN1)	
			Allele		Allele	
			Positive-effect	Negative-effect	Positive-effect	Negative-effect
Mental	Neuroticism	NEU1	0.015 (2/131)	0.076 (10/131)	0 (0/131)	0.038 (5/131)
		NEU2	0.009 (1/109)	0.064 (7/109)	0.018 (2/109)	0.018 (2/109)
	Educational attainment	EDU	0.057 (218/3845)	0.019 (74/3845)	0.024 (94/3845)	0.006 (24/3845)
	Intelligence	INT	0.030 (7/231)	0.056 (13/231)	0.022 (5/231)	0.026 (6/231)
	Chronotype	CHR	0.039 (13/332)	0.021 (7/332)	0.039 (13/332)	0.027 (9/332)
			Risk	Non-risk	Risk	Non-risk
	Schizophrenia	SCZ1	0.038 (4/106)	0.085 (9/106)	0 (0/105)	0.038 (4/105)
		SCZ2	0.064 (17/267)	0.064 (17/267)	0.026 (7/267)	0.049 (13/267)
	Depression	DEP	0.028 (8/288)	0.052 (15/288)	0.014 (4/288)	0.049 (14/288)
Physical			Positive-effect	Negative-effect	Positive-effect	Negative-effect
	Body mass index (BMI)	BMI	0.030 (71/2386)	0.048 (114/2386)	0.017 (41/2386)	0.020 (48/2386)
	Height	HEI	0.056 (670/11872)	0.054 (640/11872)	0.025 (301/11872)	0.022 (260/11872)
	Vitamine D	VID	0.026 (2/78)	0.051 (4/78)	0.038 (3/78)	0.013 (1/78)
			Risk	Non-risk	Risk	Non-risk
	Type 2 Diabetes	T2D	0.037 (5/134)	0.037 (5/134)	0.022 (3/134)	0.030 (4/134)
	Migraine	MIG	0.052 (6/116)	0.043 (5/116)	0.026 (3/116)	0.026 (3/116)