American Journal of Medical Genetics Part B (Neuropsychiatric Genetics) 127B:104–112 (2004)

Linkage and Association Analysis at the Serotonin Transporter (*SLC6A4*) Locus in a Rigid-Compulsive Subset of Autism

J.L. McCauley,[1] L.M. Olson,[1] M. Dowd,[4] T. Amin,[1] A. Steele,[2,3] R.D. Blakely,[2,3] S.E. Folstein,[4] J.L. Haines,[1,2] and J.S. Sutcliffe[1,2]*

[1]*Program in Human Genetics, Department of Molecular Physiology and Biophysics, Vanderbilt University, Nashville, Tennessee*
[2]*Center for Molecular Neuroscience, Vanderbilt University, Nashville, Tennessee*
[3]*Department of Pharmacology, Vanderbilt University, Nashville, Tennessee*
[4]*Department of Psychiatry, Tufts University and New England Medical Center, Boston, Massachusetts*

Autism is a complex genetic neurodevelopmental disorder in which affected individuals display deficits in language, social relationships, and patterns of compulsive and stereotyped behaviors and rigidity. Linkage analysis in our dataset of 57 New England and 80 AGRE multiplex autism families reveals a multipoint heterogeneity LOD (HLOD) score of 2.74 at D17S1871 in 17q11.2. Analysis of phenotypic subsets shows an increased HLOD of 3.62 in families with compulsive behaviors and rigidity. The serotonin transporter locus (*SLC6A4*) maps nearby and is considered a functional candidate gene in autism and obsessive-compulsive disorder. We genotyped an insertion/deletion polymorphism in the promoter (5-HTTLPR), and seven single nucleotide polymorphisms (SNPs) across the 38-kb transcriptional unit. Transmission disequilibrium (TD) analysis reveals nominal association at a SNP in intron 5 ($P = 0.02$) as well as 5-HTTLPR ($P = 0.01$), corresponding to over-transmission of the short allele. TD analysis in the rigid-compulsive subset shows no evidence for association. Intermarker linkage disequilibrium was determined. All SNPs define a single haplotype block, while 5-HTTLPR lies 5′ to this block. Three SNPs are sufficient to detect all common alleles ($\geq 5\%$) in this >26-kb block. Analysis of haplotypes for these markers demonstrates no evidence for association to autism. These data indicate that a common allele within the coding region of *SLC6A4* is not responsible for the observed linkage. However, the presence of heterogeneous disease variants within the block or the existence of a common disease-associated allele either upstream or downstream of this block is possible. In fact, such variants may well account for linkage to 17q11.2 in our families. © 2003 Wiley-Liss, Inc.

KEY WORDS: autism; linkage disequilibrium; haplotype; phenotypic subsetting

INTRODUCTION

Autism [MIM 209850] is a neuropsychiatric disorder characterized by deficits in reciprocal social interaction, language, and patterns of repetitive or stereotyped behaviors and interests [Rapin, 1997]. Autism typically presents with developmental abnormalities before 3 years of age and results in life-long difficulties for affected individuals. A wealth of evidence indicates that autism is a complex genetic disorder resulting from oligogenic inheritance of an unknown number of susceptibility alleles [Folstein and Rosen-Sheidley, 2001]. Up to fifteen genes may contribute overall to risk for autism, with locus heterogeneity resulting in different families possessing a different constellation of susceptibility alleles [Pickles et al., 1995; Risch et al., 1999].

Serotonin (5-hydroxytryptamine or 5-HT) has long been considered an attractive candidate system for involvement in autism etiology, given well-replicated observations of elevated platelet serotonin in a subset of probands and first-degree relatives [Schain and Freedman, 1961; Anderson et al., 1987; Abramson

Grant sponsor: National Institute of Mental Health; Grant numbers: MH55135, MH61009.

*Correspondence to: J.S. Sutcliffe, Ph.D., Department of Molecular Physiology and Biophysics, Vanderbilt University, Nashville, TN 37232-0615.
E-mail: james.s.sutcliffe@vanderbilt.edu

Received 8 May 2003; Accepted 29 August 2003
DOI 10.1002/ajmg.b.20151

et al., 1989]. The serotonin transporter (*SLC6A4*) in particular has received significant attention because selective serotonin reuptake inhibitors (SSRIs), which preferentially target the transporter, are effective in treating anxiety, rituals, and aggression in autism and related disorders, such as obsessive-compulsive disorder (OCD) [Gordon et al., 1993; McDougle et al., 1996; Vaswani et al., 2003]. Multiple association studies have been performed at this locus using two common variants: 5-HTTLPR, an insertion/deletion polymorphism in the promoter, and a variable number tandem repeat (VNTR) polymorphism located in intron 2. Both markers have been proposed to exhibit allelic differences in transcription of *SLC6A4* message [Heils et al., 1996; Lesch et al., 1996; Fiskerstrand et al., 1999; MacKenzie and Quinn, 1999]. Additionally, these variants have been examined in a number of other neuropsychiatric phenotypes and SSRI treatment response (reviewed in [Veenstra-VanderWeele et al., 2000]). Rare coding variants have been described for *SLC6A4*, although it is not clear whether any of these have phenotypic consequences [Glatt et al., 2001; Hahn and Blakely, 2002].

Transmission disequilibrium test (TDT) studies in autism have failed to yield any consistent findings of allelic association at *SLC6A4*. Cook and colleagues reported preferential transmission of the short allele at 5-HTTLPR in autism [Cook et al., 1997], while others have reported similar results for the long allele [Klauck et al., 1997; Tordjman et al., 2001; Yirmiya et al., 2001]. Other groups have failed to detect association [Maestrini et al., 1999; Persico et al., 2000]. One detailed study reported by Kim et al. describes high-density SNP genotyping across the *SLC6A4* locus and subsequent TDT analysis [Kim et al., 2002]. They replicate previous findings of nominally significant association of a haplotype containing 5-HTTLPR and the VNTR, as well as the VNTR marker alone. Additionally, they identify significant linkage disequilibrium (LD) at other SNPs (near the 5′-end of the gene) in their dataset of autism families.

Growing evidence indicates that defining phenotypic subsets may improve power to detect genetic effects in complex disorders [Buxbaum et al., 2001; CLSA, 2001; Alarcon et al., 2002; Shao et al., 2002, 2003; Nurmi et al., 2003]. In order to leverage phenotypic heterogeneity in autism for genetic studies, we have sought to define subsets using traits that (a) vary from one autistic person to the next; (b) are present, sometimes in milder form, in non-autistic family members significantly more often than controls; and (c) aggregate in particular autism families. Towards this end, we recently completed a principal components analysis of items common to the Autism Diagnostic Interview (ADI) and its revision (ADI-R) and identified six phenotypic subsets, for which significant correlation of items within a given factor was observed [Tadevosyan-Leyfer et al., 2003]. These subsets are (1) spoken language, (2) social intent, (3) compulsions and rigidity, (4) developmental milestones, (5) savant skills, and (6) sensory aversions. For five of the six factor subsets, there was significant intersibling correlation in multiplex families, suggesting that they are genetically relevant. Social intent was not significantly correlated, but became marginally so when age was covaried.

To evaluate the utility of these phenotypic subsets for genetic studies, we analyzed linkage data from a recently completed genomic screen of multiplex families. Using a categorical autism diagnosis, this analysis showed suggestive linkage to chromosome 17. Examination in the ADI subsets, however, revealed increased linkage in the rigid-compulsive families. This finding underscores the potential importance of utilizing trait-based phenotypic subsets for genetic studies of complex disease. Given that (1) the *SLC6A4* locus maps under the linkage peak and (2) the fact that the rigid-compulsive dimension of autism and other psychiatric disorders is ameliorated by SSRIs, we examined this compelling candidate for evidence of allelic association. Our approach to this study included the use of markers previously shown to associate with autism [Kim et al., 2002], and detailed characterization of LD and haplotype structures across the locus. We identify nominal but not highly significant association to autism at *SLC6A4*, however, no association was seen in the rigid-compulsive subset. We conclude that a common variant in the *SLC6A4* coding region is not responsible for the observed linkage. However, the existence of multiple rare coding variants in *SLC6A4* is consistent with potential heterogeneous risk alleles in autism and OCD, and such a scenario would explain our linkage results.

MATERIALS AND METHODS

Families

The sample for the linkage study consisted of 137 multiplex families. A total 57 multiplex families were recruited from the Tufts/New England Medical Center (NEMC) and 80 affected sib-pair families were purchased from the AGRE consortium (www.agre.org). Association studies were conducted using a dataset of 123 multiplex families, all of which were included in the linkage study. All probands were at least 4 years of age and were clinically assessed with the ADI or ADI-R and most with the Autism Diagnostic Observation Schedule (ADOS). In multiplex families at least one sib had to meet ADI algorithm criteria for an autism diagnosis, while additional siblings may be on the broader autism spectrum. Probands were excluded from the study if they had a known medical or neurological condition suspected to be associated with their autistic phenotype (e.g., fragile X syndrome). The procedures for clinically evaluating affected individuals for the AGRE families have been previously described [Liu et al., 2001]. The study was undertaken after Institutional Review Board review and approval.

Molecular Analysis

DNA was isolated from peripheral blood or lymphoblastoid cells using the PureGene kit according to the manufacturer's recommendations (Gentra Systems, Minneapolis, MN). 5-HTTLPR genotypes were determined using size discrimination of PCR products on 3%

NuSeive (3:1) agarose (FMC Bioproducts; Rockland, ME) gels. The short allele corresponds to a product of 484 bp, while the longer allele is 528 bp; amplifying PCR primers have been described previously [Cook et al., 1997]. SNPs from the dbSNP and Celera SNP databases were selected based on their map position, minor allele frequency, and previous findings of allelic association. Database reference numbers and other details for markers are cited in Table I. PCR assays were developed and optimized to amplify an ~200-bp region flanking SNPs. Individual SNPs were genotyped by either fluorescent polarization template-directed dye terminator incorporation assay (FP-TDI) or TaqMan™. PCR primers and probes for assays are listed in Table II. This information is unavailable for marker seven, which was obtained from Applied Biosystems (Foster City, CA) as an Assay-On-Demand™.

For FP-TDI genotyping, PCR reaction volumes were 8 μl, employing 10 ng genomic DNA template, 0.2 μM primers, 125 μM dNTPs, and Applied Biosystems AmpliTaq Gold DNA polymerase and buffer (Applied Biosystems). Cycling conditions included an initial denaturation at 95°C for 10 min, followed by 50 cycles of 94°C for 15 sec, optimal annealing temperature (T_A °C) for 30 sec, and 72°C for 15 sec, and a final extension at 72°C for 10 min. FP-TDI analysis was performed using materials supplied in commercially-available Acycloprime™ kits according to the manufacturer's published protocols (Perkin-Elmer Lifesciences, Boston, MA) and as described elsewhere [Hsu et al., 2001]. Samples were analyzed using a VICTOR$_2$™ multilabel plate reader instrument (Perkin-Elmer Life Sciences).

For TaqMan genotyping assays [Holloway et al., 1999], reactions were performed in a 5 μl volume according to manufacturer's recommendations (Applied Biosystems). Cycling conditions included an initial denaturation at 95°C for 10 min, followed by 50 cycles of 92°C for 15 sec, 60°C for 1 min. Samples were analyzed using an ABI 7900HT Sequence Detection System.

Statistical Analyses

Genotype data from chromosome 17 microsatellite markers spaced at ~10 cM intervals were included in linkage analyses. The deCODE genetic map [Kong et al., 2002] was used to provide genetic distances for these markers. For a few markers not present on the deCODE map, genetic map location was determined by relating the deCODE and Marshfield (http://research.marshfieldclinic.org/genetics/) maps. Microsatellite genotype data were analyzed following checks for quality control and Mendelian inconsistencies. HLOD values were calculated for the sample as a whole and for each subset under recessive and dominant models using GENEHUNTER-PLUS [Kong and Cox, 1997]. Disease allele frequencies were estimated to be $q = 0.01$ or 0.1 for dominant and recessive models, respectively, and a penetrance value of 0.5 was used for the analysis.

For each of the six factors, derived from items common to the ADI and ADI-R, families were subdivided into two groups. For factors (including rigid-compulsive behaviors) whose scores were unimodally distributed across the overall dataset, the mean factor score was calculated for the two (sometimes three) probands, and families were split into two groups at the median of the mean probands' scores for that factor. The two groups for each factor correspond to families who had scores higher than the mean on the phenotypic domain in question ("positive" families) and families with scores below the mean ("negative" families). Linkage was analyzed in both positive and negative subsets for each factor as described above. The rigid-compulsive cluster comprised both "current" and "ever" codings for the following ADI-R items: stereotyped utterances, unusual preoccupations, compulsions/rituals, resistance to trivial changes in the environment, and unusual attachment to objects. The rigid-positive group ($n = 70$) had a mean score on this factor of 0.39 (±0.11), while the rigid-negative subset ($n = 67$) had a mean score of 0.17 (±0.06). A simulation study to calculate an empiric *P*-value for the rigid-compulsive linkage result was performed using SIMULATE (http://linkage.rockefeller.edu/ott/simulate.htm) to randomly draw, in each of 10,000 replicates, 70 (corresponding to the number of rigid-positive) families from the overall dataset. Linkage for all markers was calculated separately for this group and the remainder of families under both dominant and recessive models using GENEHUNTER-PLUS.

Genotype data from 5-HTTLPR and the following SNPs were used in LD analysis: rs2066713, rs2020936, rs2020937, rs2020942, rs140700, hCV7911143, rs1042173 (dbSNP and Celera). Initial analysis of SNP genotype data involved quality control checks consisting of verification of internal controls and assessment of

TABLE I. *SLC6A4* Markers

Marker no.	Marker type	*SLC6A4* region	dbSNP rs#/ Celera hCV#	Alleles[a]	Minor allele frequency	Intermarker distance (kb)
1	Insertion/deletion	Promoter 1A	5-HTTLPR	528(l)/484(s)	0.45	~12.5
2	SNP	Intron 1A	rs2066713	C/T	0.41	0.851
3	SNP	Intron 1A	rs2020936	T/C	0.16	0.057
4	SNP	Intron 1A	rs2020937	T/A	0.41	3.843
5	SNP	Intron 2	rs2020942	G/A	0.40	3.525
6	SNP	Intron 5	rs140700	G/A	0.07	4.857
7	SNP	Intron 8	hCV7911143	T/C	0.43	13.521
8	SNP	3′-UTR	rs1042173	T/G	0.41	

[a]Major/minor allele.

TABLE II. PCR and Genotyping Primers

Marker no.		Primer sequences	Product size (bp)	T_A (°C)	Genotype assay
1	F[a]	5′-GGCGTTGCCGCTCTGAATGC-3′	484/528	63	Size discrimination
	R[b]	5′-GAGGGACTGAGCTGGACAACCAC-3′			
2	F	5′-ACTGCTCACTGCTGCTGCTAAATG-3′	127	58	FP
	R	5′-GCATCACCCAAGCGTTCCC-3′			
	FP-F[c]	5′-TTGCTTCTGAGATGGACCGCATTTCCCTTC-3′			
3	F	5′-GCCAGGCAGTAGCATAAATGGT-3′	84	51	TaqMan (AbD[e])
	R	5′-CAAACACCACTCAGAAGGATATGAA-3′			
	VIC[d]	VIC-AGAGCG**G**TCTCCATAA-MGB-NFQ			
	FAM[d]	FAM-AAGAGCG**A**TCTCCA-MGB-NFQ			
4	F	5′-CATATCCTTCTGAGTGGTGTTTGC-3′	181	56	FP
	R	5′-AATTTTAAAGGGATCGATTGTTGC-3′			
	FP-F	5′-TGTTTGCATTCTTGAGCCTGGGG-3′			
5	F	5′-AGGAAGGCCATCACGAGAACAC-3′	187	58	FP
	R	5′-CCTGCAGCCTGAGTTTTTAGCCTA-3′			
	FP-F	5′-AACACATGGTTTTATTCTCGAGCC-3′			
6	F	5′-TGCATAGTGGGCTCAGAGGTAGT-3′	151	55	FP
	R	5′-GGAGGTGGGTGAATGGATGTC-3′			
	FP-F	5′-TGATCTTTCTGCCACACCACCTC-3′			
7	—	Unavailable	Unknown	—	TaqMan (AoD[f])
8	F	5′-GTAGGAGAGAACAGGGATGCTATC-3′	153	52	FP
	R	5′-CACACTATTTTTCATTTTAGCTTCTTACA-3′			
	FP-R[c]	5′-AGGTTCTAGTAGATTCCAGCAATAAAATT-3′			

[a]Forward PCR primer.
[b]Reverse PCR primer.
[c]Forward (F) or reverse (R) FP-TDI extension primer.
[d]VIC and FAM are flourescent labels for TaqMan allelic discrimination probes.
[e]AbD is an Assay-By-Design from ABI.
[f]AoD is an Assay-On-Demand from ABI.

Mendelian inconsistencies, followed by final haplotype consistency analyses using Simwalk2 [Sobel and Lange, 1996]. Conformity with anticipated Hardy–Weinberg equilibrium expectations was established, and SNPs were examined for intermarker LD using both the delta and D′ measures calculated by the GOLD (Graphical Overview of Linkage Disequilibrium [Abecasis and Cookson, 2000]) software package. TD in autism families was determined using the PDT statistic, a variant of the transmission disequilibrium test (TDT), developed for use with general pedigrees [Martin et al., 2000]. Common haplotypes were determined using TRANSMIT [Clayton, 1999], and analysis of TD was performed using three SNPs (5-HTT-5, 5-HTT-6, and 5-HTT-7) that define all common (≥5%) haplotypes. Results were considered significant at the nominal level for markers or haplotypes with $P < 0.05$. Calculation of the power to detect an association in any dataset is dependent on the underlying model, which remains unknown. However, to develop a general idea of power, we calculated the power of the PDT to detect an association assuming an underlying complex disease model with an autism risk allele frequency of 0.04 and sibling relative risk (λ_S) of 1.25. We further assumed that the autism risk allele is in a high level of disequilibrium with the tested polymorphism, a reasonable assumption with this focused study. The power to detect an effect with this sample size is approximately 75%.

Visualization tools for alignment (VISTA) analysis were performed via web-based submission (http://www.gsd.lbl.gov/vista/) of human and mouse genomic sequence for *SLC6A4* [Mayor et al., 2000].

RESULTS

To facilitate the identification of genes underlying the etiology of autism, we recently completed a 10-cM genomic screen of multiplex families (Sutcliffe et al., unpublished observation). We detected suggestive evidence for linkage on chromosome 17 (Fig. 1), with a maximum multipoint HLOD of 2.74 at D17S1871 (~50 cM) under a dominant model in our dataset of families (n = 137). In an exploratory analysis, we used the six ADI factors as a basis for subsetting our families for linkage analysis. We detected a substantial increase in linkage in the subset of families in which probands are relatively affected (positive) by rigid-compulsive behaviors (Fig. 1). In the rigid-positive families (n = 70), the dominant HLOD increased to 3.62 at D17S1294 (~53 cM), adjacent to D17S1871 in the genomic screen panel. The HLOD for the remaining 67 families was 0.02 at this location. A Monte Carlo simulation was performed to calculate the empirical significance of this result. In each of 10,000 replicates, 70 families from the overall autism dataset were selected at random for a rigid-positive control group and the remainder as a rigid-negative control. Dominant and recessive HLODs were calculated for all chromosome 17 markers for each group. This permutation test revealed that the true *P*-value was 0.02, indicating it to be very unlikely that we obtained the HLOD of 3.62 by chance.

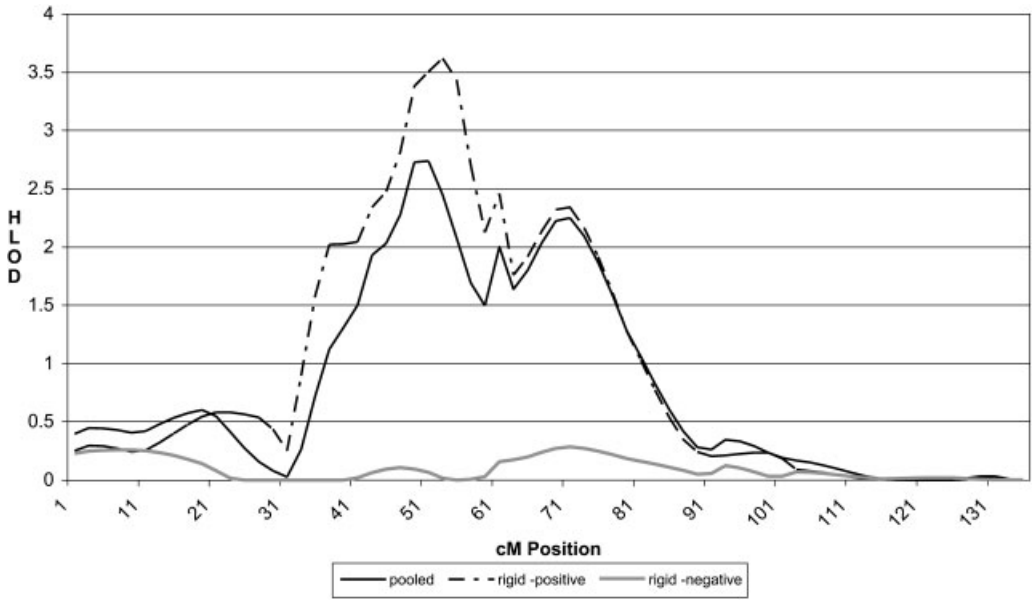

Fig. 1. Linkage analysis of autism for chromosome 17. Dominant multipoint HLOD scores are plotted for the overall (pooled) autism dataset (n = 137), the rigid-positive (relatively affected; n = 70), and the rigid-negative (relatively unaffected; n = 67). The overall dataset has a peak multipoint HLOD of 2.74 at D17S1871, and this increases to 3.62 in the rigid-positive subset.

We noted that the serotonin transporter (SLC6A4) locus maps adjacent to D17S1294. Given the long-standing hypothesis that 5-HTT is involved in serotonin-related autism etiology, we directly tested whether common alleles at this gene were associated with autism in our dataset. 5-HTTLPR and seven SNPs were selected as markers for this study providing an average marker density of ~5 kb across the transcriptional unit (see Table I). Criteria for marker selection included previous evidence for association [Kim et al., 2002], minor allele frequencies, and uniform spacing across the region. Markers were genotyped by PCR and size discrimination (5-HTTLPR), FP-TDI, or TaqMan in a dataset of 123 multiplex families that represent a subset of the 137 families analyzed for linkage.

Genotypes at individual markers did not deviate from expectations of Hardy–Weinberg equilibrium (data not shown). In order to discover all common alleles at this locus, we characterized intermarker LD to permit identification of haplotype blocks and corresponding haplotypes. Intermarker LD was assessed using GOLD, and markers 2-8 were found to be in strong LD with one another (Fig. 2). By contrast, 5-HTTLPR was in relatively weak LD ($D' < 0.5$) with these markers. These data suggest that at least two distinct haplotype blocks span the promoter and transcriptional unit of the *SLC6A4* locus.

Individual marker association analysis for both the overall (n = 123) and rigid-compulsive subset (n = 64) was tested using PDT. Nominal association ($P < 0.05$) was seen at 5-HTTLPR ($P = 0.01$) and 5-HTT-6 ($P = 0.02$) in the overall dataset (n = 123). We observed an over-transmission to affected individuals of the short allele at 5-HTTLPR (Table III). None of these markers demonstrate evidence of association in the rigid-compulsive subset.

Haplotype frequencies and transmissions were assessed using TRANSMIT for the haplotype block defined by markers 2-8. We identified four haplotypes having estimated allele frequencies > 5% that represent ~98% of the haplotypes in our sample. The common haplotypes may be differentiated by genotyping only three of the seven SNPs in the block. Using SNPs 5-HTT-5, 5-HTT-6, and 5-HTT-7 to represent these haplotypes, association analysis was performed on the overall dataset using TRANSMIT (Table IV). No significant evidence for association was detected with this combination of markers. Given the absence of association in the full autism dataset, we did not test the rigid-compulsive subset.

In order to relate potential functional sequences at *SLC6A4* to available genetic markers, we compared

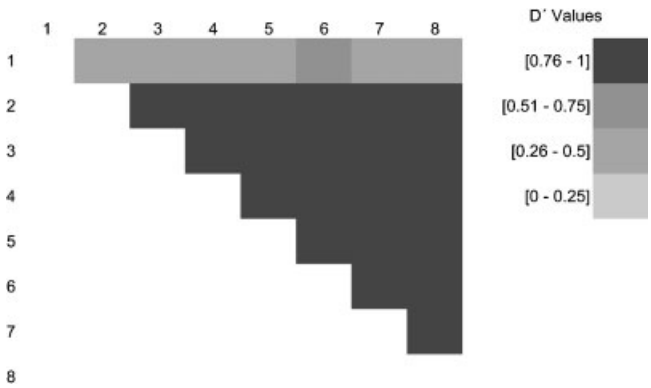

Fig. 2. Intermarker linkage disequilibrium (LD) at *SLC6A4*. LD was measured for all markers using GOLD. D′ values are represented by shaded boxes to display intermarker LD relationships. The major haplotype block includes markers 2-8. 5-HTTLPR, for which nominal association is seen, is only in weak LD ($D' < 0.5$) with other markers and is not located in the main haplotype block.

TABLE III. PDT Analysis of the *SLC6A4* Locus

Marker no.	Overall PDT (*P*)	Allele	T[a]	NT[b]	Rigid subset PDT (*P*)
1	0.01	484(s)	221	180	0.07
		528(l)	237	278	
2	0.46	T	175	188	0.37
		C	291	278	
3	0.29	C	77	89	0.82
		T	387	375	
4	0.50	A	177	189	0.50
		T	287	275	
5	0.63	A	175	184	0.88
		G	291	282	
6	0.02	A	29	49	0.10
		G	435	415	
7	0.22	C	205	184	0.38
		T	253	274	
8	0.29	G	202	185	0.53
		T	260	277	

[a]Number of alleles transmitted to affected individuals.
[b]Number of alleles not transmitted to affected individuals.

human and mouse genomic sequences using VISTA [Mayor et al., 2000]. DNA sequence 20-kb upstream and downstream of the transcriptional unit was included, so that likely regulatory regions would be represented. This analysis revealed evolutionarily conserved coding and non-coding regions. Figure 3 highlights the positions of SNP markers and the promoter variant relative to exons and conserved sequence. In addition to coding sequences, several regions of non-coding conservation are detected both inside and outside of the main haplotype block.

DISCUSSION

It is well established that autism is one of the most genetic of neuropsychiatric disorders, and that multiple genes are involved in its etiology. Two complementary approaches for identifying genes underlying complex genetic disorders are (1) identification of genomic regions more frequently inherited in common by affected individuals in a family through genomic linkage screens and (2) analysis of genes in candidate systems suspected to be involved in the disease based on altered biology. Linkage-oriented efforts to identify susceptibility genes tend to become studies of positional candidates, once a region of linkage is identified. These two approaches are unified in this study, where an excellent functional candidate gene is found in a very strong region of linkage in our dataset.

In this report, we identify *SLC6A4* as a positional and functional candidate in a region of linkage on chromosome 17. Given the genetic heterogeneity in autism and other complex disorders, the need for approaches to identify genetically more homogeneous subsamples is becoming increasingly apparent. An alternative to the use of a categorical autism diagnosis to define affection status is the application of individual traits comprising the broader phenotype. To this end, we selected families that were relatively affected for the previously defined ADI factors and used them for linkage analysis. We observed increased linkage in a rigid-compulsive subset, and a simulation study suggests this result is significant ($P = 0.02$). However, the simulation does not account for testing multiple phenotypic subsets, and Bonferroni correction would render this result non-significant. The potential for *SLC6A4* as a candidate is strengthened by (1) the nature of the phenotypic subset in which we observe increased linkage and (2) the efficacy of SSRIs in treating aspects of the autism phenotype such as those present in this subset and related phenotypes (e.g., obsessive-compulsive disorder). While we believe it is likely that these ADI-based subsets are genetically relevant and will thus provide increased power to detect risk alleles, we recognize that splitting our dataset reduces overall power to detect effects.

To our knowledge, this is the first report describing intermarker LD data and haplotype structures for this locus. We identified a single block spanning the coding region, and three SNPs will represent all common haplotypes. We did not detect association at *SLC6A4* in the main haplotype block corresponding to SNPs 2-8, although we saw nominal but not highly significant

TABLE IV. TRANSMIT Haplotype Analysis for SNPs 5-6-7

2-3-4-5-6-7-8 Haplotype	Frequency	Observed transmissions	Expected transmissions	χ^2	Global χ^2	*P*
C - T - T - G - G - C - G	0.43	215.5	203.0	2.23		
C - C - T - G - A - T - T	0.08	31.0	41.0	5.75		
T - T - A - A - G - T - T	0.37	175.0	178.2	0.12		
C - C - T - G - G - T - T	0.10	53.5	53.7	0.01		
Overall TRANSMIT analysis					8.00	0.09

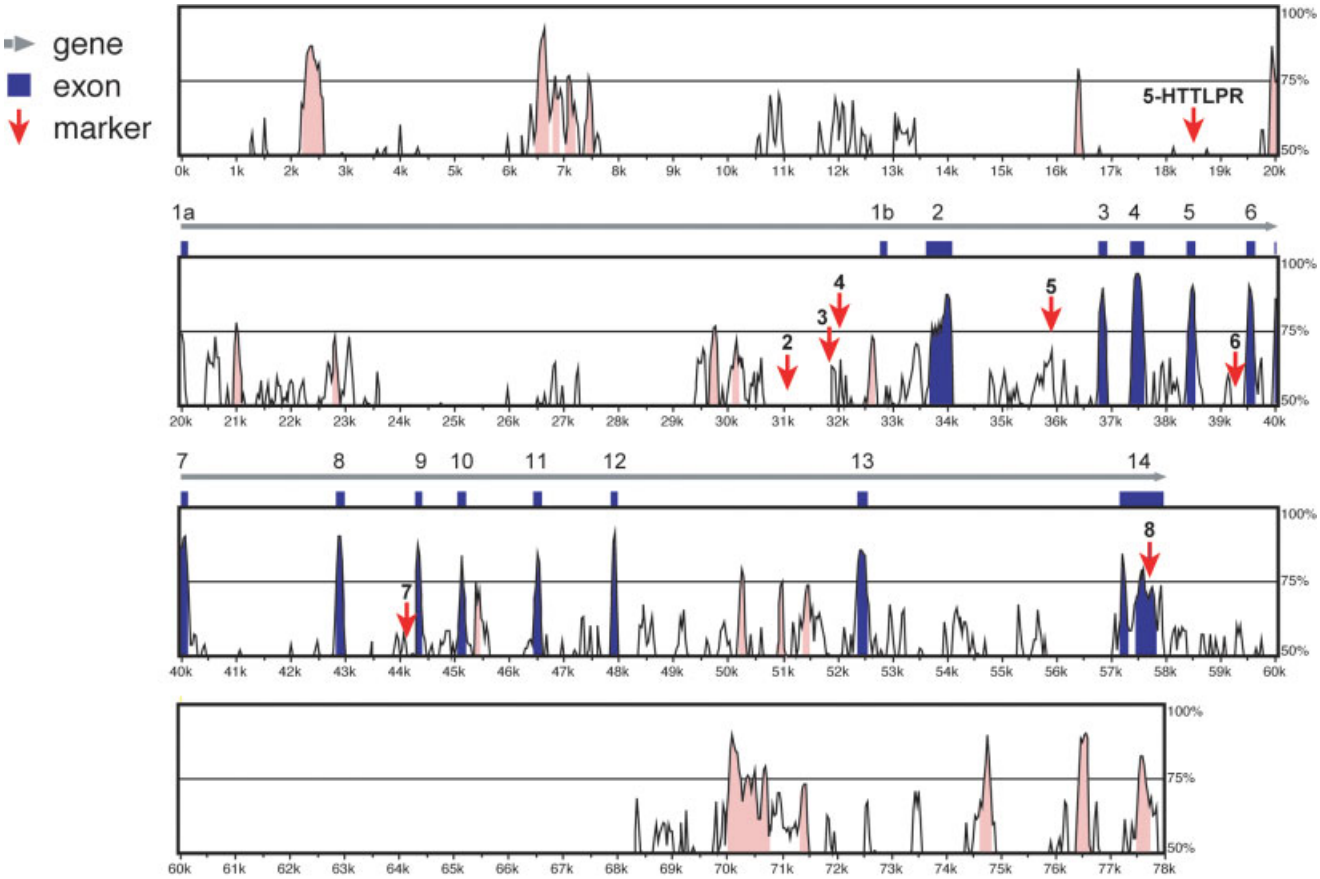

Fig. 3. Evolutionary conservation at *SLC6A4*. Output from VISTA analysis of the *SLC6A4* transcriptional unit (indicated by gray arrow above the plot) is shown, with regions of non-coding sequence conservation (>75% identity) highlighted by pink shading and coding homology by blue shading. The position of markers is indicated by red vertical arrows. [Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

association at 5-HTTLPR, 5′ to this block. Thus, we conclude that a common disease-associated variant does not exist in this block. Future studies will examine the possibility of heterogeneous and/or rare variants in the main haplotype block in our dataset. Additionally, regions 5′ and 3′ to this block will be tested for the presence of either common or heterogeneous variants. We will be informed by the location of conserved and potentially functional sequences identified by VISTA to guide these efforts. We are cautious in interpreting the modest association at 5-HTTLPR. In addition to the magnitude of the result in relation to linkage, the association data have not been corrected for multiple comparisons. Bonferroni correction would not be appropriate in this situation, given the high degree of LD between markers and thus non-independent nature of the corresponding tests. If the association at 5-HTTLPR is biologically meaningful, one possibility is that the short allele is present on, but not in perfect LD with, a haplotype containing a functional variant located 5′ to the coding region. Alternatively, 5-HTTLPR may interact with another variation located at *SLC6A4*, presumably influencing gene expression.

These results must be interpreted in the light of the significant association to autism in parent-child trio families reported by Cook and colleagues [Kim et al., 2002]. Comparison across study is aided by a commonality of genetic markers but is complicated by the difference in family structure analyzed. For 5-HTTLPR, we replicate their previous finding of excess transmission of the short allele [Cook et al., 1997]. With the exception of 5-HTTLPR, markers demonstrating association in the Kim et al. report do not show a similar effect in our dataset. The possibility of biases exists towards different genetic mechanisms in multiplex families compared to singleton cases. These populations may harbor overlapping but significantly different constellations of autism-susceptibility alleles. Considering the serotonin transporter in particular, it would be interesting to know whether affected individuals in multiplex and trio families differ in response to SSRIs. The number of transmissions analyzed in this study is slightly larger than that tested in the Kim et al. paper, so the current study is not relatively limited by power. Ultimately, such detailed studies in independent datasets will be required to make a determination of how or if these data may be generalized across autism families.

Single marker association studies can be powerful in their ability to detect genetic effects. However, it is unlikely that a given polymorphic marker will be a

disease-susceptibility variant. Rather, a positive finding of association indicates linkage disequilibrium to a nearby variant. We argue that characterization of all common haplotypes at a locus is important for meaningful association analysis. The absence of positive findings may simply be the result of failing to test the relevant allele. In this regard, potential susceptibility alleles may exert their effect through perturbations in gene expression, and relevant sequences controlling expression may lay distant 5′ or 3′ to the transcriptional unit. This point highlights the utility of identifying conserved sequences around a gene locus under study, so that potential functional sequences may be included in efforts to detect disease-associated variants. Evaluation of a candidate gene must ultimately be able to detect the presence of either common or heterogeneous disease-associated variants. Allelic association studies permit identification of *common* variants at a disease-susceptibility locus. However, the absence of association does not exclude the possibility of heterogeneous variants or rare mutations as underlying increased risk associated with a gene. Therefore, a thorough examination of a candidate gene must supplement allelic association analysis with direct screening to identify sequence variants. Such an effort represents an important future direction for continued analysis of this gene in autism and related disorders.

ACKNOWLEDGMENTS

This work was supported by grants MH55135 to S.E.F. and MH61009 to J.S.S., a predoctoral fellowship to J.L.M. from the National Alliance for Autism Research, and a grant from the Hobbs Society, Vanderbilt John F. Kennedy center to B.D.B. We acknowledge the invaluable assistance of the Vanderbilt Program in Human Genetics DNA Resources and Data Analysis Cores. We acknowledge the resources purchased from the Autism Resource Genetic Exchange (AGRE), and we thank the participating New England and AGRE families for making this study possible.

REFERENCES

Abecasis GR, Cookson WO. 2000. GOLD—Graphical overview of linkage disequilibrium. Bioinformatics 16(2):182–183.

Abramson RK, Wright HH, Carpenter R, Brennan W, Lumpuy O, Cole E, Young SR. 1989. Elevated blood serotonin in autistic probands and their first-degree relatives. J Autism Dev Disord 19(3):397–407.

Alarcon M, Cantor RM, Liu J, Gilliam TC, Geschwind DH, AGRE Consortium. 2002. Evidence for a language quantitative trait locus on chromosome 7q in multiplex autism families. Am J Hum Genet 70(1): 60–71.

Anderson GM, Freedman DX, Cohen DJ, Volkmar FR, Hoder EL, McPhedran P, Minderaa RB, Hansen CR, Young JG. 1987. Whole blood serotonin in autistic and normal subjects. J Child Psychol Psychiatry 28(6):885–900.

Buxbaum JD, Silverman JM, Smith CJ, Kilifarski M, Reichert J, Hollander E, Lawlor BA, Fitzgerald M, Greenberg DA, Davis KL. 2001. Evidence for a susceptibility gene for autism on chromosome 2 and for genetic heterogeneity. Am J Hum Genet 68(6):1514–1520.

Clayton D. 1999. A generalization of the transmission/disequilibrium test for uncertain-haplotype transmission. Am J Hum Genet 65(4):1170–1177.

Collaborative Linkage Study of Autism. 2001. Incorporating language phenotypes strengthens evidence of linkage to autism. Am J Med Genet 105(8):539–547.

Cook EH, Jr., Courchesne R, Lord C, Cox NJ, Yan S, Lincoln A, Haas R, Courchesne E, Leventhal BL. 1997. Evidence of linkage between the serotonin transporter and autistic disorder. Mol Psychiatry 2(3):247–250.

Fiskerstrand CE, Lovejoy EA, Quinn JP. 1999. An intronic polymorphic domain often associated with susceptibility to affective disorders has allele dependent differential enhancer activity in embryonic stem cells. FEBS Lett 458(2):171–174.

Folstein SE, Rosen-Sheidley B. 2001. Genetics of autism: Complex aetiology for a heterogeneous disorder. Nat Rev Genet 2(12):943–955.

Glatt CE, DeYoung JA, Delgado S, Service SK, Giacomini KM, Edwards RH, Risch N, Freimer NB. 2001. Screening a large reference sample to identify very low frequency sequence variants: Comparisons between two genes. Nat Genet 27(4):435–438.

Gordon CT, State RC, Nelson JE, Hamburger SD, Rapoport JL. 1993. A double-blind comparison of clomipramine, desipramine, and placebo in the treatment of autistic disorder. Arch Gen Psychiatry 50(6):441–447.

Hahn MK, Blakely RD. 2002. Monoamine transporter gene structure and polymorphisms in relation to psychiatric and other complex disorders. Pharmacogenomics J 2(4):217–235.

Heils A, Teufel A, Petri S, Stober G, Riederer P, Bengel D, Lesch KP. 1996. Allelic variation of human serotonin transporter gene expression. J Neurochem 66(6):2621–2624.

Holloway JW, Beghe B, Turner S, Hinks LJ, Day IN, Howell WM. 1999. Comparison of three methods for single nucleotide polymorphism typing for DNA bank studies: Sequence-specific oligonucleotide probe hybridisation, TaqMan liquid phase hybridisation, and microplate array diagonal gel electrophoresis (MADGE). Hum Mutat 14(4):340–347.

Hsu TM, Chen X, Duan S, Miller RD, Kwok PY. 2001. Universal SNP genotyping assay with fluorescence polarization detection. Biotechniques 31(3):560, 562, 564–568, passim.

Kim SJ, Cox N, Courchesne R, Lord C, Corsello C, Akshoomoff N, Guter S, Leventhal BL, Courchesne E, Cook EH, Jr. 2002. Transmission disequilibrium mapping at the serotonin transporter gene (SLC6A4) region in autistic disorder. Mol Psychiatry 7(3):278–288.

Klauck SM, Poustka F, Benner A, Lesch KP, Poustka A. 1997. Serotonin transporter (5-Htt) gene variants associated with autism. Hum Mol Genet 6(13):2233–2238.

Kong A, Cox NJ. 1997. Allele-sharing models: LOD scores and accurate linkage tests. Am J Hum Genet 61(5):1179–1188.

Kong A, Gudbjartsson DF, Sainz J, Jonsdottir GM, Gudjonsson SA, Richardsson B, Sigurdardottir S, Barnard J, Hallbeck B, Masson G, Shlien A, Palsson ST, Frigge ML, Thorgeirsson TE, Gulcher JR, Stefansson K. 2002. A high-resolution recombination map of the human genome. Nat Genet 31(3):241–247.

Lesch KP, Bengel D, Heils A, Sabol SZ, Greenberg BD, Petri S, Benjamin J, Muller CR, Hamer DH, Murphy DL. 1996. Association of anxiety-related traits with a polymorphism in the serotonin transporter gene regulatory region. Science 274(5292):1527–1531.

Liu J, Nyholt DR, Magnussen P, Parano E, Pavone P, Geschwind D, Lord C, Iversen P, Hoh J, Ott J, Gilliam TC, The Autism Genetic Resource Exchange C. 2001. A genomewide screen for autism susceptibility loci. Am J Hum Genet 69(2):327–340.

MacKenzie A, Quinn J. 1999. A serotonin transporter gene intron 2 polymorphic region, correlated with affective disorders, has allele-dependent differential enhancer-like properties in the mouse embryo. Proc Natl Acad Sci USA 96(26):15251–15255.

Maestrini E, Lai C, Marlow A, Matthews N, Wallace S, Bailey A, Cook EH, Weeks DE, Monaco AP. 1999. Serotonin transporter (5-HTT) and gamma-aminobutyric acid receptor subunit beta3 (GABRB3) gene polymorphisms are not associated with autism in the IMGSA families. The International Molecular Genetic Study of Autism Consortium. Am J Med Genet 88(5):492–496.

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(1):146–154.

Mayor C, Brudno M, Schwartz JR, Poliakov A, Rubin EM, Frazer KA, Pachter LS, Dubchak I. 2000. VISTA: Visualizing global DNA sequence alignments of arbitrary length. Bioinformatics 16(11):1046–1047.

McDougle CJ, Naylor ST, Cohen DJ, Volkmar FR, Heninger GR, Price LH. 1996. A double-blind, placebo-controlled study of fluvoxamine in adults with autistic disorder [see comments]. Arch Gen Psychiatry 53(11): 1001–1008.

Nurmi EL, Dowd M, Tadevosyan-Leyfer O, Haines JL, Folstein SE, Sutcliffe JS. 2003. Exploratory subsetting of autism families based on savant skills improves evidence of genetic linkage to 15q11-q13. J Am Acad Child Adolesc Psychiatry 42(7):856–863.

Persico AM, Militerni R, Bravaccio C, Schneider C, Melmed R, Conciatori M, Damiani V, Baldi A, Keller F. 2000. Lack of association between serotonin transporter gene promoter variants and autistic disorder in two ethnically distinct samples. Am J Med Genet 96(1):123–127.

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(3):717–726.

Rapin I. 1997. Autism. N Engl J Med 337(2):97–104.

Risch N, Spiker D, Lotspeich L, Nouri N, Hinds D, Hallmayer J, Kalaydjieva L, McCague P, Dimiceli S, Pitts T, Nguyen L, Yang J, Harper C, Thorpe D, Vermeer S, Young H, Hebert J, Lin A, Ferguson J, Chiotti C, Wiese-Slater S, Rogers T, Salmon B, Nicholas P, Petersen PB, Pingree C, McMahon W, Wong DL, Cavalli-Sforza LL, Kraemer HC, Myers RM. 1999. A genomic screen of autism: Evidence for a multilocus etiology. Am J Hum Genet 65(2):493–507.

Schain RJ, Freedman D. 1961. Studies on 5-hydroxyindole metabolism in autistic disorder and other mentally retarded children. J Pediatrics 58:315–320.

Shao Y, Cuccaro ML, Hauser ER, Raiford KL, Menold MM, Wolpert CM, Ravan SA, Elston L, Decena K, Donnelly SL, Abramson RK, Wright HH, DeLong GR, Gilbert JR, Pericak-Vance MA. 2003. Fine mapping of autistic disorder to chromosome 15q11-q13 by use of phenotypic subtypes. Am J Hum Genet 72(3):539–548.

Shao Y, Raiford KL, Wolpert CM, Cope HA, Ravan SA, Ashley-Koch AA, Abramson RK, Wright HH, DeLong RG, Gilbert JR, Cuccaro ML, Pericak-Vance MA. 2002. Phenotypic homogeneity provides increased support for linkage on chromosome 2 in autistic disorder. Am J Hum Genet 70(4):1058–1061.

Sobel E, Lange K. 1996. Descent graphs in pedigree analysis: Applications to haplotyping, location scores, and marker-sharing statistics. Am J Hum Genet 58(6):1323–1337.

Tadevosyan-Leyfer O, Dowd M, Mankoski R, Winklosky B, Putnam S, McGrath L, Tager-Flusberg H, Folstein SE. 2003. A principal components analysis of the autism diagnostic interview-revised. J Am Acad Child Adolesc Psychiatry 42(7):864–872.

Tordjman S, Gutknecht L, Carlier M, Spitz E, Antoine C, Slama F, Carsalade V, Cohen DJ, Ferrari P, Roubertoux PL, Anderson GM. 2001. Role of the serotonin transporter gene in the behavioral expression of autism. Mol Psychiatry 6(4):434–439.

Vaswani M, Linda FK, Ramesh S. 2003. Role of selective serotonin reuptake inhibitors in psychiatric disorders: A comprehensive review. Prog Neuropsychopharmacol Biol Psychiatry 27(1):85–102.

Veenstra-VanderWeele J, Anderson GM, Cook EH, Jr. 2000. Pharmacogenetics and the serotonin system: Initial studies and future directions. Eur J Pharmacol 410(2–3):165–181.

Yirmiya N, Pilowsky T, Nemanov L, Arbelle S, Feinsilver T, Fried I, Ebstein RP. 2001. Evidence for an association with the serotonin transporter promoter region polymorphism and autism. Am J Med Genet 105(4): 381–386.