ORIGINAL ARTICLE

Convergent Evidence for 2′,3′-Cyclic Nucleotide 3′-Phosphodiesterase as a Possible Susceptibility Gene for Schizophrenia

Timothy R. Peirce, PhD; Nicholas J. Bray, PhD; Nigel M. Williams, PhD; Nadine Norton, PhD; Valentina Moskvina, PhD; Anna Preece, Bsc; Vahram Haroutunian, PhD; Joseph D. Buxbaum, PhD; Michael J. Owen, PhD, FRCPsych; Michael C. O'Donovan, PhD, FRCPsych

Context: Convergent data make 2′,3′-cyclic nucleotide 3′-phosphodiesterase (*CNP*) a candidate gene for schizophrenia. Reduced expression has been reported in the schizophrenic brain. The *CNP* gene maps to a region to which we have reported linkage to schizophrenia. Mice in which the *CNP* gene has been knocked out display central nervous system pathological characteristics reminiscent of some features observed in schizophrenia. 2′,3′-Cyclic nucleotide 3′-phosphodiesterase is used as a marker of myelin-forming cells and is detectable in cells of oligodendrocyte lineage throughout life. Because *CNP* is thought to be important for oligodendrocyte function, altered expression has been interpreted as supportive of the hypothesis that altered oligodendrocyte function may be an etiological factor in schizophrenia. However, it is unclear whether the observed changes in the schizophrenic brain are primary or secondary.

Objectives: To determine if *CNP* expression is influenced by DNA polymorphisms and to verify if these polymorphisms are associated with schizophrenia.

Design: Allele-specific messenger RNA expression assay and genetic association studies.

Setting: Unrelated subjects were ascertained from secondary psychiatric inpatient and outpatient services.

Participants: We used brain tissue from 60 anonymous individuals with no known psychiatric disorder; a case-control sample of 708 white individuals from the United Kingdom meeting *DSM-IV* criteria for schizophrenia matched for age, sex, and ethnicity to 711 blood donor controls; and a pedigree with DNA from 6 affected siblings and 1 parent, showing evidence for linkage to *CNP*.

Main Outcome Measures: Association between allele and gene expression. Association between allele and schizophrenia.

Results: The exonic single nucleotide polymorphism rs2070106 was associated with *CNP* expression ($P<.001$). Compatible with underexpression of *CNP* messenger RNA in schizophrenia, the lower-expressing A allele was significantly associated with schizophrenia ($P=.04$) in the case-control sample. All affected individuals in the linked pedigree were homozygous for the lower-expression allele, providing independent support for the association ($P=.03$).

Conclusions: Our data support the hypothesis that reduced *CNP* expression in the schizophrenic brain is relevant to disease etiology and therefore provide support for the general hypothesis that altered oligodendrocyte function is an etiological factor in schizophrenia.

Arch Gen Psychiatry. 2006;63:18-24

Author Affiliations: Department of Psychological Medicine (Drs Peirce, Bray, Williams, Norton, Owen, and O'Donovan and Ms Preece) and Biostatistics and Bioinformatics Unit (Dr Moskvina), School of Medicine, Cardiff University, Cardiff, Wales; Department of Psychiatry, Mount Sinai School of Medicine, New York, NY (Drs Haroutunian and Buxbaum); Mental Illness Research, Education and Clinical Centers, Bronx Veterans Affairs Medical Center, New York (Dr Haroutunian).

METHODS FOR TESTING genes using association designs in general offer more power to detect small effects than methods based on linkage, but until genomewide association scans become economically realistic, such analyses are dependent on prior functional and/or positional support for a disease locus. Unfortunately, our current knowledge of the functional pathological characteristics of schizophrenia so far only provides broad clues about the pathological mechanisms that might be involved, and therefore, it is difficult to define a group of candidate genes small enough to be comprehensively studied.

Global surveys of messenger RNA (mRNA) expression levels offer an alternative approach to linkage for selecting candidate genes and/or pathophysiological pathways for a disease in the absence of a prior hypothesis linking the function of that gene or pathway to the disease etiology.[1] Although the large number of transcripts to be examined, as well as the numerous potential sources of experimental and sampling variance that accompany the use of complex tissues obtained post mortem, poses significant analytic and design problems, recent studies suggest that

robust and etiologically relevant changes in gene expression can be detected even in tissues as complex as the human postmortem brain. Examples include the identification of altered expression of the neuregulin receptor erbB3 in the schizophrenic brain,[2,3] a finding whose relevance is supported by the evidence implicating variation in the gene encoding its ligand, neuregulin 1, in susceptibility to schizophrenia.[4] Others include the identification of the G protein receptor kinase 3 gene as a candidate gene for bipolar disorder[5] and the regulator of G-protein signaling 4 as a candidate gene for schizophrenia,[6] findings that have received supportive evidence from follow-up genetic studies.[7,8] In the cases of regulator of G-protein signaling 4 and the G protein receptor kinase 3 gene, each was selected for genetic analysis because, as well as displaying altered expression, each mapped to a putative region of linkage for the relevant disorder. Thus, each candidate gene was supported by data from 2 non–hypothesis based methods, an approach that has been termed *convergent functional genomics*.[5]

Recently, microarray studies have reported the down-regulation of genes related to oligodendrocyte function and myelination in the schizophrenic brain compared with control subjects.[2,3,9] Among the mRNAs down-regulated was that encoded by the 2′,3′-cyclic nucleotide 3′-phosphodiesterase gene (*CNP*), a finding confirmed at the protein level by Flynn and colleagues.[10] These observations are of interest because *CNP* maps to 17q21.2, a region of the genome in which we have recently observed genomewide significant evidence for linkage (logarithm odds score, 8.32; genomewide empirical P value $\leq .02$) to schizophrenia in a single pedigree.[11] Moreover, given evidence for altered myelination and oligodendrocyte function in schizophrenia,[12] *CNP* is also a plausible functional candidate gene.

2′,3′-Cyclic nucleotide 3′-phosphodiesterase is widely used as a marker protein of myelin-forming glial cells. In brain development, CNP is detectable in cells of oligodendrocyte lineage and is maintained in mature oligodendrocytes throughout life.[13] 2′,3′-Cyclic nucleotide 3′-phosphodiesterase is associated with noncompacted myelin regions, absent from compacted myelin, and also found in oligodendrocyte cytoplasm.[10] Although the precise function of *CNP* in oligodendrocytes is unclear, recent evidence suggests that it interacts with mitochondria and cytoskeletal proteins and may act to promote microtubule assembly or act as a membrane anchor for tubulin.[14] Intriguingly, a recent animal study showed that CNP-deficient mice display a reduction in overall brain size, enlarged ventricles, and corpus callosum atrophy, all of which are consistent with pathological features observed in schizophrenia.[13]

That altered *CNP* expression can be observed and replicated in small samples of cases and controls suggests a schizophrenia-related influence on *CNP* expression that is probably too large and homogeneous to attribute entirely to genetic variation at the *CNP* locus itself. If altered *CNP* expression is relevant to pathogenesis, it is probably as a final common pathway resulting from multiple trans-acting genetic and environmental risk factors. Alternatively, the replicated evidence for the association between altered *CNP* expression and schizophrenia in small samples might suggest that the former is a common consequence of the latter, for example, a result of its treatment.

In this study, we applied molecular genetic approaches in a bid to distinguish between the hypotheses mentioned earlier. The rationale underpinning our study was that if it is true that altered *CNP* expression influences schizophrenia susceptibility, any direct influences on *CNP* expression resulting from a *cis*-acting polymorphism at the *CNP* locus will have a similar effect, albeit probably with a small genetic effect size. Thus, we sought evidence that the expression of *CNP* is influenced by at least 1 *cis*-acting polymorphism using a highly quantitative allele-specific expression assay.[15] Second, we sought direct and indirect evidence for association between *CNP* and schizophrenia in a large sample of schizophrenic cases and controls. Third, we examined the sequence of *CNP* in a family showing evidence for 17q linkage to schizophrenia for the presence of alleles that might represent rare alleles of large effect size. Fourth, we undertook de novo polymorphism discovery across the complete genomic sequence (including introns) of *CNP* and examined all variants for evidence of association with schizophrenia. The analyses presented show that *CNP* is indeed subject to *cis*-acting influences on gene expression, that an allele associated with lower *CNP* expression is also associated with schizophrenia, and that while all affected members of the pedigree showing evidence for linkage to schizophrenia were less identical by descent on 17q than suggested by our original linkage analysis, all were homozygous for the less common putative risk allele, a finding that is unlikely to occur simply by chance.

METHODS

SAMPLE

Our case-control sample consisted of 708 white subjects with schizophrenia from the United Kingdom and Ireland (487 men and 221 women), matched for age, sex and ethnicity to 711 blood donor controls (478 men and 233 women). All patients had a consensus diagnosis of schizophrenia according to *DSM-IV* criteria made by 2 independent raters following a semistructured interview by trained psychiatrists or psychologists using the Schedules for Clinical Assessment in Neuropsychiatry interview[16] and review of case records. High levels of reliability ($\kappa > 0.8$) were achieved between raters for diagnoses. All cases were screened to exclude substance-induced psychotic disorder or psychosis due to a general medical condition. The mean (SD) age at first psychiatric contact was 23.6 (7.7) years and the mean (SD) age at ascertainment was 41.8 (13.5) years. The blood donor controls were not specifically screened for psychiatric illness but individuals were not taking regular prescribed medications. In neither country are blood donors remunerated even for expenses, and as a result, blood donors are not enriched for substance abusers or indigents who may have relatively high rates of psychosis. Ethics committee approval was obtained in all regions where patients were recruited, and informed written consent was obtained from all participants. The pedigree showing evidence for linkage to 17q[11] consisted of a single sibship of 6 individuals with *DSM-IV* schizophrenia (diagnosed as described earlier) from whom DNA was available. DNA was also available from 1 parent of unknown diagnostic status but who was believed to be unaffected.

LABORATORY METHODS

Primers to amplify exons, promoters, flanking sequences, and introns were designed based on alignment of mRNA sequence (NM_033133) and the corresponding genomic sequence (NT_010755.14). Polymerase chain reaction amplification and mutation screening were performed using denaturing high-performance liquid chromatography and, where indicated, sequencing, using protocols we have described extensively.[17,18]

The sample for mutation screening consisted of 14 unrelated white subjects from the United Kingdom meeting *DSM-IV* criteria for schizophrenia, each of whom had at least 1 affected sibling. At a later stage, we also screened all available individuals from the linked pedigree.

ALLELE-SPECIFIC EXPRESSION ASSAY

The principle of the assay is that in heterozygous carriers of a *cis*-acting polymorphism that affects the transcription or stability of a species of mRNA, the quantity of mRNA originating from each gene copy will be unequal. The simplest method for distinguishing between mRNA molecules originating from each copy of a pair of autosomal genes is to use a polymorphism within the mRNA sequence as a copy-specific tag.[15] It is then possible to apply quantitative methods of allele discrimination to mRNA samples originating from individuals who are known to be heterozygous for the marker polymorphism in order to measure relative copy expression.[15] Factors that can influence the measured amount of total expression of a specific gene (eg, tissue preparation, mRNA quality, drug exposure, preagonal state, hormones, effects secondary to regulatory polymorphisms in other genes) will, in the absence of a *cis*-acting allele-specific interaction, influence the amount of mRNA originating from each chromosome equally and are therefore controlled for by this assay. While not allowing the measurement of total RNA abundance, the allelic expression assay does have the major advantage of allowing the detection of small polymorphic *cis*-acting influences even in the face of large trans-acting influences.

Postmortem brain tissue was derived from the frontal, parietal, or temporal cortex of 60 white, European, unrelated, anonymized human adults of whom 50 were from the United Kingdom (the Medical Research Council London Neurodegenerative Diseases Brain Bank, London, England) and 10 were from Sweden (Department of Clinical Neuroscience, Karolinska Institute, Stockholm). All were free from psychiatric or neurological disorder at the time of death. Genomic DNA and RNA were extracted from each individual tissue sample, with subsequent deoxyribonuclease treatment of RNA as described.[15] Heterozygotes for the transcribed marker single-nucleotide polymorphism (SNP) were identified by genotyping genomic DNA from all subjects. Allelic expression was estimated as described by Bray et al.[15,19] Briefly, deoxyribonuclease-treated RNA samples were subject to reverse-transcription polymerase chain reaction and primer extension with allele-specific dye-terminator incorporation using the SNaPshot kit (Applied Biosystems, Foster City, Calif), and the relative levels of the products representing each transcribed allele were measured on a capillary sequencer. Samples were assayed using primers based on single exonic sequence, capable of amplifying either genomic DNA or complementary DNA (cDNA). The cDNA samples were assayed alongside the corresponding heterozygote genomic DNA, which represents a perfect 1:1 ratio of the 2 alleles. The ratio obtained from genomic DNA thus provided a correction factor for any inequalities in the efficiency of allelic representation specific to each cDNA assay.[20] Absence of genomic DNA in the RNA extracts was confirmed by including the RNA samples that had not been reverse transcribed. Analysis of heterozygous samples was performed as 2 separate experiments. In each experiment, 2 cDNA samples were assayed for each heterozygous individual (as 2 separate reverse-transcription reactions), alongside the corresponding genomic DNA sample.

Individual genotyping was performed using AcycloPrime reagents (PerkinElmer Life Science Products, Boston, Mass). Genotypes were scored blind to affection status, and each plate of samples contained a mixture of equal numbers of cases and controls.

The DNA pools from 691 of the cases and 710 of the controls were also genotyped using the SNaPshot primer extension (Applied Biosystems) protocol described by us in detail by Norton et al.[21]

INDIRECT ASSOCIATION ANALYSIS

A grid of SNPs (minimum density of 1 SNP every 5 kilobases [kb]) across *CNP*, spanning the genomic sequence from 4 kb 5′ to the transcription start site through to 1 kb 3′ to the end of transcription was identified from the CHIP bioinformatics database.[22] A linkage disequilibrium (LD) approach was used to supplement our direct gene analyses because genetic variation altering expression may occur in genomic sequences whose locations are unpredictable with reference to the exonic sequence.

STATISTICAL ANALYSIS

Contingency tables were used to test genotypes and alleles for association with schizophrenia and to calculate odds ratios. All analyses were 2-tailed. Power estimations were calculated using the Web-based program provided by the Department of Statistics, University of California, Los Angeles.[23] The D′ and r^2 estimates of LD were calculated using the Web program Haploview.[24] Haplotype frequency analysis was performed using EH plus[25] with a permutation test.[26] A low-redundancy set of SNPs for haplotype analysis was selected using an entropy method developed in house (V.M., unpublished data, 2003). Entropy as a measure of haplotypic diversity is calculated as:

$$E = -\sum_{i=1}^{k} f_i \log_2 f_i ,$$

where f_i is the frequency of the i-th haplotype, and k is the number of distinct haplotypes in the sample. The haplotype frequencies f_i were obtained using the Expectation maximization-algorithm realized in the EH plus software.[25,26] The basic principle of defining tag markers is that the larger the value of entropy, the more informative is the set of markers forming the haplotype. For m biallelic markers (eg, SNPs), the number of possible haplotypes is 2^m. The full amount of information for these m markers is:

$$E_{full} = -\sum_{i=1}^{2^m} f_i \log_2 f_i ,$$

where f_i are the m markers' haplotype frequencies. Calculating entropies for each subset of these m markers and dividing them by E_{full}, we obtain the relative amount of information (values between 0 and 1) provided by a particular subset of markers compared with the full set of markers. By this method, optimal haplotype-tagging SNPs can be selected even in the absence of a clear haplotypic block structure. A similar approach is described in Ackerman et al.[27] Entropy calculations were based on genotypes of all the markers in 96 unrelated control subjects. These individuals were not part of the matched control sample and are not included in the analysis.

RESULTS

Initial analysis of the *CNP* exonic sequence of 14 unrelated individuals with schizophrenia revealed 4 novel variants (–351 G>A, 4926 G>A, 7230 G>A, and 7284 G>A) and 2 variants already lodged in the CHIP bioinformatics database (rs4432296 and rs2070106). Single nucleotide polymorphism labeling is based on genomic sequence (NT_010755.14) and follows the labeling protocol suggested.[28] Of these, 3 were exonic (rs2070106, 7230 G>A, and 7284 G>A), although none encode an amino-acid change. Pooled analysis revealed that rs2070106 had the highest minor allele frequency, and this SNP was subsequently selected as a marker for allelic expression analysis.

Of the 60 subjects for whom brain was available, 25 individuals were heterozygous for SNP rs2070106 and therefore informative for analysis. Allelic expression analysis of these individuals revealed that in each case, the A allele was expressed at a lower level than the G allele (**Figure**). The data points from each individual cDNA sample showed good reproducibility, with a mean coefficient of variation of 0.07. When data from all the individuals were combined, the A allele was significantly underexpressed ($P<.001$), a finding that was replicated in a repeat analysis ($P<.001$). Across the 2 experiments, the mean reduction was 24%.

Pooled genotyping revealed a trend for association between the lower-expression A allele of rs2070106 and schizophrenia ($P=.06$). Individual genotyping confirmed this finding at a level that meets conventional levels of significance ($P=.04$; odds ratio, 1.2 [95% confidence interval, 1.0-1.40]). Genotypic analysis revealed no stronger evidence (**Table 1**), nor did post hoc tests stipulating AA homozygotes as the risk genotype ($P=.14$; odds ratio, 1.3 [95% confidence interval 0.9-1.9]).

To look for evidence of a functional locus in LD with rs2070106, we performed a more detailed analysis of the *CNP* locus based on all the SNPs we previously detected supplemented by a grid of markers from databases at a minimum density of 1 SNP per 5 kb (**Table 2**). The optimal approach to LD mapping remains a subject of controversy, and therefore we applied 2 approaches. The first was to genotype all the earlier mentioned markers in pools on the premise that reduced *df* means that single-locus LD analysis based on a fairly dense set of markers may be more powerful than multilocus analysis if genotypes at 1 of the tested markers correlate strongly with those at the undetected locus. The second approach was based on haplotype analysis where our aim was to type a subset of the SNPs that could retrieve more than 95% of the haplotype diversity across the gene that would have been captured by genotyping all the earlier mentioned SNPs based on the entropy method (see the "Methods" section). This analysis revealed that by typing 3 markers (*CNP* –351, rs11296, and rs4796751) in addition to rs2070106, we could extract 98% of the haplotypic diversity that would be provided by genotyping all the markers. No additional evidence for association was obtained with any 2-, 3-, or 4-marker haplotype (data not shown). All genotypes were in Hardy-Weinberg equilibrium. The LD relationships between markers are given in **Table 3**.

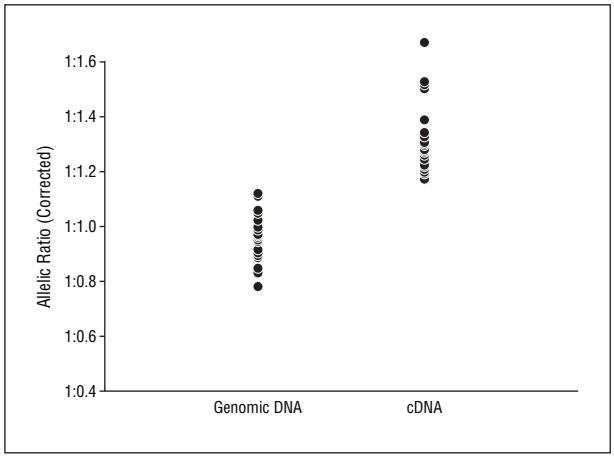

Figure. Corrected genomic and complementary DNA (cDNA) ratios for rs2070106. Comparison between the observed corrected genomic ratios and the corrected cDNA ratios assayed for 2′,3′-cyclic nucleotide 3′-phosphodiesterase (n=25). Data are expressed as the mean of the ratio of A:G for 2 measurements of each genomic DNA sample and 4 measurements of each cDNA sample.

Table 1. Individual Genotyping Results for rs2070106

	Genotypic			Allelic	
	GG	GA	AA	G	A
No. of cases	289	320	74	898	468
No. of controls	328	306	59	962	424
χ^2 Test	4.4 (2 *df*)			4.3 (1 *df*)	
P value	.11			.04	

To determine if variation at *CNP* was at least in part responsible for our previous report of linkage to 17q,[11] we sequenced the exons of *CNP* in all 7 available family members. Additional sequence variants present in the family are presented in Table 2. None changed the predicted amino-acid sequence of the encoded CNP. Based on the microsatellite data,[11] we had expected that all affected siblings would show identity by descent for both maternal and paternal chromosomes. Therefore, with respect to any 1 *CNP* marker, affected siblings would be either all homozygous or all heterozygous. However, for some markers (eg, rs4432296, *CNP* 4926, *CNP* 7230, and *CNP* 7284) homozygous and heterozygous affected subjects were observed, indicating that among the siblings, 3 parental chromosomes were represented with respect to the *CNP* locus. The availability of DNA allows us to determine that each affected subject received the same paternal copy of *CNP*, and therefore, we can conclude that for the *CNP* locus, both maternal chromosomes were represented in the affected siblings. Despite this, all the affected siblings in the family were homozygous for the minor lower-expression A allele at rs2070106. The prior probability that the single paternal chromosome transmitted would carry this allele is equal to the minor allele frequency (0.31) while that of the mother being homozygous for this allele is the square of the allele frequency (approximately 0.1). Conditional on the known pattern of chromosomal transmissions in this family, the prior probability that by chance both parents would have geno-

Table 2. Pooled and Individual Genotyping Results for CNP SNPs

SNP No.	Pooled Genotyping Data: Allele Frequency in Cases	Pooled Genotyping Data: Allele Frequency in Controls	Pooled Genotyping Data: Difference	Pooled Genotyping Data: χ^2 Test	Pooled Genotyping Data: *P* Value	Individual Genotyping Data: Goldenpath Position (HG 16, July 2003)†	Individual Genotyping Data: Allelic χ^2 Test	Individual Genotyping Data: Allelic *P* Value
rs4796750 G/A*	0.7 (G)	0.69 (G)	0.01	0.3	.58	40 487 709		
rs8080978 C/T§	0.8 (C)	0.78 (C)	0.02	1.7	.20	40 489 982		
rs8080979 C/G§	0.95 (C)	0.95 (C)	0	0.2	.63	40 489 984		
CNP –1148 AG Ins/Del§	0.8 (Ins)	0.78 (Ins)	0.02	0.6	.45	40 490 919		
CNP –351 G/A*	0.78 (G)	0.78 (G)	0	0.005	.94	40 491 717	0.717	.40
rs4432296 T/C*	0.67 (T)	0.68 (T)	0.01	0.02	.89	40 491 972		
rs8078650 T/G‡	0.8 (T)	0.79 (T)	0.01	0.53	.47	40 493 936		
CNP 1944 A/T‡	0.97 (A)	0.97 (A)	0	0.006	.94	40 494 011		
rs4258677 A/G*	0.88 (A)	0.89 (A)	0.01	0.04	.84	40 494 439		
CNP 3235 A/G‡	0.72 (A)	0.7 (A)	0.02	0.9	.34	40 495 303		
CNP 3582 C/T‡	0.81 (C)	0.8 (C)	0.01	0.18	.67	40 495 650		
CNP 3625 T/C‡	0.85 (T)	0.83 (T)	0.02	2.1	.15	40 495 693		
CNP 4926 A/G§	0.71 (A)	0.69 (A)	0.02	1.1	.30	40 496 994		
rs2070106 G/A*	0.72 (G)	0.78 (G)	0.06	3.5	.06	40 499 029	4.227	.04
CNP 7230 G/A§	0.78 (G)	0.8 (G)	0.02	1.2	.27	40 499 298		
CNP 7284 G/A§	0.87 (G)	0.87 (G)	0	0.2	.67	40 499 352		
rs11296 T/C*	0.9 (T)	0.89 (T)	0.01	0.3	.58	40 500 225	0.91	.34
rs4796751 C/T*	0.82 (C)	0.81 (C)	0.01	0.3	.58	40 500 702	0.001	.97
CNP 8962 C/T‡	0.95 (C)	0.96 (C)	0.01	1.5	.22	40 501 030		

Abbreviations: CNP, 2′,3′-cyclic nucleotide 3′-phosphodiesterase; kb, kilobase; SNP, single-nucleotide polymorphism.
*Detected during the initial gene screening or studied in the original 5-kb marker grid.
†Refers to the particular version of the Goldenpath database used and the position of the polymorphisms within this version.
‡ Detected by sequencing of the remaining genomic sequence of *CNP*.
§Detected during sequencing of the exons in the family 702 pedigree.

Table 3. Results of Linkage Disequilibrium Analysis*

SNP ID	rs4796750	rs8080978	rs8080979	*CNP* –1148	*CNP* –351	rs4432296	rs8078650	*CNP* 1944	rs4258677	*CNP* 3235	*CNP* 3582	*CNP* 3625	*CNP* 4926	rs2070106	*CNP* 7230	*CNP* 7284	rs11296	rs4796751	*CNP* 8962
rs4796750		1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	0.96	0.94	1.0	0.92	0.12	0.81	1.0	1.0	1.0
rs8080978	0.57		1.0	1.0	1.0	1.0	1.0	0.22	1.0	1.0	0.97	0.96	0.92	0.84	0.09	1.0	1.0	0.2	1.0
rs8080979	0.1	0.16		1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	0.13	0.09	1.0	1.0	1.0	1.0
CNP –1148	0.13	0.08	0.01		1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	0.23	0.34	0.14	1.0	1.0
CNP –351	0.58	0.97	0.17	0.08		1.0	1.0	0.25	1.0	1.0	0.97	0.96	1.0	0.83	0.28	1.0	1.0	0.27	1.0
rs4432296	1.0	0.57	0.09	0.13	0.58		1.0	1.0	1.0	1.0	0.96	0.94	0.97	0.92	0.01	0.82	1.0	1.0	1.0
rs8078650	0.59	0.93	0.18	0.08	1.0	0.59		0.3	1.0	0.96	0.96	0.95	1.0	0.83	0.39	0.65	1.0	0.25	1.0
CNP 1944	0.01	0	0	0.1	0	0.01	0		1.0	1.0	0.28	0.09	1.0	0.61	0.41	0.44	1.0	1.0	1.0
rs4258677	0.24	0.03	0.01	0.03	0.04	0.24	0.03	0		1.0	1.0	1.0	1.0	1.0	0.03	1.0	1.0	0.93	1.0
CNP 3235	0.98	0.59	0.1	0.13	0.59	0.97	0.58	0.01	0.25		0.96	0.95	1.0	0.92	0.17	0.8	1.0	1.0	1.0
CNP 3582	0.52	0.93	0.17	0.08	0.93	0.52	0.93	0	0.04	0.53		1.0	1.0	0.83	0.34	1.0	1.0	0.22	1.0
CNP 3625	0.39	0.72	0.01	0.06	0.69	0.38	0.64	0	0.03	0.4	0.75		1.0	1.0	0.37	1.0	1.0	0.14	1.0
CNP 4926	0.98	0.49	0.1	0.13	0.56	0.93	0.57	0.01	0.24	0.95	0.55	0.41		0.93	0.2	0.82	1.0	1.0	1.0
rs2070106	0.21	0.1	0	0.13	0.1	0.21	0.1	0.01	0.06	0.21	0.1	0.11	0.23		0.2	0.15	1.0	1.0	0.13
CNP 7230	0	0	0	0.05	0.01	0	0.01	0	0	0	0.01	0.01	0.01	0.01		0.13	1.0	0.05	0.09
CNP 7284	0.06	0.05	0.01	0.08	0.05	0.06	0.02	0.03	0.02	0.06	0.05	0.04	0.06	0.01	0		0.89	0.53	1.0
rs11296	0.01	0.01	0	0	0.01	0.01	0.01	0	0	0.01	0.01	0.01	0.01	0.01	0.08	0		1.0	1.0
rs4796751	0.46	0.03	0.23	0.06	0.06	0.44	0.05	0.01	0.47	0.45	0.04	0	0.44	0.11	0	0.01	0.01		1.0
CNP 8962	0.1	0.16	1.0	0.01	0.17	0.09	0.18	0	0.01	0.1	0.17	0.01	0.1	0	0	0.01	0	0.23	

Abbreviations: CNP, 2′,3′-cyclic nucleotide 3′-phosphodiesterase; SNP, single-nucleotide polymorphism.
*Values shown above the diagonal are D′ and values shown below are r^2.

types that would allow all affected siblings to be homozygous for the lower-expression allele is 0.03. If we do not allow for linkage (in our view, incorrectly), given the known parental genotypes (father heterozygous, mother uninformative homozygous), the corresponding probability is $0.5^6 = 0.015$.

We also screened the whole 11 kb of *CNP* genomic sequence to look for further polymorphisms in our 14 unrelated patients with schizophrenia and the family 702 pedigree. This resulted in an additional 9 polymorphisms being detected in the *CNP* sequence. All were genotyped through the pooled-association sample, but none showed any evi-

dence of association with schizophrenia (Table 2). The SNPs were also genotyped in 96 unrelated control subjects to calculate D′ and r^2 values (Table 3).

COMMENT

Given previous reports of altered *CNP* expression in the schizophrenic brain, its location within a region to which we have previously reported linkage, and the observation of central nervous system pathological features in *CNP*-knockout mice reminiscent of findings reported in schizophrenia, we examined the specific hypothesis that sequence variants within *CNP* that influence its expression are associated with schizophrenia. Altered gene expression can be a consequence of *cis*- or *trans*-acting effects. For genetic analysis of a candidate gene, it is the former that is of relevance, and therefore, we first sought to confirm that *CNP* gene expression is in fact influenced by 1 or more *cis*-acting polymorphisms. Our allelic expression analysis clearly confirms the presence of such variants. The observation that in every sample, the A allele of marker rs2070106 is underexpressed compared with the G allele suggests that either the marker SNP itself is responsible for the altered expression or that by virtue of strong LD, haplotypes in which allele G is in phase with the lower-expression allele are uncommon.

Within the context of previous expression studies of the postmortem schizophrenic brain,[2,3] our observation that allele A of rs2070106 is associated with lower *CNP* expression leads to a highly specific hypothesis, namely, that this locus is associated with schizophrenia. Moreover, unusual for candidate-gene studies, not only are we able to specify a specific SNP, our allelic expression data allow us to postulate a specific allele, namely that allele A is more common in affected individuals. Although only just meeting conventional criteria for significance, our findings in the case-control sample are entirely consistent with this hypothesis. Given the highly specific nature of the prior hypothesis with respect to both the SNP and the particular allele, we suggest that this very modest degree of support is stronger than a similar finding based on association with an anonymous SNP of no known functional relevance whose association is the end point of examination of multiple SNPs in and around a candidate gene.

To determine if variation at *CNP* was at least in part responsible for our previous report of linkage to 17q,[11] we sequenced the exons of *CNP* in all family 702 pedigree members. This did not reveal any conspicuous sequence variants that might account for the linkage. Furthermore, the results showed that the *CNP* region on 17q had a lower level of identity by descent sharing than was previously predicted by our linkage study. Nevertheless, all of the affected siblings in the family were homozygous for the lower-expression A allele at marker rs2070106. Conditional on the genotypes and transmission patterns of the parental chromosomes, this observation can be expected to occur as an incidental finding around 3 times in 100 and provides fully independent support for *CNP* as a susceptibility gene for schizophrenia. However, we would also stress that, with a rate in the general population of around 10%, homozygosity for the putative susceptibility allele in *CNP* is not a highly penetrant recessive genotype and cannot alone account for our linkage finding in this family.

Further direct and indirect association analysis based on a dense map of markers, and, ultimately, on resequencing the full genomic sequence of *CNP*, revealed no further evidence for association between schizophrenia, either at the single marker or haplotype level. This suggests that if our finding is correct, marker rs2070106 is likely to be the susceptibility variant per se.

In conclusion, we have shown that *CNP* expression is under the influence of at least 1 *cis*-acting polymorphism. We have also shown that, consistent with previous findings of low *CNP* expression in the schizophrenic brain, an allele associated with lower *CNP* expression is also associated with schizophrenia. Our findings, while supported by a strong directional prior hypothesis and by observations in a family showing linkage to the region, are modest and certainly require independent replication. Nevertheless, they provide support for the specific hypothesis that reduced *CNP* expression in the schizophrenic brain may be etiologically relevant to schizophrenia. Moreover, given the context in which *CNP* is a functional candidate gene for schizophrenia, our data also provide support for the more general hypothesis that altered myelination and/or oligodendrocyte function may play a role in schizophrenia etiology.

Submitted for Publication: December 10, 2004; final revision received May 5, 2005; accepted May 23, 2005.
Correspondence: Michael J. Owen, PhD, FRCPsych, Department of Psychological Medicine, School of Medicine, Cardiff University, Heath Park, Cardiff CF14 4XN, Wales (owenmj@groupwise.cf.ac.uk).
Funding/Support: This work was funded by a component grant of a National Institute of Mental Health (Bethesda, Md)–funded Silvio O. Conte Centre for the Neuroscience of Mental Disorders coordinated by Kenneth Davis, MD, of Mount Sinai School of Medicine, New York, NY, and by grants from the Medical Research Council, London, England. Dr Moskvina is funded by a Higher Education Funding Council for Wales (Cardiff) initiative in biostatistics and bioinformatics.

REFERENCES

1. Williams NM, O'Donovan MC, Owen MJ. Genome scans and microarrays: converging on genes for schizophrenia? *Genome Biol*. 2002;3:REVIEWS1011.
2. Hakak Y, Walker JR, Li C, Wong WH, Davis KL, Buxbaum JD, Haroutunian V, Fienberg AA. Genome-wide expression analysis reveals dysregulation of myelination-related genes in chronic schizophrenia. *Proc Natl Acad Sci U S A*. 2001; 98:4746-4751.
3. Tkachev D, Mimmack M, Ryan M, Wayland M, Freeman T, Jones P, Starkey M, Webster M, Yolken R, Bahn S. Oligodendrocyte dysfunction in schizophrenia and bipolar disorder. *Lancet*. 2003;362:798-805.
4. Stefansson H, Sigurdsson E, Steinthorsdottir V, Bjornsdottir S, Sigmundsson T, Ghosh S, Brynjolfsson J, Gunnarsdottir S, Ivarsson O, Chou TT, Hjaltason O, Birgisdottir B, Jonsson H, Gudnadottir VG, Gudmundsdottir E, Bjornsson A, Ingvarsson B, Ingason A, Sigfusson S, Hardardottir H, Harvey RP, Lai D, Zhou M, Brunner D, Mutel V, Gonzalo A, Lemke G, Sainz J, Johannesson G, Andresson T, Gudbjartsson D, Manolescu A, Frigge ML, Gurney ME, Kong A, Gulcher JR, Peturs-

son H, Stefansson K. Neuregulin 1 and susceptibility to schizophrenia. *Am J Hum Genet*. 2002;71:877-892.
5. Niculescu AB III, Segal DS, Kuczenski R, Barrett T, Hauger RL, Kelsoe JR. Identifying a series of candidate genes for mania and psychosis: a convergent functional genomics approach. *Physiol Genomics*. 2000;4:83-91.
6. Mirnics K, Middleton FA, Stanwood GD, Lewis DA, Levitt P. Disease-specific changes in regulator of G-protein signaling 4 (RGS4) expression in schizophrenia. *Mol Psychiatry*. 2001;6:293-301.
7. Chowdari KV, Mirnics K, Semwal P, Wood J, Lawrence E, Bhatia T, Deshpande SN, B KT, Ferrell RE, Middleton FA, Devlin B, Levitt P, Lewis DA, Nimgaonkar VL. Association and linkage analyses of RGS4 polymorphisms in schizophrenia. *Hum Mol Genet*. 2002;11:1373-1380.
8. Barrett TB, Hauger RL, Kennedy JL, Sadovnick AD, Remick RA, Keck PE, McElroy SL, Alexander M, Shaw SH, Kelsoe JR. Evidence that a single nucleotide polymorphism in the promoter of the G protein receptor kinase 3 gene is associated with bipolar disorder. *Mol Psychiatry*. 2003;8:546-557.
9. Aston C, Jiang L, Sokolov BP. Microarray analysis of postmortem temporal cortex from patients with schizophrenia. *J Neurosci Res*. 2004;77:858-866.
10. Flynn SW, Lang DJ, Mackay AL, Goghari V, Vavasour IM, Whittall KP, Smith GN, Arango V, Mann JJ, Dwork AJ, Falkai P, Honer WG. Abnormalities of myelination in schizophrenia detected in vivo with MRI, and post-mortem with analysis of oligodendrocyte proteins. *Mol Psychiatry*. 2003;8:811-820.
11. Williams NM, Norton N, Williams H, Ekholm B, Hamshere ML, Lindblom Y, Chowdari KV, Cardno AG, Zammit S, Jones LA, Murphy KC, Sanders RD, McCarthy G, Gray MY, Jones G, Holmans P, Nimgaonkar V, Adolfson R, Osby U, Terenius L, Sedvall G, O'Donovan MC, Owen MJ. A systematic genomewide linkage study in 353 sib pairs with schizophrenia. *Am J Hum Genet*. 2003;73:1355-1367.
12. Davis KL, Stewart DG, Friedman JI, Buchsbaum M, Harvey PD, Hof PR, Buxbaum J, Haroutunian V. White matter changes in schizophrenia: evidence for myelin-related dysfunction. *Arch Gen Psychiatry*. 2003;60:443-456.
13. Lappe-Siefke C, Goebbels S, Gravel M, Nicksch E, Lee J, Braun PE, Griffiths IR, Nave KA. Disruption of Cnp1 uncouples oligodendroglial functions in axonal support and myelination. *Nat Genet*. 2003;33:366-374.
14. Bifulco M, Laezza C, Stingo S, Wolff J. 2' 3'-Cyclic nucleotide 3'-phosphodiesterase: a membrane-bound, microtubule-associated protein and membrane anchor for tubulin. *Proc Natl Acad Sci U S A*. 2002;99:1807-1812.
15. Bray NJ, Buckland PR, Owen MJ, O'Donovan MC. *cis*-acting variation in the expression of a high proportion of genes in human brain. *Hum Genet*. 2003;113:149-153.
16. Wing JK, Babor T, Brugha T, Burke J, Cooper JE, Giel R. SCAN: Schedules for Clinical Assessment in Neuropsychiatry. *Arch Gen Psychiatry*. 1990;47:589-593.
17. Jones AC, Austin J, Hansen N, Hoogendoorn B, Oefner PJ, Cheadle JP, O'Donovan MC. Optimal temperature selection for mutation detection by denaturing HPLC and comparison to single-stranded conformation polymorphism and heteroduplex analysis. *Clin Chem*. 1999;45:1133-1140.
18. Austin J, Hoogendoorn B, Buckland P, Speight G, Cardno A, Bowen T, Williams N, Spurlock G, Sanders R, Jones L, Murphy K, McCarthy G, McGuffin P, Owen MJ, O'Donovan MC. Comparative sequencing of the proneurotensin gene and association studies in schizophrenia. *Mol Psychiatry*. 2000;5:208-212.
19. Bray NJ, Buckland PR, Williams NM, Williams HJ, Norton N, Owen MJ, O'Donovan MC. b: a haplotype implicated in schizophrenia susceptibility is associated with reduced *COMT* expression in human brain. *Am J Hum Genet*. 2003;73:152-161.
20. Hoogendoorn B, Norton N, Kirov G, Williams N, Hamshere ML, Spurlock G, Austin J, Stephens MK, Buckland PR, Owen MJ, O'Donovan MC. Cheap, accurate and rapid allele frequency estimation of single nucleotide polymorphisms by primer extension and DHPLC in DNA pools. *Hum Genet*. 2000;107:488-493.
21. Norton N, Williams NM, Williams HJ, Spurlock G, Kirov G, Morris DW, Hoogendoorn B, Owen MJ, O'Donovan MC. Universal, robust, highly quantitative SNP allele frequency measurement in DNA pools. *Hum Genet*. 2002;110:471-478.
22. CHIP bioinformatics tools. Available at: http://snpper.chip.org. Accessed 2004.
23. UCLA Department of Statistics Web site. Available at: http://calculators.stat.ucla.edu. Accessed 2004.
24. Haploview. Available at: http://www.broad.mit.edu/personal/jcbarret/haploview/index.php. Accessed 2004.
25. Terwilliger J, Ott J. *Handbook of Human Genetic Linkage*. Baltimore, Md: The Johns Hopkins University Press; 1994.
26. Zhao JH, Curtis D, Sham PC. Model-free analysis and permutation tests for allelic associations. *Hum Hered*. 2000;50:133-139.
27. Ackerman H, Usen S, Mott R, Richardson A, Sisay-Joof F, Katundu P, Taylor T, Ward R, Molyneux M, Pinder M, Kwiatkowski DP. Haplotypic analysis of the TNF locus by association efficiency and entropy. *Genome Biol*. 2003;4:R24.
28. den Dunnen JT, Antonarakis SE. Nomenclature for the description of human sequence variations. *Hum Genet*. 2001;109:121-124.