The Psychosis Susceptibility Gene ZNF804A: Associations, Functions, and Phenotypes

As the first gene to have achieved genome-wide significance for psychosis, ZNF804A has predictably been a subject of intense research activity. We review the evidence to date for the association between schizophrenia and the original risk variant rs1344706 identified as well as additional common and rare variants at this locus. We describe the still scant literature on the biological function of ZNF804A and discuss the efforts being made to characterize and refine the associated phenotype using imaging and neuropsychological approaches. We conclude that ZNF804A is robustly, if modestly, associated with schizophrenia risk, with much work still remaining to elucidate its role in schizophrenia biology.

Key words: ZNF804A/schizophrenia/genetics

Introduction

Genome-wide association studies (GWASs) have rapidly become the standard method for discovery of common risk variants, with more than 450 GWAS completed across disorders since 2005, implicating >2000 single nucleotide polymorphisms (SNPs) in phenotypic variation (http://www.genome.gov/gwastudies/; for a review, see Ku et al). A key feature of the GWAS approach is that it provides an opportunity to study common DNA variation across the genome independent of a priori knowledge about either genome location or disease etiology. This has led to identification of novel, and often surprising, risk genes whose functions may fall outside the known biology of a disorder. Early schizophrenia (SZ) GWAS based on DNA pooling and then individual genotyping failed to find evidence of any variant that achieved genome-wide significance for the disorder—ZNF804A rs1344706. Since then ZNF804A has received substantial support from replication efforts. This review examines the evidence for ZNF804A (OMIM: 612282) as a psychosis susceptibility factor, describes the initial work undertaken to establish the function of this gene, and examines potential directions for further study of the gene.

Identification of Common Variants at ZNF804A Achieving Genome-Wide Significance for Psychosis

The SNP rs1344706 is located in an intron of ZNF804A that maps to a short region of conserved mammalian sequence on chromosome 2q32.1. Based on an initial discovery sample of 642 cases and 2937 controls, rs1344706 was one of 12 loci achieving significance of \( P < 1 \times 10^{-8} \) in the study by O’Donovan and colleagues. Three of these 12 loci were supported in replication samples of 6666 cases and 9897 controls. Of these, rs1344706 provided the greatest statistical evidence of support in a meta-analysis of all samples (\( P = 1.61 \times 10^{-7} \); OR = 1.12). The finding for rs1344706 was more significant when the phenotype was broadened to include bipolar disorder (\( P = 9.96 \times 10^{-9} \)).

Published replication studies of the association between ZNF804A and SZ are presented in table 1. Three independent SZ GWAS each supplied independent support for this SNP with the latter study subsequently being extended and providing an even more significant association. Riley et al sought to both replicate the association at rs1344706 and extend the analysis of ZNF804A with 11 other linkage disequilibrium (LD)-tagging SNPs that capture common variation across the gene in an Irish case-control sample of 1021 cases and 626 controls. Despite the relatively low power of these samples to detect association given the reported
ORs of ~1.10, significant association with the risk allele was observed for rs1344706. Three other SNPs were also associated, the strongest signal at rs7597593 (P = .001; OR = 1.28). Zhang et al also found significant support for rs1344706 association in a Han Chinese sample of 566 patients and 574 healthy controls (P = .0008; OR = 1.32). This study also found a trend-level association (P = .058) in a small family-based trio study (n = 101) with an overtransmission of the risk allele to SZ probands.

**Fine Mapping of the ZNF804A Locus**

The range of SNP coverage in the original O’Donovan et al’. GWAS left open the possibility that rs1344706 was a tagging SNP for the causative variant at this locus rather than representing the true functional variant itself. Addressing this, Williams et al. sought to localize the association signal at this locus through a process of genomic resequencing and fine-scale LD mapping. One hundred and seventy-six SNPs were analyzed across the gene to give near complete coverage of common variants (96% of SNPs with minor allele frequency >0.01 at r² > .9). After detailed association analysis, rs1344706 remained the most strongly associated marker in the gene. In addition, a meta-analysis of rs1344706 in 21 274 SZ and BP cases and 38 675 controls again supported the association between rs1344706 and both schizophrenia (P = 2.5 × 10⁻¹¹; OR 1.10, 95% CI: 1.07–1.14) and schizophrenia and bipolar disorder combined (P = 4.1 × 10⁻¹³; OR 1.11, 95% CI: 1.07–1.14). These data provided evidence far in excess of accepted thresholds for a SNP to be considered genome-wide significant (even dealing with the “winner’s curse” problem by excluding the original Cardiff sample from the analysis) and again positioned ZNF804A as a compelling risk gene for schizophrenia and a broader psychosis phenotype.

**Copy Number Variation at ZNF804A and SZ Risk**

There has long been speculation as to the relative contributions of common and rare genetic variation to the etiology of complex disease. Recent studies have established an important role for rare genomic deletions and duplications (copy number variants [CNVs]) in susceptibility to schizophrenia but also to other neurodevelopmental disorders. Steinberg et al investigated CNVs at the ZNF804A locus that could be associated with risk for psychiatric disorders in samples of schizophrenia, bipolar disorder, depression, or anxiety. These samples included overlapping and novel samples to those included in the GWAS study that had replicated the association with rs1344706. Steinberg et al identified 2 CNVs spanning at least part of ZNF804A in psychosis patients and none in controls (P = .013 for association with psychosis). Specifically, these consisted of a deletion in an individual with schizophrenia and a duplication in an individual with bipolar disorder. In addition, Steinberg et al also reported finding a deletion in an individual with anxiety disorder. These 3 cases contrasted with no CNVs identified at the locus in almost 40 000 controls (P = .0016). Other CNV studies

### Table 1. Published Association Results for ZNF804A Single Nucleotide Polymorphism rs1344706

<table>
<thead>
<tr>
<th>Study</th>
<th>Sample Identifier</th>
<th>Schizophrenia</th>
<th>Schizophrenia + Bipolar Disorder</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Number of Cases (f Aa)</td>
<td>Number of Controls (f Aa)</td>
</tr>
<tr>
<td>O'Donovan et al⁶</td>
<td>Combined UK</td>
<td>642 (0.66)</td>
<td>2937 (0.59)</td>
</tr>
<tr>
<td>O'Donovan et al⁶</td>
<td>Replication 1 + 2</td>
<td>6666</td>
<td>9897</td>
</tr>
<tr>
<td>O'Donovan et al⁶</td>
<td>Metaᵇ</td>
<td>7308</td>
<td>12 834</td>
</tr>
<tr>
<td>International</td>
<td>ISC minus Dublin</td>
<td>2519</td>
<td>2110</td>
</tr>
<tr>
<td>Schizophrenia</td>
<td>and Bulgaria</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Shi et al⁸</td>
<td>EA + AA</td>
<td>3967</td>
<td>3624</td>
</tr>
<tr>
<td>Steinberg et al¹⁰</td>
<td>Combined</td>
<td>5077</td>
<td>20 506 (0.59)</td>
</tr>
<tr>
<td>Riley et al¹¹</td>
<td>ICCSS</td>
<td>1021 (0.65)</td>
<td>626 (0.61)</td>
</tr>
<tr>
<td>Zhang et al¹²</td>
<td>Case-Control</td>
<td>566 (0.53)</td>
<td>574 (0.46)</td>
</tr>
<tr>
<td>Williams et al¹³</td>
<td>Combinedᶜ</td>
<td>18 945</td>
<td>38 675</td>
</tr>
</tbody>
</table>

**Note:** ICCSS, Irish case control study of schizophrenia.

ᵃFrequency of the risk A allele (f A; forward strand) only available where included in study manuscript.

ᵇThis sample represents the Combined UK and Replication 1 + 2 detailed above.

ᶜThese samples include most of the samples listed above plus other unpublished data. Full details are supplied in the manuscript.
of schizophrenia and bipolar disorder have failed to identify additional carriers of CNVs affecting ZNF804A.\textsuperscript{15–18} Finally, Cocchella et al\textsuperscript{19} recently reported a case study of an adult female patient showing facial dysmorphisms, mental retardation, and absence of speech who presented a 4.4 Mb deletion in the 2q31.2q32.3 region, which spans ZNF804A as well as \textit{NEUROD1}, \textit{PDE1A}, and \textit{ITGA4}.

Steinberg et al\textsuperscript{10} note from their review of the database of genomic variants\textsuperscript{20} that 2 CNVs involving ZNF804A exons are reported in healthy controls (one adult and one child); however, both the age (less than 18 years) and lack of detailed phenotypic characterization from a psychiatric standpoint makes it difficult to determine whether these individuals have, or will develop, pertinent neurodevelopmental phenotypes. From among the many CNV studies reported in autism to date,\textsuperscript{21–26} a duplication of the entire gene in 2 affected siblings and a partial ZNF804A duplication in another patient has been reported. While, as Steinberg et al\textsuperscript{10} note, it is difficult to make strong inferences about the relationship between these CNVs and psychiatric disability, overlaps with other disorders including autism and anxiety disorders make a specific identification between ZNF804A and schizophrenia unlikely. This is not dissimilar to the common variant at this locus, association of which extended beyond the schizophrenia phenotype. However, it is yet to be established conclusively if the common variant is a risk factor for other neuropsychiatric phenotypes.

**Functional Role of ZNF804A**

One of the goals of gene identification in GWAS studies is to develop novel insights into disease biology. The example of the ZNF804A gene highlights some of the current limitations of this approach. Consisting of 4 exons and transcribing a protein of 1210 amino acids, ZNF804A is known to be brain expressed and contains a C2H2-type domain associated with the zinc-finger protein family but is currently a protein of unknown function. Proteins with this zinc-finger domain were originally identified as DNA-binding molecules with a role in transcription but have diverse interactions with many molecules including RNA and proteins. Bioinformatic analysis of the conserved mammalian sequence around rs1344706 suggests the presence of transcription factor-binding sites. Riley et al\textsuperscript{11} in their analysis suggested that the 2 alleles result in differential prediction of 2 brain-expressed transcription factors, Myt1L zinc-finger protein and the POU3F1/Oct-6 POU domain transcription factor, both of which are involved in oligodendrocyte differentiation and proliferation. Alternatively, the mouse homologue of ZNF804A, zfp804a, has recently emerged as a target for HOXC8 suggesting that the gene may also be involved in the regulation of early neurodevelopment.\textsuperscript{27} Much work is still to be done to understand the role of this gene before understanding of its role in etiology can emerge.

The functional mechanism by which the risk allele contributes to etiology also remains to be determined. Williams and colleagues\textsuperscript{13} examined genotype and lymphoblastoid expression data from the GeneVar database and identified that rs1344706 was significantly associated with expression of ZNF804A mRNA, and the risk allele was associated with higher expression. They then measured the relative expression of each parental copy of ZNF804A in postmortem brain mRNA taken from 34 individuals heterozygous carriers of a proxy for the rs1344706 risk SNP (rs4667001, D’ = 1). They determined that risk allele was associated with a 1.13-fold (SD 0.08) increase in ZNF80A expression. This finding is compatible with the report by Riley and colleagues\textsuperscript{11} that the risk allele is associated with higher ZNF804A expression in human brain. Williams et al\textsuperscript{13} went further and identified that though the risk allele rs1344706 is generally carried by a higher ZNF804A expression haplotype, it doesn’t seem to be the eQTL responsible for higher expression. This finding, and the observation of a deletion at the ZNF804A locus in schizophrenia, suggests that more as yet unidentified variants at the locus may be involved in susceptibility and that the mechanism involved is more complex than simply upregulation of gene expression.

**Brain Imaging and Neuropsychological Studies of ZNF804A**

Changes in brain structure and function are core features of schizophrenia and may better represent underlying pathophysiology than clinical diagnostic categories.\textsuperscript{28} Whether or not the measurement of these deficits increases power to detect association with psychiatric risk genes,\textsuperscript{29} they allow investigations that extend beyond statistical associations studies of broad illness phenotypes toward a delineation of the specific effects of risk alleles on brain structure and function. The utility of this approach has already been demonstrated with other candidate genes for schizophrenia (including \textit{DISC1}, \textit{NRG1}, \textit{DTNBPI}, \textit{DAOA} [also known as \textit{G72}], and \textit{RGS4}; discussed in O’Donovan et al\textsuperscript{30}).

The first such study of ZNF804A by Esslinger et al\textsuperscript{31} investigated the influence of rs1344706 cortical activity within, and connectivity between, brain regions often associated with abnormalities in schizophrenia during working memory (n-back task) and emotion recognition task performance in a sample of 115 healthy controls. No differences in regional activation were associated with genotype. By contrast, differences in functional connectivity—a measure of the correlation in activity between 2 brain regions—were observed. Specifically, Esslinger et al\textsuperscript{31} observed reduced connectivity in the dorsolateral prefrontal cortex both between and within hemispheres. They also found increased connectivity...
between the hippocampal formation (HF) and the dorsolateral prefrontal cortex and between the amygdala and the HF, orbitofrontal cortex, and prefrontal cortex. Given the widespread evidence of reduced brain function in SZ and the evidence of ZNF804A as a SZ risk, Esslinger et al.\textsuperscript{31} concluded that this pattern of altered connectivity represented a deleterious effect on brain function. This was despite the fact that no behavioral differences were observed, although this might have been influenced by the sample size and the fact that only healthy participants were included.

A recent study from our group by Walters et al.\textsuperscript{32} on the neuropsychological effects of rs1344706 offers an alternative account of ZNF804A’s effect on cognition. This study sought to investigate neuropsychological performance in patients and healthy controls on cognitive functions typically impaired in schizophrenia—general intelligence, episodic memory, working memory, and attentional control. We found, and then replicated, evidence that carriers of the risk allele had significantly “better” cognitive performance than nonrisk allele carriers in patients but not healthy controls. Of note, the cognitive functions involved—working memory and episodic memory—implicate precisely those cortical regions—dorsolateral prefrontal cortex and HF—implicated in the Esslinger et al.\textsuperscript{31} study. This counterintuitive response—found only in patients and not in healthy participants—was interpreted to suggest that ZNF804A was associated with a psychosis phenotype in which cognitive performance was relatively less impaired by comparison with other SZ phenotypes characterized by greater cognitive impairment. This hypothesis was based in part on earlier evidence that ZNF804A was associated with a broader psychosis phenotype that included bipolar disorder, for which cognitive deficits are a less significant feature. To test this hypothesis, Walters et al.\textsuperscript{32} reran the association analysis between ZNF804A and SZ excluding low IQ cases. As lower IQ cases were excluded the association with SZ strengthened. These data seemed to support the idea that ZNF804A may be delineating a SZ subtype in which cognitive deficits are a less significant feature. Recent structural imaging data from our group provides further support for this view of relatively spared neurocognitive function: in patients but not controls, the ZNF804A risk allele was associated with relatively larger hippocampal volumes than noncarriers (G. Donohoe, E. Rose, T. Frodl, D. Morris; I. Spoletini, F. Adriano, S. Bernardini; C. Caltagirone, P. Bossu, M. Gill, A. Corvin, G. Spalletta, unpublished data).

Evidence that ZNF804A may be less important for the cognitive deficits associated with SZ than the perceptual, and social deficits are also suggested in recent study by Walter et al.\textsuperscript{32} Based on an overlapping sample to that reported in Esslinger et al.\textsuperscript{31} Walter et al.\textsuperscript{33} investigated cortical activation and connectivity associated with ZNF804A during performance on a theory of mind (ToM) task (which measured participant’s ability to infer mental state). A significant risk allele dose effect was found for activations of regions implicated in ToM function—the dorsomedial prefrontal cortex and the temporoparietal cortex. The authors also found differences in activation in the left inferior prefrontal associated with the ZNF804A risk allele which they attributed to social information processing difficulties generally. These deficits are at least consistent with the notion of ZNF804A as associated with social more than cognitive deficits, although confirmation of this hypothesis is likely to require investigation in clinical samples as well as the healthy participants samples reported here.

Further Clinical Studies of ZNF804A

Evidence that the association between ZNF804A and psychosis is strengthened when patients with bipolar disorder are included together with the data reviewed above suggesting that ZNF804A may be associated with a less cognitively impaired phenotype, but one in which deficits in social cognition are a feature, raises important questions about the clinical phenotype being implicated. For example, might ZNF804A be particularly associated with affective components of the broader psychosis phenotype? A recent study from our group investigated this question in a large samples (n = 820) of patients with schizophrenia, schizoaffective, and bipolar disorder based on a principal components analysis of symptom severity scores.\textsuperscript{34} ZNF804A genotype was significantly associated with increased mania-related symptom severity but only explained ~1% of variance in these symptoms. Although other studies will be required to answer this question more fully, these data do not suggest a clinically identifiable phenotype associated with this risk variant.

Conclusions

Since ZNF804A’s identification as the first genome-wide associated common variant for psychosis, several additional risk variants for schizophrenia have been identified.\textsuperscript{7–9} Among these, the substantial association evidence that has accumulated for ZNF804A positions, it as a compelling risk gene for schizophrenia and the broader psychosis phenotype. Reported ORs for ZNF804A are modest (~1.1) by comparison with non-psychiatric illnesses (eg, type II diabetes) and unlikely to be diagnostically informative, but this remains typical among psychiatric diseases. Despite an extensive search for other functional variants at this locus, rs1344706 remains the most strongly associated variant. Further investigation of rare variants, particularly of smaller CNV’s, remains to be undertaken at this locus.

Some clues about the gene’s function have been gathered from cognitive neuroscience studies in which


ZNF804A rs1344706 has been associated with altered functional connectivity, relatively less impaired neuropsychological performance, and reduced activation during measures of social cognition. Consistent with the original finding that ZNF804A is associated with a broad psychosis phenotype, these data may indicate that ZNF804A is associated with a less severe psychosis phenotype—further studies will help confirm whether or not this is the case. The biological understanding of ZNF804A remains unclear: We still know little about how this gene increases illness liability, whether by effects on DNA transcription regulation or downstream involvement in pathways critical to embryonic brain development or some other mechanism. As with other SZ risk variants, elucidating this biological role in animal and human studies represent important next steps in understanding SZ pathophysiology.

Funding
Research funding from Science Foundation Ireland (Gen437); The Health Research Board (HO1366); National Alliance for Research on Schizophrenia and Depression (young investigator grant to G.D.).

Acknowledgments
The Authors have declared that there are no conflicts of interest in relation to the subject of this study.

References
16. Walsh T, McClelland JM, McCarthy SE, et al. Rare structural variants disrupt multiple genes in neurodevelopmental pathways in schizophrenia. Science. 2008;320:539–543.
27. Chung HJ, Lee JY, Deocarisi CC, Min H, Kim SH, Kim MH. Mouse homologue of the schizophrenia susceptibility gene


32. Walters J, Corvin A, Owen MJ, et al. The psychosis susceptibility gene ZNF804A is associated with less impaired cognitive performance in schizophrenia. *Arch Gen Psychiatry*. 2010;76(7).
