Adaptive evolution of genes underlying schizophrenia

Bernard Crespi\textsuperscript{1,*}, Kyle Summers\textsuperscript{2} and Steve Dorus\textsuperscript{3}

\textsuperscript{1}Department of Biosciences, Simon Fraser University, Burnaby, British Columbia, Canada V5A 1S6
\textsuperscript{2}Department of Biology, East Carolina University, Greenville, NC 27858-4353, USA
\textsuperscript{3}Department of Biology and Biochemistry, University of Bath, Bath BA2 7AY, UK

Schizophrenia poses an evolutionary-genetic paradox because it exhibits strongly negative fitness effects and high heritability, yet it persists at a prevalence of approximately 1% across all human cultures. Recent theory has proposed a resolution: that genetic liability to schizophrenia has evolved as a secondary consequence of selection for human cognitive traits. This hypothesis predicts that genes increasing the risk of this disorder have been subject to positive selection in the evolutionary history of humans and other primates. We evaluated this prediction using tests for recent selective sweeps in human populations and maximum-likelihood tests for selection during primate evolution. Significant evidence for positive selection was evident using one or both methods for 28 of 76 genes demonstrated to mediate liability to schizophrenia, including \textit{DISC1}, \textit{DTNBP1} and \textit{NRG1}, which exhibit especially strong and well-replicated functional and genetic links to this disorder. Strong evidence of non-neutral, accelerated evolution was found for \textit{DISC1}, particularly for exon 2, the only coding region within the schizophrenia-associated haplotype. Additionally, genes associated with schizophrenia exhibited a statistically significant enrichment in their signals of positive selection in HapMap and PAML analyses of evolution along the human lineage, when compared with a control set of genes involved in neuronal activities. The selective forces underlying adaptive evolution of these genes remain largely unknown, but these findings provide convergent evidence consistent with the hypothesis that schizophrenia represents, in part, a maladaptive by-product of adaptive changes during human evolution.

Keywords: adaptive evolution; schizophrenia; positive Darwinian selection; disease genetics

1. INTRODUCTION

Schizophrenia is characterized by delusions and auditory hallucinations, loss of coherence and logic in thought and language, and emotionality inappropriate to social context (Crow 1997; Horrobin 1998; Tamminga & Holcomb 2005). This disorder has a polygenic basis involving multiple genes of small effect (Riley & Kendler 2006), and its phenotypic effects represent one end of a continuum that grades into schizotypal cognition and to normality (Rossi & Daneluzzo 2002; Mata et al. 2003).

Schizophrenia has also been associated with creativity throughout recorded history (Claridge et al. 1990; Nettle 2001). Diverse evidence from psychological experiments, biographical survey studies and neurophysiology has been accumulating for many years in support of the hypothesis that schizotypal cognition is associated with creativity and divergent thinking (see reviews in Claridge et al. 1990; Nettle 2001, 2006a; Barrantes-Vidal 2004). An untested prediction of this hypothesis is that schizophrenia has evolved, and is maintained, in part as a maladaptive by-product of recent positive selection and adaptive evolution in humans (Crow 1997; Horrobin 1998).

To assess the hypothesis that genes associated with schizophrenia have been subject to positive selection, we analysed the molecular evolution of 76 genes, which have been inferred to increase schizophrenia risk. We used two complementary approaches. First, we used linkage disequilibrium-based methods to test for signatures of recent selective sweeps, which involve relatively large coding and non-coding genomic regions (extended haplotypes) inferred to have recently and rapidly increased in frequency in human ancestry, such that they have not yet been broken up by recombination (Biswas & Akey 2006; Sabeti et al. 2006; Voight et al. 2006). Second, we employed phylogeny-based maximum-likelihood methods (Yang 1997; Zhang et al. 2005) for detecting positive selection, in humans and related primates, which are based on the ratio of non-synonymous to synonymous substitution rates. We also review the literature from previous studies on the molecular evolution of schizophrenia-related genes along the human lineage.

2. MATERIAL AND METHODS

(a) Criteria for inclusion of genes

Schizophrenia risk is mediated by many genes, each of small effect. The unambiguous identification of such genes has been difficult, because detecting small effects requires very large samples, different combinations of alleles can mediate liability, risk alleles often vary between populations and schizophrenia exhibits considerable clinical heterogeneity: it is not a simple, unitary disorder (Tamminga & Holcomb 2005; Riley & Kendler 2006). Our criterion for inclusion of schizophrenia-linked genes was genetic association with the disorder via association studies (comparing allele frequencies

\* Author for correspondence (crespi@sfu.ca).

Electronic supplementary material is available at http://dx.doi.org/10.1098/rspb.2007.0876 or via http://www.journals.royalsoc.ac.uk.

Received 29 June 2007
Accepted 16 August 2007

\textcopyright 2007 The Royal Society
or genotype distributions between cases and controls), or via family-based transmission–disequilibrium studies that test for differential inheritance of alleles between affected and non-affected siblings. We excluded all genes that were linked with schizophrenia in single studies, which were subject to failed replication attempts. The list of 76 such genes used here (electronic supplementary material) is highly congruent with that of Schmidt-Kastner et al. (2006) and was fully assembled prior to our analyses.

The strength of the evidence for linkage of these genes to the development of schizophrenia varies considerably, and it is very likely that some proportion of these linkages will turn out to be false positives given the history of failed replications in this field and the high degree of clinical, genetic, gender and ethnic heterogeneity in this disorder. We appreciate this fact and, as this is the first analysis of its kind, we have chosen to initially be relatively permissive in the inclusion of genes. However, we are careful to highlight cases where evidence of selection is observed among genes for which especially strong and highly replicated association data exist.

We also note that the presence, in our analyses, of any genes that are eventually shown not to be associated with schizophrenia will make our test of enhanced positive selection on schizophrenia-associated genes (compared with control genes associated with neuronal activities) more, rather than less, conservative. Indeed, with regard to this test, the main issue at hand is whether or not the set of genes has been statistically linked with schizophrenia in one or more studies is more likely to actually exhibit association with schizophrenia than is a random sample of control genes. We believe that this assumption is justified.

(b) Tests for positive selection using the human HapMap

Recent positive selection can usefully be inferred by the identification of selective sweeps, whereby selection for a specific allele causes a relatively large block of surrounding DNA (an extended haplotype) to increase in frequency. This method is most effective for recent selection, as recombinaton breaks up such extended haplotypes over time, and it can detect selection on both coding and non-coding regions.

Voight et al. (2006) have developed tests for recent selective sweeps in human populations using data from the human haplotype map. For the three genotyped populations, one European, one African and one Asian, they tested for evidence of positive selection as indicated by the tendency of recently selected alleles to sweep a set of tightly linked sites to relatively high frequency. They also provide a web interface, HAPLOTTER (hg-wen.uchicago.edu/selection/haplotter.htm), which we used for testing the presence and significance of selective sweeps for specific genes and sets of contiguous genes, and for localizing inferences of selection to specific gene regions and SNPs. Our criterion for positive selection in these data was a probability value of 0.05 or lower for one or more of the three populations.

To test for a statistically increased signal of positive selection in schizophrenia-associated genes, we compared the frequency of positive selection in the set of 76 schizophrenia-associated genes with the frequency of positive selection in a set of 300 control genes that were derived from the PANTHER gene-ontology category ‘neuronal activities’ (Mi et al. 2005).

(c) PAML analyses

Analyses of historical selection on specific lineages, using the ratios of non-synonymous to synonymous amino acid substitutions for coding DNA, were carried out with the CODEML program in the PAML program package (Yang 1997; Zhang et al. 2005). This program implements maximum-likelihood methods to estimate the ratio of non-synonymous to synonymous substitution rates for each codon in a gene, across an aligned set of sequences from that gene in different species, in a phylogenetic framework (electronic supplementary material). For the analysis of each gene, we used estimates of the phylogenetic relationships of the specific taxa involved using recently published phylogenetic analyses of mammalian relationships from the literature. Only the topological relationship among the species involved was recorded for these trees. Initial branch-length estimates were determined using a preliminary CODEML run with the single $\omega$ model (M0).

We obtained orthologous mammalian gene sequences for each candidate human gene in the following manner. Orthology information was used from the Homologene database (release 50.1) at NCBI (http://www.ncbi.nlm.nih.gov/) where available. Additional mammalian orthologues were then obtained using reciprocal best BLASTP hits from the GenBank protein database. Corresponding nucleotide sequences were obtained and in-frame CLUSTALW alignments of the coding regions were conducted. Aligned nucleotide sequences were then manually converted into PAML input files.

We used codon and branch-specific models (Zhang et al. 2005) to investigate the action of positive selection on each of the three branches in each tree: the human lineage; the human–chimpanzee lineage; and the basal-primate lineage, for each gene. In some cases, the primate lineage was the same as the human–chimpanzee lineage (i.e. no other primate sequence was available). These three lineages were chosen a priori because they represent periods when evolutionary changes underlying schizophrenia are likely to have been established: (i) at the origin of distinctive primate traits including increased encephalization and complex social behaviour, (ii) at the origin of large social-group sizes and complex communication in the lineage ancestral to chimpanzees and humans, and (iii) in the evolution of humans, the lineage expressing schizophrenia. The human lineage is most salient for analyses of positive selection in schizophrenia-associated genes, but the human–chimpanzee and primate lineage may also provide useful insights in that these lineages presumably underwent relatively strong selection for aspects of social cognition.

For the analyses of positive selection, we used ‘model A’ as implemented in CODEML, which detects selection on specific codons along specified branches of a phylogenetic tree (Zhang et al. 2005). Model A allows four categories of selection on codon sites in a sequence: two categories ($\omega_1$ and $\omega_0$) with uniform selection across all branches ($\omega_1 = 1$ and $0 < \omega_0 < 1$; estimated from the data) and two categories for which selection pressure differs in one or a few pre-specified ‘foreground’ branches where selection is assumed to have changed.

The method finds the model with the maximum likelihood, given the data. If the maximum-likelihood model includes a category of sites with an $\omega > 1$, then this provides evidence that positive selection has acted on those sites along the specific lineage analysed. To test the significance of the results, the log likelihood of the maximum-likelihood model can be compared to similar models lacking selection using a log-likelihood ratio.
Table 1. Genes with significant evidence of selection in human populations.

<table>
<thead>
<tr>
<th>gene</th>
<th>population</th>
<th>( p )</th>
<th>localized signal (gene(s))</th>
</tr>
</thead>
<tbody>
<tr>
<td>APC</td>
<td>Asian</td>
<td>0.021</td>
<td>—</td>
</tr>
<tr>
<td>BRD1</td>
<td>European</td>
<td>0.013</td>
<td>—</td>
</tr>
<tr>
<td>CAPON</td>
<td>European</td>
<td>0.036</td>
<td>2</td>
</tr>
<tr>
<td>CLDN5</td>
<td>European</td>
<td>0.042</td>
<td>—</td>
</tr>
<tr>
<td>DGCR2 (DGeorge syndrome region gene 2)</td>
<td>Asian</td>
<td>0.014</td>
<td>2</td>
</tr>
<tr>
<td>DTNBPI* (dystrobrevin-binding protein 1)</td>
<td>European</td>
<td>0.003</td>
<td>1</td>
</tr>
<tr>
<td>ENTH (epsin 1)</td>
<td>European</td>
<td>0.027</td>
<td>1</td>
</tr>
<tr>
<td>ERBB4 (erythroblastic leukaemia oncogene homolog 4)</td>
<td>Asian</td>
<td>0.044</td>
<td>1</td>
</tr>
<tr>
<td>GABRA1 (( \gamma )-aminobutyric acid receptor)</td>
<td>European</td>
<td>0.035</td>
<td>1</td>
</tr>
<tr>
<td>GABRA1</td>
<td>Asian</td>
<td>0.050</td>
<td>1</td>
</tr>
<tr>
<td>GRM3* (glutamate receptor 3)</td>
<td>Asian</td>
<td>0.005</td>
<td>2</td>
</tr>
<tr>
<td>PCMI (pericentriolar material 1)</td>
<td>European</td>
<td>0.049</td>
<td>1</td>
</tr>
<tr>
<td>PCMI</td>
<td>African</td>
<td>0.044</td>
<td>1</td>
</tr>
<tr>
<td>PIK3C3 (phosphatidylinositol 3-kinase)</td>
<td>African</td>
<td>0.037</td>
<td>1</td>
</tr>
<tr>
<td>SLC6A4 (serotonin transporter)</td>
<td>European</td>
<td>0.035</td>
<td>—</td>
</tr>
<tr>
<td>SLC6A4</td>
<td>Asian</td>
<td>0.014</td>
<td>—</td>
</tr>
<tr>
<td>SYNGR1 (synaptogyrin)</td>
<td>Asian</td>
<td>0.031</td>
<td>—</td>
</tr>
</tbody>
</table>

* Association with schizophrenia is especially strong and well replicated.

Two genes showing significance in the PAML analyses, NOTCH4 and TNXB, were marginally non-significant in HAPLOTTER analyses (NOTCH4: 0.094 European; TNXB: 0.06 European), and three genes showed marginal non-significance in the HAPLOTTER analysis (HOPA: 0.060 Asian; NOGO: 0.064 African; TPH1: 0.0505 Asian).

3. RESULTS

(a) HapMap analyses of positive selection in recent human evolution

HapMap-based analyses provided evidence of recent positive selection on 14 of the genes associated with schizophrenia (table 1 and electronic supplementary material). For six of these genes, the signal of selection was localized specifically to the focal gene, and for three genes, only one adjacent gene also showed significance \((p<0.05)\). Of particular interest is the identification of a particularly strong and localized focus of selection for DTNBPI, which exhibits one of the strongest genetic and functional associations with schizophrenia of any gene analysed to date (Riley & Kendler 2006).

Of the 14 genes inferred as selected from the HapMap data, four displayed significant evidence for recent selection in more than one human population, while the rest were found to be significant in only one population. This is suggestive of extremely recent selection acting on these genes since the time of human demographic expansion out of Africa, and it is also generally consistent with the presence of variation among human populations in the genetic basis of schizophrenia (e.g. Bulayeva et al. 2007).

The proportion of schizophrenia genes inferred as selected in HAPLOTTER (18.4%, 14 of 76) was significantly higher than the proportion of neuronal activities control genes inferred as selected (9.0%, 27 of 300, \( \chi^2 = 5.52, p = 0.019 \)). Similar results were obtained when we restricted the analyses to the subset of schizophrenia genes \((N=44)\) for which differences in gene expression levels or gene-associated neuroanatomy have been documented between schizophrenia and controls (electronic supplementary material): 9 (20%) of 44 positively selected compared with 27 (9%) of 300 for neuronal activities controls \((\chi^2 = 5.37, p = 0.02)\).

These statistical comparisons with control genes indicate that schizophrenia-associated genes may show an enriched signal of positive selection in recent human evolution, and that the evidence of selection is often heterogeneous across human populations. Additionally, 12 genes, AHI1, CHL1, GRIA4, GRIN2A, GRIN2B, NOS-1, PDLIM5, PL2G4A, RELN, SYN2, TNXB and TPH1, exhibited non-significance in gene-wide analyses using HAPLOTTER, but had one or more single nucleotide polymorphisms that were inferred to have been subject to positive selection.
Table 2. Significant maximum-likelihood results on the human lineage for genes associated with schizophrenia.

<table>
<thead>
<tr>
<th>gene</th>
<th>model</th>
<th>ω (proportion)</th>
<th>λ</th>
<th>−2Δλ</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>NRG1a (neuregulin 1)</td>
<td>MA</td>
<td>999 (0.050)</td>
<td>−2010.46</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td></td>
<td>MÅofix</td>
<td>−2012.50</td>
<td>4.07</td>
<td>&lt;0.05</td>
<td></td>
</tr>
<tr>
<td>PLP1 (proteolipid protein 1)</td>
<td>MA</td>
<td>482.65 (0.004)</td>
<td>−1630.56</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td></td>
<td>MÅofix</td>
<td>−1636.11</td>
<td>11.08</td>
<td>&lt;0.01</td>
<td></td>
</tr>
<tr>
<td>TNXB (tenascin XB)</td>
<td>MA</td>
<td>999 (0.006)</td>
<td>−38734.00</td>
<td>21.90</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td></td>
<td>MÅofix</td>
<td>−38745.00</td>
<td>—</td>
<td>—</td>
<td></td>
</tr>
<tr>
<td>TP53 (tumour protein p53)</td>
<td>MA</td>
<td>25.38 (0.007)</td>
<td>−5115.85</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td></td>
<td>MÅofix</td>
<td>−5118.93</td>
<td>3.78</td>
<td>&lt;0.05</td>
<td></td>
</tr>
</tbody>
</table>

Table 3. Significant maximum-likelihood results in the human and chimpanzee or primate lineages for genes associated with schizophrenia.

<table>
<thead>
<tr>
<th>gene</th>
<th>lineage</th>
<th>model</th>
<th>ω (proportion)</th>
<th>λ</th>
<th>−2Δλ</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>DISC1</td>
<td>human–chimpanzee</td>
<td>MA</td>
<td>14.97 (0.024)</td>
<td>−3512.58</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td></td>
<td></td>
<td>MÅofix</td>
<td>−3514.60</td>
<td>4.02</td>
<td>&lt;0.05</td>
<td></td>
</tr>
<tr>
<td>CHRNA7</td>
<td>primate</td>
<td>MA</td>
<td>17.56 (0.002)</td>
<td>−4098.03</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td></td>
<td></td>
<td>MÅofix</td>
<td>−4099.89</td>
<td>3.72</td>
<td>&lt;0.05</td>
<td></td>
</tr>
<tr>
<td>CLDN5 (claudin-5)</td>
<td>human–chimpanzee</td>
<td>MA</td>
<td>22.91 (0.061)</td>
<td>−1371.13</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td></td>
<td></td>
<td>MÅofix</td>
<td>−1373.85</td>
<td>5.44</td>
<td>&lt;0.05</td>
<td></td>
</tr>
<tr>
<td>COMT</td>
<td>primate</td>
<td>MA</td>
<td>41.58 (0.013)</td>
<td>−2914.90</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td></td>
<td></td>
<td>MÅofix</td>
<td>−2918.90</td>
<td>8.00</td>
<td>&lt;0.01</td>
<td></td>
</tr>
<tr>
<td>NT3N</td>
<td>human–chimpanzee</td>
<td>MA</td>
<td>999 (0.006)</td>
<td>−7704.50</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td></td>
<td></td>
<td>MÅofix</td>
<td>−7709.90</td>
<td>10.88</td>
<td>&lt;0.01</td>
<td></td>
</tr>
<tr>
<td>DISC1</td>
<td>primate</td>
<td>MA</td>
<td>46.90 (0.032)</td>
<td>−7703.50</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td></td>
<td></td>
<td>MÅofix</td>
<td>−7710.10</td>
<td>13.2</td>
<td>&lt;0.01</td>
<td></td>
</tr>
<tr>
<td>NOTCH4</td>
<td>primate</td>
<td>MA</td>
<td>294.83 (0.006)</td>
<td>−16526.56</td>
<td>5.26</td>
<td>&lt;0.05</td>
</tr>
<tr>
<td></td>
<td></td>
<td>MÅofix</td>
<td>−16526.56</td>
<td>—</td>
<td>—</td>
<td></td>
</tr>
<tr>
<td>PLXNA2 (plexin A2)</td>
<td>human–chimpanzee</td>
<td>MA</td>
<td>999 (0.001)</td>
<td>−12045.56</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td></td>
<td></td>
<td>MÅofix</td>
<td>−12049.12</td>
<td>7.13</td>
<td>&lt;0.05</td>
<td></td>
</tr>
</tbody>
</table>

(a) PAML analyses of positive selection in primate evolution

PAML analyses yielded significant evidence of positive selection on four genes on the human lineage (table 2 and electronic supplementary material). Of particular interest is the identification of NRG1 (neuregulin 1) that shows notably strong genetic association with schizophrenia (Hall et al. 2006; Riley & Kendler 2006; Thomson et al. 2007). This gene shows an absence of evidence for positive selection either in the human–chimpanzee lineage, in the ancestral primate lineage leading to humans, or in both lineages (table 3).

Overall, 11 (15%) of the 76 schizophrenia-associated genes showed evidence of positive selection in the PAML analyses (tables 2 and 3), compared with 10 (8.3%) of the 120 control genes analysed (χ² = 1.83, p = 0.176). For the human lineage, a higher proportion, 4 (5.6%) of 76 of schizophrenia-associated genes, showed evidence of positive selection than did the control genes (none of 120; Fisher’s exact test, p = 0.022). However, there was no such difference for the human–chimpanzee lineage (with three positively selected control genes versus five positively selected schizophrenia-associated genes; Fisher’s exact test, p > 0.50), or for the primate-origin lineage (seven positively selected control genes versus five positively selected schizophrenia-associated genes; Fisher’s exact test, p > 0.50). If the overly conservative (Zhang et al. 2005) test using χ²-test with 1 d.f. is used, then one of the schizophrenia-associated genes, TP53, shows marginally non-significant evidence of positive selection in the human lineage, such that three genes show positive selection in this lineage (Fisher’s exact test, p = 0.057).

(b) PAML analyses of positive selection in primate evolution

PAML analyses also yielded a category of sites with an ω > 1, providing evidence that positive selection may have acted on those sites along the specific lineage analysed, for a set of genes that did not show significance under the likelihood ratio tests. These genes, and the putatively selected lineages, include AHI1 (human–chimpanzee), AKT1 (human), CAPON (human–chimpanzee), CCK-AR (human), CHRM5 (human–chimpanzee), CNTF (human), DRD4 (human–chimpanzee), DRP-2 (human), GABRB3 (human–chimpanzee), GAD1 (primate), HOXB (human), GRIN2A (human and human–chimpanzee), GRIN2B (primate), GSTM1 (human), MLC1 (human–chimpanzee), NT3N (human), PDLIM5 (human), PTGS2 (human), RGS4 (human–chimpanzee), SLC6A4 (primate), SNAP29 (human) and STOP (human–chimpanzee and primate). Table 3 shows the human–chimpanzee and primate branch results.

(c) Positive selection of DISC1

DISC1 (disrupted in schizophrenia 1), a gene with strong, well-replicated genetic ties to this disorder (Mackie et al. 2007), was inferred as positively selected in both the human and chimpanzee lineages (table 3). This gene shows an absence of evidence for positive selection either in the human–chimpanzee lineage, in the ancestral primate lineage leading to humans, or in both lineages (table 3).
human–chimpanzee and primate-origin lineages. Remarkably, this gene showed 16 amino acid sites as having evolved under the influence of positive selection, of which six are within the HEP3 haplotype that segregates with schizophrenia (see figure 1 and the following; Porteous et al. 2006). To further investigate the DISC1 coding region in the HEP3 haplotype, we conducted a sliding-window analysis and found a localized peak of accelerated evolution in this region (figure 1). DISC1 thus evolved extremely rapidly during primate evolution, with significant evidence for acceleration in the only coding region of the gene showing reproducible evidence of a genetic association with schizophrenia.

4. DISCUSSION

Our results provide evidence for a significantly stronger signal of positive selection specific to the human lineage for schizophrenia-associated genes than for a control set of neuronal activities genes, for both the HapMap-based analyses of recent selective sweeps and the PAML analyses of positive selection. Furthermore, significant evidence of positive selection was found for several genes that exhibit some of the best-replicated and best-understood genetic and functional associations with this disorder, including DISC1, DTNB1 and NRG1.

The primary, non-exclusive implications of these results are that schizophrenia may represent, in part, a maladaptive by-product of adaptive changes in human and primate evolution, and that some alleles increasing liability to the cognitive impairments associated with schizophrenia may have been selected against. These hypotheses can be evaluated by testing the same sets of polymorphisms for association with schizophrenia and for signatures of positive selection or other evolutionary-genetic processes (Mutsuddi et al. 2006). Such integrated evolutionary-genomic analyses will also provide a novel, powerful perspective into the functions of genes, alleles and neurological substrates that underlie the aetiology of schizophrenia and schizotypal cognition, since positively selected variants are expected to exhibit functional significance. Genetic and phenotypic changes along the human line are most directly salient to the evolutionary basis of schizophrenia, but changes earlier in primate evolution, as apparently exemplified by DISC1, may have generated some of the basic social-cognitive substrates that are dysregulated in this disorder.

Additional convergent evidence that selection has impacted the evolutionary underpinnings of schizophrenia comes from two sources. First, neurological studies have shown that brain areas differentially dysregulated in schizophrenia include the regions most-notably subject to differential evolutionary change along the human lineage (Randall 1998; Brüne 2004; Burns 2006). Second, Wayland et al. (2006) found that genes exhibiting positive selection for differential expression between humans and chimpanzees are differentially dysregulated in dorsolateral prefrontal and orbitofrontal cortices of individuals with schizophrenia (Wayland et al. 2006). These findings link recent evolutionary changes in human neuroanatomy and gene expression with alterations of these phenotypes in schizophrenia and demonstrate that the phenotype substrates of this disorder, as well as its genetic basis, show evidence of effects from positive selection and adaptive evolution.

(a) Evidence for positive selection from previous studies

Our maximum-likelihood protein-coding and haplotype-based analyses probe only for a subset of the possible signatures of selection on genes increasing liability to schizophrenia. Moreover, these methods detect signals over different time scales, and for different types of
positively selected variants, such that the general absence of concordance of selection identified as selected between the two types of method (HapMap and PAML) is not unexpected. For example, the great majority of positively selected alleles in humans and other mammals are non-coding (Williamson et al. 2007; Resch et al. 2007), such that recent selective sweeps are expected to involve mainly non-coding regions that will be detected only by analyses based on HapMap data.

Given that selection involving only one or few amino acid sites, selective sweeps older than HapMap data can distinguish, or balancing selection that maintains human polymorphisms, may have influenced the genetic architecture of schizophrenia, we also reviewed the literature for evidence of selection or rapid evolution of genes associated with schizophrenia risk, from previous studies. Sixteen genes that exhibit genomic and functional links with schizophrenia, ADCYAP1, AHI1, APOE, CCR5, DRD4, DTNBP1, FOX2, GABRB2, GRM3, HOPA, LRRTM1, MAOA, MCPH1, OLIG2, PDLIM5 and SLC6A4, also show evidence of positive or balancing selection in the human lineage from previous published studies, and five genes, CHRM5, GRIK4, ENTH, HLA-DRB and NPA53, show evidence of rapid evolution (table 4 and electronic supplementary material). Furthermore, Voight et al. (2006) noted an unusual preponderance of phosphatidylinositol 3-kinase signal-transduction pathway genes subject to recent selective sweeps, and evidence from functional analyses, physiology, pharmacology and genetic-association studies indicates that dysregulated activation of this pathway (via effects of genes such as AKT1, DTNBP1, ENTH, ERBB4, NRG1, PIK3C3 and RELN) represents an important risk factor for schizophrenia (Kalkman 2006). Finally, the schizophrenia-related dopamine receptor genes DRD2, DRD3 and DRD4 exhibit heterosis in some of their effects on human physiology and behavior (Comings & MacMurray 2000). These findings provide additional evidence for selection and rapid evolution of genes underlying schizophrenia, which is generally concordant with the results from PAML and HapMap analyses and localizes signatures of selection to the human lineage.

(b) Potential mechanisms mediating the evolution of genes that underlie schizophrenia

Positive selection of genes implicated in schizophrenia may be mediated by an array of neurodevelopmental, neurophysiological and psychological mechanisms. Our analyses provide complementary types of evidence for positive selection of genes underlying this disorder, but do not provide any direct information concerning how selective pressures at the phenotypic level may have impacted the evolution of these genes.

Evidence from previous work salient to the hypothesis that the adaptive evolution of genes underlying schizophrenia has involved phenotypic selection comes from studies of genetic linkages between schizophrenia-risk genes and aspects of creativity. Alleric variants of three genes associated with schizophrenia, SLC6A4, TPH1 and DRD2, have recently been associated with measures of creativity and imagination in normal population (Bachner-Melman et al. 2005; Reuter et al. 2006). Notably, the STin2.12 allele of an intronic polymorphism in the SLC6A4 gene and the A allele of the A779C polymorphism in the TPH1 gene are each associated with both increased creativity and increased risk of schizophrenia (Bachner-Melman et al. 2005; Reuter et al. 2006). Alleles of several genes associated with schizophrenia also influence measures of schizotypal cognition and ‘openness to experience’ in non-clinical populations, including COMT and HTR2A (Ott et al. 2005), MAOA (Samochowiec et al. 2004) and SLC6A4 (Golimbet et al. 2003). Description of the linkages of creativity and cognition with schizotypy and schizophrenia is provided for the four genes, SLC6A4, TPH1, COMT and DRD2, with sufficient integrative information available for detailed analysis (electronic supplementary material). We also describe here the neurological and psychological evidence that links schizotypy with aspects of creativity and divergent thinking, and epidemiological data for variation in measures of fitness between schizophrenics, first-order relatives and controls.

The mechanisms connecting schizotypal cognition and creativity with postulated fitness benefits are unclear, but may include sexual selection, creative and artistic skills, or general benefits from insight problem solving (Shaner et al. 2004; Nettle 2006a; Karimi et al. 2007). Such processes could potentially help to explain the paradoxical high heritability and persistence of schizophrenia, owing to some combination of multilocus balancing selection, antagonistic pleiotropy affecting fitness and mutation–selection balance.

We stress that enhanced creativity is not expected to be a function of schizophrenia itself (which involves profound cognitive impairments), but rather is inferred to be a function of schizotypal cognition (Claridge et al. 1990; Nettle 2001, 2006a,b; Barrantes-Vidal 2004). Indeed, alleles underlying the deleterious aspects of schizophrenia or schizotypal cognition are expected to be selected against (Keller & Miller 2006; Nettle 2006b). We also note that most studies of the cognitive correlates of genetic liability to schizophrenia have focused on documenting deficits in clinical populations, rather than testing for potentially positive aspects of cognition in healthy relatives or other individuals who carry schizophrenia-associated alleles or genotypes.

(c) Limitations and implications

Interpretation of our results is limited by the observation that few of the genes associated with increased schizophrenia risk are specific to this disorder: some such genes are also known to affect liability to bipolar disorder, and most of them influence a diverse range of neurodevelopmental and physiological processes, and exert only a small effect on schizophrenia risk that may vary among populations (Kendler 2005; Riley & Kendler 2006). Bipolar disorder and schizophrenia in particular exhibit substantial overlap in cognitive symptoms as well as their genetic basis, and links of bipolar disorder with creativity and imagination have also been repeatedly described (Claridge et al. 1990; Nettle 2001).

A second limitation of our study is that there remains considerable controversy about what constitutes adequate statistical demonstration of association between allelic variants of a specific gene and schizophrenia. Taking the most conservative approach on this question would result in too few genes for any meaningful analyses, whereas including all genes that have ever been linked with schizophrenia in a single study or population would be

Table 4. Evidence for selection and rapid evolution of genes associated with schizophrenia from previous studies.

<table>
<thead>
<tr>
<th>gene</th>
<th>evidence of selection</th>
<th>references</th>
</tr>
</thead>
<tbody>
<tr>
<td>ADCYAPI (adenylate cyclase-activating polypeptide 1)</td>
<td>positive selection</td>
<td>Bustamante et al. (2005); Wang et al. (2005); Hashimoto et al. (in press)</td>
</tr>
<tr>
<td>AHI1 (Abelson helper integration site 1)</td>
<td>rapid evolution; positive selection</td>
<td>Polland et al. (2004); Amann-Zakenstein et al. (2006)</td>
</tr>
<tr>
<td>APOE (apolipoprotein E)</td>
<td>positive and balancing selection</td>
<td>Fullerton et al. (2000); Finch &amp; Stanford (2004); Xu et al. (2006)</td>
</tr>
<tr>
<td>CCR5 (chemokine, C–C motif, receptor 5)</td>
<td>balancing selection</td>
<td>Bamshad et al. (2002); Wooding et al. (2005); Geroldi et al. (2006); Rasmussen et al. (2006)</td>
</tr>
<tr>
<td>CHRM5 (cholinergic receptor muscarinic 5)</td>
<td>rapid evolution in primate compared with rodents</td>
<td>Dorus et al. (2004)</td>
</tr>
<tr>
<td>DRD4 (dopamine receptor D4)</td>
<td>selective sweep (approx. 40 000 yrs ago)</td>
<td>Ding et al. (2002); Lung et al. (2002); Jönsson et al. (2003a); Wang et al. (2004)</td>
</tr>
<tr>
<td>DTNBPI* (dystrobrevin-binding protein 1)</td>
<td>haplotype structure suggests positive selection</td>
<td>van den Oord et al. (2003); Williams et al. (2005); Riley &amp; Kendler (2006); Tochigi et al. (2006)</td>
</tr>
<tr>
<td>ENTH (epsin 1)</td>
<td>nearest gene to human-accelerated genomic region</td>
<td>Polland et al. (2006); Tang et al. (2006); Gurling et al. (2007)</td>
</tr>
<tr>
<td>FOXP2 (forkhead box P2)</td>
<td>positive selection</td>
<td>Enard et al. (2002); Zhang et al. (2002); Sanjuán et al. (2006)</td>
</tr>
<tr>
<td>GABRB2 (γ-aminobutyric acid receptor β2)</td>
<td>positive selection</td>
<td>Lo et al. (2007a,b)</td>
</tr>
<tr>
<td>GRK4 (glutamate receptor ionotropic kainate 4)</td>
<td>rapid evolution in primate compared to rodents</td>
<td>Dorus et al. (2004); Pickard et al. (2006a)</td>
</tr>
<tr>
<td>GRM3 (glutamate receptor metabotropic 3)</td>
<td>selective sweep</td>
<td>Diller et al. (2002); Egan et al. (2004); Chen et al. (2005)</td>
</tr>
<tr>
<td>HLA-DRB (major histocompatibility complex, class II, DRβ-5)</td>
<td>rapid evolution</td>
<td>Bergström et al. (1998); Wright et al. (2001); Holjoh et al. (2003)</td>
</tr>
<tr>
<td>HOP4 (mediator of RNA polymerase II transcription)</td>
<td>selective sweep; positive selection</td>
<td>DeLisi et al. (2000); Philibert et al. (2001, 2007); Spinks et al. (2004); Bustamante et al. (2005)</td>
</tr>
<tr>
<td>LRRTM1 (leucine-rich repeat transmembrane protein 1)</td>
<td>selective sweep</td>
<td>Voight et al. (2006); Laurén (2007); Franck et al. (in press)</td>
</tr>
<tr>
<td>MAOA (monoamine oxidase A)</td>
<td>selective sweep</td>
<td>Balciuniene et al. (2001); Gilad et al. (2002); Jönsson et al. (2003b); Andrés et al. (2004)</td>
</tr>
<tr>
<td>MCPH1 (microcephalin 1)</td>
<td>positive selection</td>
<td>Evans et al. (2004, 2005); Wang &amp; Su (2004); Lencz et al. (2007)</td>
</tr>
<tr>
<td>NPAS3 (neuronal PAS domain protein 3)</td>
<td>nearest gene to human-accelerated genomic region</td>
<td>Pickard et al. (2006b); Pollard et al. (2006)</td>
</tr>
<tr>
<td>OLG1 (oligodendrocyte transcription factor 2)</td>
<td>selective sweep</td>
<td>Wright et al. (2001); Diller et al. (2002); Georgieva et al. (2006)</td>
</tr>
<tr>
<td>PDLIM5 (PDZ and LIM domain 5)</td>
<td>positive selection</td>
<td>Bustamante et al. (2005); Kato et al. (2005); Horiuichi et al. (2006)</td>
</tr>
<tr>
<td>SLC6A4 (solute carrier family 6; serotonin transporter)</td>
<td>positive selection</td>
<td>Bustamante et al. (2005)</td>
</tr>
</tbody>
</table>

* Association with schizophrenia is especially strong and well replicated. Detailed information on evidence for selection, linkages to schizophrenia and cognitive effects of each gene is provided in the electronic supplementary material. Prabhakar et al. (2006) also reported evidence for accelerated evolution of CHL1, ERBB4, GRK4, GRM8, LRRTM1 and QKI genes in the human lineage, involving human-specific substitutions in non-coding regions directly adjacent to these genes.

relatively liberal, although it would not be expected to bias analytic results towards a finding of an enhanced signal of positive selection, relative to control genes. We have taken a middle ground and used genes that have not been subject to failed replication attempts. However, it is important to note that significant evidence for selection was found in multiple genes showing the strongest, best-replicated associations with schizophrenia to date, including DISC1, DTNBPI and NRG1. We have also repeated the HapMap analyses using only the 44 genes for which expression or gene-associated neuroanatomy differences have been recorded between schizophrenics and controls, and found the same results.

An important implication of our findings is that evolutionary-genomic analyses can provide novel insights into the nature and functions of the genes that underlie the aetiologies of schizophrenia. More generally, analyses that probe for signatures of selection on genes implicated in schizophrenia may localize functional sites affecting gene expression, enzyme activity, neurocognitive phenotypes and disease risk, and thus highlight potential allelic variants salient to the cognitive changes that have driven the origin of modern humans (Crow 1997; Horrobin 1998).

We thank F. Breden and B. Voight for advice and comments and four anonymous reviewers for their insightful suggestions. This work was funded by grants from NSERC and the
REFERENCES


Ferland, R. J. et al. 2004 Abnormal cerebellar development and axonal decussation due to mutations in AHI1 in Joubert syndrome. *Nat. Genet.* 36, 1008–1011. (doi:10.1038/ng1419)


Georgieva, L. et al. 2006 Convergent evidence that oligoden-drocyte lineage transcription factor 2 (OLIG2) and interacting genes influence susceptibility to schizophrenia. *Proc. Natl Acad. Sci. USA* 103, 12 469–12 474. (doi:10.1073/pnas.0603029103)


Hohjoh, H., Ohashi, J., Takasu, M., Nishioka, T., Ishida, T. & Tokunaga, K. 2003 Recent divergence of the HLA-DRB1*04 allelic lineage from the DRB1*0701 lineage after the separation of the human and chimpanzee species. Immunogenetics 54, 856–861.


Pickard, B. S. et al. 2006a Cytogetic and genetic evidence supports a role for the kainate-type glutamate receptor gene, GRIK4, in schizophrenia and bipolar disorder. Mol. Psych. 11, 847–857. (doi:10.1038/sj.mp.4001867)


