Am. J. Hum. Genet. 76:950–966, 2005

Candidate-Gene Screening and Association Analysis at the Autism-Susceptibility Locus on Chromosome 16p: Evidence of Association at *GRIN2A* and *ABAT*

Gabrielle Barnby,[1] Aaron Abbott,[1] Nuala Sykes,[1] Andrew Morris,[1] Daniel E. Weeks,[3] Richard Mott,[1] Janine Lamb,[1] Anthony J. Bailey,[2] Anthony P. Monaco,[1] and the International Molecular Genetics Study of Autism Consortium (IMGSAC)*

[1]Wellcome Trust Centre for Human Genetics, University of Oxford, and [2]Section of Child and Adolescent Psychiatry, University Department of Psychiatry, Park Hospital for Children, Oxford, United Kingdom; and [3]Departments of Human Genetics and Biostatistics, Graduate School of Public Health, University of Pittsburgh, Pittsburgh

Autism is a highly heritable neurodevelopmental disorder whose underlying genetic causes have yet to be identified. To date, there have been eight genome screens for autism, two of which identified a putative susceptibility locus on chromosome 16p. In the present study, 10 positional candidate genes that map to 16p11-13 were examined for coding variants: *A2BP1*, *ABAT*, *BFAR*, *CREBBP*, *EMP2*, *GRIN2A*, *MRTF-B*, *SSTR5*, *TBX6*, and *UBN1*. Screening of all coding and regulatory regions by denaturing high-performance liquid chromatography identified seven nonsynonymous changes. Five of these mutations were found to cosegregate with autism, but the mutations are not predicted to have deleterious effects on protein structure and are unlikely to represent significant etiological variants. Selected variants from candidate genes were genotyped in the entire International Molecular Genetics Study of Autism Consortium collection of 239 multiplex families and were tested for association with autism by use of the pedigree disequilibrium test. Additionally, genotype frequencies were compared between 239 unrelated affected individuals and 192 controls. Patterns of linkage disequilibrium were investigated, and the transmission of haplotypes across candidate genes was tested for association. Evidence of single-marker association was found for variants in *ABAT*, *CREBBP*, and *GRIN2A*. Within these genes, 12 single-nucleotide polymorphisms (SNPs) were subsequently genotyped in 91 autism trios (one affected individual and two unaffected parents), and the association was replicated within *GRIN2A* (Fisher's exact test, $P < .0001$). Logistic regression analysis of SNP data across *GRIN2A* and *ABAT* showed a trend toward haplotypic differences between cases and controls.

Introduction

Autism is a severe neurodevelopmental disorder that affects reciprocal communication and social interaction and is associated with repetitive and stereotyped behaviors. Recent epidemiological studies have revised the population prevalence of autism upward from 4 in 10,000 to 9–11 in 10,000 (Chakrabarti and Fombonne 2001; Fombonne 2003), and the overall rate of pervasive developmental disorders is ~60 in 10,000 (Fombonne 2002). Autism is highly heritable and shows strong familial clustering, with a λ (sibling relative risk) of 75–100, although the mode of inheritance is complex (Bailey et al. 1995; Pickles et al. 1995). Autism is found in three-to-four times as many males as females, which suggests a possible sex-linked effect, but multivariate analysis of the phenotype in relatives does not support X-linked inheritance (Pickles et al. 2000). Moderate evidence of linkage to the X chromosome has been observed in two genome screens (Liu et al. 2001; Shao et al. 2002). A single mutation in each of the X-linked neuroligin genes *NLGN3* and *NLGN4* has also been reported, but the low frequency of the mutations makes it unlikely that they contribute significantly to autism etiology (Jamain et al. 2003; Stone et al. 2003; Vincent et al. 2004). Autism is associated with phenotypes of known etiologies in ~6% of cases, including fragile-X syndrome, epilepsy, neurofibromatosis, and tuberous sclerosis (Fombonne 2002). The increased comorbidity of these disorders with autism may indicate a contribution to the autism phenotype by disruption of a common neurological—although not necessarily genetic—pathway.

Recently, the IMGSAC performed a whole-genome screen of 83 affected sib pairs (ASPs). Thirteen loci that displayed evidence of linkage were subsequently genotyped for 152 sib pairs (IMGSAC 2001). The three high-

Received November 1, 2004; accepted for publication March 23, 2005; electronically published April 13, 2005.

Address for correspondence and reprints: Dr. Anthony P. Monaco, Wellcome Trust Centre for Human Genetics, University of Oxford, Roosevelt Drive, Oxford, OX3 7BN, United Kingdom. E-mail: anthony.monaco@well.ox.ac.uk

* IMGSAC members are listed in the Acknowledgments.

0002-9297/2005/7606-0005$15.00

Table 1

Candidate Genes Screened for Autism-Susceptibility Variants

Gene Symbol	Gene Name	Position[a]	GenBank Accession Number	Genomic Region (kb)	mRNA Length (bp)	No. of Exons
A2BP1	Ataxin 2–binding protein	6009132–7700843	NM_018723	420.449	2,279	16
ABAT	4-Aminobutyrate aminotransferase	8675992–8782887	BC031413	106.895	1,705	14
BFAR	Bifunctional apoptosis regulator	14634296–14670193	NM_016561	35.898	2,531	8
CREBBP	CREB-binding protein	3716569–3870712	NM_004380	154.144	8,694	32
EMP2	Epithelial membrane protein 2	10533475–10582040	NM_001424	48.566	690	4
GRIN2A	Glutamate receptor, ionotropic, NMDA 2A	9762922–10184112	NM_000833	421.191	6,293	14
MRTF-B	Myocardin-related transcription factor B	14235623–14268130	NM_014048	32.508	7,815	8
SSTR5	Somatostatin receptor 5	1068869–1069964	NM_001053	1.096	1,095	1
TBX6	T-box 6	30004582–30010709	NM_004608	6.128	2,468	9
UBN1	Ubinuclein 1	4842216–4872364	NM_016936	30.149	5,747	18

[a] From the May 2004 release of the International Human Genome Sequencing Consortium.

est-scoring loci detected were on 2q21-33 at *D2S2118,* with a multipoint maximum LOD score (MLS) of 3.74; 7q22.2-31.3 at *D7S477,* with a multipoint MLS of 3.20; and 16p11-13 at *D16S3102,* with a multipoint MLS of 2.93. Detailed analysis of sex-specific and parent-of-origin effects at these loci has recently been reported (Lamb et al. 2005). The linkage to chromosome 16p was found in male-male ASPs (MLS of 2.48 for 145 male-male ASPs vs. MLS of 0 for 74 male-female or female-female ASPs) (Lamb et al. 2005). *D16S3102* at the peak of linkage was physically mapped at 12.46 Mb in the May 2004 release of the Human Genome Sequence. Linkage overlapping the region on chromosome 16p, between 10 cM and 23 cM, has been reported by Philippe et al. (1999), with a multipoint MLS of 0.74 at 10.6 cM ($n = 51$ multiplex families). Additionally, trisomy of 16p and duplications mapping to 16p11.2 and 16p12.2 have been reported for individuals with autism traits (Finelli et al. 2004).

To identify the autism-susceptibility gene at 16p11-13, positional functional candidate genes were selected, and their coding sequences were screened by denaturing high-performance liquid chromatography (DHPLC) for 48 individuals (from multiplex families) who contributed to linkage on 16p11-13. Variants detected by DHPLC were further characterized by sequencing, and missense variants were tested for cosegregation with autism. The candidate genes that were examined in the present study are described in table 1.

In total, 38 SNPs were selected for genotyping in 239 multiplex families and were tested for association by use of the pedigree disequilibrium test (PDT). These included three SNPs in *TSC2; TSC2* maps to 16p13 and encodes one of the genes that cause tuberous sclerosis, which has been reported to be associated with autism. Following the suggestion that there is a phenotypic overlap between attention-deficit/hyperactivity disorder (ADHD) and autism, four SNPs previously analyzed in a study of ADHD that found linkage to chromosome 16p were also genotyped (Smalley et al. 2002; Ogdie et al. 2003, 2004), as shown in table 2. All SNPs were also genotyped for 192 controls, and the allele and genotype frequencies were compared with those of 239 unrelated affected individuals. SNP data were analyzed for linkage disequilibrium (LD), association with autism, and haplotype

Table 2

SNPs Genotyped in *TSC2* and below ADHD Linkage Peak

Genomic Region anddbSNP Number	Position[a]	Allele Variant
TSC2[b]:		
rs2516740	2037111	G→T
rs2074968	2050572	G→C
rs1051771	2078585	G→C
ADHD linkage peak[c]:		
rs153783	15509867	G→A
rs1125972	15514488	G→A
rs1065838	15535539	T→C
rs1107143	15594082	T→C

[a] From the May 2004 release of the International Human Genome Sequencing Consortium.

[b] The distance between *rs2516740* and *rs2074968* is 13,461 bp; the distance between *rs2074968* and *rs1051771* is 28,013 bp.

[c] The distance between *rs153783* and *rs1125972* is 4,621 bp; the distance between *rs1125972* and *rs1065838* is 21,051 bp; the distance between *rs1065838* and *rs1107143* is 58,543 bp.

transmission disequilibrium. Twelve SNPs that showed evidence of association were subsequently genotyped for 91 autism trios (one affected individual and two unaffected parents). Logistic regression analysis was performed for haplotypes from genes containing SNPs that showed evidence of association.

Description of Candidate Genes

GRIN2A (MIM 138253) and *SSTR5* are cell-surface receptors. *GRIN2A* encodes one of four N-methyl-D-aspartate (NMDA) receptor 2 subunits that form an NMDA-receptor channel. NMDA receptors have been shown to be involved in long-term potentiation (LTP), an activity associated with learning and memory. Targeted disruption of the mouse homologue of *GRIN2A, GluRε*, reduces LTP and produces other behavioral deficits (Sakimura et al. 1995). Evidence of association between a *GRIN2A* exon 5 polymorphism and ADHD has also been reported recently (Turic et al. 2004). *SSTR5* (MIM 182455) encodes a G-protein–coupled receptor for the inhibitory hormonal regulator somatostatin. There is evidence that somatostatin and its receptors are transiently expressed in developing neurons; they are thought to control cell migration. Lauritsen et al. (2003) have performed analysis of 12 polymorphisms in *SSTR5* in 79 families with autism but found no evidence of association.

Three of the genes screened are involved in transcriptional regulation. *TBX6* (MIM 602427) is a member of the T-box family that encodes transcription factors involved in the regulation of developmental processes and is expressed in discrete embryonic domains (Papapetrou et al. 1999). *UBN1* encodes a ubiquitously expressed nuclear protein found to interact with EB1 and c-Jun, which, in turn, interact with cellular promoters and transcription factors. *CREBBP* (MIM 600140) encodes a large protein of 2,442 aa and binds specifically to the protein-kinase-A–phosphorylated form of the CREB protein (Chrivia et al. 1993). *CREBBP* can activate transcription through its C-terminus and acts to augment the ability of phosphorylated CREB to activate cAMP-responsive genes, interacting directly with transcription factor IIB. Mutations in *CREBBP* are responsible for Rubenstein-Taybi syndrome (RTS [MIM 180849]), which is characterized by mental retardation and broad thumbs and halluces. RTS has been associated with autism-like features in a minority of cases (Hellings et al. 2002).

EMP2 (MIM 602334) and *BFAR* are involved in cell-cycle control. *EMP2* has high sequence similarity with peripheral myelin protein 22, suggesting that it may have a similar function in cell proliferation and cell-cell interactions. *BFAR* encodes a protein that inhibits Bax-induced and Fas-induced apoptosis (Itoh et al. 1993; Wei et al. 2001).

The remaining three candidate genes were chosen because of their expression patterns and potential connection with neurological disease. *ABAT* (MIM 137150) is responsible for the catabolism of gamma-aminobutyric acid (GABA) into succinic semialdehyde (De Biase et al. 1995). GABA is an important inhibitory neurotransmitter in the brain and is estimated to be present in 20%–50% of human synapses. GABA is involved directly and indirectly in the pathogenesis of many neurological diseases, including convulsions, which may be relevant, given that approximately one-third of individuals with autism develop epilepsy (Fombonne 2002). *A2BP1* (MIM 605104) possesses a ribonucleoprotein motif, which is highly conserved among RNA-binding proteins. *A2BP1* binds to the C-terminus of ataxin-2 and may contribute to the pathology of spinocerebellar ataxia type 2 (SCA2 [MIM 183090]) (Shibata et al. 2000). *MRTF-B* was identified from cDNA libraries of size-fractioned adult and fetal human brain (Nagase et al. 1999). *MRTF-B* is highly expressed in the amygdala, caudate nucleus, hippocampus, and fetal brain and was recently found to be involved in muscle differentiation (Selvaraj and Prywes 2003).

Material and Methods

IMGSAC Family Collection

The IMGSAC has collected family data in successive stages, and details of assessment protocols are described elsewhere (IMGSAC 2001). To summarize, after an initial screen for medical disorders and a preliminary assessment, parents were administered the Autism Diagnostic Interview–Revised (ADI-R) and the Vineland Adaptive Behavior Scales (Sparrow et al. 1984; Le Couteur et al. 1989). Probands were administered the Autism Diagnostic Observation Schedule (ADOS) (Lord et al. 2000), and psychometric data were obtained. Affected individuals also underwent a thorough medical examination so that those with tuberous sclerosis and neurofibromatosis could be excluded, and any morphological abnormalities were recorded. All participating clinical sites followed the same assessment protocol, and ADI-R and ADOS consensus-coding meetings were held regularly by clinicians. DNA was extracted from blood, buccal swabs, or cell lines by use of the Nucleon kit and standard techniques. Samples from one affected individual per multiplex family were tested for fragile-X syndrome, and, when possible, all affected individuals underwent karyotypic analysis (Lamb et al. 2005). Relevant ethical committees have reviewed this study.

The IMGSAC has collected a further 67 sibling pairs since the publication of the completed genome screen, for a total of 219 sibling pairs from 207 families. In addition, 32 families containing extended relative pairs have been collected. The male:female ratio of affected

individuals is 3.9:1. One affected individual was selected at random from each multiplex family, to create a group of 239 affected unrelated individuals for case-control analysis. A singleton sample consisting of 91 autism trios with a clinical diagnosis of autism has also been collected. The singleton sample consists of 22 trios from Denmark, 37 trios from the Netherlands, and 32 trios from the United Kingdom. Analysis of linkage data from the IMGSAC sample on chromosomes 2,7, and 16 does not indicate that country of origin significantly affects either the magnitude or location of the linkage signal or the evidence of association in the complete IMGSAC sample (authors' unpublished data) (Bonora et al. 2005).

Controls

Random controls from the European Collection of Cell Cultures (ECACC) ($n = 192$) were used for SNP typing. DNA was extracted at ECACC from lymphoblastoid cell lines derived from white randomly selected blood donors whose parents and grandparents were born in the United Kingdom or Ireland. The control samples were not characterized for autism phenotypes. Half of the controls were male and half female. Although the male:female ratio is different for cases and controls, only the presence of a very strong sex-specific allelic effect could reduce the ability to detect association. Linkage analysis based on the sex of APSs on chromosome 16 shows some evidence of a sex-specific effect, but it is unclear how this would affect the frequency of a disease-causing variant (Lamb et al. 2005).

Mutation Screening

Intron-exon boundaries were characterized *in silico* by use of National Center for Biotechnology Information LocusLink and University of California–Santa Cruz (UCSC) Golden Path (see UCSC Genome Browser Web site). The mRNA sequence was subjected to BLAST analysis, to determine or confirm genomic structure. Exon-specific primers were designed using Primer3. Large exons and promoter regions were covered by overlapping PCR products of ~500 bp. PCRs were performed in a reaction mixture containing 40 ng of DNA and final concentrations of 2.5 mM KCl; 2.0, 2.5, or 3.0 mM $MgCl_2$; 0.2 mM primer; 0.2 mM dNTPs; 0.9 U Ampli*Taq* Gold polymerase; and 0.1 U *Pfu*Turbo (Stratagene). For GC-rich products of 2.5%–10%, DMSO was used with a mix of 0.2 mM dATP, dTTP, and dCTP; 0.1 mM dGTP; and 0.1 mM 7-deaz-2′-deoxyguanosine-5′-triphosphate (Amersham Pharmacia Biotech). The PCR consisted of touchdown protocols with a maximum annealing temperature of 50°C–67.5°C (T1) and a minimum annealing temperature of 42.5°C–60°C (T2) (15 min at 95°C, followed by 14 cycles of 30 s at 95°C, 30 s at T1−0.5°C per cycle, and 30 s at 72°C, followed by 29 cycles of 30 s at 95°C, 30 s at T2, and 30 s at 72°C). PCR products were denatured and reannealed using the following protocol: 4 min at 95°C, 30 s at 95°C, and 42 cycles of 30 s at −1.6°C per cycle.

The Transgenomic WAVE DNA Fragment Analysis System was used to detect heterozygous individuals by DHPLC. Gradients were designed using WAVEMAKER Software (version 4.1) and spanned a temperature range of 51.5°C–70°C. Crude PCR products were loaded onto the DNASep Column, and elution profiles were analyzed for heteroduplex formation. Melting curves for fragments containing c^7dGTP were experimentally determined.

Variant Identification

Common variants were identified by sequencing seven heteroduplex samples and a control sample with a homoduplex profile. Low-frequency variants were sequenced for all individuals showing the heteroduplex profile and for a control sample. Sequencing was performed with BigDye Terminator mix on ABI 377 machines, in accordance with standard protocols, and variants were analyzed with Sequence Analysis and Sequence Navigator.

SNP Typing

Mass-extension assay.—Primers and mass-extension assays were designed for Sequenom SNP typing with SpectroDESIGNER. PCRs were performed in 10-μl reactions with 8 ng of DNA template and final concentrations of 2.5 mM KCl and $MgCl_2$, 200 μM dNTPs, 1 U of HotStar*Taq* Polymerase (Qiagen), and primer concentration of 0.15 mM. Multiplex PCRs were performed using up to four primer pairs per reaction; 45 cycles of 20 s at 95°C, 30 s at 56°C, and 1 min at 72°C were performed. Nonincorporated dNTPs were removed with shrimp alkaline phosphatase for 20 min at 37°C and for 5 min at 85°C. The mass-extension reaction was performed using MassEXTEND enzymes—thermosequenase, hME termination mixes, and hME extension primers; 55 cycles were performed for 5 s at 94°C, for 5 s at 52°C, and for 5 s at 72°C. Unincorporated ddNTPs and dNTPs were removed with SpectroCLEAN resin, and products were transferred to a 384 SpectroCHIP by use of the SpectroPOINT robot. The chip was read using the Bruker Biflex III Mass Spectrometer system. Sequenom Genotype Analyzer was used to manually check failed genotypes and to assign alleles, when possible. Data either were exported from SpectroTYPER using Report Generator or were directly streamed into the Wellcome Trust Centre for Human Genetics Laboratory Information Management System Integrated Genotyping System (IGS).

Restriction digest.—PCR for *ABAT* exon 2 was performed in 10-μl reactions containing 24 ng of DNA and final concentrations of 2.5 mM KCl, 2.5 mM $MgCl_2$,

0.2 mM primer, 0.2 mM dNTPs, and 0.25 U Ampli*Taq* Gold polymerase. Touchdown PCR was performed with a maximum annealing temperature of 62.5°C. PCR product from each plate was tested on agarose and was incubated with 0.1 U of *Bso*B1 and 2 μl of buffer 2 (New England BioLabs), in a total volume of 20 μl at 37°C for 6 h. Digested products were visualized on a 3% agarose gel. Digests were genotyped blind to affection status and were entered into the IGS.

Statistical Analysis

Association analysis.—The PDT was used to test for transmission disequilibrium (Martin et al. 2000) for the multiplex and singleton samples. This test is more powerful than the transmission/disequilibrium test (TDT) for the multiplex sample and analyzes informative branches of extended families with all available data. The PDT retains the within-family nature of the TDT and provides a valid test for association, even in the presence of population substructure. A test for the transmission disequilibrium of genotypes—rather than alleles—was also performed using the geno-PDT, an extension of the PDT (Martin et al. 2003*a*). Case-control association analysis was performed using related cases from the multiplex families and controls, as proposed by Slager and Schaid (2001). This test accounts for the biological relationship between cases and compares genotype frequencies between cases and controls by comparing the trend in proportions as the dosage of the risk allele increases. Genotype-frequency differences between unrelated affected cases from the multiplex and singleton samples and controls were analyzed using Fisher's exact test for count data.

Haplotype analysis.—Haplotypes were determined using PHASE v2.0.2 for groups of SNPs within candidate genes. Any genotypes that were not assigned phase with >90% certainty were disregarded. Because PHASE v2.02 reconstructs haplotypes from population genotype data, two rounds of haplotype reconstruction were performed. First, multiplex case and control data were analyzed together, and, second, singleton case and control data were analyzed together, with the result of small differences in the numbers of haplotypes constructed for controls. PHASE v2.0.2 was selected instead of GENEHUNTER 2.0 for determination of haplotypes, because PHASE v2.0.2 does not assume linkage equilibrium over short distances and can be used to generate haplotypes for individuals without parental genotypes. Haplotype analysis was performed using TRANSMIT for both multiplex and singleton data (Clayton and Jones 1999). TRANSMIT implements a TDT that analyzes the transmission of multilocus haplotypes and has good power when parental genotypes are available, as is the case in this sample (~80%) (Martin et al. 2003*b*). For each haplotype, a χ^2 test with 1 df for excess transmission is produced. A global test for association is also produced with $H - 1$ df (H = number of haplotypes for which information is available).

Logistic regression analysis.—Logistic regression analysis of haplotype data was performed with the program R (R Project for Statistical Computing), as described by Wallenstein et al. (1998). Unrelated affected individuals from the multiplex and singleton families, identical to those used in the case-control analyses, were used in the logistic regression analysis. They impose an assumption of additivity on the haplotype analysis, thus providing predictor variables that include the number of copies (0, 1, or 2) of each haplotype and that yield a logistic regression for which the outcome is either case or control. The model is constructed so that each coefficient gives the log odds ratio for disease for an individual with a single copy of the relevant haplotype and another copy of the baseline haplotype, compared with an individual with two copies of a baseline haplotype. Models were fitted in which each gene (*ABAT, CREBBP, GRIN2A,* and *TSC2*) was tested alone and jointly. Genotypes for an individual (i.e., the pairs of haplotypes) are denoted as (*x;y*).

LD mapping.—Haplotype files were analyzed using HaploXT to generate an output file describing LD (Abecasis and Cookson 2000). Individual marker-by-marker contingency tables were produced and imported into the Graphical Overview of Linkage Disequilibrium (GOLD) program. Lewontin's standardized disequilibrium coefficient D' was used (Lewontin and Kojima 1960), because it accounts for allele frequency, although, like other measures of LD, it may be artificially inflated by rare haplotypes.

Error detection.—Mendelian consistency of SNP genotype data was checked using PedCheck (O'Connell and Weeks 1998), and any inconsistent genotypes were removed. Data were prepared for statistical analysis by use of Mega2 (Mukhopadhyay et al. 1999; see the University of Pittsburgh Department of Human Genetics Web site). Genotypes flanking double recombinants, detected after running GENEHUNTER 2.0, were checked and disregarded if ambiguous. SIBMED (sib-pair mutation error detection) was also used to remove possible genotyping errors (Douglas et al. 2000).

Permutation of affection status: Fisher's exact test.—An empirical experimentwise P value for association, to allow for multiple testing across all 28 SNPs, was obtained by random permutations of case and control status. For each permutation, the minimum P value from Fisher's exact test across all SNPs was recorded. An empirical estimate of the experimentwise P value was then given by the proportion of permutations for which the minimum P value exceeded that obtained for the original

Table 3

Missense Mutations in Candidate Genes Identified by DHPLC

Gene and Exon[a]	cDNA Nucleotide	Wild-Type/ Variant Allele	Amino Acid Substitution	No. of Heterozygotes in Screened Samples (n = 48)	Minor-allele Frequency in	
					Multiplex Cases (n = 239)	Controls (n = 192)
ABAT:						
2	208	A/G	G56R	21	.457	.369
BFAR:						
3	563	G/T	M140R	6	.332	.388
6	1096	A/G	V315I	2	.018	.077
MRTF-B:						
5	1624	C/G	N543K	1	.024	.035
SSTR5:						
1	92	C/T	T333I	1	ND[b]	ND[b]
UBN1:						
3	422	A/C	A103N	1	.004	.000
9	1413	G/A	Y435C	2	.031	.027

[a] SNPs genotyped in the entire IMGSAC sample and controls for all but exon 1.
[b] ND = no data were available.

sample. The same procedure was repeated for the singleton cases and controls.

Permutation of affection status: logistic regression analysis.—To test the significance of the logistic haplotype regression analysis of the four genes, with consideration of multiple testing, PHASE v2.0.2 was used to generate "best pair" haplotypes for cases and controls; affection status was then permuted 10,000 times. An additive logistic regression model incorporating the joint haplotypic effects across *ABAT, CREBBP, GRIN2A,* and *TSC2* was fitted to the permuted data, and *P* values were recorded. The significance of the joint model was estimated as the proportion of permutations in which the regression *P* value was smaller than that observed in the joint analysis of the original sample data.

Bioinformatic Analysis

Variants found within 20 bp of intron-exon boundaries were analyzed for possible effects on splice-site signal by use of sequence analysis tools available at the Berkeley *Drosophila* Genome Project Web site (Brunak et al. 1991; Kulp et al. 1997). Variants within exons were analyzed for possible changes to exon-splice enhancers by use of the ESEfinder interface (Cartegni et al. 2003). The program Pfam was used to determine the location of coding variants within known protein domains (Bateman et al. 2004). Homology searches by use of PolyPhen were performed to determine the extent of conservation at the position of missense mutations in known sequences (Sunyaev et al. 2001). PolyPhen uses empirical rules and structural information to characterize the effect of an amino acid substitution.

Results

Failure Rates and Error Checking

Of the 38 SNPs initially selected for genotyping, 10 were disregarded because of failing multiplex PCR, failing extension reaction, poor success rate, or a weak extension assay. Additionally, the A/G SNP causing the G56R substitution in *ABAT* was genotyped by restriction enzyme digestion because of spurious heterozygotes in the mass-extension assay. Of a total of 47,646 SNP genotyping assays performed, 44,647 genotypes were generated, a 93.7% success rate. Analysis with PedCheck led to the removal of 35 genotypes from the IMGSAC data (0.08% error rate). Most of these were due to nonconservative SNP genotyping allele calls from the Sequenom Genotype Analyzer. Checking for double recombinants resulted in the removal of an additional 26 genotypes. All markers were analyzed for Hardy-Weinberg equilibrium and were not found to deviate from expected allele frequencies ($P < .01$); a small number of markers with low heterozygosity did deviate from expected allele frequencies ($P < .05$), but this is likely to be due to chance. Finally, SIBMED was run with all SNP data, and microsatellite data were generated from the completion of the IMGSAC genome screen; an additional 14 low-probability genotypes were removed.

Mutation Screening

Seven missense variants were detected during mutation screening of the coding sequence, as described in table 3. Additionally, 108 noncoding/silent variants were found, as shown in table 4. Several missense variants were found in a single family; the T333I change in *SSTR5*

Table 4

Nonmissense Variants in Candidate Genes Identified by DHPLC

Gene, Location, and cDNA Nucleotide	Major/Minor Allele	No. of Heterozygotes in Screened Samples ($n = 48$)	Minor-Allele Frequency[a] in	
			Cases	Controls
A2BP1:				
Intron 1:				
730−6	+/CTT del	2	…	…
Intron 2:				
804−47	A/C	24	…	…
804−48	A/T	24	…	…
Intron 3:				
852−20	+/T del	23	…	…
Exon 5:				
1067	C/G	2	…	…
1098	G/A	2	…	…
Intron 5:				
1182−25	A/C	4	…	…
Exon 6[b]:				
1271	A/G	1	.021	.003
Intron 7:				
1379+109	A/G	1	…	…
Intron 11:				
1673−17	−/CT ins	21	…	…
Intron 13:				
183−19	A/G	2	…	…
Intron 15[b]:				
1973+3	A/G	6	.004	0
3′ UTR:				
2102+141	G/T	23	…	…
ABAT:				
Intron 1[b]:				
102+49	T/C	15	.391	.2952
Intron 2:				
209+11	A/G	1	…	…
210−68	C/A	1	…	…
210−5	C/T	1	…	…
Exon 4[b]:				
350	C/T	10	…	…
Intron 8:				
582+14	−/CA ins	5	…	…
582+54	A/C	2	…	…
Exon 9:				
685	C/T	1	…	…
Exon 12:				
995	A/C	2	…	…
Intron 13:				
1313−120	−/G ins	10	…	…
Intron 14:				
1423+9	C/T	22	…	…
3′ UTR:				
1544+71	A/G	1	…	…
BFAR:				
Exon 1:				
50	T/C	1	…	…
Intron 1:				
89+105	G/A	4	…	…
Exon 2:				
178	T/A	1	…	…
Intron 6:				
937+84	−/GTG ins	1	…	…
3′ UTR:				
1507+794	A/G	1	…	…
1507+795	G/C	1	…	…

(*continued*)

Table 4 (continued)

Gene, Location, and cDNA Nucleotide	Major/Minor Allele	No. of Heterozygotes in Screened Samples ($n = 48$)	Minor-Allele Frequency[a] in	
			Cases	Controls
CREBBP:				
Exon 3:				
258	G/A	1	…	…
Exon 4:				
1037	T/C	3	…	…
Intron 3[b,c]:				
1995−8585	T/C	ND[d]	.005	0
Intron 4:				
1174−76	+/TTTG del	1	…	…
Intron 6:				
1530−31	G/T	3	…	…
Intron 7[b,c]:				
1730+215	G/A	ND[d]	.349	.363
1731−58	A/C	1	…	…
Exon 11:				
2152	T/C	1	…	…
Intron 11[b]:				
2311+78	C/T	3	.077	.007
Intron 13:				
2358+59	T/C	1	…	…
Intron 17[b]:				
3449+6163	C/T	23	.414	.522
Intron 20[b,c]:				
3898+3049	C/T	ND[d]	.033	.015
Intron 22[b]:				
4040−8	T/C	3	.014	.022
Exon 23:				
4097	A/C	1	…	…
Intron 26[b,c]:				
4480+42	G/A	11	.027	.015
Intron 29[b]:				
4937−19	C/G	3	.028	.019
Exon 31:				
5160	A/C	2	…	…
Exon 32:				
7411	G/A	3	…	…
3′ UTR:				
7525+26	−/C ins	2	…	…
EMP2:				
Exon 2:				
53	G/C	7	…	…
Exon 3:				
172	C/T	15	…	…
205	G/A	23	…	…
245	C/T	23	…	…
3′ UTR:				
Unidentified	Unidentified	13	…	…
GRIN2A:				
Intron 1:				
112−33	C/T	27	…	…
Intron 5:				
1318−15	G/C	1	…	…
Exon 6[b]:				
1576	T/C	21	.424	.597
Intron 7:				
1808−29	G/C	1	…	…
Intron 8:				
1964+31	−/GA ins	25	…	…
Intron 10[b]:				
2316−35	G/C	25	.447	.381
Intron 11:				
2316−21	−/GA ins	1	…	…

(*continued*)

Table 4 (continued)

Gene, Location, and cDNA Nucleotide	Major/Minor Allele	No. of Heterozygotes in Screened Samples ($n = 48$)	Minor-Allele Frequency[a] in	
			Cases	Controls
Exon 11:				
2396	G/A	2	…	…
Exon 14:				
3539	G/T	2	…	…
3′ UTR[b]:				
4705+39	C/A	2	.005	.014
4105+1213	A/G	26	.298	.510
MRTF-B:				
Exon 1:				
90	G/C	1	…	…
Exon 2:				
379	T/G	1	…	…
Intron 3:				
1475−12	G/T	2	…	…
Intron 4:				
1513+21	C/T	1	…	…
3′ UTR[b]:				
2487+67	T/C	1	…	…
2487+561	T/C	1	…	…
2487+1110	G/C	1	…	…
2487+1441	T/A	1	…	…
2487+1713	A/C	1	…	…
2487+1849	G/T	1	…	…
TBX6:				
5′ UTR:				
120+6	A/G	3	…	…
Intron 7:				
938−4	C/T	20	…	…
937+19	−/GAT ins	23	…	…
Exon 8:				
1252	A/G	1	…	…
Intron 8:				
1335+428	G/C	1	…	…
UBN1:				
5′ UTR:				
1−17	C/G	7	…	…
Intron 1:				
75−109	TA/AG	3	…	…
75−63	T/C	3	…	…
75−21	A/G	6	…	…
74+26	T/C	3	…	…
75+40	T/G	25	…	…
74+68	G/T	2	…	…
Exon 3:				
147	A/G	2	…	…
Intron 3:				
454+22	A/G	2	…	…
Intron 4:				
546+3	G/A	4	…	…
Exon 5:				
599+5	C/T	11	…	…
Intron 6:				
784+8	G/A	1	…	…
Intron 9:				
1296+11	−/G ins	1	…	…
1296+39	+/TT del	1	…	…
1425+62	C/T	3	…	…
Intron 14:				
1915+17	A/G	15	…	…
1915+53	A/G	2	…	…
Exon 15:				
2743	T/G	7	…	…
2994	T/G	6	…	…
Exon 18:				
3486	C/T	3	…	…
3519+1006	C/T	12	…	…
3′ UTR:				
3519+157	T/C	7	…	…
3519+1008	−/C ins	1	…	…
3519+1889	C/T	4	…	…
3519+2302	C/G	1	…	…
Exon 18:				
3519+2588	A/C	13	…	…
3519+2660	C/T	26	…	…

[a] No data were available unless values are shown. For cases, $n = 239$; for controls, $n = 192$.
[b] SNPs genotyped in the entire IMGSAC sample and controls.
[c] SNPs identified through dbSNP or Sequenom's RealSNP Array.
[d] ND = no data were available.

was found in an unaffected mother, and the N543K missense variant in *MRTF-B* was paternally transmitted to one of two affected sons. Bioinformatic analysis indicated that these variants do not occur in any known functional domain and that they were predicted to have a benign effect. However, there are few homologous sequences available for thorough conservation analysis of *MRTF-B*. Two missense variants were detected in *BFAR*. During screening, the first missense variant in exon 3, M140R, was found in six individuals and cosegregated with autism in these families, although three parents were found to be homozygous for arginine. The second missense variant in *BFAR*, V315I, was detected, during screening, in two individuals and cosegregated with autism in both families. Bioinformatic analysis of the missense variants in *BFAR* showed that they were likely to have only a benign effect on protein structure and that they do not occur at highly conserved residues. Two rare variants were found in *UBN1*. The A103N variant in exon 3 was detected in one family and was transmitted paternally to two affected offspring. This variant is in a region of low complexity without a strong domain identity. A variant in exon 9, Y435C, was detected in two individuals but cosegregated with autism in only one family. Although this variant was predicted to have a potentially harmful effect on protein function, no homologous sequences were available, and detailed comparative analysis was not possible.

One missense variant was found at a high frequency in *ABAT* exon 2. However, the G56R change is distant from the enzyme's catalytic site and is unlikely to hinder protein function. Sequence homology analysis also pre-

dicted that the variant is likely to be benign. All of the missense variants detected during screening, except that in *SSTR5*, were typed in the entire IMGSAC sample and 192 controls. No evidence of transmission disequilibrium was found, and no significant differences in case-control allele frequencies were observed for these missense variants, as shown in table 3. None of the variants detected by DHPLC within 20 bp of the intron/exon splice site were found to potentially affect splice signal.

Association Analysis and Logistic Regression Analysis for Candidate Genes

Twenty-eight SNPs were successfully genotyped in the IMGSAC families and controls, as shown in table 5. The family data were analyzed for allele and genotype transmission disequilibrium by use of the PDT. Fisher's exact test for count data was used to compare genotypes observed in cases and controls. Additionally, multiplex genotype data were compared with controls by use of a trend test that accounts for relatedness between affected individuals. No significant evidence of association was found with these tests either for SNPs in *A2BP1, BFAR, MRTF-B, TSC2,* and *UBN1* or for SNPs within the region of linkage with ADHD, as described in table 5. SNPs within *ABAT, GRIN2A, TSC2,* and *CREBBP* were analyzed for LD, to determine whether any haplotype blocks were present in the data. The SNPs within *ABAT* and *GRIN2A* were in strong LD with each other ($D' = 0.667$–1), whereas there was variability in LD between SNPs in *TSC2* and *CREBBP.* Because of the manner in which most of the SNPs were identified and the comparatively large inter-SNP distances, it is unlikely that any underlying haplotype-block structure is fully described by these analyses. But, given the evidence of LD between SNPs in candidate genes, haplotype transmission analysis was performed. No evidence of haplotype transmission disequilibrium was detected, using TRANSMIT, in the IMGSAC families, although an exhaustive analysis of all possible SNP combinations was not performed. There was evidence of association with SNPs genotyped in *ABAT, CREBBP,* and *GRIN2A,* and further logistic regression analysis was performed, as described below. The logistic regression analysis performed does not rely on any LD between SNPs and assesses only the risk of being affected, given a particular genotype. Haplotypes are described as sequential SNP alleles across the candidate genes, with alleles labeled as in table 5 (e.g., haplotype h12 for *ABAT* describes allele 1 at *rs1731017,* followed by allele 2 at the SNP within exon 2 of the gene).

ABAT.—Both SNPs within *ABAT* showed some evidence of association in the multiplex sample, with the trend-test and case-control analysis. The strongest evidence of association was from the SNP within *ABAT* intron 1, which, when analyzed using the trend test, indicated that allele 2 was responsible for increasing the risk of autism ($P < .001$). Fisher's exact test for count data gave a P value of .0089, which indicates a significant difference in genotype distribution between cases and controls, as shown in table 5. This result was supported by the analysis of the same SNP in the singleton sample with use of Fisher's exact test ($P = .0112$), as shown in table 5. Logistic regression of SNPs within *ABAT* did not indicate that a particular haplotype significantly increased the risk of autism, although there is an overall difference in haplotype distribution between cases, from the multiplex and singleton samples, and controls ($\chi^2 = 20.37$; $P = 1.422 \times 10^{-4}$ for multiplex cases; $\chi^2 = 6.18$; $P = .1031$ for singleton cases), as shown in table 6. An additional SNP was also genotyped in exon 4 of *ABAT* but was excluded from the final analysis, despite supporting the results described above, because of a low genotyping success rate (65.9% in the multiplex sample; 73.9% in the singleton sample).

CREBBP.—In the multiplex sample, association was found for *CREBBP* with two SNPs within introns 22 and 29, with use of the geno-PDT ($P = .0363$ and $P = .001$, respectively) and with use of the allelic PDT ($P = .0074$ and $P = .0024$, respectively). However, both these SNPs had a minor-allele frequency <6%, and the number of transmissions to affected individuals was small. There was some additional evidence of association from the trend test for the SNP in intron 17 ($P = .0103$), but these results were not replicated in the singleton sample. Conversely, SNPs within introns 7 and 20 showed evidence of association in the singleton sample ($P = .03078$ and $P = .0196$, respectively) but not in the multiplex sample. It should be noted that the minor-allele frequency was also low for the SNP in intron 20, and the P value was of marginal significance for the SNP in intron 7; it is likely these represent chance findings. These results are shown in table 5. There was no evidence from the logistic regression analysis to support the presence of a risk haplotype across *CREBBP,* as shown in table 6.

GRIN2A.—One SNP in *GRIN2A* intron 10 showed some evidence of transmission disequilibrium with the geno-PDT ($P = .013$) and the trend test ($P = .0337$). This was supported by SNPs within exon 6 and the 3′ UTR that also showed evidence of association with the trend test and case-control analysis, as shown in table 5. Although the singleton analysis does appear to support these results, closer inspection shows that the transmission of different genotypes is responsible for the significant result for the SNP within exon 6—genotype 1/1 in the singleton sample compared with genotype 1/2 in the multiplex sample, as shown in table 5. A significant difference in the distribution of haplotypes, between cases and controls, was found using logistic regression analysis

Table 5

Association Analysis of SNP Genotyping Data for Multiplex Families ($n = 239$) and Singleton Families ($n = 91$)

Genomic Region, Location, and dbSNP Number	Genotyping Success (%)	Percentage of Genotypes with Alleles									*P* for				Trend Test[a] Analysis					
		1/1 in			1/2 in			2/2 in			Multiplex		Singleton		No. of Sibling Pairs/ Controls	*P* for High-Risk Allele(s)			Fisher's Exact Test of Sample	
		Multiplex Cases	Singletons	Controls	Multiplex Cases	Singletons	Controls	Multiplex Cases	Singletons	Controls	Geno-PDT	Allelic PDT	Geno-PDT	Allelic PDT		1 and 2[b]	1[c]	2[c]	Multiplex	Singleton
A2BP1:																				
Exon 6	89.1	95.8	ND	99.4	4.2	ND	.6	0	ND	0			ND	ND	219/172				.0265	ND
Intron 15	54.3	0	ND	0	.8	ND	0	99.2	ND	100			ND	ND	103/113					ND
ABAT:																				
Intron 1:																				
rs1731017	76.8	37	47.3	49.7	42.5	33.8	40.8	20.4	18.9	9.5					142/147	.0061		.0113	.0089	.0112
Exon 2:	93.3	21	5.7	13.5	46.2	54	43.2	32.9	40.2	43.2					212/185		.0257		.0483	
BFAR:																				
Exon 3	91.6	44.5	ND	37.5	43.6	ND	47.8	11.8	ND	14.7			ND	ND	209/184					ND
Exon 6	84.5	0	ND	.6	3.6	ND	4.2	96.4	ND	95.2			ND	ND	186/167					ND
CREBBP:																				
Intron 3:																				
rs3025702	97	0	ND	0	.9	ND	0	99.1	ND	100			ND	ND	244/187					ND
Intron 7:																				
rs130021	94.5	42.5	39.3	40.6	45.3	40.4	50.3	12.3	20.2	9.1					210/187					.03078
Intron 11:																				
rs130002	55.3	.6	0	0	2.5	1.1	1.4	96.9	98.9	98.6					127/74					
Intron 17:																				
rs886528	89.2	17.1	14.6	27.2	45.5	48.3	41	37.4	37.1	31.8					166/173			.0103		
Intron 20:																				
rs130025	90.5	0	0	0	6.5	2.3	2.9	93.5	97.7	97.1			.0196	.0196	219/172					
Intron 22:																				
rs3025684	98.1	97.3	100	95.7	2.7	0	4.3	0	0	0	.036	.0074			225/186					
Intron 26:																				
rs129967	92.8	0	0	0	5.4	0	2.9	94.6	100	97.1					227/175					
Intron 29:																				
rs130008	94.9	0	1.1	0	5.6	4.5	3.7	94.4	94.3	96.3	.001	.0024			221/187					
GRIN2A:																				
Exon 6	93.7	18.1	47.7	35.6	75.8	43.2	58.8	6.2	9.1	5.6					239/177	.0013		<.0001	.0003	.0486
Intron 10	96.3	20	54.7	14.5	75.7	37.2	81.2	4.3	8.1	4.3	.013				242/186			.0337		2.32×10^{-12}
3′ UTR:	96.3	99.1	ND	97.3	.9	ND	2.7	0	ND	0			ND	ND	240/186					ND
rs1014531	75.5	8.9	19.7	26	78.5	49.3	50.4	12.6	31	23.6					218/127		.0214	<.0001	2.9×10^{-7}	.0439
MRTF-B:																				
Exon 5	75.9	0	ND	0	4.7	ND	6.8	95.3	ND	93.2			ND	ND	155/177					ND
TSC2:																				
5′ UTR:																				
rs2516740	93.5	62.6	ND	63.5	31.3	ND	30.7	6.1	ND	5.8			ND	ND	219/189					ND
Intron 10:																				
rs2074968	93.7	17.4	ND	15	48.2	ND	52.8	34.4	ND	32.2			ND	ND	231/180					ND
3′ UTR:																				
rs1051771	82.1	0	ND	0	2.1	ND	6.8	97.9	ND	93.2		.0143	ND	ND	169/162	.0023	.0022		.0346	ND
UBN1:																				
Exon 3	83.8	100	ND	0	0	ND	0	0	ND	0			ND	ND	183/161					ND
Exon 9	96.8	93.9	ND	94.7	5.7	ND	5.3	.4	ND	0			ND	ND	233/187					ND
ADHD linkage peak:																				
14–15 Mb[d]:																				
rs153783	92.1	54.6	ND	51.6	37.7	ND	37.4	7.7	ND	11.1			ND	ND	202/190					ND
rs1125972	96.8	88.7	ND	86.6	10.4	ND	12.9	.9	ND	.5			ND	ND	240/186					ND
rs1065838	93.7	8.4	ND	8.4	40.2	ND	41.6	51.4	ND	50			ND	ND	201/190					ND
rs1107143	96.1	38.2	ND	39.2	46.5	ND	46.2	15.4	ND	14.5			ND	ND	237/186					ND

Note.—Results shown only when $P < .05$. ND = no data were available.

[a] Trend test performed on multiplex sample.

[b] Additive allele effect.

[c] Dominant allele effect.

[d] Below linkage peak identified by Smalley et al. (2002).

Table 6

Haplotype Logistic Regression Analysis

Gene and Haplotype	No. of Multiplex Cases[a]	No. of Multiplex Controls[b]	Log Relative Risk	SE	P (>|z|)	χ^2	P	df
ABAT:						20.37	1.4226×10^{-4}	3
h11[c]	18	7		.8226	.1246			
h12	162	85	−.1536	.4397	.0727			
h21	225	219	−.726	.4209	.0845			
h22	15	3	.7212	.7466	.334			
CREBBP:						16.52	.0355	8
h12122222[c]	7	3		1.3986	.242			
h21121212	5	0	14.6899	649.2732	.982			
h21221212	4	5	−.9564	.9743	.326			
h22112222	4	2	−.1661	1.1172	.882			
h22121212	71	85	−.9602	.7079	.175			
h22121222	28	32	−.9925	.7698	.197			
h22122212	118	93	−.7084	.7217	.326			
h22122222	51	38	−.662	.732	.366			
Other[d]	10	2	.6144	1.0274	.558			
GRIN2A:						62.05	1.7232×10^{-11}	6
h1111[c]	34	42		.4716	.0103[e]			
h1112	5	3	1.4334	.8411	.0883			
h1122	168	73	.8251	.3143	.0086[f]			
h2111	217	85	1.5479	.306	4.21×10^{-7}[f]			
h2121	2	20	−2.5906	.791	.001[b]			
h2122	13	7	.4442	.5331	.4047			
Other[d]	5	2	1.3741	.9345	.1414			
TSC2:						1.63	.8033	4
h112[c]	135	103		.22961	.22			
h122	178	142	−.0364	.1733	.834			
h212	13	6	.4427	.5116	.387			
h222	45	33	.012	.2326	.959			
Other[d]	3	4	−.5625	.783	.472			
	No. of Singleton Cases[a]	**No. of Singleton Controls[b]**	**Log Relative Risk**	**SE**	**P (>|z|)**	**χ^2**	**P**	**df**
ABAT:						6.18	.1031	3
h11[c]	2	7		1.6134	.2746			
h12	55	106	.4878	.8411	.562			
h21	120	244	.5369	.8129	.509			
h22	7	3	2.122	1.0626	.0458			
CREBBP:						9.77	.1346	6
h12122222[c]	4	2		1.5728	.422			
h21221212	0	5	−16.7436	1,072.0759	.988			
h22121212	35	86	−1.0608	.8039	.187			
h22121222	12	33	−1.2422	.8471	.143			
h22122212	49	96	−1.013	.8081	.21			
h22122222	30	40	−.6101	.8348	.465			
Other[d]	2	6	−1.4596	1.1374	.199			
GRIN2A:						37.05	5.8528×10^{-7}	5
h1111[c]	22	42		.5113	.138			
h1122	38	72	−.2564	.3701	.4885			
h2111	72	83	.7756	.3248	.0169[f]			
h2121	0	20	−17.5343	877.517	.9841			
h2122	4	7	−.4984	.6939	.4726			
Other[d]	0	4	−16.6823	1,972.8826	.9933			

[a] $n = 239$.
[b] $n = 192$.
[c] Referent haplotype.
[d] Haplotypes combined when <5.
[e] Haplotype has a protective effect $P < .05$.
[f] Haplotype has a deleterious effect $P < .05$.

for both the multiplex and singleton samples ($\chi^2 = 62.05$; $P = 1.7232 \times 10^{-11}$) compared with controls ($\chi^2 = 37.05$; $P = 5.8528 \times 10^{-7}$). Haplotype h2111—genotypes $x = (2,1,1,1)$, $y = (1,1,1,1)$—significantly increased the chances of being affected ($P = 4.21 \times 10^{-7}$), as did haplotype h1122—$x = (1,1,2,2)$,$y = (1,1,1,1)$—although not as strongly ($P = .0086$), as shown in table 6. The effect of the h2111 haplotype was replicated in the singleton sample but at a less significant level ($P = .0169$). Additionally, haplotype h2121—genotypes $x = (1,2,1,2)$, $y = (1,1,1,1)$—is found at a rate of 20/232 in controls, compared with 2/580 in multiplex cases, and appears to decrease the chances of being affected ($P = .001$).

Statistical significance of association and logistic regression analysis.—A simple Bonferroni correction, to reduce the chance of type 1 error to 0.05 across all tests, gives a significance threshold of .00714, if we consider that seven tests were performed: geno-PDT, allelic PDT, trend test for high-risk alleles 1 and 2, trend test for high-risk allele 1, and trend test for high-risk allele 2, Fisher's exact test, and logistic regression analysis. To reduce type 1 error to 0.001 across all tests, a Bonferroni significance threshold of .000142 is required. This represents a highly conservative correction, given that the PDTs and trend tests are correlated, although a number of the results in table 5 do exceed these significance levels.

Permutation tests, which provide more-accurate significance thresholds, were performed (see the "Material and Methods" section). For the single-point analysis with Fisher's exact test across 28 SNPs, the minimum P value of 2.9×10^{-7} for *rs1014531* in *GRIN2A* in the multiplex case-control analysis was smaller than any obtained in 10,000 permutations, for an empirical P value of $< 10^{-4}$. Similarly, for the joint logistic haplotype regression of all four genes, the observed χ^2 statistic of 102.9 with 34 df (nominal $P = 7.4 \times 10^{-9}$) was greater than for all 10,000 permutations; that is, an empirical P value of $< 10^{-4}$.

Separate haplotype logistic analysis of the four genes indicates that *GRIN2A* has by far the most significant contribution. A test comparing the fit of *GRIN2A* alone with that of all four genes gave a χ^2 value of 47.44 with 25 df ($P < .0043$), which suggests that, whereas the *GRIN2A* haplotypes are the principal contributors, there may be some effect associated with the *ABAT* haplotypes.

Discussion

Seven missense variants were detected during the screening of 10 positional candidate genes on chromosome 16p11-13 for autism. Although a number of these variants cosegregated with autism, both their presence at similar frequencies in controls and the lack of evidence of transmission disequilibrium suggest that they are unlikely to make a significant contribution to the etiology of autism.

A large number of noncoding and silent mutations were detected, but these are not predicted to alter splice-site recognition sequences. However, bioinformatic analysis alone may be insufficient to detect the effects of intronic and silent variants on splicing and transcriptional efficiency, as recently reported for *CFTR* and *SLC22A4* (Pagani et al. 2003; Tokuhiro et al. 2003), and it is possible that novel regulatory elements could be disrupted. A lack of missense mutations in coding sequence does not preclude the presence of harmful noncoding variants within a candidate gene. In fact, for a multigenic disorder such as autism, more subtle changes in gene regulation and expression are anticipated. This has been found for *PAD14,* a gene associated with rheumatoid arthritis (Suzuki et al. 2003). Suzuki et al. (2003) failed to detect any missense mutations in 48 affected individuals but found a functional SNP haplotype affecting mRNA stability.

The 28 SNPs genotyped in the IMGSAC families are unlikely to provide sufficient coverage of the genomic regions to conclusively exclude genes by association mapping. This was particularly true for genes such as *A2BP1,* which has a genomic region of 1,690 kb, from which only two SNPs with low heterozygosities were genotyped. Although a number of the SNPs genotyped in this study had low heterozygosities, it has been argued that they are useful for detecting association between disease variants with more-recent origins, despite reduced informativeness (Pritchard and Cox 2002). Additionally, even though most of the SNPs were successfully genotyped in 85% of individuals, the SNP-genotyping success rates had a range of 54.3%–98.1%. Although this failure rate could reduce confidence in the data used for analysis, Sequenom traces were manually inspected for markers with low success rates, and low-probability genotypes were removed when identified by software programs.

The association analysis performed in this study attempts to extract the maximum information from the data available and is, in some respects, exploratory. Family-based and case-control association analyses were performed for single-marker and multimarker haplotypes. Although this requires multiple testing of the data, it has been suggested that different approaches have varying power, depending on disease and marker-allele frequency. A family-based analysis benefits when parental genotypes are used as controls but may yield little informative data if marker heterozygosity is low and parental data is missing. Conversely, a case-control analysis may have more power at varying allele frequencies for the number of individuals genotyped but could provide an artifactual result because of popu-

lation-specific differences between cases and controls (Risch and Teng 1998). It has also been suggested that the added information from haplotype analysis of SNPs has more power than single-marker analysis (Zaykin et al. 2002). A number of procedures have been suggested for correcting for multiple testing, including Bonferroni correction, but there is as yet little consensus as to an "ideal statistical framework" either for reporting *P* values or for analysis of SNP data (Roeder et al. 2005). We have presented raw *P* values uncorrected for multiple testing, but the significance of reported *P* values should be viewed with caution, given the number of statistical tests performed.

Empirical significance levels were determined by permutation of affection status for statistical tests involving cases and controls. These levels suggest that the results for the SNPs within *GRIN2A* for *rs1014531* ($P = 2.9 \times 10^{-7}$) in the multiplex case-control analysis and for intron 10 ($P = 2.32 \times 10^{-12}$) in the singleton case-control analysis are significant results. Additionally, the logistic regression analysis suggests *GRIN2A* haplotypes are significant contributors to genetic risk for autism. Determining empirical *P* values for the family-based tests for association, such as the PDT, requires correlations between transmitted alleles within pedigrees to be preserved, and "[i]t is not clear whether a simple permutation procedure can be developed for this purpose in general pedigrees" (Martin et al. 2000, p. 153). Additionally, simple permutation of affection status is inappropriate for the trend test, because the correlation between alleles in related cases must be preserved (Slager and Schaid 2001). Interpreting the results from these statistical tests is challenging, but, on balance, the strongest evidence of association with autism comes principally from *GRIN2A,* with *ABAT* accounting for the remaining association.

Both SNPs genotyped in *ABAT* have genotypes associated with autism in the case-control analysis of the multiplex sample, with use of Fisher's exact test for count data (Fisher's exact test provides more power than the PDT, given the sample sizes). The result for the intron 1 SNP was also replicated in the singleton sample. The result is further supported by the logistic regression analysis, which indicates a difference in haplotype distribution between cases and controls. *ABAT* is a good positional functional candidate for autism, especially given its involvement in the catabolism of GABA, which is known to be involved in neurological disease (Wong et al. 2003). Other genes involved in this pathway have been studied elsewhere, including the GABA receptor subunit $\beta 3$ (*GABRB3*) on 15q11-13, following reports of chromosomal abnormalities in autistic individuals. A number of studies have been published on the role of this pathway in autism (reviewed by Veenstra-VanderWeele and Cook [2004]).

Screening the coding sequence of *ABAT* revealed one high-frequency missense mutation, but this does not occur in a highly conserved position and was less strongly associated with autism, compared with the SNP within intron 1. It seems likely that any variants influencing the expression of *ABAT* in individuals with autism are located outside of the coding sequence within regulatory regions.

No missense variants were detected by DHPLC in *GRIN2A*. However, SNPs within *GRIN2A* gave strong evidence of association in the case-control analysis of the multiplex and singleton samples. There was also some evidence of transmission disequilibrium with the SNP in intron 10 in PDT results. Additionally, the logistic regression analysis found a significant difference in haplotype distribution between cases and controls. *GRIN2A* is an interesting candidate gene for autism, given its role in learning and memory and the evidence from mouse models of a complex behavioral phenotype (Bannerman et al. 2004). There is also evidence, from postmortem analysis, of expression differences in the glutamate system between cerebella of subjects with autism and of controls (Purcell et al. 2001). Variants within *GRIN2A* may also confer increased risk for ADHD, which indicates involvement in behavioral disorders (Turic et al. 2004).

The candidate-gene analysis presented here is substantial but not completely exhaustive. Candidate genes were screened for coding variants, and association analysis was performed, but additional SNPs with inter-SNP distances of 5–10 kb may be required to conclusively discount these genes having a role in autism susceptibility (Kruglyak 1999). Increasing the density of SNPs across candidate genes, on the basis of known patterns of LD, to capture the maximum information about haplotypic diversity is now increasingly feasible, as data become available from the International HapMap Project. Prior knowledge of the LD structure of a region will allow a more-intelligent and -efficient approach to association mapping in studies of candidate genes and perhaps will lead to less-tentative statistical analysis and more-definitive conclusions (Zondervan and Cardon 2004). A combination of strategies that include association mapping and complementary statistical and bioinformatic analysis will be needed to determine the significance of the results for *GRIN2A* and *ABAT,* as well as replication in independent samples. The design of subsequent studies should also include careful consideration of the control samples used for allele- and genotype-frequency comparisons, in terms of their ethnicity, sex, and independence. Chromosome 16p has been identified by a number of genome screens for autism and is likely to contain an autism-susceptibility variant. Although the 16p locus has been less consistently replicated than the chromosome 7 locus, it remains an important

area for future research (Lamb et al. 2002). However, as with all multifactorial disorders, progressing from a broad linkage peak to identifiable functional genetic variants remains a significant challenge.

Acknowledgments

This work would not have been possible without the cooperation of the subjects, their families, and the many referring professionals. This work was funded by the U.K. Medical Research Council, The Wellcome Trust, BIOMED 2 grant CT-97-2759, European Community Fifth Framework grant QLG2-CT-1999-0094, Telethon-Italy grant GGP030227, Fondazione Cassa di Risparmio di Bologna, the Janus Korczak Foundation, Deutsche Forschungsgemeinschaft, Fondation France Télécom, Conseil Régional Midi-Pyrénées, Danish Medical Research Council, Sofiefonden, the Beatrice Surovell Haskells Fund for Child Mental Health Research of Copenhagen, Danish Natural Science Research Council grant 9802210, and National Institutes of Health (NIH) grants MO1 RR06022 GCRC NIH, NIH K05 MH01196, K02 MH01389, and U19 HD35482. A.J.B is the Cheryl and Reece Scott Professor of Psychiatry. A.P.M. is a Wellcome Trust Principal Research Fellow. Members of the IMGSAC follow and are grouped by institution. In the United Kingdom: Nicola Mathews, Jeremy Parr, Simon Wallace, and A.J.B. (Section of Child and Adolescent Psychiatry, University Department of Psychiatry, Park Hospital for Children, Oxford); G.B., Elena Bonora, J.L., Angela Marlow, N.S., and A.P.M. (Wellcome Trust Centre for Human Genetics, University of Oxford, Oxford); Patrick Bolton and Michael Rutter (Social, Genetic and Development Research Centre, Institute of Psychiatry, London); Gillian Baird (Newcomen Centre, Guy's Hospital, London); Stephen Abbs, Zoe Docherty, Caroline Ogilvie, and Pamela Warburton (South East Thames Regional Genetics Centre, Division of Medical and Molecular Genetics, Guy's Hospital, London); Tom Berney, Ann Le Couteur, and Helen McConachie (School of Clinical Medical Sciences, University of Newcastle, Newcastle); Petrus de Vries and Emma Weisblatt (Developmental Psychiatry Section, University of Cambridge Clinical School, Cambridge); Catherine Aldred, Julie-Anne Wilkinson, and Jonathan Green (Academic Department of Child Psychiatry, University of Manchester, Manchester); and Andrew Pickles (School of Epidemiology and Health Science, University of Manchester, Manchester); Philippa Bracegirdle, Ros Packer, and Bryan Bolton (ECACC Health Protection Agency, Porton Down). In Italy: Elena Bacchelli, Francesca Blasi, Simona Carone, and Elena Maestrini (Dipartimento di Biologia Evoluzionistica Sperimentale, University of Bologna, Bologna). In the Netherlands: Maretha De Jonge, Chantal Kemner, Margreet Scherpenisse, Karlijn Steggehuis, and Herman Van Engeland (Department of Child and Adolescent Psychiatry, Academisch Ziekenhuis Utrecht [AZU], Utrecht); and Ineke den Hartog (Department of Child Psychiatry, AZU, Utrecht). In Germany: Sabine M. Klauck, Bärbel Felder, Claudia Schuster, and Annemarie Poustka (Deutsches Krebsforschungszentrum, Molecular Genome Analysis, Heidelberg); and Sven Bölte, Sabine Feineis-Matthews, Nora Uhlig, Dorothea Rühl, Gabriele Schmötzer, and Fritz Poustka (Department of Child and Adolescent Psychiatry, J.W. Goethe-Universität, Frankfurt). In France: Eric Fombonne, Jeanne Fremolle-Kruck, Carine Mantoulan, and Kerstin Wittemeyer (Hôpital la Grave, Toulouse); and Bernadette Rogé (Centre d'Études et de Recherches en PsychoPathologie, Maison de la Recherche, Université de Toulouse le Mirail, Toulouse). In Denmark: Ester Ulsted Sorensen, Birgitte Viskum, Rodney Cotterill, Karen-Brondum-Nielsen, Gunna Eriksen, Demetrious Haracopos, Torben Isager, and Lennart Pedersen (Centre for Autisme, Bagsvaerd). In Greece: Katerina Papanikolaou and John Tsiantis (Agia Sophia Children's Hospital, Athens). In the United States: D.E.W. (Department of Human Genetics, University of Pittsburgh, Pittsburgh); Jeff Salt, Stephen Guter, Bennett Leventhal, and Edwin Cook (Department of Psychiatry, University of Chicago, Chicago); Stanley Nelson and Susan Smalley (Center for Neurobehavioral Genetics, University of California at Los Angeles [UCLA], Los Angeles); Jennifer Levitt (Child Psychiatry Division, UCLA Psychiatric Institute, Los Angeles); Christina Corsello and Catherine Lord (University of Michigan Autism and Communicative Disorders Center, Ann Arbor); and Joel Bregman, Ami Klin, Kathy Koenig, Rosalind Oti, and Fred Volkmar (Yale University, New Haven).

Electronic-Database Information

Accession numbers and URLs for data presented herein are as follows:

Berkeley *Drosophila* Genome Project, http:///www.fruitfly.org/seq_tools/splice.html (for SpliceSite analysis)
dbSNP Home Page, http://www.ncbi.nlm.nih.gov/SNP/index.html
ECACC, http://www.ecacc.org.uk/
ESEfinder, http://rulai.cshl.edu/tools/ESE/
GenBank, http://www.ncbi.nlm.nih.gov/Genbank/ (for *A2BP1* [accession number NM_018723], *ABAT* [accession number BC031413], *BFAR* [accession number NM_016561], *CREBBP* [accession number NM_004380], *EMP2* [accession number NM_001424], *GRIN2A* [accession number NM_000833], *MRTF-B* [accession number NM_014048], *SSTR5* [accession number NM_001053], *TBX6* [accession number NM_004608], and *UBN1* [accession number NM_016936])
IGS, http://bioinformatics.well.ox.ac.uk/project-lims.shtml
IMGSAC, http://www.ox.ac.uk/~maestrin/iat.html
International HapMap Project, http://www.hapmap.org/
LocusLink, http://www.ncbi.nih.gov/entrez/query.fcgi?db=gene
Online Mendelian Inheritance in Man (OMIM), http://www.ncbi.nlm.nih.gov/Omim/ (for *GRIN2A, SSTR5, TBX6, CREBBP,* RTS, *EMP2, ABAT, A2BP1,* and *SCA2*)
Pfam, http://www.sanger.ac.uk/Software/Pfam/
PolyPhen, http://tux.embl-heidelberg.de/ramensky
Primer3, http://www-genome.wi.mit.edu/cgi-bin/primer/primer3_www.cgi
R Project for Statistical Computing, http://www.r-project.org/
SIBMED program, http://www.sph.umich.edu/
UCSC Genome Browser, http://genome.ucsc.edu/
University of Pittsburgh Department of Human Genetics, http://watson.hgen.pitt.edu/ (for Mega2, version 2.5)

References

Abecasis GR, Cookson WO (2000) GOLD—graphical overview of linkage disequilibrium. Bioinformatics 16:182–183

Bailey A, Le Couteur A, Gottesman I, Bolton P, Simonoff E, Yuzda E, Rutter M (1995) Autism as a strongly genetic disorder: evidence from a British twin study. Psychol Med 25: 63–77

Bannerman DM, Deacon RM, Brady S, Bruce A, Sprengel R, Seeburg PH, Rawlins JN (2004) A comparison of GluR-A-deficient and wild-type mice on a test battery assessing sensorimotor, affective, and cognitive behaviors. Behav Neurosci 118:643–647

Bateman A, Coin L, Durbin R, Finn RD, Hollich V, Griffiths-Jones S, Khanna A, Marshall M, Moxon S, Sonnhammer EL, Studholme DJ, Yeats C, Eddy SR (2004) The Pfam protein families database. Nucleic Acids Res 32 D138–D141

Bonora E, Lamb JA, Barnby G, Sykes N, Moberly T, Beyer KS, Klauck SM, et al (2005) Mutation screening and association analysis of six candidate genes for autism on chromosome 7q. Eur J Hum Genet 13:198–207

Brunak S, Engelbrecht J, Knudsen S (1991) Prediction of human mRNA donor and acceptor sites from the DNA sequence. J Mol Biol 220:49–65

Cartegni L, Wang J, Zhu Z, Zhang MQ, Krainer AR (2003) ESEfinder: a Web resource to identify exonic splicing enhancers. Nucleic Acids Res 31:3568–3571

Chakrabarti S, Fombonne E (2001) Pervasive developmental disorders in preschool children. JAMA 285:3093–3099

Chrivia JC, Kwok RP, Lamb N, Hagiwara M, Montminy MR, Goodman RH (1993) Phosphorylated CREB binds specifically to the nuclear protein CBP. Nature 365:855–859

Clayton D, Jones H (1999) Transmission/disequilibrium tests for extended marker haplotypes. Am J Hum Genet 65:1161–1169

De Biase D, Barra D, Simmaco M, John RA, Bossa F (1995) Primary structure and tissue distribution of human 4-aminobutyrate aminotransferase. Eur J Biochem 227:476–480

Douglas JA, Boehnke M, Lange K (2000) A multipoint method for detecting genotyping errors and mutations in sibling-pair linkage data. Am J Hum Genet 66:1287–1297

Finelli P, Natacci F, Bonati MT, Gottardi G, Engelen JJ, de Die-Smulders CE, Sala M, Giardino D, Larizza L (2004) FISH characterisation of an identical (16)(p11.2p12.2) tandem duplication in two unrelated patients with autistic behaviour. J Med Genet 41:e90

Fombonne E (2002) Epidemiological trends in rates of autism. Mol Psychiatry (Suppl 2) 7:S4–S6

——— (2003) Epidemiological surveys of autism and other pervasive developmental disorders: an update. J Autism Dev Disord 33:365–382

Hellings JA, Hossain S, Martin JK, Baratang RR (2002) Psychopathology, GABA, and the Rubinstein-Taybi syndrome: a review and case study. Am J Med Genet 114:190–195

IMGSAC (2001) A genomewide screen for autism: strong evidence for linkage to chromosomes 2q, 7q, and 16p. Am J Hum Genet 69:570–581

Itoh N, Tsujimoto Y, Nagata S (1993) Effect of bcl-2 on Fas antigen-mediated cell death. J Immunol 151:621–627

Jamain S, Quach H, Betancur C, Rastam M, Colineaux C, Gillberg IC, Soderstrom H, Giros B, Leboyer M, Gillberg C, Bourgeron T, and the Paris Autism Research International Sibpair Study (2003) Mutations of the X-linked genes encoding neuroligins NLGN3 and NLGN4 are associated with autism. Nat Genet 34:27–29

Kruglyak L (1999) Prospects for whole-genome linkage disequilibrium mapping of common disease genes. Nat Genet 22:139–144

Kulp D, Haussler D, Reese MG, Eeckman FH (1997) Integrating database homology in a probabilistic gene structure model. Pac Symp Biocomput 232–244

Lamb JA, Barnby G, Bonora E, Sykes N, Bacchelli E, Blasi F, Maestrini E, Broxholme J, Tzenova J, Weeks D, Bailey AJ, Monaco AP (2005) Analysis of IMGSAC autism susceptibility loci: evidence for sex limited and parent of origin specific effects. J Med Genet 42:132–137

Lamb JA, Parr JR, Bailey AJ, Monaco AP (2002) Autism: in search of susceptibility genes. Neuromolecular Med 2:11–28

Lauritsen MB, Nyegaard M, Betancur C, Colineaux C, Josiassen TL, Kruse TA, Leboyer M, Ewald H (2003) Analysis of transmission of novel polymorphisms in the somatostatin receptor 5 (SSTR5) gene in patients with autism. Am J Med Genet B Neuropsychiatr Genet 121:100–104

Le Couteur A, Rutter M, Lord C, Rios P, Robertson S, Holdgrafer M, McLennan J (1989) Autism diagnostic interview: a standardized investigator-based instrument. J Autism Dev Disord 19:363–387

Lewontin R, Kojima K (1960) The evolutionary dynamics of complex polymorphisms. Evolution 14:458–472

Liu J, Nyholt DR, Magnussen P, Parano E, Pavone P, Geschwind D, Lord C, Iversen P, Hoh J, Autism Genetic Resource Exchange Consortium, Ott J, Gilliam TC (2001) A genomewide screen for autism susceptibility loci. Am J Hum Genet 69:327–340

Lord C, Risi S, Lambrecht L, Cook EH Jr, Leventhal BL, DiLavore PC, Pickles A, Rutter M (2000) The autism diagnostic observation schedule-generic: a standard measure of social and communication deficits associated with the spectrum of autism. J Autism Dev Disord 30:205–223

Martin ER, Bass MP, Gilbert JR, Pericak-Vance MA, Hauser ER (2003*a*) Genotype-based association test for general pedigrees: the genotype-PDT. Genet Epidemiol 25:203–213

Martin ER, Bass MP, Hauser ER, Kaplan NL (2003*b*) Accounting for linkage in family-based tests of association with missing parental genotypes. Am J Hum Genet 73:1016–1026

Martin ER, Monks SA, Warren LL, Kaplan NL (2000) A test for linkage and association in general pedigrees: the pedigree disequilibrium test. Am J Hum Genet 67:146–154

Mukhopadhyay N, Almasy L, Schroeder M, Mulvihill WP, Weeks DE (1999) Mega2, a data-handling program for facilitating genetic linkage and association analysis. Am J Hum Genet Suppl 65:A436

Nagase T, Ishikawa K, Kikuno R, Hirosawa M, Nomura N, Ohara O (1999) Prediction of the coding sequences of unidentified human genes. XV. The complete sequences of 100 new cDNA clones from brain which code for large proteins in vitro. DNA Res 6:337–345

O'Connell JR, Weeks DE (1998) PedCheck: a program for identification of genotype incompatibilities in linkage analysis. Am J Hum Genet 63:259–266

Ogdie MN, Fisher SE, Yang M, Ishii J, Francks C, Loo SK, Cantor RM, McCracken JT, McGough JJ, Smalley SL, Nelson SF (2004) Attention deficit hyperactivity disorder: fine mapping supports linkage to 5p13, 6q12, 16p13, and 17p11. Am J Hum Genet 75:661–668

Ogdie MN, Macphie IL, Minassian SL, Yang M, Fisher SE, Francks C, Cantor RM, McCracken JT, McGough JJ, Nelson SF, Monaco AP, Smalley SL (2003) A genomewide scan for attention-deficit/hyperactivity disorder in an extended sample: suggestive linkage on 17p11. Am J Hum Genet 72: 1268–1279

Pagani F, Stuani C, Tzetis M, Kanavakis E, Efthymiadou A, Doudounakis S, Casals T, Baralle FE (2003) New type of disease causing mutations: the example of the composite exonic regulatory elements of splicing in CFTR exon 12. Hum Mol Genet 12:1111–1120

Papapetrou C, Putt W, Fox M, Edwards YH (1999) The human TBX6 gene: cloning and assignment to chromosome 16p11.2. Genomics 55:238–241

Philippe A, Martinez M, Guilloud-Bataille M, Gillberg C, Rastam M, Sponheim E, Coleman M, Zappella M, Aschauer H, Van Maldergem L, Penet C, Feingold J, Brice A, Leboyer M, and the Paris Autism Research International Sibpair Study (1999) Genome-wide scan for autism susceptibility genes. Hum Mol Genet 8:805–812

Pickles A, Bolton P, Macdonald H, Bailey A, Le Couteur A, Sim CH, Rutter M (1995) Latent-class analysis of recurrence risks for complex phenotypes with selection and measurement error: a twin and family history study of autism. Am J Hum Genet 57:717–726

Pickles A, Starr E, Kazak S, Bolton P, Papanikolaou K, Bailey A, Goodman R, Rutter M (2000) Variable expression of the autism broader phenotype: findings from extended pedigrees. J Child Psychol Psychiatry 41:491–502

Pritchard JK, Cox NJ (2002) The allelic architecture of human disease genes: common disease-common variant…or not? Hum Mol Genet 11:2417–2423

Purcell AE, Jeon OH, Zimmerman AW, Blue ME, Pevsner J (2001) Postmortem brain abnormalities of the glutamate neurotransmitter system in autism. Neurology 57:1618–1628

Risch N, Teng J (1998) The relative power of family-based and case-control designs for linkage disequilibrium studies of complex human diseases I. DNA pooling. Genome Res 8:1273–1288

Roeder K, Bacanu SA, Sonpar V, Zhang X, Devlin B (2005) Analysis of single-locus tests to detect gene/disease associations. Genet Epidemiol 28:207–219

Sakimura K, Kutsuwada T, Ito I, Manabe T, Takayama C, Kushiya E, Yagi T, Aizawa S, Inoue Y, Sugiyama H, Mishina M (1995) Reduced hippocampal LTP and spatial learning in mice lacking NMDA receptor epsilon 1 subunit. Nature 373:151–155

Selvaraj A, Prywes R (2003) Megakaryoblastic leukemia-1/2, a transcriptional co-activator of serum response factor, is required for skeletal myogenic differentiation. J Biol Chem 278:41977–41987

Shao Y, Wolpert CM, Raiford KL, Menold MM, Donnelly SL, Ravan SA, Bass MP, McClain C, von Wendt L, Vance JM, Abramson RH, Wright HH, Ashley-Koch A, Gilbert JR, DeLong RG, Cuccaro ML, Pericak-Vance MA (2002) Genomic screen and follow-up analysis for autistic disorder. Am J Med Genet 114:99–105

Shibata H, Huynh DP, Pulst SM (2000) A novel protein with RNA-binding motifs interacts with ataxin-2. Hum Mol Genet 9:1303–1313

Slager SL, Schaid DJ (2001) Evaluation of candidate genes in case-control studies: a statistical method to account for related subjects. Am J Hum Genet 68:1457–1462

Smalley SL, Kustanovich V, Minassian SL, Stone JL, Ogdie MN, McGough JJ, McCracken JT, MacPhie IL, Francks C, Fisher SE, Cantor RM, Monaco AP, Nelson SF (2002) Genetic linkage of attention-deficit/hyperactivity disorder on chromosome 16p13, in a region implicated in autism. Am J Hum Genet 71:959–963

Sparrow S, Balla D, Cicchetti D (1984) Vineland Adaptive Behavior Scales. Circle Pines, MN

Stone JL, Tsai S, Merriman B, Cantor-Chui RM, Geschwind D, Nelson SF, Consortium A (2003) High density SNP based association study of 21 candidate genes in autism spectrum disorder. Paper presented at The American Society of Human Genetics, Los Angeles, November 4–8

Sunyaev S, Ramensky V, Koch I, Lathe W 3rd, Kondrashov AS, Bork P (2001) Prediction of deleterious human alleles. Hum Mol Genet 10:591–597

Suzuki A, Yamada R, Chang X, Tokuhiro S, Sawada T, Suzuki M, Nagasaki M, Nakayama-Hamada M, Kawaida R, Ono M, Ohtsuki M, Furukawa H, Yoshino S, Yukioka M, Tohma S, Matsubara T, Wakitani S, Teshima R, Nishioka Y, Sekine A, Iida A, Takahashi A, Tsunoda T, Nakamura Y, Yamamoto K (2003) Functional haplotypes of PADI4, encoding citrullinating enzyme peptidylarginine deiminase 4, are associated with rheumatoid arthritis. Nat Genet 34:395–402

Tokuhiro S, Yamada R, Chang X, Suzuki A, Kochi Y, Sawada T, Suzuki M, Nagasaki M, Ohtsuki M, Ono M, Furukawa H, Nagashima M, Yoshino S, Mabuchi A, Sekine A, Saito S, Takahashi A, Tsunoda T, Nakamura Y, Yamamoto K (2003) An intronic SNP in a RUNX1 binding site of SLC22A4, encoding an organic cation transporter, is associated with rheumatoid arthritis. Nat Genet 35:341–348

Turic D, Langley K, Mills S, Stephens M, Lawson D, Govan C, Williams N, Van Den Bree M, Craddock N, Kent L, Owen M, O'Donovan M, Thapar A (2004) Follow-up of genetic linkage findings on chromosome 16p13: evidence of association of N-methyl-D aspartate glutamate receptor 2A gene polymorphism with ADHD. Mol Psychiatry 9:169–173

Veenstra-VanderWeele J, Cook EH (2004) Molecular genetics of autism spectrum disorder. Mol Psychiatry 9:819–832

Vincent JB, Kolozsvari D, Roberts WS, Bolton PF, Gurling HM, Scherer SW (2004) Mutation screening of X-chromosomal neuroligin genes: no mutations in 196 autism probands. Am J Med Genet B Neuropsychiatr Genet 129B:82–84

Wallenstein S, Hodge SE, Weston A (1998) Logistic regression model for analyzing extended haplotype data. Genet Epidemiol 15:173–181

Wei MC, Zong WX, Cheng EH, Lindsten T, Panoutsakopoulou V, Ross AJ, Roth KA, MacGregor GR, Thompson CB, Korsmeyer SJ (2001) Proapoptotic BAX and BAK: a requisite gateway to mitochondrial dysfunction and death. Science 292:727–730

Wong CG, Bottiglieri T, Snead OC 3rd (2003) GABA, gamma-hydroxybutyric acid, and neurological disease. Ann Neurol (Suppl 6) 54:S3–S12

Zaykin DV, Westfall PH, Young SS, Karnoub MA, Wagner MJ, Ehm MG (2002) Testing association of statistically inferred haplotypes with discrete and continuous traits in samples of unrelated individuals. Hum Hered 53:79–91

Zondervan KT, Cardon LR (2004) The complex interplay among factors that influence allelic association. Nat Rev Genet 5:89–100