

UC Irvine

UC Irvine Previously Published Works

Title

Evaluation of the chromosome 2q37.3 gene CENTG2 as an autism susceptibility gene

Permalink

<https://escholarship.org/uc/item/9m33n261>

Journal

American Journal of Medical Genetics Part B Neuropsychiatric Genetics, 136B(1)

ISSN

1552-4841

Authors

Wassink, Thomas H

Piven, Joseph

Vieland, Veronica J

et al.

Publication Date

2005-07-05

DOI

10.1002/ajmg.b.30180

Copyright Information

This work is made available under the terms of a Creative Commons Attribution License, available at <https://creativecommons.org/licenses/by/4.0/>

Peer reviewed

Evaluation of the Chromosome 2q37.3 Gene *CENTG2* as an Autism Susceptibility Gene

Thomas H. Wassink,^{1*} Joseph Piven,² Veronica J. Vieland,^{1,3,4} Laura Jenkins,¹ Rebecca Frantz,¹ Christopher W. Bartlett,³ Rhinda Goedken,³ Deb Childress,² M. Anne Spence,⁵ Moyra Smith,⁵ and Val C. Sheffield⁶

¹Department of Psychiatry, University of Iowa Carver College of Medicine, Iowa City, Iowa

²Neurodevelopmental Disorders Research Center and Department of Psychiatry, University of North Carolina, Chapel Hill, North Carolina

³Center for Statistical Genetics Research, University of Iowa College of Public Health and University of Iowa Carver College of Medicine, Iowa City, Iowa

⁴Program in Public Health Genetics, University of Iowa College of Public Health, Iowa City, Iowa

⁵Department of Pediatrics, University of California, Irvine, California

⁶Department of Pediatrics and the Howard Hughes Medical Institute, University of Iowa Carver College of Medicine, Iowa City, Iowa

Autism is a highly heritable neurodevelopmental syndrome with a complex genetic etiology for which no disease genes have yet been definitively identified. We ascertained three subjects with autism spectrum disorders and chromosome 2q37.3 terminal deletions, and refined the deletion breakpoint regions using polymorphism mapping and fluorescence in situ hybridization (FISH) probes. We then genotyped polymorphic markers downstream from the breakpoint region in a sample of autism affected sibling pair families. Both the chromosomal breakpoints and linkage analyses focused our attention on the gene centaurin gamma-2 (*CENTG2*), an attractive candidate gene based also on its function and pattern of expression. We therefore assessed *CENTG2* for its involvement in autism by (1) screening its exons for variants in 199 autistic and 160 non-autistic individuals, and (2) genotyping and assessing intra-genic polymorphisms for linkage and linkage disequilibrium (LD). The exon screen revealed a Ser → Gly substitution in one proband, an Arg → Gly substitution in another, and a number of additional variants unique to the autism families. No unique variants were found in the control subjects. The genotyping produced strong evidence for linkage from two intronic polymorphisms, with a maximum two-point HLOD value of 3.96 and a posterior probability of linkage (PPL) of 51%. These results were contradicted, however, by substantially weaker evidence for linkage from multi-point analyses and by no evidence of LD. We conclude, therefore, that 2q37.3 continues to be a region of interest for autism susceptibility, and that *CENTG2* is an intriguing

candidate gene that merits further scrutiny for its role in autism. © 2005 Wiley-Liss, Inc.

KEY WORDS: linkage; linkage disequilibrium; autistic disorder; chromosomal abnormalities

INTRODUCTION

Autism (MIM 209850) is a neurodevelopmental syndrome consisting of deficits in social interaction and communication, specific ritualistic-repetitive behaviors, and a characteristic course [Rutter et al., 1993]. Onset is typically first noted at the age when relevant behaviors such as speech and complex social interactions become observable, and symptoms generally continue throughout life [Gillberg, 1993]. Family and twin studies have demonstrated that the predisposition to develop autism is largely genetically determined, with a sibling relative risk between 50 and 100 and a heritability estimate of at least 90% [Szatmari et al., 1998]. These same studies also show that the mode of inheritance for autism is likely to be complex, characterized by multiple genetic variants that confer susceptibility by interacting in as yet unknown patterns [Pickles et al., 1995; Risch et al., 1999; Szatmari, 1999]. While the identification of these variants has thus far largely been thwarted by such complexity, an array of increasingly sophisticated approaches are being brought to bear on the search.

The investigations presented here integrate cytogenetic, statistical, molecular, and clinical methodologies in a focused examination of chromosome 2q37.3. Our attention was drawn to this region for two reasons. First, we identified terminal 2q37 deletions in three autistic individuals, a finding consonant with a growing number of similar cases [Burd et al., 1988; Gorski et al., 1989; Stein et al., 1992; Fisher et al., 1994; Conrad et al., 1995; Phelan et al., 1995; Friedman et al., 1997; Viot-Szoboszlai et al., 1998; Ghaziuddin and Burmeister, 1999; Smith et al., 2001; Wolff et al., 2002; Daniel et al., 2003; Casas et al., 2004; Lukusa et al., 2004]. Second, in follow-up analyses of our initial genome-wide screen [CLSA, 1999], we identified a modest linkage signal close to the 2q telomere in a subgroup of families with severe language impairment (maximum multipoint heterogeneity LOD score of 1.13 at 252 cM; unpublished data).

We therefore, as described in this report, proceeded to investigate these initial findings more intensively. Since few of the reported 2q37.3 chromosomal abnormalities have been delineated with precision, we began by demarcating the deletion breakpoints in our cases using polymorphism and fluorescence in situ hybridization (FISH) mapping. We also genotyped polymorphic markers along 2q37 in a large sample

This article contains supplementary material, which may be viewed at the American Journal of Medical Genetics website at <http://www.interscience.wiley.com/jpages/1552-4841/suppmat/index.html>.

*Correspondence to: Thomas H. Wassink, M.D., University of Iowa Carver College of Medicine, Psychiatry Research/MEB, Iowa City, Iowa 52242. E-mail: thomas-wassink@uiowa.edu

Received 14 October 2004; Accepted 31 January 2005

DOI 10.1002/ajmg.b.30180

of affected sibling pair families. Results from both of these approaches focused our efforts on the gene centaurin gamma-2 (*CENTG2*), which codes for a cytoskeletal protein involved in multiple neuronal processes in the developing and mature brain [Jackson et al., 2000]. We screened the coding sequence of *CENTG2* for variants and genotyped numerous intragenic polymorphisms that were tested for linkage and linkage disequilibrium (LD). While the chromosomal deletion cases support the growing interest in 2q37.3, the results of the *CENTG2* screen were mixed with both supportive and confuting evidence regarding its potential contribution to autism susceptibility.

SUBJECTS AND METHODS

Recruitment and Assessment of Chromosome Deletion Cases

Deletion probands 1 and 2 were recruited through an ongoing study of chromosomal abnormalities in autism. Subjects are recruited from throughout the United States at a clinical and data collection center at the University of North Carolina (UNC), Chapel Hill. Subjects must have a DSM-IV diagnosis of autism, Asperger's syndrome, or pervasive developmental disorder not otherwise specified (PDD NOS), and are excluded if they have associated medical conditions (e.g., tuberous sclerosis or fragile X syndrome), evidence of gross central nervous system injury, or have suffered severe perinatal events. All accessible medical records are acquired, and blood is drawn that is used for cytogenetic tests and DNA extraction. If a chromosomal abnormality is detected, we attempt to perform the more detailed assessment described in the affected sibling pair ascertainment below. Deletion proband 3 was referred to us by collaborators at the University of California, Irvine (UCI). The ascertainment and evaluation of this subject has been described in detail in a previous report [Smith et al., 2001].

Chromosome Deletion Breakpoint Delineation

For all three deletion probands, the deletions had been identified clinically by karyotyping, and additional molecular delineation had been performed for proband 3. The focus of the cytogenetic work for this study, therefore, was to refine the delineation of the breakpoint intervals using polymorphic markers and FISH probes. We first selected markers within 2q37 from the Marshfield [Broman et al., 1998] and NCBI UniSTS genetic maps. These markers were genotyped in the probands and their parents, and the results examined for evidence of loss of heterozygosity (LOH). Parental DNA was not available for proband 2, so polymorphisms from this individual were simply examined for the presence of two alleles. When known markers were exhausted, we used the utility tandem repeats finder [Benson, 1999] to identify additional tandem repeats for PCR. Repeats found to be polymorphic were used to assess the deletions, with this process continuing until the breakpoint gaps were maximally narrowed.

For the FISH probe analysis, bacterial artificial chromosome (BAC) clones across the region of interest were identified through NCBI physical map. BAC clones thus identified were ordered from Research Genetics, Huntsville, AL. Individual clones were cultured, and DNA was isolated and used for polymerase chain reaction (PCR) to confirm that each clone was positive for the corresponding gene sequences. Purified BAC clone DNA was labeled with spectrum green or spectrum orange using the Vysis labeling system. Labeled BAC clone DNA was hybridized to human Cot 1 DNA to block repetitive sequences. Thereafter, it was hybridized to metaphase and interphase nuclei on prepared slides. Following the wash-

ing of slides according to standard protocols, slides were examined using direct fluorescence microscopy at appropriate wavelengths.

Patient Ascertainment and Sample Used to Screen *CENTG2* and 2q37.3

Examination of 2q37.3 markers for linkage and LD was performed in a sample of affected sibling pair families ascertained under a previously described protocol [CLSA, 1999]. Briefly, families were recruited from the Midwest and Mid-Atlantic United States through data collection sites at the University of Iowa (UIA) and the University of North Carolina (UNC). All affected individuals were at least 3 years old and were required to meet the ADI-R [Lord et al., 1994] and the ADOS [Lord et al., 1989] or ADOS-G [Lord et al., 2000] criteria for autism. Affecteds were excluded if they had fragile X syndrome (based on fragile X DNA testing) or any neurological or medical condition thought to be associated with autism. DNA was extracted from whole blood using standard techniques, and this sample provided 117 families with at least two affected siblings. These families, as per the CLSA protocol, were split into two sets. The 58 families in Set 1 have been analyzed previously in a genome-wide linkage screen [CLSA, 1999]. That study reported on a total of 75 families, 15 of which are no longer available. Genotyping in the 59 Set 2 families has not been previously reported.

Exon screening was performed using DNA from 199 independent autistic individuals from the Iowa sample and a comparison group of 160 unrelated Iowans of similar ethnicity who did not have autism. DNA for 169 of the autistic individuals came from the UIA and UNC sites, while DNA for an additional 30 affecteds was graciously provided by Dr. Susan Folstein from Tufts University-New England Medical Center, who had ascertained these individuals using the CLSA protocol described above. All individuals (or, when appropriate, their guardians) provided written, informed consent for participation in this study.

CENTG2 Coding Sequence Screening

The 4112 bp *CENTG2* cDNA contains 2415 bp of coding sequence across 17 exons, with two alternatively spliced exons (11B and 12B in Fig. 2) present in shorter cDNA. Exons, including flanking splice-junction sequences, were screened with a total of 27 amplicons (primers available upon request), none of which exceeded 250 bp in length, using single strand conformational polymorphism (SSCP) analysis. All amplicons were screened in the entire autism and comparison groups.

PCR was performed with 20 ng of genomic DNA amplified in a reaction mixture containing 1.0 L of PCR buffer (100 μ M Tris-HCl (pH8.8), 500 μ M KCl, 15 μ M MgCl₂, 0.01% gelatin (w/v)), 200 μ M each of dATP, dCTP, dGTP, and dTTP, 2.5 pM of each primer, and 0.05 U of Taq DNA polymerase, increased to a final volume of 10.0 μ L with water. Samples were initially denatured at 94°C for 3 min, followed by 35 cycles of 94°C for 30 sec, 50–62°C for 30 sec (see Table I), and 72°C for 30 sec. PCR products were electrophoresed on 6% non-denaturing polyacrylamide gels at 20 W for approximately 3 hr at room temperature while being cooled by a fan. The gels were then silver stained using standard protocols [Bassam et al., 1991]. Amplicons containing SSCP shifts were forward and reverse sequenced on an Applied Biosystems (ABI, Foster City, CA) model 377 automated sequencer using dye terminator chemistry. Sequence data were analyzed using Sequencher 4.0 (Gene Codes, Ann Arbor, MI).

The only exception to this SSCP-based screening was for exon 1. This section of *CENTG2* was highly homologous to *CENTG2* pseudogenes, hindering the development of unique amplicons small enough for sensitive SSCP assays. We there-

TABLE I. Deletion Breakpoint Delineation

BAC FISH			Deletion proband 1		Deletion proband 2		Deletion proband 3	
Marker	Probes ^a	MB ^b	Marker	FISH	Marker	FISH	Marker	FISH
D2S1363		226.86	■ ■	—	■ ■	—	■ ■	—
D2S1349		228.01	—	—	hmz	—	—	—
ATA20H03		228.29	—	—	■ ■	—	—	—
D2S427		232.03	■ ■	—	■ ■	—	■ ■	—
D2S1279		233.78	■ ■	—	■ ■	—	—	—
D2S1357		235.93	—	—	■ ■	—	—	—
AC010148.AAT	RP11-367B19	236.02	—	—	—	■	■ ■	■ ■
D2S336		236.06	■ ■	—	—	—	hmz	—
D2S407		236.31	—	—	hmz	—	hmz	—
D2S2973		236.32	—	—	hmz	—	hmz	—
AC092576.AAAT	RP11-7N16	236.53	—	—	—	■	hmz	■
AC093785.AT	RP11-212M1	236.57	—	—	—	—	■	■
AC093785.AT2		236.60	—	—	—	—	■	—
AC012305.ATT		236.61	■ ■	—	—	■	—	■
AC012305.ATTT		236.63	■ ■	—	—	■	—	—
AC012305.CTTT	RP11-84G18	236.68	■ ■	—	—	—	hmz	—
AC012305.GT1		236.72	■ ■	—	—	—	—	hmz
AC012305.GT2		236.74	■ ■	—	—	—	hmz	—
AC012305.CA2		236.77	■	—	—	—	■	—
AC073989.CA	RP11-473L20	236.88	■	—	—	■	—	■
AC073989.GT		236.90	■	—	—	—	—	hmz
D2S2202		236.92	■	—	—	—	■	—
D2S1397	RP11-485M18	237.08	■	—	—	—	■	■
D2S1359		238.16	■	—	—	hmz	—	—

■ ■, heterozygous; ■, deletion; hmz, homozygous but uninformative; —, not typed.

^aShading shows approximate extent of region covered by BAC clones to which FISH probes were hybridized.

^bPhysical map locations based on UCSC July, 2003 Assembly. Shading shows proximal portion of *CENTG2* covered by deletion screening.

fore forward and reverse sequenced a unique 324 bp fragment spanning this exon in 48 affected individuals and 48 control subjects.

Identification and Genotyping of Intra-Genic and Telomeric Polymorphisms

Polymorphisms within and downstream from *CENTG2* were identified by searching SNP, STS, and other polymorphism databases, and by using TRF. PCR was performed as described above, and PCR products were electrophoresed on 6% denaturing polyacrylamide gels at 60 W for ~2 hr. Gels were visualized using silver staining and read blindly with respect to sample status by two independent raters with discrepancies resolved by re-genotyping. SNPs were genotyped using TaqMan[®] Pre-Designed SNP Genotyping Assays and Custom TaqMan SNP Genotyping Assays, run on a StrataGene Mx3000P qPCR thermocycler.

Analysis of Polymorphisms for Linkage and LD

Two-point and multi-point heterogeneity LOD (HLOD) scores were calculated in the same manner as in our genome-wide linkage screen, using the MLINK and HOMOG sub-routines of the LINKAGE package [Lathrop et al., 1985] under one simple recessive model (disease allele frequency = 0.10, penetrance = 50%) and one simple dominant model (disease allele frequency of 0.04, penetrance = 40%). Multi-point analyses were calculated using all available markers. Allele frequencies and marker heterozygosities were determined from parental genotypes. Normal allele penetrance was set to zero (no phenocopies). We also performed two-point and multi-point analysis of the markers using the posterior probability of linkage (PPL) statistic, which integrates out parameters of the trait model, thereby providing a largely model-free analysis

[Vieland, 1998; Logue et al., 2003]. The PPL is calculated on the straightforward probability scale (0,1), with 1 indicating certain linkage and 0 indicating no linkage. The prior probability of linkage is set to 0.02, so that a PPL > 0.02 indicates some degree of evidence in favor of linkage. Of import to the language-based subgroup analyses described below, the PPL enables results from multiple samples or subgroups of families to be integrated en bloc sequentially. This protects homogeneity in those groups from the potentially diluting effect of simply merging all families into one large sample [Vieland et al., 2001]. Lastly, Genehunter 2.1 [Kruglyak et al., 1996] was used to calculate the TDT statistic for all polymorphisms using one randomly selected affected sibling from each family.

Subgroup Analyses

Subsequent to our genome-wide screen, we attempted to address the apparent heterogeneity of autism by splitting the ASP families into two groups based on their language characteristics. Families were classified as *language impaired* if neither autistic sibling had developed phrase speech by 36 months, and otherwise as *language normal* (the average normal onset of phrase speech is 18 months). Analyses of the two resultant subgroups at our two strongest loci (7q and 13q) revealed that the predominance of the linkage signals arose from the *language impaired* families [CLSA, 2001], a pattern that has since been replicated [Buxbaum et al., 2001; Alarcon et al., 2002; Shao et al., 2002]. In addition, we had assessed, by direct questioning, whether the ASP parents had a history of developmental language difficulties, including delayed onset of speech, trouble learning to read, or persistent trouble with spelling. Parents who experienced any of these difficulties were classified as *language impaired* (affected) and otherwise as *language normal* (unaffected). Incorporation of this data into the analyses further strengthened the evidence for linkage at

both loci. We thus provided support for the hypotheses that (1) language-based classification defines more homogeneous subgroups of autistic families, and (2) parental language phenotype is in some way genetically related to autism. In light of this, we analyzed the markers genotyped for this study in a similar manner, examining the evidence for linkage and LD in the pooled Set 1 and Set 2 samples and in the two language based subgroups from each set (for a total of four groups), and incorporating parental language phenotypes.

RESULTS

Clinical Assessments of Deletion Proband

Detailed clinical assessments of the deletion probands are provided in online supplementary material. Briefly, probands 1 and 3 met ADI-R and ADOS-G criteria for autistic disorder; proband 2 has not been assessed with these instruments, but has been given a DSM-IV diagnosis of PDT, NOS. Proband 1 and 2 have mild mental retardation, while proband 3 has a normal IQ.

Deletion Breakpoint Delineation in the Deletion Proband

For proband 1, previous karyotyping of this child had revealed a single terminal deletion, 46,XY,del(2)(q37.1) which was not found in karyotypes of the parents. Subsequent genotyping in this study of a total of 35 polymorphisms revealed that the maternal chromosome was deleted and led to delineation of the breakpoint to a 33 kb region within intron 1 of *CENTG2* (Table I and Fig. 2).

Karyotyping of proband 2 had revealed a de novo 2q37 telomeric deletion: 46,XY,del(2q)(q37.1q37.3). While parental DNA was not available for this individual, combined FISH and marker data narrowed the deletion breakpoint to a 100–150 kb interval ~500 kb upstream from *CENTG2* (Table I and Fig. 2). The deletion also appeared to be terminal, as no markers or

FISH probes demonstrated two chromosomes telomeric to the breakpoint.

For proband 3, previous karyotypes of the patient and parents had detected a de novo 2q37 deletion with the breakpoint interval narrowed to a span of ~5 MB using BAC FISH probes and polymorphic markers (Smith ref). The interval was further narrowed in this study using 45 polymorphic markers and additional FISH probes. The markers reduced the breakpoint interval to a 600 kb region approximately 200 kb upstream from the proband 1 breakpoint interval. The additional FISH probes confirmed the marker results, narrowing the interval to ~500 kb and placing the breakpoint ~200–500 kb upstream from *CENTG2* (Table I and Fig. 2).

CENTG2 Coding Sequence Screen

Results of the coding sequence screen are shown in Table II. Sixteen variants were identified: four non-synonymous and six synonymous coding sequence variants, and six intronic variants. Of the non-synonymous variants, S82G, R745G, and P801T (see Fig. 1) were found in one autism family each. The probands containing S82G and R745G were from trios, while P801T was found in both the affected and unaffected sibs from a discordant pair. S82 and R745 are both highly conserved amino acids within, respectively, ras and ankyrin functional domains; P801, conversely, is not in a known functional domain. The fourth non-synonymous variant, I618V (rs2034648), was common in both the autism and control populations and, while itself not conserved, is within the ArfGap functional domain.

Of the synonymous coding sequence variants, four were common in both the autism and control samples, while two—T1222C and C2686T—were found in one autism family each; no variants were found that were unique to the control subjects. T1222C was found in a trio proband, while the T allele of C2686T was transmitted from a mother to both of her affected children. Similarly, the intron 12 AG deletion was found in one ASP family in which it was transmitted from the father to both

TABLE II. Results of *CENTG2* Coding Sequence Screen

Coding sequence variants					
Exon	cDNA variant	Amino acid effect	Functional domain	dbSNP	Comment ^a
1	C703T	None, N43N	—	rs8178993	Common
2	A781G	None, P210P	—	ss9807226	Common
3	A818G	S83G	Ras	—	One trio proband
6	T1222C	None, N217N	—	—	One trio proband
8	C1451A	None, R294R	—	—	ASP father and one sib
8	C1528T	None, T319T	—	rs2292708	Common
15	A2423G	I618V	ArfGap	rs2034648	Common
16	C2686T	None, D705D	—	—	One ASP family (mother and both affecteds)
17	A2805G	R745G	Ankyrin	—	One trio proband
17	C2923T	None, P784P	—	ss9807225	Common
17	C2972A	P801T	—	—	One trio proband and unaffected sib
Intronic variants					
Intron	BAC	Nucleotide within BAC	Variant	dbSNP	Comment ^a
1	AC012305	81238	A/C	rs3738998	Common
9	AC074874	107882	T/A	ss9807227	Common
10	AC079400	50563	T indel	ss9807228	common
12	AC079400	110517	AG deln	—	One ASP family (father and both affecteds)
12	AC019047	45468	T indel	rs2123511	Common
16	AC019047	151211	C/T	rs2278884	Common

^a“Common” indicates variants with minor allele frequencies > 10%.

		↓S82G	↓I618V	↓R745G	↓P801T
CENTG2	HS	VGNLASGKSAL . . .	LDDWPIELIK . . .	VDTVARDAHG . . .	NSSGRVPTII
CENTG2	MM	VGNLASGKSAL . . .	LDDWPMEELIK . . .	VDVMARDAHG . . .	NSSGRAPSVI
CENTG2	XL	VGNLASGKSAL . . .	LDDWPPEELIK . . .	VDVMARDFHG . . .	GGSGLMPTLI
CENTG1	HS	LGDARSGKSSL . . .	LDDWPRELTL . . .	ADVAARDAQG	
CENTG3	HS	VGNLSGKSAL . . .	LDDWPPELLA . . .	ADVAARDAQG	
CENTG1	MM	LGDVRSKGSSL . . .	LDDWPRELTL . . .	VDVRSRDARG	
CENTG3	MM	VGNLSGKSAL . . .	LDDWPPELLA . . .	VDVRSRDARG	
domain		Ras	ArfGap	ArfGap	none

Fig. 1. Non-synonymous *CENTG2* coding sequence variants. S82G and R745G were found in one autism trio proband each. I618V was common in both autistic individuals and psychiatrically normal controls. P801T was found in both of a discordant sib pair. Red letters indicate conserved amino acids in known functional domains of *CENTG2*. [Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

affected children. Phenotypic information available on this father revealed multiple characteristics similar to the core features of autism, including a rigid personality style and deficits in social interactions and friendships, while the mother was phenotypically normal.

Analysis of Polymorphisms for Linkage and LD

Sixteen polymorphisms were identified, genotyped, and analyzed for evidence of linkage and LD (Table III and Fig. 2; primers available upon request). Two of these, D2S2968 and D2S427, had been genotyped in half the families for our initial linkage screen, and were thus genotyped in the remainder of the sample for this study.

Two-point linkage analyses revealed, under the recessive model, strong evidence in favor of linkage at two markers: D2S1397 and CENTG2.32112 (Table IV and Fig 3). For both markers, the evidence for linkage arose predominantly from the language impaired subsets of families. For D2S1397, the evidence for linkage arose primarily from the Set 1 families, while for CENTG2.32112 the evidence arose primarily from the Set 2 families. Six of the remaining markers within *CENTG2* produced LOD scores >1.0 and/or PPLs greater than 0.02, suggesting some evidence in favor of linkage.

In contrast to the two-point results, however, multi-point analyses provided little support for the presence of linkage. Three-point, four-point, and all-marker analyses were universally low across the entire region of interest (Fig. 3). In addition, the single-sibling TDT analyses failed to identify evidence for LD at any of these markers, and no evidence of

allelic instability was detected at either of the CAG trinucleotide repeats.

DISCUSSION

The genetic and phenotypic complexity that typifies psychiatric disorders necessitates integration of multiple approaches in the attempt to identify susceptibility genes. In the current study, the delineation of breakpoint intervals for three terminal deletions of chromosome 2q37.3 in individuals with autism or autism spectrum disorders provides support for this as a region that is likely to harbor an autism susceptibility gene. The chromosomal studies were then complemented by subsequent linkage analyses of 2q37 that took into account phenotypic heterogeneity. These analyses focused our attention on the gene *CENTG2*, which we therefore examined by screening its coding sequence for genetic variants and by examining intra-genic polymorphic markers for linkage and LD. These studies combined to produce a mixed picture of the relationship between *CENTG2* and autism susceptibility.

The chromosome 2q37.3 terminal deletions that we describe join a growing body of similar reports in the literature (for a recent review of 2q terminal deletions, see [Casas et al., 2004]). While the most common syndrome associated with such abnormalities is albright's hereditary osteodystrophy (AHO) [Phelan et al., 1995; Bijlsma et al., 1999; Casas et al., 2004], the presence of autism or autistic traits is also common [Burd et al., 1988; Gorski et al., 1989; Stein et al., 1992; Fisher et al., 1994; Conrad et al., 1995; Phelan et al., 1995; Friedman et al., 1997; Viot-Szoboszlai et al., 1998; Ghaziuddin and Burmeister, 1999; Smith et al., 2001; Wolff et al., 2002; Daniel et al., 2003; Casas et al., 2004; Lukusa et al., 2004]. The specific chromosomal sub-band locations of the breakpoints do not appear to be predictably associated with autistic features [Casas et al., 2004], though only two such cases pursued breakpoint delineation beyond karyotyping [Giardino et al., 2001; Daniel et al., 2003], and both demonstrated breakpoint boundaries in close proximity to those presented in this study.

The first described a child with multiple medical problems, abnormal facies, no AHO phenotype, moderate developmental delay, and multiple autistic features [Daniel et al., 2003]. Genotyping of polymorphic markers in the proband and parents showed that the first deleted marker was D2S2202, which is in intron 3 of *CENTG2* (Fig. 2). A second report of interest described a family with multiple still births and a child with dysmorphic facial features, multiple physical anomalies, and mild MR [Giardino et al., 2001]. Karyotyping of this child revealed a subtelomeric translocation t(2;16)(q37;q24), which FISH studies refined to 46,XY,der(2)t(2;16)(q37.3;q24.3), producing partial monosomy for distal 2q37. The mother, maternal grandfather, and a subsequent conceptus were found to have balanced t(2;16) translocations. The affected child was quite young at the time of the report and appears not to have

TABLE III. 2q37.3 Polymorphisms That Were Genotyped

Polymorphism	Type	Heterozygosity ^a	MB
D2S427	Micro	0.79	232.03
CAGL234	Micro	0.65	236.71
CENTG2.32112	Mini	0.30	236.72
CENTG2.71457	Mini	0.50	236.76
CENTG2.TAAA	Micro	0.64	236.83
CENTG2.214374	Mini	0.47	236.90
CENTG.GTTG	Micro	0.74	236.92
hcv1543812	SNP	0.34	236.95
D2S1397	Micro	0.68	237.08
rs4663627	SNP	0.45	237.08
CENTG2.GGAT	Micro	0.88	237.09
rs765880	SNP	0.48	237.23
D2S2968	Micro	0.65	237.86
2Q112721.GTTT	Micro	0.66	238.59
2Q16776.AAGG	Micro	0.88	239.14
2Q17028.GATA	Micro	0.76	240.18

Primers available in supplementary online material shading indicates markers within *CENTG2*.

^aHeterozygosity was determined using all parental genotypes.

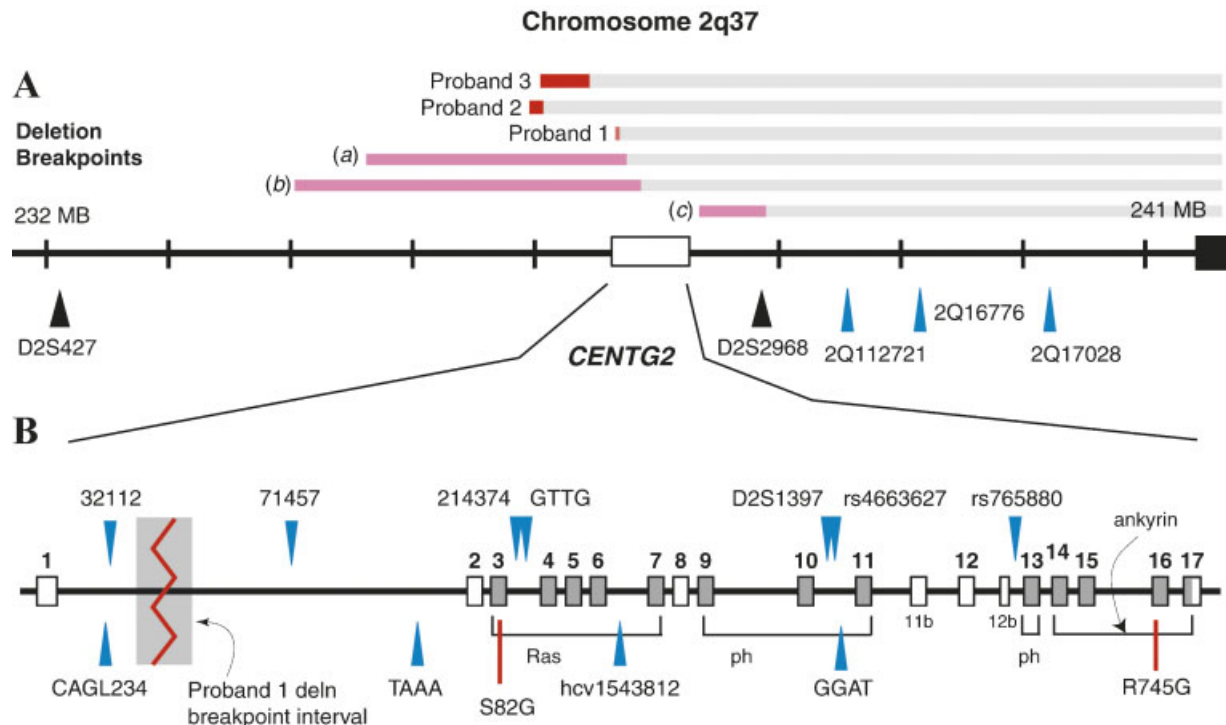


Fig. 2. Chromosome 2q37 and *CENTG2* findings. **A:** Deletion breakpoint intervals for deletion probands 1, 2, and 3. Red bars indicate breakpoint intervals, and gray bars indicate extent of the deletions. Pink bars show deletion breakpoints for three cases from the literature with autistic features, 2q37.3 deletions, and some degree of molecular delineation of the breakpoint intervals; (a) is from [Giardino et al., 2001], (b) is from [Daniel et al., 2003], and (c) is from [Giardino et al., 2003]. Black arrows indicate markers previously genotyped in part; blue arrows indicate markers

developed and genotyped for the current study. **B:** *CENTG2* structure and analyses. Numbered boxes indicate exons; shading in exons indicates functional domains (named underneath). Red straight lines show locations of two non-synonymous coding sequence variants. Blue arrows indicate markers within *CENTG2* that were genotyped for the current study. [Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

been assessed for autism, and no phenotypic information was provided for the mother or grandfather. Additional FISH studies were performed using YAC clones, and the first deleted clone (746H11) contained exon 1 of *CENTG2* (Fig. 2). Thus

these two anomalies appear to interrupt 2q37 near the deletions that we describe.

The deletions described in this and these previous reports cause monosomy for 7–10 MB of genomic material that con-

TABLE IV. Two-Point Linkage Results

		Maximum HLODs in family subgroups ^a						
MARKER	MB Pos	Family Set 1			Family Set 2			PPL
		Language impaired n = 42	Language normal n = 16	Pooled n = 58	Language impaired n = 37	Language normal n = 22	Pooled n = 59	
D2S427	232.03	1.83	0.00	1.02	0.99	0.02	0.77	0.04
CAGL234	236.71	0.97	0.14	0.84	0.28	0.62	0.71	0.02
CENTG2.32112	236.72	1.48	0.08	0.66	3.51	1.15	2.41	0.37
Centg271457	236.76	0.09	0.01	0.06	0.61	0.67	0.69	0.02
CENTG2.TAAA	236.83	0.42	0.00	0.15	0.43	1.14	0.96	0.02
Centg21437	236.90	0.25	0.09	0.20	1.18	0.81	1.87	0.02
CENTG2.GTTG	236.92	1.33	0.01	0.37	0.92	1.66	0.86	0.03
hcv1543812	236.95	1.89	0.02	1.20	0.75	0.08	0.56	0.05
D2S1397	237.08	3.96	0.58	4.03	0.85	0.36	0.18	0.51
rs4663627	237.08	1.69	0.19	1.68	0.58	0.63	0.57	0.04
CENTG2.GGAT	237.09	1.30	0.60	1.79	0.48	0.52	0.86	0.03
rs765880	237.23	1.13	0.05	0.91	0.70	0.08	0.50	0.03
D2S2968	238.01	0.81	0.00	0.15	0.62	0.15	0.35	0.02
2Q112721.GTTT	238.59	0.50	0.39	0.79	0.65	0.54	0.75	0.03
2Q16776.AAGG	239.14	0.50	0.11	0.61	0.72	0.52	0.88	0.02
2Q17028.GATA	240.18	0.96	0.14	0.68	0.46	0.63	1.08	0.02

Shading indicates markers within *CENTG2*.

^aHLODs are reported for a simple recessive model (disease allele frequency = 0.10, penetrance = 50%).

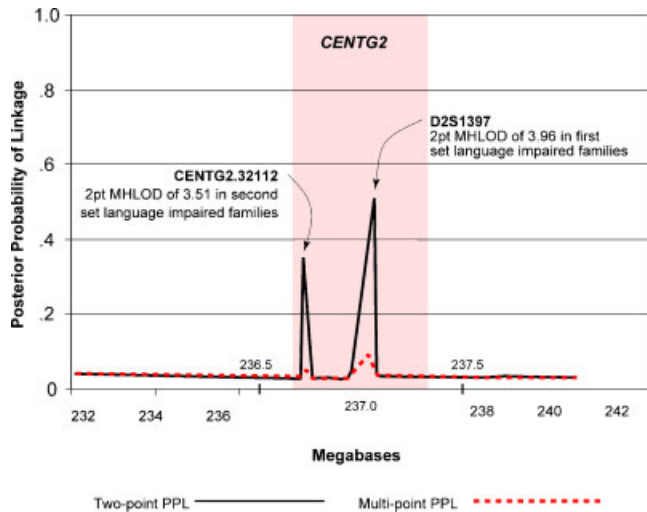


Fig. 3. Linkage analyses of 2q37 and *CENTG2* markers. The black line shows two-point PPLs and the dashed red line shows multi-point PPLs using all markers. The pink shaded area demarcates the boundaries of *CENTG2*. Note that the scale for the area containing *CENTG2* is expanded compared to the rest of the graph. [Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

tains over 100 genes and RNA transcripts. We therefore genotyped markers across the interval to test for linkage and/or LD. In this initial pass, we found minimal evidence for linkage downstream from *CENTG2* that contrasted with strong evidence for linkage within the gene. In addition, *CENTG2* was suggested as a candidate gene because of its function and pattern of expression. Centaurin proteins (at least 10 have been identified in humans) are involved in phosphatidylinositol 3,4,5-trisphosphate (PI 3-phosphate) and ADP ribosylation factor signaling pathways [Jackson et al., 2000]. Both of these pathways influence neuronal processes such as cell growth and survival, differentiation, metabolism, transcription, vesicular trafficking, and cytoskeletal organization [Fruman et al., 1998; Chavrier and Goud, 1999; Leever et al., 1999; Rameh and Cantley, 1999; Roth, 1999; Donaldson and Jackson, 2000; Rodgers and Theibert, 2002]. In line with this, the PI 3-phosphate pathway is activated by neurotrophic, differentiation, and growth factors [Rodgers and Theibert, 2002]. *CENTG2* is expressed in multiple regions of fetal and adult brain, and centaurin proteins are found in neuronal growth cone vesicles and in synapses of developing and mature neurons [Kreutz et al., 1997; Sedehizade et al., 2002; Aggensteiner and Reiser, 2003]. Lastly, the gene coding for *centaurin alpha-1*, a relative of *CENTG2*, is located within the neurofibromatosis (NF) micro-deletion region on chromosome 17q11.2, and NF is associated with an elevated rate of autism.

We therefore genotyped additional markers within *CENTG2* and screened the coding sequence for variations. The genotyping identified two markers that produced strong evidence for linkage based on two-point heterogeneity LOD scores and PPLs. The signal at D2S1397 arose primarily from the language impaired subgroup of the Set 1 families, while the signal at *CENTG2.32112* arose primarily from the language impaired subgroup of the Set 2 families. The majority of the remaining *CENTG2* markers provided at least a modest degree of supportive evidence for linkage. Multi-point analyses, however, produced results that stood in stark contrast to the two-point results with LOD scores and PPLs near zero across the entire interval. These negative results were consistent for three-point, four-point, and all marker analyses.

The discrepancy between the two- and multi-point results is striking and, while we do not know with certainty its cause, a

number of possibilities exist. The first is genotyping error; we rigorously examined our genotypes, however, using multiple methods for detecting and removing errors, and do not believe that errors in our genotyping could account for the degree of disparity between the results. A second possibility is that the two-point scores were false positives that were revealed as such by the greater informativeness of the multi-points. Conversely, however, the two-point linkage signals may be detecting a real linkage that is lost in the multi-points. Multi-point analyses are generally assumed to more accurately assess linkage due to their use of all available genetic information. While true for “perfect” datasets, imperfections that characterize real data such as missing parents, map errors, missing genotypes, and genotyping errors—even when these all occur at low frequencies—have been shown to inordinately degrade multi-point, as compared to two-point, linkage signals [Sullivan et al., 2003]. Thus it may be that the two-point analyses more accurately represent the evidence for linkage in our CLSA sample. The finding of strong linkage in distinct sets of families at two different markers in close proximity to each other seems an unlikely chance occurrence, and subsequent marker-to-marker analyses found no evidence for disequilibrium between D2S1397 and *CENTG2.32112*.

The coding sequence screen of *CENTG2* produced similarly mixed results. We found, in two independent autistic subjects, non-synonymous variants that altered single amino acids—glycine for serine in one case (S82G) and glycine for arginine (R745G) in another—that were both in conserved regions of functional domains (Fig. 1). We also found one non-synonymous variant in both sibs of a discordant pair, though the affected amino acid was not in a known functional domain. Lastly, we found numerous synonymous and intronic variants, a number of which were unique to the autism sample and segregated with the disorder within the sib-pair families; no unique variants were found in the non-autistic control subjects.

These variants could be interpreted to implicate *CENTG2*. S82G and R745G substitute significantly different amino acids in presumably sensitive regions of the protein, and it has become clear that synonymous and intronic variants can alter multiple aspects of gene expression and processing. Conversely, non-synonymous variants, while more likely to be deleterious, often occur without any associated known effect on phenotype. Furthermore, we are limited in our ability to assess the effect of these variants by our typically small pedigrees and by the daunting task of assessing their functional significance.

In conclusion, we present a constellation of evidence derived from investigations assessing the contribution of chromosome 2q37 and the gene *CENTG2* to autism susceptibility. The terminal deletions that we report provide further support for 2q37.3 as a region of interest in autism genetics, and the molecular breakpoint delineations refine the proximal interval boundary. Results from the examination of *CENTG2*, however, are mixed, with some supporting and others refuting the gene’s involvement in autism heritability. Thus, while *CENTG2* merits further scrutiny as an autism disease gene, the numerous other plausible candidate genes in the region must be considered for examination as well.

ACKNOWLEDGMENTS

The authors thank the patients and families that contributed DNA to this study, and all individuals involved in the Collaborative Linkage Study of Autism (CLSA). Thanks also to Dr. Susan Folstein for providing additional DNA samples of individuals with autism. We also are grateful to John Beck, Gretel Beck, and Ann Nalley for their technical assistance. This work was supported by the following grants: U54 MH066418, K02-MH01568, and MH 55284 to Dr. Piven;

5K21-MH01338, K02-MH01432, and MH 52841 to Dr. Vieland; and K08-MH62123-01 and R01-NS43550 to Dr. Wassink. Dr. Sheffield is an Investigator of the Howard Hughes Medical Institute.

REFERENCES

- Aggensteiner M, Reiser G. 2003. Expression of the brain-specific membrane adapter protein p42(IP4)/centaurin alpha, a Ins(1,3,4,5)P(4)/PtdIns(3,4,5)P(3) binding protein, in developing rat brain. *Brain Res Dev Brain Res* 142(1):77–87.
- Alarcon M, Cantor RM, Liu J, Gilliam TC, Geschwind DH. 2002. Evidence for a language quantitative trait locus on chromosome 7q in multiplex autism families. *Am J Hum Genet* 70(1):60–71.
- Bassam BJ, Caetano-Anolles G, Gresshoff PM. 1991. Fast and sensitive silver staining of DNA in polyacrylamide gels. *Anal Biochem* 196(1):80–83.
- Benson G. 1999. Tandem repeats finder: A program to analyze DNA sequences. *Nucleic Acids Res* 27(2):573–580.
- Bijlsma EK, Aalfs CM, Sluiter S, Oude Luttikhuis ME, Trembath RC, Hoovers JM, Hennekam RC. 1999. Familial cryptic translocation between chromosomes 2qter and 8qter: Further delineation of the albright hereditary osteodystrophy-like phenotype. *J Med Genet* 36(8):604–609.
- Broman KW, Murray JC, Sheffield VC, White RL, Weber JL. 1998. Comprehensive human genetic maps: Individual and sex-specific variation in recombination. *Am J Hum Genet* 63(3):861–869.
- Burd L, Martsolf JT, Kerbeshian J, Jalal SM. 1988. Partial 6p trisomy associated with infantile autism. *Clin Genet* 33(5):356–359.
- Buxbaum JD, Silverman JM, Smith CJ, Kilifarski M, Reichert J, Hollander E, Lawlor BA, Fitz-gerald 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.
- Casas KA, Mononen TK, Mikail CN, Hased SJ, Li S, Mulvihill JJ, Lin HJ, Falk RE. 2004. Chromosome 2q terminal deletion: Report of 6 new patients and review of phenotype-breakpoint correlations in 66 individuals. *Am J Med Genet* 130(4):331–339.
- Chavrier P, Goud B. 1999. The role of ARF and Rab GTPases in membrane transport. *Curr Opin Cell Biol* 11(4):466–475.
- CLSA. 1999. An autosomal genomic screen for autism. *Am J Med Genet* 88(6):609–615.
- CLSA. 2001. Incorporating language phenotypes strengthens evidence of linkage to autism. *Am J Med Genet* 105:539–547.
- Conrad B, Dewald G, Christensen E, Lopez M, Higgins J, Pierpont ME. 1995. Clinical pheno-type associated with terminal 2q37 deletion. *Clin Genet* 48(3):134–139.
- Daniel A, Baker E, Chia N, Haan E, Malafiej P, Hinton L, Clarke N, Ades L, Darmanian A, Callen D. 2003. Recombinants of intrachromosomal transposition of subtelomeres on chromosomes 1 and 2: A cause of minute terminal chromosomal imbalances. *Am J Med Genet* 117A(1):57–64.
- Donaldson JG, Jackson CL. 2000. Regulators and effectors of the ARF GTPases. *Curr Opin Cell Biol* 12(4):475–482.
- Fisher AM, Ellis KH, Browne CE, Barber JC, Barker M, Kennedy CR, Foley H, Patton MA. 1994. Small terminal deletions of the long arm of chromosome 2: Two new cases. *Am J Med Genet* 53(4):366–369.
- Friedman BD, Gorski JL, Hall BD, Brothman A, Carey JC, Fletjer WL. 1997. Deletion of chromosomal region 2q37: Clinical phenotype in eight cases. *Am J Hum Genet* 61:A98(545).
- Fruman DA, Meyers RE, Cantley LC. 1998. Phosphoinositide kinases. *Annu Rev Biochem* 67:481–507.
- Ghaziuddin M, Burmeister M. 1999. Deletion of chromosome 2q37 and autism: A distinct sub-type? *J Autism Dev Disord* 29(3):259–263.
- Giardino D, Finelli P, Gottardi G, Clerici D, Mosca F, Briscioli V, Larizza L. 2001. Cryptic sub-telomeric translocation t(2;16)(q37;q24) segregating in a family with unexplained still-births and a dysmorphic, slightly retarded child. *Eur J Hum Genet* 9(12):881–886.
- Giardino D, Finelli P, Gottardi G, De Canal G, Della Monica M, Lonardo F, Scarano G, Larizza L. 2003. Narrowing the candidate region of albright hereditary osteodystrophy-like syndrome by deletion mapping in a patient with an unbalanced cryptic translocation t(2;6)(q37.3;q26). *Am J Med Genet* 122A(3):261–265.
- Gillberg C. 1993. Autism and related behaviours. *J Intellect Disabil Res* 37(Pt 4):343–372.
- Gorski JL, Cox BA, Kyine M, Uhlmann W, Glover TW. 1989. Terminal deletion of the long arm of chromosome 2 in a mildly dysmorphic hypotonic infant with karyotype 46,XY,del(2)(q37). *Am J Med Genet* 32(3):350–352.
- Jackson TR, Kearns BG, Theibert AB. 2000. Cytohesins and centaurins: Mediators of PI 3-kinase-regulated Arf signaling. *Trends Biochem Sci* 25(10):489–495.
- Kreutz MR, Bockers TM, Sabel BA, Hulser E, Stricker R, Reiser G. 1997. Expression and sub-cellular localization of p42IP4/centaurin-alpha, a brain-specific, high-affinity receptor for inositol 1,3,4,5-tetrakisphosphate and phosphatidylinositol 3,4,5-trisphosphate in rat brain. *Eur J Neurosci* 9(10):2110–2124.
- Kruglyak L, Daly MJ, Reeve-Daly MP, Lander ES. 1996. Parametric and nonparametric linkage analysis: A unified multipoint approach. *Am J Hum Genet* 58(6):1347–1363.
- Lathrop GM, Lalouel JM, Julier C, Ott J. 1985. Multilocus linkage analysis in humans: Detection of linkage and estimation of recombination. *Am J Hum Genet* 37(3):482–498.
- Leevers SJ, Vanhaesebroeck B, Waterfield MD. 1999. Signalling through phosphoinositide 3-kinases: The lipids take centre stage. *Curr Opin Cell Biol* 11(2):219–225.
- Logue MW, Vieland VJ, RJ G, RR C. 2003. Bayesian analysis of a previously published genome screen for panic disorder reveals new and compelling evidence for linkage to chromosome 7. *Am J Med Genet, Neuropsychiatr Genet* 121B(1):95–99.
- Lord C, Rutter M, Goode S, Heemsbergen J, Jordan H, Mawhood L, Schopler E. 1989. Autism diagnostic observation schedule: A standardized observation of communicative and social behavior. *J Autism Dev Disord* 19(2):185–212.
- Lord C, Rutter M, Le CA. 1994. Autism diagnostic interview-revised: A revised version of a diagnostic interview for caregivers of individuals with possible pervasive developmental disorders. *J Autism Dev Disord* 24(5):659–685.
- 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(3):205–223.
- Lukusa T, Vermeesch JR, Holvoet M, Fryns JP, Devriendt K. 2004. Deletion 2q37.3 and autism: Molecular cytogenetic mapping of the candidate region for autistic disorder. *Genet Couns* 15(3):293–301.
- Phelan MC, Rogers RC, Clarkson KB, Bowyer FP, Levine MA, Estabrooks LL, Sevenson MC, Dobyns WB. 1995. Albright hereditary osteodystrophy and del(2)(q37.3) in four unrelated individuals. *Am J Med Genet* 58(1):1–7.
- Pickles A, Bolton P, Macdonald H, Bailey A, Le CA, 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.
- Rameh LE, Cantley LC. 1999. The role of phosphoinositide 3-kinase lipid products in cell function. *J Biol Chem* 274(13):8347–8350.
- Risch M, Spiker D, Lotspeich L, Nouri N, Hinds D, Hallmayer J, Kalaydjieva L, McCague P, Dimiceli S, Pitts T, et al. 1999. A genomic screen of autism: Evidence for a multilocus etiology. *Am J Hum Genet* 65:493–507.
- Rodgers EE, Theibert AB. 2002. Functions of PI 3-kinase in development of the nervous system. *Int J Dev Neurosci* 20(3–5):187–197.
- Roth MG. 1999. Lipid regulators of membrane traffic through the Golgi complex. *Trends Cell Biol* 9(5):174–179.
- Rutter M, Bailey A, Bolton P, Le Couteur A. 1993. Autism: Syndrome definition and possible genetic mechanisms. In: Baron-Cohen S, H T-F, D C, editors. *Nature, nurture, and psychology*. Washington, DC: APA Books. pp 269–284.
- Sedehizade F, Hanck T, Stricker R, Horstmayer A, Bernstein HG, Reiser G. 2002. Cellular expression and subcellular localization of the human Ins(1,3,4,5)P(4)-binding protein, p42(IP4), in human brain and in neuronal cells. *Brain Res Mol Brain Res* 99(1):1–11.
- Shao Y, Raiford KL, Wolpert CM, Cope HA, Ravan SA, Ashley-Koch AA, Abramson RK, Wright HH, DeLong RG, Gilbert JR, et al. 2002. Phenotypic homogeneity provides increased support for linkage on chromosome 2 in autistic disorder. *Am J Hum Genet* 70(4):1058–1061.

- Smith M, Escamilla JR, Filipek P, Bocian ME, Modahl C, Flodman P, Spence MA. 2001. Molecular genetic delineation of 2q37.3 deletion in autism and osteodystrophy: Report of a case and of new markers for deletion screening by PCR. *Cytogenet Cell Genet* 94(1-2):15-22.
- Stein CK, Del Signore C, Bellinger M, Bryke CR. 1992. Deletion of 2q37-a new syndrome? *Am J Hum Genet* 51:A310(1223).
- Sullivan PF, Neale BM, Neale MC, van den Oord E, Kendler KS. 2003. Multipoint and single point non-parametric linkage analysis with imperfect data. *Am J Med Genet B Neuropsychiatr Genet* 121(1):89-94.
- Szatmari P. 1999. Heterogeneity and the genetics of autism. *J Psychiatry Neurosci* 24(2):159-165.
- Szatmari P, Jones MB, Zwaigenbaum L, MacLean JE. 1998. Genetics of autism: Overview and new directions. *J Autism Dev Disord* 28(5):351-368.
- Vieland VJ. 1998. Bayesian linkage analysis, or: How I learned to stop worrying and love the posterior probability of linkage. *Am J Hum Genet* 63(4):947-954.
- Vieland VJ, Wang K, Huang J. 2001. Power to detect linkage based on multiple sets of data in the presence of locus heterogeneity: Comparative evaluation of model-based linkage methods for affected sib pair data. *Hum Hered* 51(4):199-208.
- Viot-Szoboszalai G, Amiel J, Doz F, Prieur M, Couturier J, Zucker JN, Henry I, Munnich A, Ve-kemans M, Lyonnet S. 1998. Wilms' tumor and gonadal dysgenesis in a child with the 2q37.1 deletion syndrome. *Clin Genet* 53(4):278-280.
- Wolff DJ, Clifton K, Karr C, Charles J. 2002. Pilot assessment of the subtelomeric regions of children with autism: Detection of a 2q deletion. *Genet Med* 4(1):10-14.