Archival Report

Biological Psychiatry

Genetic Markers of Human Evolution Are Enriched in Schizophrenia

Saurabh Srinivasan, Francesco Bettella, Morten Mattingsdal, Yunpeng Wang, Aree Witoelar, Andrew J. Schork, Wesley K. Thompson, Verena Zuber, The Schizophrenia Working Group of the Psychiatric Genomics Consortium, The International Headache Genetics Consortium, Bendik S. Winsvold, John-Anker Zwart, David A. Collier, Rahul S. Desikan, Ingrid Melle, Thomas Werge, Anders M. Dale, Srdjan Djurovic, and Ole A. Andreassen

ABSTRACT

BACKGROUND: Why schizophrenia has accompanied humans throughout our history despite its negative effect on fitness remains an evolutionary enigma. It is proposed that schizophrenia is a by-product of the complex evolution of the human brain and a compromise for humans' language, creative thinking, and cognitive abilities.

METHODS: We analyzed recent large genome-wide association studies of schizophrenia and a range of other human phenotypes (anthropometric measures, cardiovascular disease risk factors, immune-mediated diseases) using a statistical framework that draws on polygenic architecture and ancillary information on genetic variants. We used information from the evolutionary proxy measure called the Neanderthal selective sweep (NSS) score.

RESULTS: Gene loci associated with schizophrenia are significantly ($p = 7.30 \times 10^{-9}$) more prevalent in genomic regions that are likely to have undergone recent positive selection in humans (i.e., with a low NSS score). Variants in brain-related genes with a low NSS score confer significantly higher susceptibility than variants in other brain-related genes. The enrichment is strongest for schizophrenia, but we cannot rule out enrichment for other phenotypes. The false discovery rate conditional on the evolutionary proxy points to 27 candidate schizophrenia susceptibility loci, 12 of which are associated with schizophrenia and other psychiatric disorders or linked to brain development.

CONCLUSIONS: Our results suggest that there is a polygenic overlap between schizophrenia and NSS score, a marker of human evolution, which is in line with the hypothesis that the persistence of schizophrenia is related to the evolutionary process of becoming human.

Keywords: Evolution, GWAS, Human, Neanderthal, Polygenic, Schizophrenia

http://dx.doi.org/10.1016/j.biopsych.2015.10.009

Schizophrenia affects approximately 1% of the world's population and has accompanied humans through much of our recorded history (1–6). This seemingly human-specific disorder is characterized by hallucinations and delusions (often involving language), thought disorders, and higher order cognitive dysfunctions. The mechanisms of schizophrenia are not well understood, but its heritability is high, between 60% and 80% (7), and the fecundity of affected people is reduced (8). Nevertheless, the prevalence of the disease seems to remain stable across generations, giving rise to the yet unresolved "evolutionary enigma" of schizophrenia (3,4,9,10). Large variations in incidence across populations argue for environmental causes. However, by using standard, precisely drawn diagnostic criteria, the variation in incidence can be reduced (11). Classic explanations include a single, partially dominant gene with low penetrance giving slight physiologic advantages (12); balanced selection, where the gene variants conferring risk of the disease provide an advantage in particular environments; and hitchhiking, where disease variants are passed along with advantageous neighboring gene variants. Newer studies have focused on the polygenic nature of schizophrenia and have attributed the prevalence of the disease to the sporadic nature of complex disorders (13).

Archaeological and paleontologic evidence points to the appearance of various hominid forms such as *Homo habilis*, *Homo erectus*, *Homo neanderthalensis* (Neanderthals), and modern *Homo sapiens* (humans) over 2.5 million years from the Lower Paleolithic Age to the Neolithic Age. It is debated whether the emergence of the "modern human" was a morphologic or a behavioral process, a one-time event or a continuous process of adaptation and assimilation of different forms. Even as morphologic changes stopped, behavioral changes continued, rapidly leading to the ultimate success of humans (14).

Over the Pleistocene period, we see the appearance of specialized tools, the introduction of decorative arts, burial

 ISSN: 0006-3223

practices (15), and possibly the development of language (16). Research suggests that language acquisition played an important role in shaping the brain, helping humans to think abstractly and be more creative, but it also made humans vulnerable to psychiatric disorders such as schizophrenia (17). Changes that contributed to the ability to think more creatively and to improve executive function (18) could have also harbored susceptibility to this pathology (19). However, although archaeological evidence provides clues about other aspects of human evolution, it cannot offer insights into the origin of psychiatric disorders.

More recent developments in human genetics have provided unprecedented opportunities to investigate evolutionary aspects of schizophrenia. Genome-wide association studies (GWASs) have identified >100 schizophrenia risk loci and highlighted the polygenic architecture of the disease (20). The genome sequence of Neanderthals (21,22), close relatives of early modern humans, can help pinpoint the genomic regions affected by positive selection since the two species diverged. The genomic differences between the two *Homo* species may help explain specific human features and the relationship between human evolution and schizophrenia.

Several lines of evidence indicate that schizophrenia is a polygenic disorder (23,24) with a large number of risk loci, each with a small effect (20). We have recently developed statistical tools, building on an empirical Bayesian framework (25), that are specifically designed for polygenic architectures. These tools have been successfully applied to investigate several complex human phenotypes (26–32) but have not yet been used to study the evolutionary features thereof. We hypothesized that schizophrenia is the result of human polygenic adaptation (24) and investigated whether regions of the human genome, which may have undergone recent positive selection, are enriched of association with schizophrenia.

METHODS AND MATERIALS

Samples

We obtained summary statistics for ~1.0–2.5 million single nucleotide polymorphisms (SNPs) from GWASs of schizophrenia (conducted by the Psychiatric Genomics Consortium) and other phenotypes, including anthropometric measures (body mass index, height, waist-to-hip ratio), cardiovascular disease risk factors (systolic blood pressure, total cholesterol, triglycerides), immune-mediated diseases (celiac disease, Crohn's disease, rheumatoid arthritis, ulcerative colitis), and other psychiatric and central nervous system disorders (attention-deficit/hyperactivity disorder, Alzheimer's disease, bipolar disorder, and multiple sclerosis) (Supplemental Table S1). These studies included ~1.3 million phenotypic observations, although overlap between samples makes the number of unique subjects lower.

Neanderthal Selective Sweep Score

The Neanderthal selective sweep (NSS) score is obtained through alignment of human, Neanderthal, and primate consensus sequences (21,33) and is downloadable from the UCSC Genome Browser website (34) (http://genome.ucsc.edu; ntSssZScorePMVar track [S-scores]), developed and maintained by the University of California, Santa Cruz. This track consists of two entries per SNP (*z*-score + SD) and (*z*-score − SD). The NSS score provides a likelihood index of positive selection in humans sometime after the divergence of humans and Neanderthals (21,33) by measuring the relative abundance of ancestral/nonancestral (i.e., aligned/nonaligned with primate consensus) alleles in these two lineages. A negative NSS score indicates scarcity of nonancestral alleles in Neanderthals compared with modern humans and therefore possible positive selection in the latter. The (*z*-score + SD) entries in the genome track represent an upper bound on the statistic and are therefore conservative measures of positive selection likelihood. These were extracted for all SNPs in the GWASs of interest (Supplemental Table S1) and follow the distribution illustrated in Supplemental Figure S1. The (*z*-score + SD) entries, termed NSS scores, were used as ancillary information or covariates in the enrichment analyses. Using the NSS scores, the authors of the two articles on the Neanderthal genome identified regions of the human genome that are significantly likely to have undergone recent positive selection. The same analyses performed directly using the NSS scores were also performed using linkage disequilibrium (LD) weighted scores (see Analytical Approach) measuring affiliation to these regions.

Brain Genes

To control the enrichment analyses for affiliation to brain genes, we identified genes with a known function in the brain using information from the National Center for Biotechnology Information (http://www.ncbi.nlm.nih.gov/gene). The query "human brain" in *Homo sapiens* revealed 2494 genes (March 2015). For comparison, we also used the list of brain genes from Kang *et al*. (35), which includes 1415 genes selected based on expression in various neural cells. The LD weighted procedure (see Analytical Approach) applied to the above-mentioned NSS regions was applied to these genes, yielding brain genes LD weighted affiliation scores.

Analytical Approach

We employed a genetic enrichment method developed to dissect the genetic architecture of complex traits (26,28,29,32). Specifically, we investigated the enrichment of associations concurrent with the NSS score selection index in a covariate-modulated statistical approach (36). We investigated whether SNPs with a low NSS score and therefore in regions possibly subjected to positive selection in humans are more likely associated with schizophrenia or other phenotypes. All statistical analyses were carried out with a covariate-modulated enrichment analysis package developed on R (www.r-project.org) and MATLAB (www.mathworks.se/products/matlab/) programming platforms.

Quantile-Quantile and Fold Enrichment Plots. Quantile-quantile (Q-Q) plots are designed to compare two distributions; we compared the nominal *p* value distribution with the empirical distribution. In the presence of null relationships only, the nominal *p* values form a straight line on a Q-Q plot when plotted against the empirical distribution. We plotted