

Research article

Open Access

Examination of *NRCAM*, *LRRN3*, *KIAA0716*, and *LAMB1* as autism candidate genes

Holli B Hutcheson^{1,2}, Lana M Olson¹, Yuki Bradford¹, Susan E Folstein³, Susan L Santangelo⁴, James S Sutcliffe¹ and Jonathan L Haines*¹

Address: ¹Center for Human Genetics Research and Department of Molecular Physiology and Biophysics, Vanderbilt University Medical Center, Nashville, TN, 37027, U.S.A, ²Laboratory of Genomic Diversity, National Cancer Institute, Frederick Cancer Research and Development Center, Frederick, MD, 21702, U.S.A, ³Department of Psychiatry, New England Medical Center/Tufts University School of Medicine, Boston, MA, 02111, U.S.A and ⁴Psychiatric & Neurodevelopmental Genetics Unit, Massachusetts General Hospital, Boston, MA, 02129, U.S.A

Email: Holli B Hutcheson - hutcheso@ncifcrf.gov; Lana M Olson - olson@chgr.mc.vanderbilt.edu; Yuki Bradford - bradford@chgr.mc.vanderbilt.edu; Susan E Folstein - SFolstein@tufts-nemc.org; Susan L Santangelo - ssantana@hsph.harvard.edu; James S Sutcliffe - james.s.sutcliffe@vanderbilt.edu; Jonathan L Haines* - jonathan@chgr.mc.vanderbilt.edu

* Corresponding author

Published: 05 May 2004

Received: 20 November 2003

BMC Medical Genetics 2004, 5:12

Accepted: 05 May 2004

This article is available from: <http://www.biomedcentral.com/1471-2350/5/12>

© 2004 Hutcheson et al; licensee BioMed Central Ltd. This is an Open Access article: verbatim copying and redistribution of this article are permitted in all media for any purpose, provided this notice is preserved along with the article's original URL.

Abstract

Background: A substantial body of research supports a genetic involvement in autism. Furthermore, results from various genomic screens implicate a region on chromosome 7q31 as harboring an autism susceptibility variant. We previously narrowed this 34 cM region to a 3 cM critical region (located between D7S496 and D7S2418) using the Collaborative Linkage Study of Autism (CLSA) chromosome 7 linked families. This interval encompasses about 4.5 Mb of genomic DNA and encodes over fifty known and predicted genes. Four candidate genes (*NRCAM*, *LRRN3*, *KIAA0716*, and *LAMB1*) in this region were chosen for examination based on their proximity to the marker most consistently cosegregating with autism in these families (D7S1817), their tissue expression patterns, and likely biological relevance to autism.

Methods: Thirty-six intronic and exonic single nucleotide polymorphisms (SNPs) and one microsatellite marker within and around these four candidate genes were genotyped in 30 chromosome 7q31 linked families. Multiple SNPs were used to provide as complete coverage as possible since linkage disequilibrium can vary dramatically across even very short distances within a gene. Analyses of these data used the Pedigree Disequilibrium Test for single markers and a multilocus likelihood ratio test.

Results: As expected, linkage disequilibrium occurred within each of these genes but we did not observe significant LD across genes. None of the polymorphisms in *NRCAM*, *LRRN3*, or *KIAA0716* gave $p < 0.05$ suggesting that none of these genes is associated with autism susceptibility in this subset of chromosome 7-linked families. However, with *LAMB1*, the allelic association analysis revealed suggestive evidence for a positive association, including one individual SNP ($p = 0.02$) and three separate two-SNP haplotypes across the gene ($p = 0.007$, 0.012 , and 0.012).

Conclusions: *NRCAM*, *LRRN3*, *KIAA0716* are unlikely to be involved in autism. There is some evidence that variation in or near the *LAMB1* gene may be involved in autism.

Background

Autism is a severe neuro-developmental disorder that manifests itself during the first three years of life and persists throughout a patient's lifetime. Because of the frequency with which it occurs, its severity, and its impact on children and families, autism is a major public health concern. It is estimated to occur in ~1/300 births, affecting males three times more often than females. Autistic individuals display impairments in sociability, communication, and also demonstrate repetitive and/or obsessive-compulsive behaviors [1-6].

A substantial body of data supports a genetic involvement in the etiology of autism. Further data suggests that at least one gene for autism lies on chromosome 7q22-31 [7-10]. Previously, the Collaborative Linkage Study of Autism (CLSA) narrowed the 34 cM 7q22-31 critical region to a 3 cM region between D7S496 and D7S2418.) [11]. Four genes within this region, Neuronal Cell Adhesion Molecule (NRCAM) (XM_027222); Leucine Rich Repeat Protein Neuronal 3 (LRRN3) (XM_045261); KIAA0716 (NM_014705); and Laminin Beta-1 (LAMB1) (NM_002291) were chosen for further study. This choice was based equally on their proximity to the microsatellite marker (D7S1817) most consistently cosegregating in this group of families, their tissue expression patterns, and their likely biological relevance to the autism disease process as described below.

The NRCAM gene encodes the Nr-CAM protein that is expressed in structures in the developing brain, including the floor plate. In this neuronal region, Nr-CAM has been implicated in axonal guidance through interaction with TAG-1/axonin-1 [12,13]. Additionally, when it is presented as a substrate in *in vitro* studies, Nr-CAM induces neurite outgrowth from dorsal root ganglia neurons [13]. Nr-CAM also serves as a receptor for several different neuronal recognition molecules. In its role as a receptor, it is active during nervous system development in several different regions including the spinal cord, the visual system, and the cerebellum [12].

Studies in *Drosophila* demonstrate that many members of the LRR family provide an essential role in target recognition, axonal pathfinding, and cell differentiation during neural development [14,15]. NLRR-3, the murine ortholog of human LRRN3, was isolated using a human brain cDNA fragment encoding an LRR as a probe against a mouse brain cDNA [16]. NLRR-3 mRNA is expressed abundantly in the brain and has very little expression in other tissues. Its expression is developmentally regulated and is confined to the nervous system. Its molecular structure and its expression pattern suggest that the NLRR-3 protein plays a role in the development and maintenance of the murine nervous system through protein-protein

interactions [17,18]. These murine studies further strengthen the relevance of the *Drosophila* results in that these LRR proteins could have similar, if not the same, integral functions in mammalian neural development. Other studies have also implicated NLRR-3 as an important component of the murine pathophysiological response to brain injury [17]. Since LRRN3 is located within this autism candidate region, and shares a high degree of homology to NLRR-3, it was considered a strong autism candidate gene.

Based on an hypothesis that large cDNAs (> 4 kb) encoding large proteins (>50 kDA) in brain are likely to play an important role in mammals, studies to identify such novel genes have been performed [19]. During the course of these analyses, the KIAA0716 cDNA was identified. Although little is known about KIAA0716 and its protein product, homology relationships suggest it is involved in cell signaling and communication based on its similarity to KIAA0299, which encodes the Dedicator of Cytokinesis 3 (DOCK3) protein. The DOCK proteins bind to the Src-Homology 3 (SH3) domain of the v-crk sarcoma virus CT10 oncogene homolog (CRK) protein and play an important role in signaling from focal adhesions [20].

Laminin is a large molecular weight glycoprotein present in basement membranes. Most laminin molecules are comprised of an α chain and two β chains that are assembled into a cruciform structure held together by disulfide bonds. Three types of β chains (β 1 chain, β 2 chain, and β 3 chain) have been described and are thought to contribute to the functional variability of the members of the laminin family. The laminin-B1 (LAMB1) gene encodes the laminin B1 chain protein which is present at low levels in serum and is involved in cell attachment and chemotaxis presumably through its binding to the laminin receptor (expressed highly in the brain) (reviewed in [21]). Data from *Xenopus* studies demonstrate that laminin-1-beta is important in axonal guidance during embryonic development when it is co-expressed with netrin. More specifically, when netrin and laminin-1-beta are coexpressed, axons are repulsed into the areas where only netrin, and not laminin-1-beta, is present [22].

Based on their putative biological functions as well as their genetic locations within our previously reported 3 cM chromosome 7 autism critical region [11], these four neuronally expressed genes presented themselves as intriguing autism candidate genes. Thirty six intronic and exonic single nucleotide polymorphisms (SNPs) roughly spaced at 10 kb intervals throughout these candidate genes were analyzed via both single locus (Pedigree Disequilibrium Test-PDT) and multilocus (TRANSMIT) approaches. Additionally, Linkage Disequilibrium (LD)

Table 1: Primers used for candidate geneSNP assays.

Gene	Assay ID	Forward Primer	Reverse Primer	Product Size	Major Allele Frequency
NRCAM	RS427615	ttgaggcaaccaatagaatg	cctgaccattttcccttcag	203	0.85
NRCAM	RS381828	tcctgggatttactgcttgaa	tctgtaatgaaagcaaacatga	178	0.79
NRCAM	RS414077	ctgccttggaaggtgaggat	cggaattttctggaaaaaca	177	0.89
NRCAM	RS401433	ccctagaagtcccgcctttcc	gtggcaccacaaaactgaag	183	0.75
NRCAM	RS1043895	ccctagaagtcccgcctttcc	gtggcaccacaaaactgaag	183	0.78
NRCAM	RS405945	atgtagtctgcaccatcc	gctctcagtctgggactct	173	0.75
NRCAM	RS1269620	tgatcttgccggctctagtt	catggtaccttgactcttctca	176	0.80
NRCAM	RS1269627	actgacagaggcttgcacaa	tggaaaggcctccataaaaa	160	0.80
NRCAM	RS1269629	ccccacagcttgacttctct	gggtgggagcatgagagaga	124	0.54
NRCAM	RS2072546	acttactggatggcgga	ttttctggaccgtaatgg	202	0.79
NRCAM	RS1269634	gcttctttttcggcatta	ctgcagactccgcttctca	212	0.57
LRRN3	RS280282	ttgaagagcaacatgggta	gatgattgggaaaatcaa	193	0.98
LRRN3	RS280306	tgtaccatgtgtaaaactgtaa	gcagaagcatcctaacaacagg	206	0.89
LRRN3	RS280309	actggctggagccacattat	gaaaatgggagaggtgaggat	149	0.95
LRRN3	RS280310	caaaggcattgcttctgaaa	tgtcctgtttctctgcctct	203	0.87
KIAA0716	RS1362155	tcctgacctcaggcgatactc	tgctaggtgagagatagcagagaa	185	0.96
KIAA0716	RS1859746	gccatagggaacaatcacca	ccatgcccttatatgcttgg	193	0.65
KIAA0716	RS757219	gctcaaacatgagattgc	attacgtgggcaagcaagg	195	0.90
KIAA0716	RS867915	tccaattcactctacctctg	ccttgctgctgctcctta	169	0.91
KIAA0716	RS718845	catggttggagcctgtgtt	tgaaaaccagacacgacctca	168	0.67
KIAA0716	RS1859025	aaggtttggaaaggcaggt	tcccatggtccttgtatca	224	0.72
KIAA0716	CV2541541RPT.	tgtgggtaatttcacaattgg	ttcccatcttccagctgtc	180	0.31
KIAA0716	CV2541544	acaggagcccccaataaat	gacccaaatatcagaccaaaa	178	0.77
KIAA0716	CV339112	ggactttgtcctggagttg	cagtttctcacctgttttc	174	0.80
LAMB1	CVC8856543	ggtgatgtttttttgaagagca	ggaaggagaagtccttact	160	0.67
LAMB1	CV11428541	cggacttctcttccagctct	gggtgggtcttttcacacagg	225	0.67
LAMB1	CV2193685	agggctaagtcctcattctcc	caagatccaaagagcctcca	154	0.67
LAMB1	CV2193686	gatgagaggaggaggaagc	ttcctgtgttttcaaggagg	184	0.50
LAMB1	CV11428543	cctcaggcctagatggtgag	tcctacctccctctcttgc	177	0.60
LAMB1	CV2193689	ctgttttgggggttctttt	cttccctctgattggacagt	145	0.61
LAMB1	CV3268606	cacagaccttggtcaagcag	agacctttccaagcctgtt	201	0.63
LAMB1	CV2193712	ggctgggatcaactctttca	tccagcttccaaaaaggga	200	0.88
LAMB1	CV152495	tgatctcaagcattctcctg	gaggattcactccccaaa	176	0.74
LAMB1	CV2193721	caagtacagtgagggtgat	tctgggaacaattcctggag	205	0.88
LAMB1	CV8428961	tgacagatttaccacttagcc	ggcctttcatcaggcaaat	226	0.60
LAMB1	CV1091266	gaggcctcacagtcataggc	gtggcatgcagttgagtac	190	0.83
LAMB1	CV2193735	aagcgcctgtcattgctg	aaggcctgattgtgtgtg	153	0.52

relationships between all tested SNPs were examined within the four candidates.

Methods

Sample composition

All probands met algorithm criteria of the Autism Diagnostic Interview-Revised (ADI-R) [23] and were at least four years old. All probands were also assessed with the Autism Diagnostic Observation Schedule (ADOS) [24] or a later revision (ADOS-G). Affected sibling pair (ASP) and trio families were recruited. Affected individuals were excluded if they had fragile X syndrome, tuberous sclerosis, or any other medical condition known to be associated with autism.

Families were recruited from three regions of the United States (Midwest, New England, and mid-Atlantic states) through three clinical data collection sites: the University of Iowa, Tufts University-New England Medical Center, and Johns Hopkins University. The Institutional Review Boards at each institution approved this study and appropriate informed consent was obtained from all subjects.

The sample was comprised of the 30 nuclear CLSA chromosome 7 families as previously described.[11]. Both parents were available for genotyping in all but two families. Twenty-nine families had two affected children while one family had three affected children. Briefly, each of these families had evidence for linkage within a 34 cM region of chromosome 7q22-31 with LOD scores = 0.55.

Table 2: LAMB1 exonic screening primer pairs. Capital letters indicate sequence within an exon. SNP assays were already run in the initial association studies using SNPs in exons 6 and 12.

Exon #	Forward Primer	Reverse Primer	Product Size (bp)
1	ccctccatttcctgtcg	gcaacctgaaaaccacct	440
2	ttccatccagtcctctcc	gcgcccttctctcttgaa	298
3	cgcttcctatctggattgct	ccaactctgttgccctgtcct	314
4	tccaggcagtcaatctgtca	aggagcttcatggggtttt	356
5	atccattttgaaaggggaca	tgccagatgtctgttccag	409
6	CV1091266 already run		
7	gagcagatcaccgagggt	atgactgtgaggcctctgtt	556
8	aggtgcacagaggccttact	tttcatctcttgatgaccttc	252
9	aaaatgtccctcaacacaga	tggagaaactggctattcat	210
10	tcccacacagtgagtacatt	ttcttgactgtctcttgca	247
11	caaacacaagaaaactccaaa	gtgctttgattcttccagg	243
12	CV2193721 already run		
13	ggaagactgaggacaagagca	catgtttcacgtttgtctgc	144
14	ctgccacacactcccact	actcagctgatgtggctct	191
15	cctcctccaccagtccacta	aataaccattgttgcctatgc	220
16	atatgaagccgttgagctg	aggcagtggtgggctaactat	217
17	ccagaagcaatctgcactta	tggatgcatgaaggtgactt	200
18	actcagagccagaacagcag	gccgttctcttcttcacag	274
19	ctgccacaatggctgagta	ctggcctctttaccaca	224
20	gtctgctcagccagatca	gcagttcctgttgcctgt	312
21	ctccc aaagtatgcacacga	tggcctgttctctctgtga	218
22	tgggcagaataaatgtgcatc	tttggtttatcgggtgaca	271
22	tcattctaagaatgggcagaa	tgacacgacagaccagaag	172
23	cctgagctcataaaggatca	tgctgtgtattgggagtg	309
24	atcaagtcaatgagagaacg	cctgcttgaccaaggtctgt	162
25	accatgccaggaaatctttg	gctagcattgcttattcccttc	477
26	tgggactattgcacacaaa	acgtgaaagtctcccttgg	273
26	tgatcctcattgtgcccaga	cagatattcgggtgagcat	159
27	cacatcccaaaccttacCA	tgcttgagatgacattacca	259
28	ctgaggcatcactgagtt	gcctgatccttggcttgtt	249
29	cagagaatcacacagaa caaatg	tcattttgagccatttcca	281
30	ccccttagtgccatgtcttt	ggggacatcttctgtttca	258
31	aggtgtgttccattcatgt	tctgggtcatttcaggaagg	281
32	aggagactgggtgctccttt	ttgctggaaaattgacatgc	259
33	aaaaaagcctgtgtgaaaagacc	ccacttcttctgctcaaatgtgg	219
34	agggattcatcaaatcagtg	ggcatccatagtttataaagtga	433

The affected sibling pairs also shared a 3 cM region between D7S496-D7S2418 (120 cM-123 cM) to a greater extent than any other chromosome 7 region.

For exon screening analysis, twelve unrelated affected individuals from these families were chosen because they shared one or both of the two overtransmitted SNP haplotypes (CV2193686/CV3268606 and CV1091266/CV2193735).

Gene characterization and SNP choice

To determine the size and exon/intron structure of the four candidate genes, the Celera, Ensembl, and NCBI databases were used [25-27]. SNPs were chosen either from the NCBI dbSNP database [28] or generated through use of the Celera Discovery System and Celera Genomics'

associated databases. An "RS" number indicates SNPs chosen from dbSNP while a "CV" number designates SNPs chosen from Celera RefSNP. When available, SNPs located in coding regions were chosen for analysis. In general, a 10 kilobase (kb) spacing of SNPs was sought to achieve as complete coverage as possible for a thorough analysis of each candidate gene.

Primer design

Once SNPs were chosen, polymerase chain reaction (PCR) primers were designed for single stranded conformational polymorphism (SSCP) analysis using the Primer 3 web-based PCR primer design software [29]. As part of the primer design process, each forward and reverse primer from the recommended pairs was compared against the NCBI non-redundant (NR) and high throughput genomic

Table 3: Candidate gene association results (results given as nominal *p*-values)

Marker	PDT	Transmit	H-W Test
NRCAM427615	0.80	0.11	1.00
NRCAM381828	0.25		0.61
NRCAM414077	0.16	0.39	0.75
NRCAM401433	1.00		0.88
NRCAM1043895	0.81	0.07	0.88
NRCAM405945	0.08		0.87
NRCAM1269620	0.43	0.78	0.85
NRCAM1269627	0.54		0.71
NRCAM1269629	0.35	0.15	1.00
NRCAM2072546	0.87		0.80
NRCAM1269634	0.16		0.95
LRRN280310	0.09	0.04	0.70
LRRN280309	0.32		0.75
LRRN280306	0.16	0.06	0.77
LRRN280282	1.00		0.63
KIAA1362155	0.32	0.10	0.68
KIAA1859746	0.14		0.35
KIAA757219	0.19	0.01	0.90
KIAA867915	0.18		0.97
KIAA718845	0.92	0.70	0.96
KIAA1859025	0.90		0.92
KIAACV2541541RPT	0.81		0.88
KIAA2541544	0.70		0.53
KIAA339112	0.35		0.94
LAMB8856543	0.67	0.26	0.45
LAMB11428541	0.37		0.62
LAMB2193685	0.35	0.05	0.47
LAMB2193686	0.12		0.69
LAMB11428543	0.47	0.01	0.76
LAMB2193689	0.10		0.68
LAMB3268606	0.17	0.10	0.95
LAMB2193712	0.26		0.78
LAMB152495	0.46	0.52	0.89
LAMB2193721	0.51		0.90
LAMB8428961	0.32	0.57	0.83
LAMB1091266	0.89		0.78
LAMB2193735	0.02		1.00

sequence (HTGS) sequence databases to determine whether or not it was sufficiently specific for accurate PCR amplification in the desired region. In some cases, it was impossible to design a set of primers that were specific enough for the SNP region (high numbers of BLAST hits in many genomic regions), and new SNPs were chosen for primer design and analysis. Table 1 lists the primer sets, product sizes, and allele frequencies for each of the SNP assays.

SNP assay testing and optimization

Two or more different SNP primer sets in the same 10 kb region were designed and subjected to a preliminary round of SSCP screening using 2 different conditions (15 Watts for 4 hours and 28 Watts for 1.5 hours at 4°C) to

ensure that the assays worked and to make a rough estimation of allele frequencies in the test population. In the preliminary screening, 14 unrelated control individuals (28 chromosomes) were tested for each polymorphism. An assay was considered successful, as well as sufficiently polymorphic, when two or more individuals had a consistently different banding pattern in the testing phase and the assay was clear (i.e. a banding pattern with only three possible genotypes representing a single SNP). After all of the assays were initially tested, successful assays spaced at regular intervals throughout the candidate genes were chosen for analysis.

Table 4: Haplotype transmission for two-SNP haplotypes with nominal P < 0.05

Markers	Haplotype	Observed	Expected	χ^2
NRCAM405945/1269620	1/1	38.0	31.4	5.4
	1/2	0.0	0.0	NA
	2/1	48.3	55.9	4.9
	2/2	25.7	24.7	0.1
LRR280310/280309	1/1	0.1	1.1	2.2
	1/2	91.4	95.3	3.1
	2/1	7.3	4.5	4.0
	2/2	13.2	11.1	1.3
KIAA1859746/757219	1/1	34.7	33.5	0.1
	1/2	14.6	9.6	7.5
	2/1	54.3	61.0	3.4
	2/2	8.4	7.9	0.1
KIAA757219/867915	1/1	77.0	85.6	8.1
	1/2	12.0	8.9	2.4
	2/1	21.0	14.2	7.4
	2/2	2.0	3.3	1.2
LAMB11428543/2193689	1/1	34.8	35.9	0.1
	1/2	3.0	4.6	1.5
	2/1	2.2	7.3	10.2
	2/2	74.0	66.1	4.5
LAMB2193689/3268606	1/1	0.3	4.3	8.4
	1/2	36.2	39.2	0.9
	2/1	47.7	38.7	5.2
	2/2	29.8	31.8	0.3
LAMB1091266/2193735	1/1	34.2	42.3	8.5
	1/2	62.3	54.8	6.5
	2/1	14.8	13.4	0.3
	2/2	2.7	3.4	0.7

SNP genotyping

For the majority of the SNP genotyping work, PCR reactions were set up according to standard procedures, denatured at 95°C for 5 minutes, and subjected to SSCP on 0.5X MDE™ gels at 4°C. Individual SSCP gels were run from 15–28 Watts for 1.5–5 hours depending on the previous test run. Gels were stained with Sybr® Gold nucleic acid gel stain (Molecular Probes) and scanned on a Hitachi FMBIOII fluoroimager (Hitachi Instruments, San Jose, CA).

For assay RS280310, denaturing high performance liquid chromatography (dHPLC) was performed using previously published methods on the Transgenomic WAVE™ [30]. In a WAVE™ analysis, all homozygous individuals (regardless of their genotype) will yield a single dHPLC peak whereas heterozygous individuals will yield 2–4 dHPLC peaks based on the melting temperature of the PCR fragment heteroduplex or homoduplex and the actual run conditions. Hence, initially it was impossible for us to distinguish between an A/A homozygote and a C/C homozygote. To unambiguously identify homozygous genotypes, we chose three samples with homozygous

peak patterns and sequenced them to determine whether or not they were A/A or C/C for the particular SNP. Next, we pooled (in a 1:1 ratio) each of the unknown homozygotes with a known (sequenced) homozygote (A/A) and ran these samples on the WAVE™ again. Obtaining a single peak indicated an A/A homozygote. Obtaining four peaks indicated a C/C homozygote for this SNP. To cross-validate the two genotyping methods, selected PCR amplified samples from the RS280310 assay were also run under SSCP conditions and blindly scored. Results were consistent for every sample.

Microsatellite genotyping

During the characterization work of KIAA0716, we identified a novel polymorphic CA dinucleotide repeat within intron 6 (the Celera genomic sequence contained (CA)₁₈, while the NCBI sequence contained (CA)₂₂). A PCR assay was designed, characterized, and applied to the dataset. Primers were designed, controls were sized, and PCR amplification and analysis was performed using standard procedures. Briefly, PCR products were denatured and electrophoresed on 6% (Polyacrylamide Gel Electrophoresis, PAGE) denaturing gels. The gels were then stained

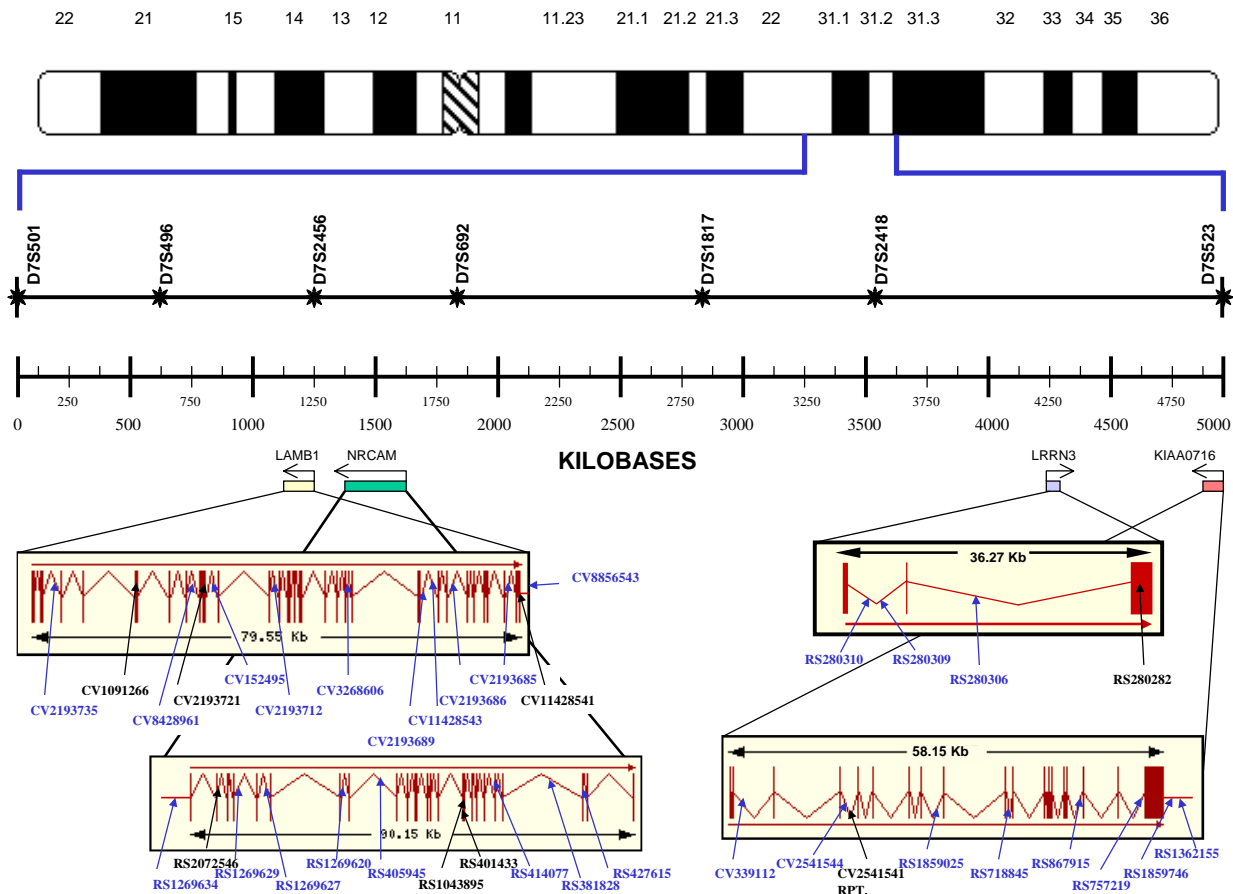


Figure 1
Gene structures for four candidate genes including SNP positions. Exonic SNPs are given in black, intronic SNPs in blue.

with Sybr® Gold nucleic acid gel stain (Molecular Probes, Eugene, OR) and scanned on a Hitachi FMBIO II fluorimager (Hitachi Instruments, San Jose, CA) using the appropriate filter for detection and visualization of the Sybr® Gold stained PCR fragments.

Exon screening

For LAMB1 exon screening analysis, denaturing high performance liquid chromatography (dHPLC) was performed via previously published methods using the Transgenomic WAVE™ [30]. First, PCR primers were designed for exonic amplification as previously described. Second, the assays were tested for specificity via PCR amplification and agarose gel electrophoresis. Third, the exonic assays were amplified in 12 unrelated affected individuals and run on the WAVE™ under denaturing conditions. Any samples with a discrepant dHPLC peak pattern were reamplified and subjected to dye-primer or dye-ter-

minator sequencing using both forward and reverse primers. Table 2 lists the primer sets and PCR product size for each of the exon screening assays.

Genotype error checking

Genotypes were checked for Mendelian consistency, SNP haplotypes were constructed using SimWalk v2.0 [31], plotted in Cyrillic (version 2.1. Oxford, Cherwell Scientific Publishing), and all recombination events were identified. Haplotype reconstruction was based on minimizing the number of recombination events on a chromosome. In cases where apparent excess recombination was observed, gels were reread. If necessary, the SNP assay was rerun to determine the most accurate genotype. Once all of the data corrections were made, the haplotypes generated from the SNP genotype information were compared to the previously generated microsatellite

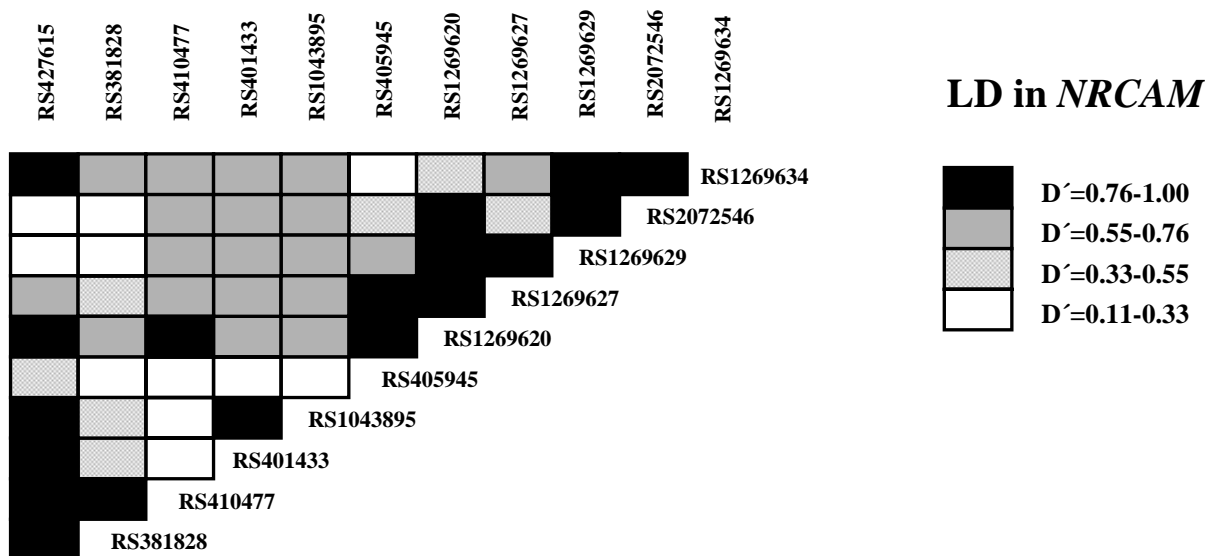


Figure 2 Linkage Disequilibrium results for NRCAM. Each block represents the amount of LD between the indicated pair of SNPs.

marker haplotypes within and around the candidate genes. In all cases, the haplotypes were consistent.

Genotype frequencies of each of the various SNPs were also analyzed to determine that they conformed to Hardy-Weinberg equilibrium based on the observed allele frequencies.

Pedigree disequilibrium test (PDT)

Since a small number of families in our dataset were missing one parent or included half-siblings (one family), we used the PDT [32] to test for association between autism susceptibility and the genotyped SNPs. The PDT is an extension of the Transmission Disequilibrium Test (TDT) [33] that allows inclusion of extended families to test for allelic association. It has been shown [34] that the PDT is a valid and unbiased test of association even when linkage exists.

Transmit

Haplotype analysis offers a valuable tool for investigating associations between disease loci and multiple markers [35]. Therefore, TRANSMIT [36] was applied to adjacent two-SNP genotypes to determine whether any of these haplotypes were preferentially transmitted. Three-SNP haplotype analyses were not attempted due to the small overall sample size. All sampled individuals were included in the analyses.

Linkage disequilibrium (LD)

LD was assessed for the SNPs using the Graphical Overview of Linkage Disequilibrium (GOLD) program [37] and characterized using the D' statistic [38].

Results

In the subset of CLSA families chosen for increased sharing of the 4.5 Mb region of 7q31, we genotyped and ana-

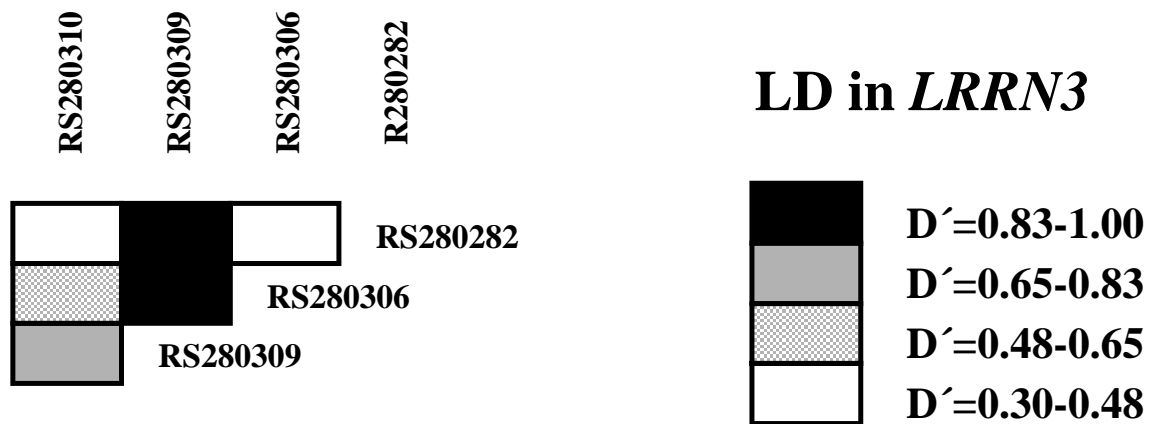


Figure 3
Linkage Disequilibrium results for LRRN3. Each block represents the amount of LD between the indicated pair of SNPs.

lyzed one microsatellite marker, 29 intronic SNPs, and seven exonic SNPs (average spacing of 10 kb) derived from the Celera RefSNP and NCBI dbSNP databases. Figure 1 gives the intron/exon structures and relative SNP positions for each gene.

Once the genotype information was gathered and error checking was performed, association analyses were carried out to determine whether these genes are associated with autism susceptibility. The results from the allelic and haplotypic association analyses are given in Table 3. None of the individual polymorphisms in or surrounding *NRCAM*, *LRRN3*, or *KIAA0716* demonstrated evidence for allelic association to autism. However, this analysis did reveal suggestive evidence for allelic association between *LAMB1* and autism susceptibility for one SNP (CV2193735) located within intron 3 ($p = 0.02$). Examination of the two-SNP haplotypes produced marginally

significant results for at least one combination in *NRCAM*, *LRRN3*, and *KIAA0716*. However, three different two-SNP combinations within *LAMB1* (CV11428543/CV2193689 $p = 0.007$ both within intron 25; CV2193689/CV3268606 $p = 0.012$ within intron 25 and intron 23; and CV1091266/CV2193735 $p = 0.012$ within introns 6 and 3) displayed strongly significant results (Table 4). All SNPs were in Hardy-Weinberg Equilibrium.

Examination of marker-to-marker linkage disequilibrium identified discrete blocks of LD present within each of the genes. With this set of SNPs, *NRCAM* has four such blocks of LD defined by its 11 SNPs. The SNPs within *LAMB1* displayed very strong LD in three blocks with a somewhat weaker level of LD among a majority of the SNPs. Within *LRRN3*, there is one SNP (RS280309) that seems to be in strong LD with the majority of the other SNPs. The 8 SNPs in *KIAA0716* also comprise definite LD blocks although

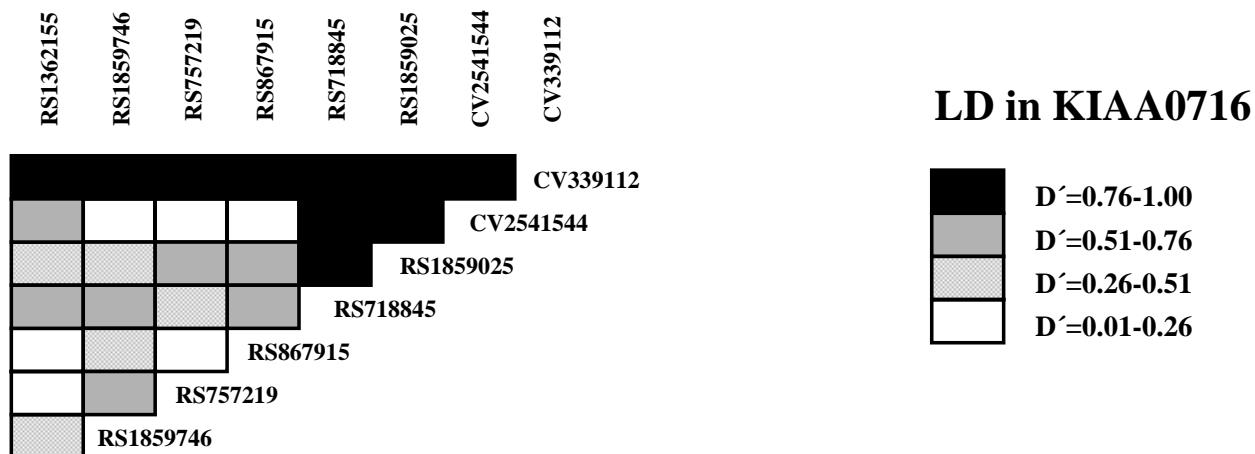


Figure 4
Linkage Disequilibrium results for KIAA0716. Each block represents the amount of LD between the indicated pair of SNPs.

they are not as tightly linked as those in *NRCAM* and *LAMB1*. LD did not extend across these genes in the broader region (data not shown). The LD results are shown in Figures 2,3,4,5.

After observing the multiple suggestive allelic association results in *LAMB1*, we screened its 34 exons for susceptibility variants using 12 unrelated affected individuals who shared one or both of the overtransmitted haplotypes (CV11428543/CV2193689-2/2 and CV1091266/CV2193735-1/2). All of the exons were PCR-amplified in these individuals and screened for variations via dHPLC on the Transgenomic WAVE™. Whenever one of the exonic assays displayed a discrepant dHPLC peak pattern in any of the individuals, the exon was reamplified and sequenced in the discrepant individual and a control with forward and reverse primers.

13 of the 34 exonic assays had discrepant WAVE™ patterns in at least one tested individual. 12 of the 13 discrepant assays resulted in detectable SNPs while a SNP was not observed via sequencing in the other discrepant assay. Six of these detectable SNPs were in intronic flanking regions (exonic assays 1, 2, 4, 5, 26, and 34) while the other six were exonic. Four of the exonic SNPs resulted in synonymous codon substitutions (exonic assays 10, 15, 23, and 31) and the other two exonic SNPs created nonsynonymous amino acid changes (exonic assays 20 and 22). The G/A SNP in exon 20 at mRNA position 2915 creates a nonconserved glycine to serine change (G544S). A restriction assay was utilized to determine if this SNP was over-represented in unrelated autistic individuals taken from the CLSA dataset (N = 90) compared to unrelated CEPH controls (N = 80). No significant difference in the allele frequencies between the two groups was found ($\chi^2 = 0.2177$; $P = 0.64$). The exon 22 SNP is an A/G change at

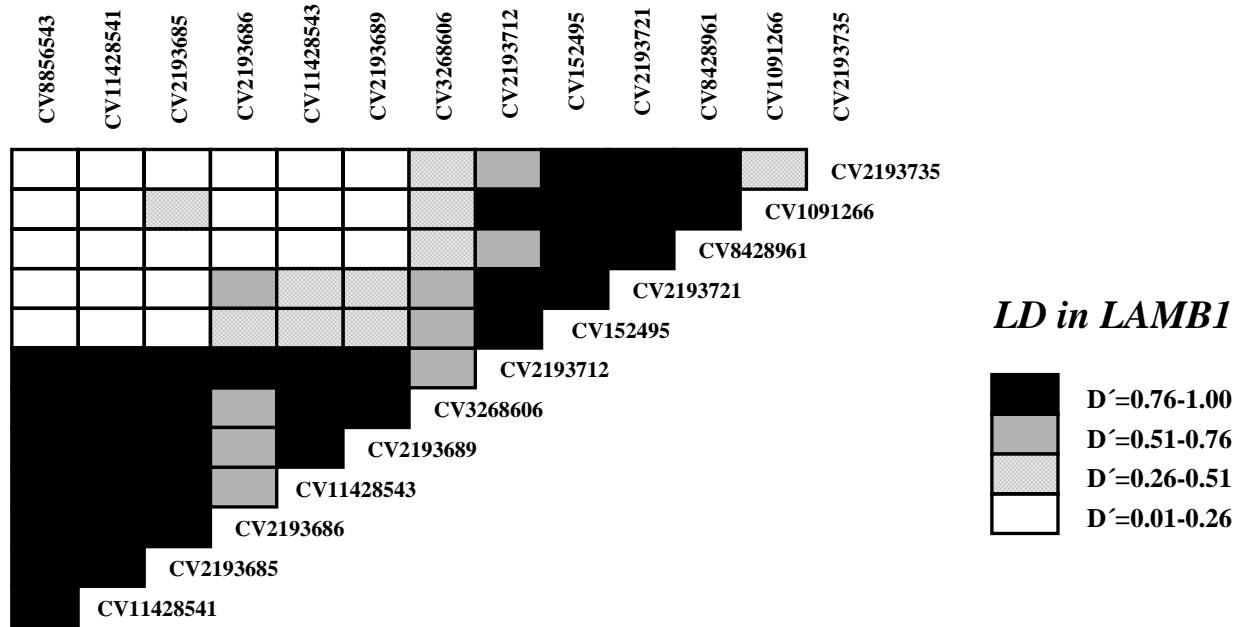


Figure 5 Linkage Disequilibrium results for LAMB1. Each block represents the amount of LD between the indicated pair of SNPs.

mRNA position 3402 creating a conserved glutamine to arginine change (Q706R). We assayed this SNP via SSCP in unrelated autistic probands taken from the CLSA and Autism Genetic Resource Exchange (AGRE) datasets (N = 215) and unrelated CEPH controls (N = 73) to determine if there was a difference in allele frequencies. We did not observe a significant difference in allele frequencies between the two groups ($\chi^2 = 0.00 ; p = 1.00$).

Discussion

Using a positional candidate gene approach we chose to examine *LAMB1*, *NRCAM*, *KIAA0716*, and *LRRN3* for association to autism susceptibility.

Since it has been shown that using one or two SNPs is insufficient for a thorough candidate gene/disease association analysis [35], we chose multiple SNPs placed at regular intervals throughout these genes for our study.

Intronic and exonic SNPs were chosen to ensure that a susceptibility variant or a variant in LD with a true susceptibility allele would be detected. From this work, we observed some evidence for association between *LAMB1* and autism, including one individual SNP (CV2193735) located within intron 3 ($p = 0.02$) and three separate two-SNP haplotypes (CV11428543/CV2193689, CV2193689/CV3268606, and CV1091266/CV2193735) across the gene's transcriptional unit ($p = 0.007, 0.012, \text{ and } 0.012$ respectively).

These results have not been corrected for multiple testing since it is still unclear as to what level of correction should be applied in an association study such as this. A Bonferroni correction ($p < .004$ correcting on 12 tests and $p < .002$ correcting on 24 tests) is too stringent to apply to these data since these tests are not all based on independent data points. However, not correcting at all for multiple

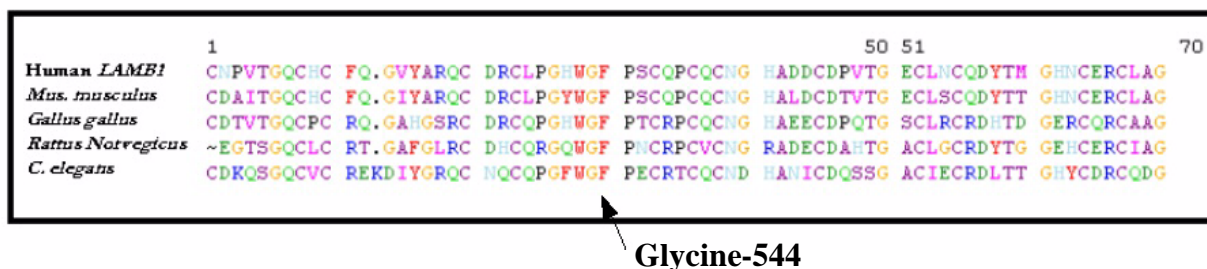


Figure 6
Amino acid conservation among various organisms for human *LAMB1* 544G.

tests will invariably lead to a high number of false positives in any study. The proper approach toward correction for multiple comparisons has yet to be resolved. Therefore, we provide the nominal results to allow the reader to decide the level of error correction to apply.

Recently, it has been shown that LD varies markedly over different chromosomal regions and distances. Furthermore, average LD measures cannot be accurately predicted from one chromosomal region to another [39-41]. Hence, we defined the pattern of LD to ensure that we had adequate SNP coverage for our association study and achieved a sufficient SNP spacing for an association examination to be performed. Given the recent studies showing the tendency for LD to occur in "blocks" of DNA that can range from ~5–100 kb [42,43] and the SNP coverage that was achieved in this study, it was hardly surprising that we observed distinct blocks of LD within these genes.

During the course of the *LAMB1* exon screening, 2 SNPs that led to amino acid changes were discovered in exons 20 and 22. The nonconserved amino acid change in exon 20 results in a glycine to serine substitution. This glycine is located within one of the Laminin Epidermal Growth Factor-like domains of the protein and has been evolutionarily conserved in various species including *Mus musculus*, *Gallus gallus*, *Rattus Norvegicus*, and *C. elegans* (Figure 6). Given the extent of the amino acid change from the conformationally important glycine, which can confer a large degree of local flexibility on its polypeptide, to serine, another residue that can have a large effect on the local polypeptide conformation, and the extent of evolutionary conservation of this particular amino acid region among various species, it was somewhat surprising to us that there was no significant difference in allele frequencies between the autism and control groups. Regardless, even if this substitution changes the protein

conformation or alters protein activity, we conclude that it does not play a significant role in susceptibility to autism. The other amino acid change that we observed through exonic screening was within exon 22 and resulted in a conserved glutamine to arginine change. This glutamine is located within the myosin tail of the protein and is not evolutionarily conserved across species (data not shown).

Conclusions

Extensive SNP genotyping in three genes within the autism candidate 7q31 region, NRCAM, LRRN3, and KIAA0716 did not reveal any genomic variation associated with autism. However, some evidence of association with a multi-locus haplotype in LAMB1 was observed. Although exon screening did not discover a common variation that alters the LAMB1 protein product in autistic individuals, it is possible that disease susceptibility could also be conferred from a variation in the gene's regulatory region or from an intronic variant that impairs or alters splicing. Thus LAMB1 remains a viable candidate gene and may be associated with autism susceptibility in a subset of autistic patients. Further testing of our genetic findings in other datasets is required to definitively confirm or negate these results.

Competing interests

None declared.

Authors' contributions

HBH carried out the molecular analyses, helped design the study, and draft the manuscript. LMO and YB helped carry out the statistical analyses. SEF ascertained the clinical material, helped with study design, and draft the manuscript. SLS helped to ascertain the clinical material and to coordinate the clinical aspects of the study. JSS helped in the overall design of the study, helped design the molecular studies, and helped draft the manuscript. JLH oversaw the design and implementation of the overall study, the molecular analyses, the statistical analyses, and drafting of the paper. All authors read and approved the final manuscript.

Acknowledgments

We would like to thank the numerous parent and family support groups that have assisted us in our studies. We would also like to offer our appreciation to all members of the CLSA, especially, for graciously allowing us to utilize the CLSA families ascertained by our collaborators at the University of North Carolina and University of Iowa. We are most grateful, though, to the subjects and their families for participating in this study and therefore granting us the privilege of performing these studies. This work was supported in part by the following grants: MH55135, KO2-MH01568, MH61009, and MH55284.

References

1. World Health Organization: *International Statistical Classification of Diseases and Related Health Problems, 1989 Revision* Geneva: World Health Organization; 1992.
2. Yeargin-Allsopp M, Rice C, Karapurkar T, Doernberg N, Boyle C, Murphy C: **Prevalence of autism in a US metropolitan area.** *JAMA* 2003, **289**:49-55.
3. Bailey A, Phillips W, Rutter M: **Autism: towards an integration of clinical, genetic, neuropsychological, and neurobiological perspectives.** *J Child Psychol Psychiatry* 1996, **37**:89-126.
4. Cohen DJ, Volkmar FR: *Handbook of Autism and Pervasive Development Disorders* 2nd edition. New York, John Wiley and Sons; 1997.
5. Volkmar FR, Szatmari P, Sparrow SS: **Sex differences in pervasive developmental disorders.** *J Autism Dev Disord* 1993, **23**:579-591.
6. McLennan JD, Lord C, Schopler E: **Sex differences in higher functioning people with autism.** *J Autism Dev Disord* 1993, **23**:217-227.
7. The International Molecular Genetic Study of Autism: **Further characterization of the autism susceptibility locus AUTS1 on chromosome 7q.** *Hum Mol Genet* 2001, **10**:973-82.
8. Bradford Y, Haines J, Hutcheson H, Gardiner M, Braun T, Sheffield V, Cassavant C, Huang W, Wang K, Vieland V, Folstein S, Santangelo S, Piven J: **Incorporating language phenotypes strengthens evidence of linkage to autism.** *Am J Med Genet* 2001, **105**:539-547.
9. 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, Leboyer M, van Maldergerme L: **Genome-wide scan for autism susceptibility genes. Paris Autism Research International Sibpair Study [published erratum appears in Hum Mol Genet 1999 Jul;8(7):1353].** *Hum Mol Genet* 1999, **8**:805-812.
10. Ashley-Koch A, Wolpert CM, Menold MM, Zaeem L, Basu S, Donnelly SL, Ravan SA, Powell CM, Qumsiyeh MB, Aylsworth AS, Vance JM, Gilbert JR, Wright HH, Abramson RK, DeLong GR, Cuccaro ML, Pericak-Vance MA: **Genetic studies of autistic disorder and chromosome 7.** *Genomics* 1999, **61**:227-236.
11. Hutcheson HB, Bradford Y, Folstein SE, Gardiner MB, Santangelo S, Sutcliffe JS, Haines JL: **Defining the Autism Minimum Candidate Gene Region on Chromosome 7.** *Am J Med Genet* 2003, **117B**:90-96.
12. Lustig M, Erskine L, Mason CA, Grumet M, Sakurai T: **Nr-CAM expression in the developing mouse nervous system: ventral midline structures, specific fiber tracts, and neuropilar regions.** *J Comp Neurol* 2001, **434**:13-28.
13. Lustig M, Sakurai T, Grumet M: **Nr-CAM promotes neurite outgrowth from peripheral ganglia by a mechanism involving axonin-I as a neuronal receptor.** *Dev Biol* 1999, **209**:340-351.
14. Battye R, Stevens A, Perry RL, Jacobs JR: **Repellent signaling by Slit requires the leucine-rich repeats.** *J Neurosci* 2001, **21**:4290-4298.
15. Liang Y, Annan RS, Carr SA, Popp S, Mevissen M, Margolis RK, Margolis RU: **Mammalian homologues of the Drosophila slit protein are ligands of the heparan sulfate proteoglycan glypican-I in brain.** *J Biol Chem* 1999, **274**:17885-17892.
16. Taniguchi H, Tohyama M, Takagi T: **Cloning and expression of a novel gene for a protein with leucine-rich repeats in the developing mouse nervous system.** *Brain Res Mol Brain Res* 1996, **36**:45-52.
17. Ishii N, Wanaka A, Tohyama M: **Increased expression of NLRR-3 mRNA after cortical brain injury in mouse.** *Brain Res Mol Brain Res* 1996, **40**:148-152.
18. Fukamachi K, Matsuoka Y, Ohno H, Hamaguchi T, Tsuda H: **Neuronal leucine-rich repeat protein-3 amplifies MAPK activation by epidermal growth factor through a carboxyl-terminal region containing endocytosis motifs.** *J Biol Chem* 2002, **277**:43549-43552.
19. Nagase T, Ishikawa K, Suyama M, Kikuno R, Miyajima N, Tanaka A, Kotani H, Nomura N, Ohara O: **Prediction of the coding sequences of unidentified human genes. XI. The complete sequences of 100 new cDNA clones from brain which code for large proteins in vitro.** *DNA Res* 1998, **5**:277-286.
20. Hasegawa H, Kiyokawa E, Tanaka S, Nagashima K, Gotoh N, Shibuya M, Kurata T, Matsuda M: **DOCK180, a major CRK-binding protein, alters cell morphology upon translocation to the cell membrane.** *Mol Cell Biol* 1996, **16**:1770-1776.
21. Powell SK, Kleinman HK: **Neuronal laminins and their cellular receptors.** *Int J Biochem Cell Biol* 1997, **29**:401-414.

22. Hopker VH, Shewan D, Tessier-Lavigne M, Poo M, Holt C: **Growth-cone attraction to netrin-1 is converted to repulsion by laminin-1.** *Nature* 1999, **401**:69-73.
23. Lord C, Rutter M, LeCouteur A: **Autism diagnostic interview-revised: A revised version of a diagnostic interview for carers of individuals with possible pervasive developmental disorders.** *J Autism Dev Disord* 1994, **24**:659-685.
24. Lord C, Rutter M, Goode S, Heemsbergen J, Jordan H, Mawhood L, Schopler E: **Autism diagnostic observation schedule. A standardized observation of communicative and social behavior.** *J Autism Dev Disord* 1989, **19**:185-212.
25. **The Celera SNP database** [<http://cds.celera.com/cds>]
26. **The National Center for Biotechnology Information** [<http://www.ncbi.nlm.nih.gov/>]
27. **The Ensembl Genome Server** [<http://www.ensembl.org/>]
28. **The National Center for Biotechnology Information SNP database** [<http://www.ncbi.nlm.nih.gov/SNP/>]
29. **Primer 3 Software Site** [http://frodo.wi.mit.edu/cgi-bin/primer3/primer3_www.cgi]
30. Kuklin A, Munson K, Gjerde D, Haefele R, Taylor P: **Detection of single-nucleotide polymorphisms with the WAVE DNA fragment analysis system.** *Genet Test* 1997, **1**:201-206.
31. Sobel E, Lange K: **Descent graphs in pedigree analysis: applications to haplotyping, location scores, and marker-sharing statistics.** *Am J Hum Genet* 1996, **58**:1323-1337.
32. Martin ER, Monks SA, Warren LL, Kaplan NL: **A test for linkage and association in general pedigrees: the pedigree disequilibrium test.** *Am J Hum Genet* 2000, **67**:146-154.
33. Spielman RS, McGinnis RE, Ewens WJ: **Transmission test for linkage disequilibrium: the insulin gene region and insulin-dependent diabetes mellitus (IDDM).** *Am J Hum Genet* 1993, **52**:506-516.
34. Martin ER, Bass MP, Hauser ER: **Correlation between linkage and association tests in families.** *Am J Hum Genet* 2001, **69**:511.
35. Martin ER, Lai EH, Gilbert JR, Rogala AR, Afshari AJ, Riley J, Finch KL, Stevens JF, Livak KJ, Slotterbeck BD, Slifer SH, Warren LL, Conneally PM, Schmechel DE, Purvis I, Pericak-Vance MA, Roses AD, Vance JM: **SNPping away at complex diseases: analysis of single-nucleotide polymorphisms around APOE in Alzheimer disease.** *Am J Hum Genet* 2000, **67**:383-394.
36. Clayton D: **A Generalization of the Transmission/Disequilibrium Test for Uncertain-Haplotype Transmission.** *Am J Hum Genet* 1999, **65**:1170-1177.
37. Abecasis GR, Cookson WO: **GOLD – graphical overview of linkage disequilibrium.** *Bioinformatics* 2000, **16**:182-183.
38. Lewontin RC: **The interaction of selection and linkage.I. General considerations; heterotic models.** *Genetics* 1964, **49**:49-67.
39. Abecasis GR, Noguchi E, Heinzmann A, Traherne JA, Bhattacharyya S, Leaves NI, Anderson GG, Zhang Y, Lench NJ, Carey A, Cardon LR, Moffatt MF, Cookson WO: **Extent and distribution of linkage disequilibrium in three genomic regions.** *Am J Hum Genet* 2001, **68**:191-197.
40. Reich DE, Cargill M, Bolk S, Ireland J, Sabeti PC, Richter DJ, Lavery T, Kouyoumjian R, Farhadian SF, Ward R, Lander ES: **Linkage disequilibrium in the human genome.** *Nature* 2001, **411**:199-204.
41. Stephens JC, Schneider JA, Tanguay DA, Choi J, Acharya T, Stanley SE, Jiang R, Messer CJ, Chew A, Han JH, Duan J, Carr JL, Lee MS, Koshy B, Kumar AM, Zhang G, Newell WR, Windemuth A, Xu C, Kalbfleisch TS, Shaner SL, Arnold K, Schulz V, Drysdale CM, Nandabalan K, Judson RS, Ruano G, Vovis GF: **Haplotype variation and linkage disequilibrium in 313 human genes.** *Science* 2001, **293**:489-493.
42. Daly MJ, Rioux JD, Schaffner SF, Hudson TJ, Lander ES: **High-resolution haplotype structure in the human genome.** *Nat Genet* 2001, **29**:229-232.
43. Jeffreys AJ, Kauppi L, Neumann R: **Intensely punctate meiotic recombination in the class II region of the major histocompatibility complex.** *Nat Genet* 2001, **29**:217-222.

Pre-publication history

The pre-publication history for this paper can be accessed here:

<http://www.biomedcentral.com/1471-2350/5/12/prepub>

Publish with **BioMed Central** and every scientist can read your work free of charge

"BioMed Central will be the most significant development for disseminating the results of biomedical research in our lifetime."

Sir Paul Nurse, Cancer Research UK

Your research papers will be:

- available free of charge to the entire biomedical community
- peer reviewed and published immediately upon acceptance
- cited in PubMed and archived on PubMed Central
- yours — you keep the copyright

Submit your manuscript here:
http://www.biomedcentral.com/info/publishing_adv.asp

