Convergent evidence that *oligodendrocyte lineage transcription factor 2* (*OLIG2*) and interacting genes influence susceptibility to schizophrenia

Lyudmila Georgieva*, Valentina Moskvina*, Tim Peirce*, Nadine Norton*, Nicholas J. Bray*, Lesley Jones*, Peter Holmans*†, Stuart MacGregor†, Stanley Zammit*, Jennifer Wilkinson*, Hywel Williams*, Ivan Nikolov*†, Nigel Williams*, Dobril Ivanov*, Kenneth L. Davis‡, Vahram Haroutunian‡§, Joseph D. Buxbaum‡, Nick Craddock*, George Kirov*, Michael J. Owen*¶, and Michael C. O'Donovan*¶

*Department of Psychological Medicine and †Biostatistics and Bioinformatics Unit, Cardiff University School of Medicine, Heath Park, Cardiff CF14 4XN, United Kingdom; ‡Department of Psychiatry, Mount Sinai School of Medicine, New York, NY 10021; and §Mental Illness Research, Education and Clinical Centers, Bronx Veterans Affairs Medical Center, New York, NY 10468

Edited by Akira Sawa, Johns Hopkins University School of Medicine, Baltimore, MD, and accepted by the Editorial Board June 26, 2006 (received for review April 13, 2006)

Abnormal oligodendrocyte function has been postulated as a primary etiological event in schizophrenia. *Oligodendrocyte lineage transcription factor 2 (OLIG2)* encodes a transcription factor central to oligodendrocyte development. Analysis of *OLIG2* in a case-control sample ($n = \approx$1,400) in the U.K. revealed several SNPs to be associated with schizophrenia (minimum $P = 0.0001$, genewide $P = 0.0009$). To obtain independent support for this association, we sought evidence for genetic interaction between *OLIG2* and three genes of relevance to oligodendrocyte function for which we have reported evidence for association with schizophrenia: *CNP*, *NRG1*, and *ERBB4*. We found interaction effects on disease risk between *OLIG2* and *CNP* (minimum $P = 0.0001$, corrected $P = 0.008$) for interaction with *ERBB4* (minimum $P = 0.002$, corrected $P = 0.04$) but no evidence for interaction with *NRG1*. To investigate the biological plausibility of the interactions, we sought correlations between the expression of the genes. The results were similar to those of the genetic interaction analysis. *OLIG2* expression significantly correlated in cerebral cortex with *CNP* ($P < 10^{-7}$) and *ERBB4* ($P = 0.002$, corrected $P = 0.038$) but not *NRG1*. In mouse striatum, *Olig2* and *Cnp* expression also was correlated, and linkage analysis for trans-effects on gene expression suggests that each locus regulates the other's expression. Our data provide strong convergent evidence that variation in *OLIG2* confers susceptibility to schizophrenia alone and as part of a network of genes implicated in oligodendrocyte function.

association | oligodentrocyte/myelin-related genes

Schizophrenia is a major psychiatric disorder characterized by disturbances of perception, emotion, social functioning, and cognition. Its etiology includes a strong heritable component (1), but despite some successes in identifying susceptibility genes (2) the fundamental pathophysiology remains uncertain.

Global surveys of mRNA expression can offer insights into potential pathophysiological pathways, even in tissues as complex as postmortem human brain. A prominent example is the identification of altered expression of *ERBB3* in schizophrenia by independent groups (3–5), a finding of likely pathophysiological relevance given that its ERBB3 is one of two receptors that directly bind neuregulin 1, whose cognate gene (*NRG1*) is strongly implicated as a susceptibility gene for schizophrenia (2, 6).

One of the most widely replicated groups of genes with altered expression in schizophrenia relate to oligodendrocyte function and myelination, oligodendrocytes being the myelinating cells in the brain (3–5, 7–12). These data are compatible with the considerable evidence for altered myelination and oligodendrocyte function in schizophrenia (13). There is therefore a strong rationale to target for genetic analysis oligodendrocyte/myelination related (OMR) genes. Here, we report strong data concerning a key OMR target, *oligodendrocyte lineage transcription factor 2* (*OLIG2*).

OLIG2 is a basic helix–loop–helix transcription factor and, among OMR genes, is an excellent candidate gene for schizophrenia. First, several studies report reduced *OLIG2* mRNA in the postmortem schizophrenic brain (4, 11, 12). Second, OLIG2 might affect the expression of many other OMR genes because it influences precursor (15–18) as well as fully matured (19) oligodendrocytes and is both necessary and sufficient for the genesis of oligodendrocytes and myelination (20–22). If altered OMR gene expression points to an etiological mechanism in schizophrenia, a parsimonious model is that susceptibility variants occur in one or a few OMR genes and that many of the other changes are secondary. Given its role as a master regulator of oligodendrocyte lineages (20, 23), *OLIG2* is a prime candidate for hosting susceptibility variants with wide-ranging secondary effects on *OMR* gene expression.

We initially sought evidence for association to schizophrenia by genotyping a dense panel of markers across *OLIG2* in pooled DNA samples. Associated markers were then individually genotyped to confirm the association. We repeated the process after *de novo* polymorphism discovery to identify SNPs showing the strongest evidence for association. We also sought supportive evidence for the hypothesis by looking for epistasis (i.e., genetic interaction) between *OLIG2* and *2′,3′cyclic nucleotide 3′-phosphodiesterase* (*CNP*), another OMR gene critical for oligodendrocyte function for which we previously reported evidence for genetic association with schizophrenia (24). Moreover, given the strong evidence that *neuregulin 1* (*NRG1*) is a susceptibility gene for schizophrenia (6), the (weaker) evidence for the gene encoding the tyrosine kinase NRG1 receptor ERBB4 (25, 26) and the evidence that, among NRG1's many functions, it influences oligodendrocyte development and maturation (27) via erbB signaling (28), we also sought evidence for a functional relationship between *OLIG2* and these genes by looking for epistatic effects on risk. Finally, we sought functional evidence to corroborate the genetic interactions we observed by testing for correlations between the expression of *OLIG2* and *CNP*, *NRG1*, and *ERBB4* in postmortem human brain samples and in mouse brain. Our data show that *OLIG2* is associated with schizophrenia, that interactive effects on disease risk exist between *OLIG2* and both *CNP* and *ERBB4*, and that

Conflict of interest statement: No conflicts declared.

This paper was submitted directly (Track II) to the PNAS office. A.S. is a guest editor invited by the Editorial Board.

Abbreviations: OMR, oligodendrocyte/myelination related; CEU, Caucasian European Utah; LD, linkage disequilibrium.

¶To whom correspondence may be addressed. E-mail: odonovanmc@cf.ac.uk or owenmj@cf.ac.uk.

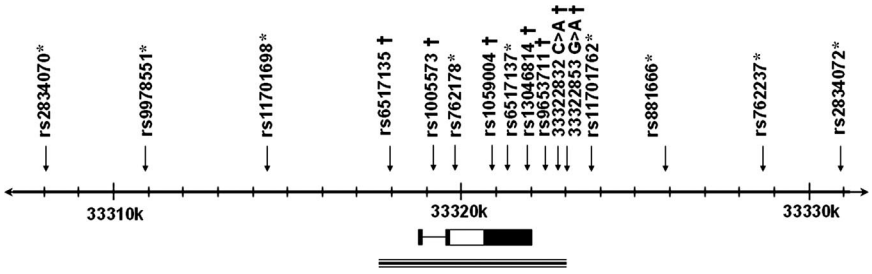

Fig. 1. Schematic of OLIG2 and marker position relative to position (in basses) given in HapMap. Phase 1 (*) and 2 (†) markers. The bottom line depicts region screened for polymorphisms.

OLIG2 expression is correlated with that of *CNP* and *ERBB4* in human brain and with *CNP* in mouse brain. Finally, linkage analysis for trans-effects on gene expression suggest that *OLIG2* and *CNP* mutually regulate each other. Our data provide strong and convergent evidence that variation in *OLIG2* confers susceptibility to schizophrenia alone and as part of a network of genes implicated in myelination and oligodendrocyte function.

Results

Association Analysis. In phase 1, we genotyped in case and control pools nine public database SNPs spanning *OLIG2* (3.2 kb) plus ≈10 kb both 5′ and 3′. The positions of these markers are indicated by the asterisks in Fig. 1 and the results of the analyses by the asterisks in Table 1. Three markers (rs2834070, rs762178, and rs881666) yielded significant evidence ($P \leq 0.05$). Individual genotyping (indicated by the asterisks in Table 2) confirmed significant association for two of these markers, with the strongest evidence for rs762178 ($P = 0.0005$).

In phase 2, DNA from 14 people with schizophrenia was screened for polymorphisms. We examined the full *OLIG2* genomic sequence plus ≈2 kb of the 5′ flanking sequence and 1 kb of 3′ flanking sequence (Fig. 1). This analysis revealed 23 additional SNPs. We were unable to design pooling assays for two markers (rs1005573 and 33322853 G→A). We discarded 16 SNPs because of undetectable minor allele frequencies in pools. None of the SNPs was nonsynonymous or was predicted to change splicing. The pooled analyses of the remaining five markers are listed in Table 1, and their positions are indicated in Fig. 1 (daggers). Three of the SNPs were associated in pools: rs1059004 ($P = 0.0001$), rs13046814 ($P = 0.013$), and rs9653711 ($P = 0.007$).

To identify nonredundant markers for individual genotyping, we individually genotyped in the 30 reference Caucasian European Utah (CEU) parent–offspring trios being used by the HapMap (www.hapmap.org) all of the informative markers plus the two markers we had intended to analyze but for which we had been unable to design robust pooling assays (Table 4, which is published as supporting information on the PNAS web site). Perfect linkage disequilibrium (LD) ($r^2 = 1$) was observed between rs1059004 and rs9653711, so only the former was taken forward for genotyping. Thus, of the significant SNPs from phase 2, we genotyped rs13046814 and rs1059004. We also typed the two SNPs that we had been unable to interrogate by pooling.

Individual genotyping essentially confirmed the data from pooled samples (Table 2, indicated by daggers). The strongest evidence was obtained with rs1059004 ($P = 0.0001$). To correct for multiple testing, we estimated the effective number of independent SNPs (*Meff*) by using the method of Nyholt (29) in the CEU data set of all 16 informative SNPs. *Meff* was estimated at 9.04, giving a Bonferroni-corrected threshold for association of $P = 0.0057$, which was surpassed by three markers (Table 2), with the strongest evidence from rs1059004 (Bonferroni correction for $Meff = 9$, $P = 0.0009$). Haplotype analyses revealed no stronger evidence for association (data not shown). All genotypes were in Hardy–Weinberg equilibrium for cases and controls except for rs1059004 in cases only ($P = 0.011$); however, given 14 Hardy–Weinberg equilibrium tests, this finding could be attributable to chance. Genotyping error as a cause is effectively excluded. First, our genotyping error with the Sequenom system is ≈0.3%, and, in the present study, all CEU genotypes matched those in HapMap. Second, rs762178, which is in strong LD with rs1059004 (case–control sample, $r^2 = 0.93$, $D' = 1$), gives a similar association with two entirely independent genotyping platforms. Third, several

Table 1. Summarized pooled genotyping

SNP ID	Position	Distance to next SNP, bases	Nucleotide change	MAF		Difference	χ^2	*P*
				Cases	Controls			
rs2834070*	33307848	—	T/G	0.390	0.336	0.055	9.67	0.002
rs9978551*	33310829	2,981	G/C	0.067	0.070	−0.003	0.08	0.78
rs11701698*	33314341	3,512	C/A	0.215	0.189	0.026	2.87	0.09
rs6517135†	33317967	979	G/A	0.129	0.147	−0.018	1.90	0.17
rs1005573†	33319112	1,145	G/A	Pooling	Failed	—	—	—
rs762178*	33319797	685	T/C	0.391	0.461	−0.070	12.94	0.0003
rs1059004†	33320859	1,062	C/A	0.425	0.500	−0.076	16.31	0.0001
rs6517137*	33321175	316	G/A	0.100	0.099	0.002	0.02	0.88
rs13046814†	33321773	598	G/T	0.241	0.277	−0.036	4.65	0.03
rs9653711†	33322345	572	C/G	0.425	0.375	0.050	7.36	0.007
33322832 C→A†	33322832	487	A/C	0.054	0.051	0.002	0.06	0.80
33322853 G→A†	33322853	21	A/G	Pooling	Failed	—	—	—
rs11701762*	33323622	769	T/C	0.137	0.116	0.021	3.06	0.08
rs881666*	33325662	2,040	C/G	0.430	0.392	0.039	4.37	0.04
rs762237*	33328573	2,911	C/T	0.384	0.367	0.017	0.87	0.35
rs2834072*	33330860	2,287	G/A	0.485	0.483	0.002	0.08	0.78

Nucleotide change shows the minor allele first. MAF, minor allele frequency in pools of 648 subjects with schizophrenia and 712 controls; —, no data obtained.

*Phase I.

†Identified *de novo*.

Table 2. Individual genotyping

Marker	Distance, bp	Nucleotide change	No. of individuals, cases/controls	Alleles MAF, cases/controls	$\chi 2$	P
rs2834070*	0	A/C	634/685	0.342/0.298	5.99	0.014
rs1005573†	11,264	C/T	660/710	0.302/0.347	6.284	0.012‡
rs762178*	685	A/G	652/705	0.386/0.452	12.13	0.0005
rs1059004†	1,062	C/A	646/704	0.400/0.470	13.43	0.0001
13046814†	914	G/T	662/710	0.393/0.335	10.06	0.002
33322853 G→A†	1,080	A/G	658/698	0.197/0.192	0.10	0.751‡
rs881666*	2,809	G/C	656/700	0.440/0.405	3.36	0.067

Nucleotide change shows the minor allele first. MAF, minor allele frequency.
*Phase I.
†Identified *de novo*.
‡No pooling assay.

markers in weaker LD with rs1059004 also are significantly associated. Fourth, rs1059004 (and several other markers) also give strong evidence for association in pools.

Interaction Analyses. To seek independent evidence for the hypothesis of altered OMR function in schizophrenia, we looked for evidence for epistatic effects on disease risk between *OLIG2* and other OMR-related genes previously associated with schizophrenia in our sample. Of 70 possible pairwise interactions between *OLIG2* and *CNP*, six were nominally significant, with the strongest evidence ($P = 0.0001$) for interaction additional to main effects being between *OLIG2* rs1005573 and *CNP* rs10540926. Permutation testing gave a corrected gene–gene-wide P value of 0.008. Inspection of the 3×3 genotype contingency tables constructed from cases and controls revealed that the genotypes at each locus were independent in controls (χ^2 5.5, 4 df, $P = 0.238$) but not in cases ($\chi^2 = 15.1$, 4 df, $P = 0.004$).

No evidence for interaction was found between the Icelandic haplotype of *NRG1* and *OLIG2*. However, for *OLIG2* and *ERBB4*, the main-effects-plus-interaction-terms model was significantly superior to a main-effects-only model for several pairings, with the strongest evidence coming from *ERBB4* rs6723461 and *OLIG2* rs1005573 ($P = 0.002$). The gene–gene-wide evidence for interaction remained significant ($P = 0.04$). The genotype contingency tables again revealed that the genotypes at each locus were independent in controls ($P = 0.334$) but not cases ($P = 0.001$).

To visualize the interaction terms, the odds ratios for each genotype–genotype pairing were calculated. Several individual odds ratios were significant in each analysis (Tables 5 and 6, which are published as supporting information on the PNAS web site).

We previously found no evidence for population stratification within the samples based on the distribution of P values obtained from genotyping pooled samples for >300 SNPs (30). We also tested for evidence of substructure in approximately one-third of our sample with STRUCTURE (31) by using 97 SNPs scattered across the genome and 1,000 SNPs targeted to regions on chromosomes 10 and 22. No evidence was found under presumed subpopulation numbers, $k = 1, 2, 3, 4$, and 5 (unpublished data). To ensure the absence of levels of substructure that might influence interaction analysis, we set up 16,500 interaction tests by using a browser-based interaction tool (Genetic Association Interaction Analysis web application, available at www.bbu.cf.ac.uk/html/research/biostats.htm). We avoided interactions between genes we postulate as OMR but included genes that others have proposed to be involved in schizophrenia susceptibility (e.g., *COMT*, *GRM3*, *AKT1*, and *G72*). Because there may be true interactions between these genes, we expect this analysis to be conservative. However, the observed distribution of P values is approximately as expected under the null hypothesis (Table 7, which is published as supporting information on the PNAS web site), indicating no appreciable influence of hidden population structure on the analysis.

Analysis of Gene Expression. To seek a biological mechanism that might underpin the interaction, we looked for correlation between mRNA levels. Human brain (67 human motor or prefrontal association cortex, 33 caudate, and 70 cerebellum) Affymetrix U133A and B GeneChip expression data (32) (Gene Expression Omnibus accession no. GSE3790) were acquired for two probe sets for *OLIG2*, five probe sets for *NRG1*, three for *ERBB4*, and one for *CNP*. The probe sets, correlations with *OLIG2*, and levels of significance adjusted by permutation for all pairs of probe sets tested for each pair of genes in cortex are given in Table 3. The data for cerebellum and caudate are available in Tables 8 and 9, which are published as supporting information on the PNAS web site.

Expression of *OLIG2* for both probe sets was positively and highly significantly correlated with that of *CNP*, even allowing for multiple testing in cortex ($P < 10^{-7}$). A similar pattern was seen in caudate and cerebellum, with probe *OLIG2 213825_at* showing a stronger and statistically more significant correlation in each tissue. For one probe set each of *OLIG2* and *ERBB4*, expression was

Table 3. Correlations in gene expression in human cerebral cortex between *OLIG2* and *NRG1*, *ERBB4*, and *CNP*

Probe set IDs	OLIG2 213824_at Correlation	OLIG2 213824_at P	OLIG2 213825_at Correlation	OLIG2 213825_at P
NRG1 206237_s_at	0.27	0.19	0.02	1.00
NRG1 206343_s_at	−0.05	0.99	−0.06	1.00
NRG1 208230_s_at	0.13	1.00*	−0.003	1.00†
NRG1 208231_at	0.11	0.97	0.11	0.98
NRG1 208241_at	−0.001	1.00	−0.05	1.00
ERBB4 206794_at	0.06	1.00	−0.04	1.00
ERBB4 214053_at	−0.23	0.29‡	−0.10	0.95
ERBB4 233498_at	−0.36	0.015§	−0.09	0.96
CNP 208912_s_at	0.60	$<10^{-7}$¶	0.73	$<10^{-7}$¶
OLIG2 213824_at	n/a	n/a	0.46	$<10^{-7}$¶

Shown are Pearson correlation values and corrected P values for all correlations in each gene–gene pairing. Correcting for all tests in cortex, experiment-wide significances for correlations are *OLIG2*/*NRG1* ($P = 0.33$), *OLIG2*/*ERBB4* ($P = 0.04$), and *OLIG2*/*CNP* ($P < 10^{-7}$). n/a, not applicable.
*Uncorrected single test [$P = 0.31$, compare cerebellum (significant for this test) and caudate (not significant)].
†Uncorrected single test [$P = 0.98$, compare cerebellum and caudate (both significantly correlated but in opposite directions for this test)].
‡Uncorrected single test [$P = 0.06$, compare cerebellum and caudate (both significantly negatively correlated for this test)].
§Uncorrected single tests in cerebellum and caudate not significant for this test.
¶Significant in all samples.

negatively correlated in cortex (Table 3), but probe set *ERBB4 233498_at* matches mRNA *AK024204*, which is located in intron 1 of *ERBB4* and does not contribute to any yet known *ERBB4* transcript. That probe pair did not significantly correlate in either caudate or cerebellum (Tables 8 and 9), although, in cerebellum, a similar trend was seen ($r = -0.21$, uncorrected $P = 0.08$). However, the data from the *OLIG2* probe that negatively correlates with the signal at *ERBB4 233498_at* in cortex also shows a significant negative correlation with *ERBB4 214053_at* in cerebellum and caudate [respectively, $r = -0.41$ and $r = -0.44$ and (corrected for all probe pairings between genes) $P = 0.002$ and $P = 0.05$]. This pairing in cortex also shows a negative correlation ($r = -0.23$) (Table 3) and a strong trend (uncorrected), with $P = 0.06$. Probe set *ERBB4 214053_at* matches an *ERBB4* transcript that has an extended 3′ UTR (*ACEview* transcript *ERBB4.aAug05*) relative to the reference sequence.

We did not observe significant correlation between *OLIG2* and *NRG1* expression in cortex.

That the data differ between different probe sets of *ERBB4* is not surprising, because *ERBB4* has multiple transcripts. However, the discrepancies between the *OLIG2* probes requires explanation because this gene is reported to have a single transcript. To try to explain the results, we undertook 3′ RACE with human brain mRNA. The results revealed the use of an alternative polyadenylation signal in human cerebral cortex resulting in two 3′UTRs. For the shorter transcript, base 2061 is the final base before polyadenylation (thus the sequence ends AATTAAAAGGCAGTTGCTGTGGAAAAAA), a finding compatible with the existence of a putative polyadenylation signal at bases 2,040–2,046 of reference sequence NM_005806.2. Probe set *OLIG2* 213824_at is represented in both mRNA species, whereas *OLIG2* 213825_at is represented only in the longer of the two mRNA species.

Exploration of the BXD mouse expression databases in WebQTL (www.genenetwork.org) also revealed correlations between *Cnp1* and *Olig2* expression in whole brain [Integrative Neuroscience Initiative on Alcoholism (INIA) Brain mRNA M430 (Apr05) robust multiarray (RMA)], hippocampus [Hippocampus Consortium M430v2 BXD (Dec05) RMA Database], and striatum [INIA Brain mRNA M430 (Apr05) RMA Database], where the correlation was strongest ($r = 0.64$, $P = 6 \times 10^{-5}$). Moreover, when we used the *Olig2* mRNA as a phenotype for linkage in striatum, we detected two linkage peaks that met the criteria for genome-wide significant linkage, with the strongest evidence (likelihood ratio statistic = 17.9) on mouse chromosome 11 maximizing 4Mb from *Cnp*. Conversely, when we used the *Cnp* probe set (1418980_a_at) that shows the strongest correlation with *Olig2* expression as the phenotype, one of the strongest linkages (likelihood ratio statistic ≈ 6) maximized within 3 megabases of *Olig2* on chromosome 16, although no genome-wide significant linkages were observed. Only in hippocampus did we find significant evidence for correlation between *Olig2* and *Erbb4*, but, in contrast to the human data, the correlation was positive ($r = 0.28$, $P = 0.016$). Because we looked in four data sets (whole brain, hippocampus, cerebellum, and striatum) this finding would fail to be significant when corrected for multiple testing. No data were available for any *Nrg1* transcript.

Discussion

Most schizophrenia research has assumed the neuron to be the primary locus of the molecular pathology, but in recent years abnormal glial (astrocyte, oligodendrocyte, and microglia) function has also been proposed (14, 33). Several lines of evidence now support the hypothesis that abnormal oligodendrocyte function and myelination occurs in schizophrenia (13). Of primary relevance to the present paper are the repeated reports of reduced expression of OMR genes in postmortem brains of people with schizophrenia (3–5, 11). Other data favoring the hypothesis include the observations of morphologic abnormalities and the reduced number of oligodendrocytes in postmortem brains of people with schizophrenia (34–37) and neuroimaging studies indicating changes in white matter volume and organization (13).

As a transcription factor central in the life cycle of oligodendrocytes and a gene whose expression has been shown in several studies to be reduced in schizophrenic brain, *OLIG2* is a plausible OMR candidate gene for schizophrenia. *OLIG2* maps to 21q22.11. There are several reports of deletion in this region in people with schizophrenia (see ref. 38) and of a low risk of schizophrenia in people with trisomy 21 (39). The potential for chromosomal rearrangements to contribute to the identification of a schizophrenia locus has recently been illustrated, although it has been observed in only two individuals (40).

Individual genotyping revealed that, among several associated markers, rs1059004 had a nominal significance level of $P = 0.0001$, easily surpassing the experiment-wide threshold of 0.0057 and corresponding to a conservatively corrected P value (Bonferroni correction of *Meff* = 9) of 0.0009 (41, 42). Moreover, we have obtained further evidence in support of association between *OLIG2* and schizophrenia from our studies of genetic interaction.

We previously reported evidence for association between *NRG1* (43), *CNP* (24), and *ERBB4* (25) in schizophrenia. The exact function of CNP in brain is unknown, but it has an important role in oligodendrocyte development, regulating oligodendrocyte process outgrowth during myelination, and in axonal maintenance (44). Among its many functions, NRG1 influences oligodendrocyte development and maturation through ERBB signaling (27). There is no direct evidence for physical interaction between the products of these genes and that of *OLIG2*. However, given that all of these genes play a roll in oligodendrocyte function, we looked for genetic interaction and correlated expression between *OLIG2* and these putative schizophrenia susceptibility genes to seek supportive evidence for *OLIG2*.

It is striking that our genetic interaction data and the expression correlation data show similar patterns. Thus, *OLIG2* shows both significant evidence of interaction affecting disease risk and correlation in expression with *CNP* and *ErbB4* but not *NRG1*. The findings with *CNP* are particularly strong and are robust to correction for all gene–gene interactions tested (all SNP pairs tested in *OLIG2*/*CNP* and a total of three gene–gene tests), whereas, for *ERBB4*, the findings are significant only at a gene–gene-wide threshold. Because the tests for interaction between the genes allow for main effects, the interaction evidence provides independent support for those genes as susceptibility genes for schizophrenia.

From the human expression data, we only infer a broad functional relationship between the genes. This relationship may even be limited to their expression being dominated by a particular cell type rather than close functional relationships. However, the observation of a consistently stronger and more robustly significant correlation between *CNP* and the long mRNA form of *OLIG2* as accessed by *OLIG2 213825_at* compared with that of all (known) *OLIG2* mRNA as measured by *OLIG2 213824_at* suggests a more specific relationship between the expression of the two genes than that arising simply from variation in cell number. It is also of interest that, in an earlier study (12), both *CNP* and *OLIG2* mRNA levels correlated with the methylation status of *SOX10*, again suggesting an intimate link between the coordinated expression of both genes. Finally, our analysis of the WebQTL expression and linkage data in mouse also support the hypothesis of a functional relationship between *Olig2* and *Cnp*. Thus, in striatum, not only does *Olig2* and *Cnp* expression correlate, but the region of the genome showing the strongest evidence for linkage to *Olig2* mRNA levels is only 4 megabases away from the *Cnp* locus. Moreover, given the focused nature of the hypothesis and the rest of the data presented here, that even a weak linkage peak for *Cnp* expression maps within a few megabases of *Olig2* is clearly of interest. Taken together, these correlation and linkage data suggest that the expression of *Olig2* and *Cnp* are functionally related and that each locus might mutually regulate the other.

Although statistically weaker, the negative correlation between *ERBB4* and *OLIG2* in all three human brain regions studied is intriguing given that the effects of neuregulin through ERBB signaling on oligodendrocyte lineage development depend on the pattern of ERBB receptor expression, with ERBB4 signaling being inhibitory (28). Thus our expression correlation data are consistent with the known biology of NRG1/ERBB4 signaling. However, we have been unable to find any strong supportive evidence for coordinated expression between *Olig2* and *Erbb4* in mouse, and the evidence we did observe was for positive correlation between expression of the two genes.

One of the major confounders of case–control studies is hidden population structure and inadequate matching of cases and controls. We have addressed this problem (ref. 30 and unpublished data), including a specific test for hidden structure in the sample that might have yielded an inflated type 1 error rate for gene–gene interactions. Thus, we performed a large number of analyses with which to approximate the null distribution of interaction *P* values. No inflation in type 1 error was observed, which strongly suggests that our findings are not explained by population stratification.

The data presented here provide a coherent and strong statistical case for the hypothesis that *OLIG2* is a susceptibility gene for schizophrenia, but the mechanistic inferences to be drawn are yet unclear. In general, conclusions about genetic models cannot easily be drawn from statistical interactions (45). Also, in this data set, the markers showing greatest evidence for gene–gene interaction are not the same as those showing the most significant main effects at each locus. However, even when all markers in a gene have been examined, the marker showing the strongest evidence for an association may not actually be the functional variant *per se*, because the evidence of association for a marker allele in high LD with the true risk variant varies not only on the population frequency of the risk variant but also on chance fluctuations in the frequencies of both the risk variant in the sample and any non-risk haplotypes carrying the risk-tagging marker allele. In the present study, the strong LD between the markers showing the strongest main effects and those showing the strongest interacting effects (Table 4) suggests that it is at least possible that each analysis simply varies in the extent to which it extracts information originating from the same as yet unmeasured functional SNP.

In support of the hypothesis of a single functional locus, haplotypes constructed from the main risk and main interacting alleles at *OLIG2* and *CNP* are significantly associated with schizophrenia to a similar extent as those main effect loci in each gene (data not shown). However, we must consider the alternative explanation that the markers showing main effects and the markers showing interactive effects capture information from different functional variants. Given that the effects of *OLIG2* on oligodendrocytes require complex patterns of activation, inactivation, and timing (23), it is possible that different alleles result in nonoptimal expression during different developmental stages, some of which might also require *CNP* and/or *ERBB4* expression, whereas others do not. This scenario might well give rise to alleles showing main effects and other interactive effects.

It is likely that human genetic studies cannot answer fundamental mechanistic questions of this nature. Nevertheless, although noting that replication is always essential as are further studies of potentially overlapping phenotypes beyond the core operational definitions (46), our data provide strong and convergent evidence that variation in *OLIG2* confers susceptibility to schizophrenia alone and as part of a network of genes implicated in myelination and oligodendrocyte function. As such, these findings suggest that disturbances in oligodendrocyte function are central to the pathogenesis of schizophrenia as recently proposed (13). One caveat is that, although *OLIG2* expression is necessary for oligodendrocyte development, it is also involved in several stages of astrocyte development (23) and in suppressing neurogenesis (47). It is therefore still possible that underpinning our findings is interplay of oligodendrocyte, neuronal, and other glial disease mechanisms (14).

Methods

Subjects. Our case sample consisted of 673 unrelated subjects with schizophrenia (455 males). All were white, born in the U.K. or Ireland, and had a consensus diagnosis of schizophrenia according to Diagnostic and Statistical Manual of Mental Disorders IV criteria. Full details concerning diagnostic practices and demographics are reported elsewhere (24). The 716 controls (482 males) were blood donors matched to cases for age, sex, and ethnicity. The use of unscreened controls does not affect power for disorder with the prevalence of schizophrenia (48). Donors were not taking regular medication and were not remunerated for expenses. Any phenotype enrichment in donors is likely to be for altruism and better than average health, not for socioeconomically disadvantaged groups who may have relatively high rates of psychiatric disorder. It is therefore unlikely that such enrichment would have influenced our results. Multicenter and Local Research Ethics Committee approval was obtained. All participants gave written informed consent.

Genotyping. The SNaPshot (Applied Biosystems, Foster City, CA) protocol for pooled DNA genotyping has been described in detail (49).

Genotyping call rates in cases and controls were, respectively, 97.5% and 97.7%. Genotyping was performed with the Amplifluor (Serologicals, Temecula, CA), MassARRAY, and iPlex systems (Sequenom, San Diego, CA) according to manufacturers' recommendations. Assays were optimized in 30 CEU trios. All plates contained cases, controls, blanks, and CEU samples. Genotypes were called in duplicate blind to sample identity and blind to the other rater. Assays were only considered suitable if, during optimization, our own data from CEU individuals were identical to those in the HapMap.

Polymorphism Detection. Primers were designed based on alignment of mRNA sequence (NM_005806) and genomic sequence (NT_011512). We used a denaturing high-performance liquid chromatography protocol followed by sequencing to screen 14 unrelated white subjects with schizophrenia (all from the U.K.), of whom 10 had at least one copy of the risk allele at rs1059004. In a blind analysis, this protocol had 100% sensitivity (50), making it unlikely that any of the screening subjects had a variant in the screened region that we did not detect.

Statistics. Pooled data were tested for association by using contingency tables created by multiplying twice the number of individuals represented in each pool by the estimated allele frequencies. Contingency tables were also used for single-marker case–control analysis. Haplotypes were analyzed with UNPHASED (51). Genotypes were tested for Hardy–Weinberg equilibrium using a χ^2 goodness-of-fit test. Analyses of LD between markers (r^2 and D') were performed by using Haploview (52). The number of effective independent SNPs assayed was estimated by the spectral decomposition method of Nyholt (29).

Interaction Analysis. We genotyped a further six markers in our case–control sample in addition to the four markers we genotyped earlier (24). These markers were selected by an entropy algorithm (24) to capture 95% of the genetic diversity represented by all markers we had identified by sequencing 11 kb of *CNP* genomic sequence in 14 unrelated individuals with schizophrenia. The *CNP* data are given in Table 10, which is published as supporting information on the PNAS web site. As anticipated from earlier pooled analyses (24), none of the additional markers was associated with schizophrenia.

ERBB4 and *NRG1* each span >1 megabase, which prohibited

systematic gene-based interaction studies. For *NRG1*, we focused on the Icelandic haplotype defined as alleles 1, 0, 0 of markers SNP8NRG221533, 478B14-848, and 420M9-1935, respectively, from the original study (6) as presented in this sample before (43). For *ERBB4*, analyses were based on three SNPs we had previously genotyped during a direct association study (25).

Interaction analysis was performed by using logistic regression (45). Markers were coded in terms of additive and dominance components of the genotype, and then two logistic regression models [main effects (4 df) and main effects plus interaction terms (8df)] were fitted and compared by using the likelihood ratio test (by using glm, a function in the R statistical package). The methodology for dealing with NRG1 haplotypes has been described before (43). Empirical interaction *P* values were calculated by permuting affection status and, each time, determining whether and to what extent the model with interactive terms was significantly superior to that with main effects. Then, for every possible between-gene marker–marker pairing, we divided the number of times the smallest *P* value was smaller in a simulated data set than in the specific marker–marker test by the number of simulated data sets ($n = 10{,}000$).

Gene Expression. Pearson's correlation coefficients between OMR genes were determined in brain RNA samples hybridized to Affymetrix (Santa Clara, CA) U133A and B GeneChips. Full details of this data set are given in ref. 32 (Gene Expression Omnibus accession no. GSE3790). Probe set level summaries were quantified by robust multiarray analysis (53) using the Affymetrix package (54). Gene-wide and experiment-wide significance levels were assessed by permutation analysis using 10^7 iterations. All probe sets for each gene were analyzed except those identified at only a poor level of confidence (Grade C, D, or E annotation) or with multiple hits in the genome.

We performed 3′RACE experiments by using FirstChoice total RNA (Ambion, Huntingdon, U.K.) from human frontal, temporal, and parietal cortex, with the *OLIG2*-specific primer: 3′-AGAACCACTTGTGGATTGGAA-5′. All amplified products were sequenced to confirm their identity.

We used BXD expression databases in WebQTL corresponding to whole brain, cerebellum, and striatum. Correlations between gene expression were sought by using the browser interface provided (Pearson's Product Moment option). We performed linkage mapping of mRNAs using the default options.

This work was supported by the National Institute of Mental Health Centers for the Neuroscience of Mental Disorders (K.L.D.) and by the U.K. Medical Research Council. V.M. was supported by a Research Councils U.K. fellowship. S.M. was supported by the Higher Education Funding Council for Wales.

1. Gottesman, I. I. (1991) in *Schizophrenia Genesis: The Origins of Madness* (Freeman, New York).
2. Owen, M. J., Craddock, N. & O'Donovan, M. C. (2005) *Trends Genet.* **21,** 518–525.
3. Hakak, Y., Walker, J. R., Li, C., Wong, W. H., Davis, K. L., Buxbaum, J. D., Haroutunian, V. & Fienberg, A. A. (2001) *Proc. Natl. Acad. Sci.* USA **98,** 4746–4751.
4. Tkachev, D., Mimmack, M. L., Ryan, M. M., Wayland, M., Freeman, T., Jones, P. B., Starkey, M., Webster, M. J., Yolken, R. H. & Bahn, S. (2003) *Lancet* **362,** 798–805.
5. Aston, C., Jiang, L. & Sokolov, B. P. (2004) *J. Neurosci. Res.* **77,** 858–866.
6. Stefansson, H., Sigurdsson, E., Steinthorsdottir, V., Bjornsdottir, S., Sigmundsson, T., Ghosh, S., Brynjolfsson, J., Gunnarsdottir, S., Ivarsson, O., Chou, T. T., *et al.* (2002) *Am. J. Hum. Genet.* **71,** 877–892.
7. Novak, G., Kim, D., Seeman, P. & Tallerico, T. (2002) *Brain Res. Mol. Brain Res.* **107,** 183–189.
8. Flynn, S. W., Lang, D. J., Mackay, A. L., Goghari, V., Vavasour, I. M., Whittall, K. P., Smith, G. N., Arango, V., Mann, J. J., Dwork, A. J., *et al.* (2003) *Mol. Psychiatry* **8,** 811–820.
9. Sugai, T., Kawamura, M., Iritani, S., Araki, K., Makifuchi, T., Imai, C., Nakamura, R., Kakita, A., Takahashi, H. & Nawa, H. (2004) *Ann. N.Y. Acad. Sci.* **1025,** 84–91.
10. Dracheva, S., Davis, K. L., Chin, B., Woo, D. A., Schmeidler, J. & Haroutunian, V. (2006) *Neurobiol. Dis.* **21,** 531–540.
11. Katsel, P., Davis, K. L. & Haroutunian, V. (2005) *Schizophr. Res.* **79,** 157–173.
12. Iwamoto, K., Bundo, M., Yamada, K., Takao, H., Iwayama-Shigeno, Y., Yoshikawa, T. & Kato, T. (2005) *J. Neurosci.* **25,** 5376–5381.
13. Davis, K. L., Stewart, D. G., Friedman, J. I., Buchsbaum, M., Harvey, P. D., Hof, P. R., Buxbaum, J. & Haroutunian, V. (2003) *Arch. Gen. Psychiatry* **60,** 443–456.
14. Moises, H. W., Zoega, T. & Gottesman, I. I. (2002) *BMC Psychiatry* **2,** 8.
15. Lu, Q. R., Yuk, D., Alberta, J. A., Zhu, Z., Pawlitzky, I., Chan, J., McMahon, A. P., Stiles, C. D. & Rowitch, D. H. (2000) *Neuron* **25,** 317–329.
16. Takebayashi, H., Nabeshima, Y., Yoshida, S., Chisaka, O., Ikenaka, K. & Nabeshima, Y. (2002) *Curr. Biol.* **12,** 1157–1163.
17. Zhou, Q., Wang, S. & Anderson, D. J. (2000) *Neuron* **25,** 331–343.
18. Lu, Q. R., Sun, T., Zhu, Z., Ma, N., Garcia, M., Stiles, C. D. & Rowitch, D. H. (2002) *Cell* **109,** 75–86.
19. Gokhan, S., Marin-Husstege, M., Yung, S. Y., Fontanez, D., Casaccia-Bonnefil, P. & Mehler, M. F. (2005) *J. Neurosci.* **25,** 8311–8321.
20. Rowitch, D. H., Lu, Q. R., Kessaris, N. & Richardson, W. D. (2002) *Trends Neurosci.* **25,** 417–422.
21. Copray, S., Balasubramaniyan, V., Levenga, J., de Bruijn, J., Liem, R. & Boddeke, E. (2005) *Stem Cells* **24,** 1001–1010.
22. Jakovcevski, I. & Zecevic, N. (2005) *J. Neurosci.* **25,** 10064–10073.
23. Marshall, C. A., Novitch, B. G. & Goldman, J. E. (2005) *J. Neurosci.* **25,** 7289–7298.
24. Pierce, T., Bray, N., Williams, N. M., Norton, N., Moskvina, V., Preece, A., Haroutunian, V., Buxbaum, J., Owen, M. J. & O'Donovan, M. C. (2006) *Arch. Gen. Psychiatry* **63,** 18–24.
25. Norton, N., Moskvina, V., Morris, D. W., Bray, N. J, Zammit, S., Williams, N. M., Williams, H. J., Preece, A. C., Dwyer, S., Wilkinson, J. C., *et al.* (2005) *Neuropsychiatr. Genet.* **141,** 96–101.
26. Silberberg, G., Darvasi, A., Pinkas-Kramarski, R. & Navon, R. (2006) *Am. J. Med. Genet. B* **141,**142–148.
27. Vartanian, T., Fischbach, G. & Miller, R. (1999) *Proc. Natl. Acad. Sci. USA* **96,** 731–735.
28. Sussman, C. R., Vartanian, T. & Miller, R. H. (2005) *J. Neurosci.* **25,** 5757–5762.
29. Nyholt, D. R. (2004) *Am. J. Hum. Genet.* **74,** 765–769.
30. Williams, H. J., Glaser, B., Williams, N. M., Norton, N., Zammit, S., MacGregor, S., Kirov, G. K., Owen, M. J. & O'Donovan, M. C. (2005) *Am. J. Psychiatry* **162,** 1736–1738.
31. Prtichard, J. K., Stephens, M. & Donnelly, P. (2000) *Genetics* **155,** 945–959.
32. Hodges, A., Strand, A. D., Aragaki, A. K., Kuhn, A., Sengstag, T., Hughes, G., Elliston, L. A., Hartog, C., Goldstein, D. R., Thu, D., *et al.* (2006) *Hum. Mol. Genet.* **15,** 965–977.
33. Coyle, J. T. & Schwarcz, R. (2000) *Arch. Gen. Psychiatry* **57,** 90–93.
34. Hof, P. R., Haroutunian, V., Copland, C., Davis, K. L. & Buxbaum, J. D. (2002) *Neurochem. Res.* **27,** 1193–1200.
35. Hof, P. R., Haroutunian, V., Friedrich, V. L., Jr., Byne, W., Buitron, C., Perl, D. P. & Davis, K. L. (2003) *Biol. Psychiatry* **53,** 1075–1085.
36. Uranova, N., Orlovskaya, D., Vikhreva, O., Zimina, I., Kolomeets, N., Vostrikov, V. & Rachmanova, V. (2001) *Brain Res. Bull.* **55,** 597–610.
37. Uranova, N. A., Vostrikov, V. M., Orlovskaya, D. D. & Rachmanova, V. I. (2004). *Schizophr. Res.* **67,** 269–275.
38. Murtagh, A., McTigue, O., Ramsay, L., Hegarty, A. M., Green, A. J., Stallings, R. L. & Corvin, A. (2005) *Schizophr. Res.* **78,** 353–356.
39. Collacott, R. A., Cooper, S. A. & McGrother, C. (1992) *Br. J. Psychiatry* **161,** 671–674.
40. Millar, J. K., Pickard, B. S., Mackie, S., James, R., Christie, S., Buchanan, S. R., Malloy, M. P., Chubb, J. E., Huston, E., Baillie, G. S., *et al*. (2005) *Science* **310,** 1187–1191.
41. Salyakina, D., Seaman, S. R., Browning, B. L., Dudbridge, F. & Muller-Myhsok, B. (2005) *Hum. Hered.* **60,** 19–25.
42. Nyholt, D. R. (2005) *Hum. Hered.* **60,** 61–62.
43. Williams, N. M., Preece, A., Spurlock, G., Norton, N., Williams, H. J., Zammit, S., O'Donovan, M. C. & Owen, M. J. (2003) *Mol. Psychiatry* **8,** 485–487.
44. Lee, J., Gravel, M., Zhang, R., Thibault, P. & Braun, P. E. (2005) *J. Cell Biol.* **170,** 661–673.
45. Cordell, H. J. (2002) *Hum. Mol. Genet.* **11,** 2463–2468.
46. Craddock, N., Owen, M. J. & O'Donovan, M. C. (2006) *Mol. Psychiatry* **11,** 446–458.
47. Buffo, A., Vosko, M. R., Erturk, D., Hamann, G. F., Jucker, M., Rowitch, D. & Gotz, M. (2005) *Proc. Natl. Acad. Sci. USA* **102,** 18183–18188.
48. Moskvina, V., Holmans, P., Schmidt , K. M. & Craddock, N. (2005) *Ann. Hum. Genet.* **69,** 566–576.
49. Norton, N., Williams, N. M., Williams, H. J., Spurlock, G., Kirov, G., Morris, D. W., Hoogendoorn, B., Owen, M. J. & O'Donovan, M. C. (2002) *Hum. Genet.* **110,** 471–478.
50. Jones, A. C., Austin, J., Hansen, N., Hoogendoorn, B., Oefner, P. J., Cheadle, J. P. & O'Donovan, M. C. (1999) *Clin. Chem.* **45,** 1133–1140.
51. Dudbridge, F. (2003) *Genet. Epidemiol.* **25,** 115–121.
52. Barrett, J. C., Fry, B., Maller, J. & Daly, M. J. (2005) *Bioinformatics* **21,** 263–265.
53. Irizarry, R. A., Hobbs, B., Collin, F., Beazer-Barclay, Y. D., Antonellis, K. J., Scherf, U. & Speed, T. P. (2003) *J. Biostat.* **4,** 249–264.
54. Gautier, L., Cope, L., Bolstad, B. M. & Irizarry, R. A. (2004) *J. Bioinformatics* **20,** 307–315.