Report

The *Epsin 4* Gene on Chromosome 5q, Which Encodes the Clathrin-Associated Protein Enthoprotin, Is Involved in the Genetic Susceptibility to Schizophrenia

Jonathan Pimm,[1,2] Andrew McQuillin,[1] Srinivasa Thirumalai,[5] Jacob Lawrence,[1] Digby Quested,[3,6] Nicholas Bass,[1] Graham Lamb,[2] Helen Moorey,[2] Susmita R. Datta,[1] Gursharan Kalsi,[1] Allison Badacsonyi,[4] Katie Kelly,[7] Jenny Morgan,[5] Bhaskar Punukollu,[3] David Curtis,[4] and Hugh Gurling[1]

[1]Molecular Psychiatry Laboratory, Department of Mental Health Sciences, University College London Medical School, Windeyer Institute of Medical Sciences, [2]Camden and Islington Mental Health and Social Care Trust, St. Pancras Hospital, [3]West London Mental Health Trust, Hammersmith and Fulham Mental Health Unit and St. Bernard's Hospital, and [4]Queen Mary College, University of London and East London and City Mental Health Trust, Royal London Hospital, London; [5]West Berkshire National Health Service (NHS) Trust, Reading, United Kingdom; [6]Department of Psychiatry, University of Oxford, Warneford Hospital, Oxford, United Kingdom; and [7]Gloucestershire Partnership NHS Trust, Gloucester, United Kingdom

Chromosome 5q33 is a region that has previously shown good evidence of linkage to schizophrenia, with four LOD scores >3.00 in independent linkage studies. We studied 450 unrelated white English, Irish, Welsh, and Scottish research subjects with schizophrenia and 450 ancestrally matched supernormal controls. Four adjacent markers at the 5′ end of the *Epsin 4* gene showed significant evidence of linkage disequilibrium with schizophrenia. These included two microsatellite markers, *D5S1403* ($P = .01$) and *AAAT11* ($P = .009$), and two single-nucleotide–polymorphism markers within the *Epsin 4* gene, *rs10046055* ($P = .007$) and *rs254664* ($P = .01$). A series of different two- and three-marker haplotypes were also significantly associated with schizophrenia, as confirmed with a permutation test (HapA, $P = .004$; HapB, $P = .0005$; HapC, $P = .007$; and HapD, $P = .01$). The *Epsin 4* gene encodes the clathrin-associated protein enthoprotin, which has a role in transport and stability of neurotransmitter vesicles at the synapses and within neurons. A genetically determined abnormality in the structure, function, or expression of enthoprotin is likely to be responsible for genetic susceptibility to a subtype of schizophrenia on chromosome 5q33.3.

Important genetic effects in the susceptibility to schizophrenia have now been established by twin, adoption, genetic-linkage, and allelic-association studies. Linkage studies have proven that there is heterogeneity of linkage or locus heterogeneity for different genetic subtypes of schizophrenia. We originally obtained LOD scores >3.00 showing linkage to the first schizophrenia locus (*SCZD1* [MIM 181150]) on chromosome 5 in a sample of 23 highly selected large multiplex families, by use of a limited set of RFLP markers that were not well localized but were thought to be within the cytogenetic region 5q11-12 (Sherrington et al. 1988; Kalsi et al. 1999). Later, we completed a genome linkage scan in a subsample of these families (Gurling et al. 2001) but found two positive regions of linkage to schizophrenia that were adjacent on either side of the original 5q11-12 region. In this later genome scan, we used better localized markers that were evenly spread along the whole chromosome 5. Some of the families that produced positive LOD scores at 5q11-12 in the original study (Sherrington et al. 1988) then showed linkage near the centromere on chromosome 5, with a LOD score of 2.5, and other families showed linkage to markers on the long arm in the 5q22-33 region (Gurling et al. 2001).

The more distal of the two regions we implicated was

Received December 15, 2004; accepted for publication March 4, 2005; electronically published March 25, 2005.

Address for correspondence and reprints: Dr. Hugh Gurling, Molecular Psychiatry Laboratory, Department of Mental Health Sciences, University College London Medical School, Windeyer Institute of Medical Sciences, 46 Cleveland Street, London W1T 4JF, United Kingdom. E-mail: rejuhxg@ucl.ac.uk

© 2005 by The American Society of Human Genetics. All rights reserved.

0002-9297/2005/7605-0019$15.00

at 5q22-34, where we obtained a LOD score of 3.6 (nominal $P = .0001$). The chromosome 5q22-34 region had also been found to be linked to schizophrenia in a sample of Irish families (Straub et al. 1997), with a heterogeneity LOD score of 3.35 near locus *D5S804* ($P = .0002$).

Subsequently, analysis of a single large kindred from Palau, Micronesia, resulted in a LOD score of 3.4 and showed linkage to schizophrenia with markers in the 5q22-31 region (Devlin et al. 2002). Later, a study of Finnish families also found linkage to schizophrenia, with a LOD score of 3.56, at 5q22-31 (Paunio et al. 2001). A collaborative study of the 5q22-31 region (Levinson 1999) found that four of eight independent linkage samples of families with schizophrenia resulted in LOD scores >1.00. For example, analysis of a sample of German families resulted in a LOD score of 1.8 at 5q22-31, at marker *D5S399* (Schwab et al. 1997).

These studies suggest that there must be at least two—and possibly three or more—different regions of chromosome 5 that harbor susceptibility loci for schizophrenia. The evidence implicating the more distal region of chromosome 5 at 5q33 is derived from four independent linkage studies that report LOD scores >3.00. We attempted to identify exactly which gene might be involved, by the detection of linkage disequilibrium (LD) in a case-control genetic-association study, using markers we had previously found to show maximal evidence of linkage to schizophrenia.

We recruited a sample of 450 volunteers with schizophrenia and 450 controls from London and southern England. All subjects were included only if both parents were of English, Irish, Welsh, or Scottish descent and if three of four grandparents were of the same descent. One grandparent was allowed to be of other white European origin but not of Jewish or non–European Union (EU) ancestry (on the basis of the EU countries before the recent enlargement). U.K. National Health Service (NHS) multicenter and local research ethics committee approval was obtained, and all subjects signed an approved consent form after reading an information sheet. All 450 schizophrenic research subjects had been diagnosed and assessed by NHS psychiatrists as part of routine clinical diagnosis and treatment. Research subjects with short-term drug-induced psychoses, learning disabilities, head injuries, and other symptomatic psychoses were excluded from the present study. Research subjects were selected on the basis of having an International Classification of Diseases version 10 diagnosis of schizophrenia recorded in the NHS medical case notes. That information together with personal, medical, and psychiatric histories and all other available information were taken into consideration to complete data collection and interviewing with the Schizophrenia and Affective Disorders Schedule–Lifetime Schedule (SADS-L) interview (Spitzer and Endicott 1977). The SADS-L protocol expects data items to be recorded accurately from previous medical records and does not rely on the accuracy of statements made during the interview by mentally ill research subjects. After the SADS-L interview was completed using all sources of information, the subject individuals received the diagnosis of schizophrenia in accordance with the probable level of the Research Diagnostic Criteria (RDC). The "supernormal" control subjects were also interviewed in accordance with SADS-L guidelines, by use of the initial clinical screening questions, and were selected on the basis of having no family history of schizophrenia, alcoholism, or bipolar disorder and having no past or present personal history of any RDC-defined mental disorder.

Genomic DNA was extracted from frozen whole-blood samples by use of a standard cell lysis, proteinase K digestion, and phenol/chloroform ethanol precipitation method. All DNA samples were quantified with pico green by laser fluorimetry. The first set of five markers genotyped was selected on the basis of having shown maximal linkage and minimal recombination fraction to a chromosome 5q33 schizophrenia-susceptibility locus, with the assumption of admixture, in our previous genome linkage scan in large kindreds with schizophrenia. Specific loci were not targeted by genotyping markers next to known genes. The microsatellite marker *AAAT11* was identified, formatted, and genotyped for the present study and was amplified with the primers (forward) CTAGTTGGGAGGCTGAGGTG and (reverse) AAGAGGGAGCCTCAATCTGG. Primer sequences for PCR amplification of the microsatellite marker *D5S1403* and the SNPs were obtained from ENSEMBL, the University of California–Santa Cruz (UCSC) Genome Bioinformatics Web site, or the National Center for Biotechnology Information SNP Database. PCR amplification of microsatellite markers was performed using an M13-tailed primer and a second nontailed oligonucleotide primer. A third universal M13 sequence primer labeled with infrared dye IRD 700 or IRD 800 was used to hybridize against the M13-tailed locus-specific primer. Microsatellite marker–fragment sizes were separated and visualized using dual argon laser LiCor 4200L sequencers. Genotyping was performed with the SAGA2 genotyping software and was checked by eye. Allele calling by SAGA2 was checked by a second independent person, blind to diagnosis. The genotype data were automatically stored in a database. Any genotypes that were discrepant between the two checkers were PCR amplified and were genotyped again.

SNPs were determined by the Amplifluor SNP genotyping method, as modified by K Biosciences. Of samples on each microtiter plate, 17% were duplicated to detect error and to confirm the reproducibility of genotypes. The data were then analyzed to confirm Hardy-Weinberg equilibrium (HWE). Markers with lack of HWE in the control group were rejected, and genotyping was

Table 1

Tests of Allelic Association with Schizophrenia at the *Epsin 4* Gene Locus

Marker and Sample	Location (in bp) on Chromosome 5	Distance (in bases) from Previous Marker	Observed Allele Frequencies at: Allele			Fragment Size																									χ^2	P[a]
			C	T	A	232	240	244	248	252	197	199	203	207	211	215	219	223	353	356	362	365	368	371	374	377	380	383	386	389		
rs254664:	157,246,580																														6.11	.01[a]
Control			638	242																												
Schizophrenic			522	258																												
5′ Start of *Epsin 4*	157,267,063																															
rs1186930:	157,272,308	2,984																													2.39	.12[a]
Control			483	363																												
Schizophrenic			470	302																												
rs10046055:	157,297,912	25,604																													9.83	.002[a]
Control				613	265																											
Schizophrenic				471	283																											
AAAT11:	157,301,697	3,785																													9.61	.007[b]
Control						0	50	3	471	236																						
Schizophrenic						1	63	1	392	265																						
D5S1403:	157,346,617	44,920																													11.28	.01[b]
Control											0	4	262	158	127	125	44	2														
Schizophrenic											2	3	201	162	148	140	34	0														
D5S1400:	157,439,237	92,620																													6.67	.06[c]
Control																			1	1	2	20	137	281	1	85	72	88	12	6		
Schizophrenic																			3	0	2	18	146	228	1	94	91	88	14	3		

[a] P from 2×2 χ^2, with 1 df.
[b] P from CLUMP Monte Carlo T4.
[c] P from CLUMP Monte Carlo T3.

Table 2

Results of Tests of Haplotypic Association with Schizophrenia at the *Epsin 4* Gene Locus

No. of Markers and Haplotype (Allele)	χ^2	P[a]
2 (440 controls; 396 cases):		
HapA: *rs254664* (T) and *rs10046055* (A)[b]	13.2	.004
HapB: *rs1186930* (T) and *rs10046055* (A)[b]	18.3	.0005
HapC: *rs10046055* (A) and *AAAT11* (252 bp)[b]	21.2	.007
3 (442 controls; 408 cases):		
HapD: *rs254664* (T), *rs10046055* (A), and *AAAT11* (252 bp)[b]	28.1	.01
HapE: *rs1145603* (T), *rs10046055* (A), and *AAAT11* (252 bp)[b]	28.3	.01

[a] Haplotype permutation test empirical *P*, on the basis of 99,999 permutations.
[b] Most likely to be specific alleles that show association in each haplotype.

repeated. Before association analysis, genotype data for each 96-well microtiter plate were analyzed for LD, by use of GENECOUNTING/LDPAIRS, to check that closely linked markers had consistent LD relationships on each plate.

Next, these data were analyzed for allelic association with schizophrenia by use of CLUMP, which employs an empirical Monte Carlo test of significance and which does not require correction for multiple alleles (Sham and Curtis 1995). Subtests of the CLUMP program are T1, Pearson's χ^2 statistic of the "raw" contingency table; T2, the χ^2 statistic of a table with rare alleles grouped together to prevent small expected cell counts; T3, the largest of the χ^2 statistics of 2×2 tables, each of which compares one allele with the rest grouped together; and T4, the largest of the χ^2 statistics of all possible 2×2 tables, comparing any combination of alleles with the rest. The genotypes were then analyzed for marker-to-marker LD and for haplotypic association with schizophrenia by use of GENECOUNTING/LDPAIRS, which computes D' and Cramer's V tests of LD and maximum-likelihood estimates of haplotype frequencies from phase-unknown case-control data (Zhao et al. 2000, 2002). The significance of any overall haplotype association with schizophrenia was computed using permutation testing. Fifteen genetic markers at chromosomal loci thought to be not involved in schizophrenia were genotyped in a subset of the sample (200 cases and 300 controls) and were analyzed to detect genetic heterogeneity between cases and controls, to confirm that the samples were genetically well matched. In addition, a statistical test (CHECKHET) for detecting subjects with atypical genetic background was employed (Curtis et al. 2002).

No evidence of genetic heterogeneity between cases and controls was found from using the reference markers. The CHECKHET test detected two subjects with schizophrenia with abnormal genotypes; these were excluded from further study before any chromosome 5 markers were genotyped. After five microsatellite markers—*D5S1480, D5S1507, D5S1499, D5S820,* and *D5S1403*—had been genotyped, a positive allelic association was found with *D5S1403,* as shown in table 1 (CLUMP T4, $P = .01$). Therefore, 11 additional markers close to *D5S1403* were selected and genotyped. One of these 11 markers was the microsatellite marker *AAAT11,* which was 40 kb from *D5S1403*. This marker also showed significant evidence of association with schizophrenia, with an empirical *P* value of .007. Two of four SNP markers within the *Epsin 4* gene (*rs10046055* [$P = .002$] and *rs254664* [$P = .02$]) were also found to be associated with schizophrenia. Flanking markers on each side of the *Epsin 4* gene did not show association.

The allele frequencies, positions on chromosome 5 for the *Epsin 4* gene, and markers showing significant evidence or a trend toward association with schizophrenia are shown in tables 1 and 2, in their order on chromosome 5q, in accordance with the June 2004 release of the UCSC Genome Browser. Two-marker haplotype

Table 3

Results of Pairwise LD Statistics between All Pairs of Associated Markers

	Absolute Value of D' and LD P Values[a]					
Marker	*rs254664*	*rs1186930*	*rs10046055*	*AAAT11*	*D5S1403*	*D5S1400*
rs254664		.992	.220	.395	.414	.044
rs1186930	<.00001		.242	.058	.383	.057
rs10046055	<.00001	<.00001		.883	.252	.339
AAAT11	<.00001	<.00001	<.00001		.238	.357
D5S1403	<.00001	<.00001	.006	.157		.220
D5S1400	<.00001	.856	<.00001	<.00001	<.00001	

[a] D' values shown above the diagonal; *P* values shown below diagonal.

HapA (see table 2), derived from combining the SNP markers *rs254664* and *rs10046055,* showed association with schizophrenia, on the basis of permutation testing ($P = .004$). Significant association verified through permutation testing ($P = .0005$) was also found with the SNP markers *rs1186930* and *rs10046055* when combined into haplotype HapB and with the markers *rs10046055* and *AAAT11* in HapC. Two three-marker haplotypes (HapD and HapE) that combined the microsatellite marker *ATTT11* with two of the three SNP markers *rs254664, rs10046055,* and *rs1145603* also showed significant association (permutation test $P =$ .001 and .007, respectively). Each of the markers that showed association with schizophrenia was in strong LD ($P < .00001$) with one or more of the other associated markers, as shown in table 3.

The five originally genotyped markers were selected as the closest markers to the chromosome 5q33 linkage peak obtained in our previously published genome analysis of large schizophrenic kindreds. One of these markers showed some evidence for association; hence, new markers nearby were genotyped. Individually, some of these markers produced more highly significant evidence for association. Because these markers are in LD with each other, it is not possible to state the overall significance of all the tests for association obtained from individual markers and from marker combinations, but, together, the results do indicate that a locus conferring susceptibility to schizophrenia is likely to be present in this narrowly defined region.

The schizophrenic subjects studied were routine cases from United Kingdom NHS facilities in London and southern England. They were not selected for having a positive family history of schizophrenia. An important implication of this finding is that genetic effects detected by linkage studies in large multiply affected pedigrees also apply to routinely treated patients in London and southern England NHS mental health facilities. The only gene in the chromosomal region that we implicated by LD was the *Epsin 4* gene. *Epsin 4* is a member of a class of proteins with an “epsin N-terminal homology (ENTH) domain.” This protein sequence is evolutionarily conserved and is found in association with other proteins that participate in clathrin-mediated pit formation and endocytosis and in vesicle stability (Koshiba et al. 2002). Enthoprotin is highly enriched on clathrin-coated vesicles isolated from the rat brain and is probably involved in memory by its effects on reuptake and storage of neurotransmitters (Wasiak et al. 2002). Replication studies are needed to fully establish this new genetic association with schizophrenia. Etiological base-pair changes present in the coding and control regions of the *Epsin 4* gene can now be sought by resequencing or mutation screening of genomic DNA from subjects with schizophrenia who have inherited the alleles and haplotypes on chromosome 5q33.3 that are associated with schizophrenia.

Acknowledgments

This research was funded by the Neuroscience Research Charitable Trust and by a Priory Hospital research lectureship to J.L. The authors have no commercial conflicts of interest.

Electronic-Database Information

The URLs for data presented herein are as follows:

ENSEMBL, http://www.ensembl.org/ (for SNPs and microsatellites)
National Center for Biotechnology Information SNP Database, http://www.ncbi.nlm.nih.gov/projects/SNP/ (for SNPs)
Online Mendelian Inheritance in Man (OMIM), http://www.ncbi.nlm.nih.gov/Omim/ (for *SCZD1*)
University of California–Santa Cruz (UCSC) Genome Bioinformatics, http://www.genome.ucsc.edu/ (for SNPs and microsatellites)

References

Curtis D, North BV, Gurling HM, Blaveri E, Sham PC (2002) A quick and simple method for detecting subjects with abnormal genetic background in case-control samples. Ann Hum Genet 66:235–244

Devlin B, Bacanu SA, Roeder K, Reimherr F, Wender P, Galke B, Novasad D, Chu A, TCuenco K, Tiobek S, Otto C, Byerley W (2002) Genome-wide multipoint linkage analyses of multiplex schizophrenia pedigrees from the Oceanic nation of Palau. Mol Psychiatry 7:689–694

Gurling HM, Kalsi G, Brynjolfson J, Sigmundsson T, Sherrington R, Mankoo BS, Read T, Murphy P, Blaveri E, McQuillin A, Petursson H, Curtis D (2001) Genomewide genetic linkage analysis confirms the presence of susceptibility loci for schizophrenia, on chromosomes 1q32.2, 5q33.2, and 8p21-22 and provides support for linkage to schizophrenia, on chromosomes 11q23.3-24 and 20q12.1-11.23. Am J Hum Genet 68:661–673

Kalsi G, Mankoo B, Curtis D, Sherrington R, Melmer G, Brynjolfsson J, Sigmundsson T, Read T, Murphy P, Petursson H, Gurling H (1999) New DNA markers with increased informativeness show diminished support for a chromosome 5q11-13 schizophrenia susceptibility locus and exclude linkage in two new cohorts of British and Icelandic families. Ann Hum Genet 63:235–247

Koshiba S, Kigawa T, Kikuchi A, Yokoyama S (2002) Solution structure of the epsin N-terminal homology (ENTH) domain of human epsin. J Struct Funct Genomics 2:1–8

Levinson DF (1999) Multicenter linkage study of schizophrenia candidate regions on chromosomes 5q, 6q, 10p and 13q. Am J Hum Genet Suppl 65:A259

Paunio T, Ekelund J, Varilo T, Parker A, Hovatta I, Turunen JA, Rinard K, Foti A, Terwilliger JD, Juvonen H, Suvisaari J, Arajarvi R, Suokas J, Partonen T, Lonnqvist J, Meyer J, Peltonen L (2001) Genome-wide scan in a nationwide study

sample of schizophrenia families in Finland reveals susceptibility loci on chromosomes 2q and 5q. Hum Mol Genet 10:3037–3048

Schwab SG, Eckstein GN, Hallmayer J, Lerer B, Albus M, Borrmann M, Lichtermann D, Ertl MA, Maier W, Wildenauer DB (1997) Evidence suggestive of a locus on chromosome 5q31 contributing to susceptibility for schizophrenia in German and Israeli families by multipoint affected sibpair linkage analysis. Mol Psychiatry 2:156–160

Sham PC, Curtis D (1995) Monte Carlo tests for associations between disease and alleles at highly polymorphic loci. Ann Hum Genet 59:97–105

Sherrington R, Brynjolfsson J, Petursson H, Potter M, Dudleston K, Barraclough B, Wasmuth J, Dobbs M, Gurling H (1988) Localization of a susceptibility locus for schizophrenia on chromosome 5. Nature 336:164–167

Spitzer R, Endicott J (1977) The schedule for affective disorders and schizophrenia, lifetime version, 3rd ed. New York State Psychiatric Institute, New York

Straub RE, MacLean CJ, O'Neill FA, Walsh D, Kendler KS (1997) Support for a possible schizophrenia vulnerability locus in region 5q22-31 in Irish families. Mol Psychiatry 2: 148–155

Wasiak S, Legendre-Guillemin V, Puertollano R, Blondeau F, Girard M, de Heuvel E, Boismenu D, Bell AW, Bonifacino JS, McPherson PS (2002) Enthoprotin: a novel clathrin-associated protein identified through subcellular proteomics. J Cell Biol 158:855–862

Zhao JH, Curtis D, Sham PC (2000) Model-free analysis and permutation tests for allelic associations. Hum Hered 50: 133–139

Zhao JH, Lissarrague S, Essioux L, Sham PC (2002) GENECOUNTING: haplotype analysis with missing genotypes. Bioinformatics 18:1694–1695