Published in final edited form as:

Autism Res. 2009 October ; 2(5): 258–266. doi:10.1002/aur.96.

Identification of Chromosome 7 Inversion Breakpoints in an Autistic Family Narrows Candidate Region for Autism Susceptibility

Holly N. Cukier, David A. Skaar, Melissa Y. Rayner-Evans, Ioanna Konidari, Patrice L. Whitehead, James M. Jaworski, Michael L. Cuccaro, Margaret A. Pericak-Vance, and John R. Gilbert

John P. Hussman Institute for Human Genomics, University of Miami, Miami, Florida (H.N.C., M.Y.R.-E., I.K., P.L.W., J.M.J., M.L.C., M.A.P.-V., J.R.G.); Center for Human Genetics, Duke University Medical Center, Durham, North Carolina (D.A.S.)

Abstract

Chromosomal breaks and rearrangements have been observed in conjunction with autism and autistic spectrum disorders. A chromosomal inversion has been previously reported in autistic siblings, spanning the region from approximately 7q22.1 to 7q31. This family is distinguished by having multiple individuals with autism and associated disabilities. The region containing the inversion has been strongly implicated in autism by multiple linkage studies, and has been particularly associated with language defects in autism as well as in other disorders with language components. Mapping of the inversion breakpoints by FISH has localized the inversion to the region spanning approximately 99–108.75Mb of chromosome 7. The proximal breakpoint has the potential to disrupt either the coding sequence or regulatory regions of a number of cytochrome P450 genes while the distal region falls in a relative gene desert. Copy number variant analysis of the breakpoint regions detected no duplication or deletion that could clearly be associated with disease status. Association analysis in our autism data set using single nucleotide polymorphisms located near the breakpoints showed no significant association with proximal breakpoint markers, but has identified markers near the distal breakpoint (~108–110Mb) with significant associations to autism. The chromosomal abnormality in this family strengthens the case for an autism susceptibility gene in the chromosome 7q22-31 region and targets a candidate region for further investigation.

Keywords

molecular genetics; paracentric inversion; fluorescent in situ hybridization (FISH); genome-wide association study (GWAS)

Introduction

Few clear causative mutations for autism have been identified to date, but there appear to be connections between certain chromosomal rearrangements and autism susceptibility. One review of 15 international studies reports that 7.4% of the autism patients examined had chromosomal abnormalities, with *fra(X)(q27.3)* and 15q alterations being the most commonly observed in 3.1 and 1% of subjects, respectively [Xu, Zwaigenbaum, Szatmari, & Scherer, 2004]. Such chromosomal rearrangements suggest regions for candidate gene studies, particularly when there is convergence of chromosomal abnormalities with linkage peaks.

One such case is an approximately 10Mb chromosome 7q paracentric inversion (inv 7q22.1-31.1), occurring in a multiplex autism family [Ashley-Koch et al., 1999]. The inversion is observed in a mother, and her three children—two sons diagnosed with autism and one daughter with expressive language disorder. The location of the inversion corresponds to the linkage peaks seen in multiple autism studies [Barrett et al., 1999; Hutcheson et al., 2003; International Molecular Genetic Study of Autism Consortium, 1998; Shao et al., 2002], and spans many of the candidate genes tested on 7q. Candidate genes *DLX5* and *DLX6* are near the proximal end of the inversion [Hamilton et al., 2005], while *LAMB1*, *NRCAM*, *IMMP2L*, and *DOCK4* are near the distal end [Bonora et al., 2005; Hutcheson et al., 2004; Maestrini et al., 2009]. *RELN*, also implicated as a candidate gene for autism, is located in the center of the inversion [Persico et al., 2001; Skaar et al., 2005]. Additional autistic patients carrying chromosomal rearrangements with breakpoints near either the distal or the proximal regions have recently been identified [Bayou et al., 2008; Vincent et al., 2008; Warburton et al., 2000]. Furthermore, regions on 7q have been connected to Tourette syndrome, for which co-morbidity with autism has been reported [Baron-Cohen, Mortimore, Moriarty, Izaguirre, & Robertson, 1999]. Rearrangements in this region have been seen in two different Tourette syndrome patients, one with an apparently balanced *t(7;18)* (q22-q31; q22.3) translocation [Boghossian-Sell, Comings, & Overhauser, 1996] and one with an inverted *dup(7)(q22.1-q31)* inserted at 7q31.3 [Petek et al., 2001].

To localize and map the inversion breakpoints in our previously reported multiplex family, fluorescent in situ hybridization (FISH) was used to identify bacterial artificial chromosomes (BACs) that span the disrupted breakpoints and identify genes that potentially could be disrupted. Copy number variation (CNV) analysis using oligo-nucleotide arrays were performed to look for duplications or deletions that might be associated with the breakpoint regions. Lastly, single nucleotide polymorphisms (SNPs) in the breakpoint regions were genotyped for association with autism in a large autism data set. Results indicate that the distal end of this 10Mb inversion region is associated with autism susceptibility.

Methods

Clinical

Family ascertainment and diagnostic criteria—Participants were recruited, enrolled, and sampled according to the Institutional Review Board protocols for the University of

Miami's Hussman Institute for Human Genomics (HIHG). Clinical ascertainment and diagnosis of autism were as previously described [Cuccaro et al., 2003]. Families and individuals with autism were ascertained by clinical referrals and active recruitment. DNA samples were usually obtained at the time of the proband's diagnostic examination. Only those families with probands who met both clinical criteria for autism were included in the analysis. Diagnosis of autism was confirmed by using the Autism Diagnostic Interview—Revised (ADI-R) algorithm [Lord, Rutter, & Le Couteur, 1994]. Discrepancies between the ADI-R and clinical findings were resolved by additional clinical evaluations, including the Autism Diagnostic Observation Schedule (ADOS-G) [Lord et al., 1989]. Affected individuals who did not meet the ADI-R criteria for autism were excluded from the analysis.

Inversion family 7543—A complete clinical description of the children in Family 7543 has been previously described (Fig. 1) [Ashley-Koch et al., 1999]. Briefly, individual III-1 presented at 43 months with developmental delays and behaviors suggestive of autism. Exhaustive testing showed a developmental age of 23 months. Language was estimated to be at the 12-month level and motor skills in the range 18–24 months. Follow-up evaluation at 72 months revealed that the patient had limited social awareness and demonstrated no meaningful language. On the ADI-R, he met diagnostic criteria for autism, surpassing the established clinical cutoffs for all three major domains. Testing on individual III-2 at 32 months revealed a developmental age of 16 months. Receptive language skills were estimated to be in the range 7–10 months and expressive language in the range 9–12 months. During follow-up evaluation at 72 months, the patient met diagnostic criteria for autism on the ADI-R. Individual III-3, the dizygotic twin of III-2, was examined at 33 months and found to have a developmental level of 25 months. Her receptive language skills were estimated to be in the range 15–18 months and expressive language skills in the range 16–20 months. At 72 months, the patient did not meet the algorithmic diagnosis for autism on the ADI-R. However, she did exceed clinical cutoffs for two of the three ADI-R domains for autism (communication and social reciprocity). A later clinical examination was performed, including the ADOS-G. She failed to exceed any of the established clinical cutoffs for autism on the ADOS-G, but was diagnosed with expressive language disorder.

The mother was also assessed as part of our overall clinical evaluation of the family. While she did not meet criteria for autism, she demonstrated obvious social discomfort on clinical interview including noticeable gaze avoidance. In addition, she was administered the Structured Interview for DSM-IV Personality, a semistructured interview that assesses personality disorders. While she did not meet criteria for a formal diagnosis, she reported several features, which are diagnostic of avoidant personality disorder including avoidance of activities that bring her into contact with others, social inhibition, and the belief that she is inept or unappealing.

Cytogenetic analysis of Family 7543 revealed that all three children, individuals III-1, III-2, and III-3, as well as the mother, individual II-2, possess a paracentric inversion on the long arm of chromosome 7 (Inv(7)(q22- q31.2)). The children's maternal grandmother, individual I-2, does not carry the inversion. Neither the father of the children (individual II-1), nor the maternal grandfather (individual I-1) was available for clinical or DNA analysis.

FISH Studies

FISH analyses were performed on metaphase cells from lymphoblast cell lines established using marmoset Epstein–Barr virus, phytohemagglutinin, and interleukin- 2 [Neitzel, 1986]. FISH with different BAC probes was performed according to the technique of Pinkel et al. [Pinkel, Straume, & Gray, 1986]. BACs were selected using the genome browser (Build 35) from the University of California at Santa Cruz (<http://genome.ucsc.edu>) and then obtained from a Human BAC filter library at the Roswell Park Cancer Institute (<http://genomics.roswellpark.org/human/overview.html>). BAC DNAs were isolated using the Qiagen Plasmid Purification kit and labeled using Nick Translation Labeling Kit (with either Spectrum Orange or Spectrum Green), according to the standard Vysis protocol. For each probe, at least ten metaphase cells were captured and analyzed. FISH images were captured on a Leica DMRB Fluorescent microscope and analyzed with Applied Imaging software (Cytovision). Breakpoints of the inversion were characterized by the presence of split FISH signals.

Comparative Genomic Hybridization (CGH)

A custom, fine tiling 385K NimbleGen chip was designed with probes that spanned non-repetitive regions of chromosome 7. Isothermal probes were generated with a target melting temperature of 76°C. Regions containing repeats were masked according to standard Nimblegen protocols. The mean probe interval spacing was 359 base pairs. Promega's human genomic DNA (#G1471) was used as a reference. This DNA is a pooled reference of six to seven anonymous male donors. Cy3 was used to label the experimental samples, while Cy5 labeled the reference DNA. The normalized data was analyzed using Nexus 3.1 Copy Number Program, Professional Edition (BioDiscovery, Inc.). Results were based on locations in the human genome build 18 (NCBI build 36.1). Rank segmentation was performed using a significance threshold of 10^{-6} and again at a threshold of 10^{-3} . Regions were evaluated that appeared across the four individuals that carry the inversion chromosome. Regions of interest were evaluated within the Database of Genomic Variants to determine if they were previously identified (<http://projects.tcag.ca/variation/?source=hg18>).

Genome-Wide Association Study

SNPs within our region of interest were examined in a genome-wide association study (GWAS) [Ma et al., 2009; Wang et al., 2009]. Using the Illumina Human 1M beadchip, SNPs were selected that spanned the region, but removed when SNPs had a minor allele frequency (MAF) < 0.05, Mendelian error > 4%, and HWE P -values < 10^{-6} [as described in Ma et al., 2009]. These quality control steps were calculated using PLINK software as previously described [Ma et al., 2009; Purcell, Cherny, & Sham, 2008]. The initial data set was composed of 487 Caucasian families (407 singleton families containing only one affected individual and 80 multiplex families with at least two affected individuals), with a total of 1,537 individuals being genotyped. Samples were evaluated using the Illumina Human 1M beadchip to identify regions of interest. Association was measured using the pedigree disequilibrium test (PDT) [Martin, Monks, Warren, & Kaplan, 2000]. Linkage

disequilibrium (LD) plots were generated using the HaploView 4.1 program. Tagged SNPs were identified using the pairwise tagging function in HaploView with an r^2 threshold = 0.6.

Genotyping and Association Analysis

All individual SNP genotyping was performed using Taq- Man genotyping assays according to standard protocols (Applied Biosystems, Foster City, CA) [Livak, Marmaro, & Todd, 1995]. We selected SNPs that had not yet been run in the region for the GWAS and that had a $MAF > 0.05$. In total, the association data set represents 418 Caucasian families, comprised of 123 singleton families and 295 multiplex families. Within this data set, 222 families were from MIHG and 196 families were from the Autism Genetic Research Exchange (AGRE) data sets. All AGRE families were multiplex, while the University of Miami data set included 99 multiplex and 123 singleton families. The PDT and Geno-PDT tests were used to measure association [Martin et al., 2000; Martin, Bass, Gilbert, Pericak-Vance, & Hauser, 2003].

Results

FISH Studies

FISH studies using labeled BACs as probes localized the inversion breakpoints by identifying probes that showed a single signal on the normal chromosome and two signals on the inversion chromosome. The split signal indicated that the probes crossed the breakpoint and mapped to both the proximal and distal breakpoints in the chromosomal spreads. On the proximal side of the inversion, screening started from 7q22.1, the estimated breakpoint location based on the original chromosome banding, and proceeded toward the telomere. RP11-316A24 and RP11-977H6 (Build 36.1) both demonstrated a single signal on the normal chromosome and two signals on the inversion chromosome, indicating that they spanned the inversion breakpoint (Figs. 2 and 3A). The shared region (~99,035,000–99,132,000 bases) between these BACs defines a maximum region of 97 kb that contains the breakpoint. Located just upstream of this common overlap region, BAC RP11-757A13 failed to demonstrate a split signal, thereby confirming the predicted maximum region. RP11-8E14 also shows a single signal on the inversion chromosome, indicating it does not substantially overlap the predicted breakpoint region and reducing the predicted maximum proximal breakpoint region to approximately 60 kb (~99,035,000–99,094,000 bases).

For the telomeric inversion breakpoint, FISH studies started in 7q31.2, based on the initial chromosome banding, and proceeded toward the centromere (Fig. 3B). RP11-21G9 and RP11-698J23 both showed two signals on the inversion chromosome, with the overlapping sequence defining a maximum region of 117kb (~108,593,000–108,710,000 bases) containing this breakpoint. Three BACs, RP11-138D8, RP11-159I15, and RP11-1124G20, overlapped this region and showed single signals on the inversion chromosome. The exclusion of these BACs defines a roughly 18 kb region (~108,687,000–108,704,000 bases), with the proviso that these BACs may slightly overlap the breakpoint, but not enough for two signals to be detectable by FISH.

Comparative Genomic Hybridization (CGH)

Within and near the areas of the inversion breakpoint, two sites of interest were revealed by comparative genomic hybridization. An 11,854 base pair region was identified near, but outside, the proximal breakpoint that appeared to be duplicated in the mother and all three children, but not in the grandmother (97,228,569–97,240,423). At the lower 10^{-3} threshold, all four individuals show the duplication, and at the higher 10^{-6} threshold, a single individual produced a robust signal (affected brother). This region has been previously recognized as an area of CNV, but, to date, only deletions have been reported [Kidd et al., 2008; Korbelt et al., 2007; McCarroll et al., 2006; Perry et al., 2008; Tuzun et al., 2005]. The region contains a large segmental duplication recognized to have 91.7% identity to a region on chromosome 8 (47,789,047–47,808,516). The duplicated region on chromosome 7 falls downstream of the *Tachykinin 1 (TAC1)* gene, which encodes substance P and additional neurokinins [Nawa, Kotani, & Nakanishi, 1984]. A second, smaller (3841 bp) region of interest was located near the distal breakpoint (109,232,099–109,235,940). This region appeared to be duplicated within the grandmother and deleted within the mother and three siblings. Both duplications and deletions in this region have been reported within the Database of Genomic Variants [Conrad, Andrews, Carter, Hurles, & Pritchard, 2006; de Smith et al., 2007; Kidd et al., 2008; Korbelt et al., 2007; McCarroll et al., 2006; Perry et al., 2008; Redon et al., 2006; Tuzun et al., 2005].

Association Mapping of the Breakpoint Regions

To investigate association with autism in the breakpoint regions, SNPs spanning the inversion breakpoints were genotyped in a large autism data set. We performed fine mapping across the region using the Illumina Human 1M beadchip. While SNPs did not appear to be in the proximal region with significant P -values (data not shown), there were several within the distal region that showed suggestive significance (Table I, Fig. 4). Between 108,632,498 and 108,756,291, there were eight SNPs with a P -value of less than 0.01 (rs6947253, rs6945785, rs2396076, rs6975162, rs980400, rs12705563, rs9969305, rs848839) and an additional six SNPs had P -values of less than 0.05 (rs1525021, rs2894534, rs1525029, rs1858897, rs6466273, rs6971422). The 19 SNPs within this region were separated into three distinct LD blocks with the region being defined by 6 tagged SNPs (Fig. 4). Three SNPs remain significant under the corrected Bonferroni P -value of 0.0083 (rs6947253, rs6975162, rs12705563).

Further mapping using TaqMan technology of additional SNPs in this region was performed in another autism cohort. Once again, no significantly associated SNPs were identified near the proximal end of the breakpoint; of 17 markers genotyped across the region from 98–100 Mb, none showed significance at the P -value threshold of 0.05 (data not shown). Near the predicted distal breakpoint, however, rs2106442 and rs2188443 show significance for the entire data set with PDT P -values of 0.033 and 0.039, respectively (Table II, Fig. 5). When evaluating singleton families alone, significant P -values for association were seen for six distal breakpoint SNPs, rs764113, rs4730365, rs2106442, rs2188443, rs1541338, and rs995027 (P -values of 0.025, 0.004, 0.002, 0.004, 0.014, and 0.007, respectively, Table II). Five of these SNPs create a single LD block, with rs764113 falling outside of this block (Fig. 5). The region can be defined by three tagging SNPs. All of the SNPs within the LD

block are found to be significant under the corrected Bonferroni *P*-value of 0.0167 (rs4730365, rs2106442, rs2188443, rs1541338, rs995027). None of these SNPs showed significant association for the group of multiplex families.

Discussion

The entire distal q arm of chromosome 7 is associated with autism in a broad linkage peak in larger data sets, including our own [Barrett et al., 1999; Huthcheson et al., 2003; International Molecular Genetic Study of Autism Consortium, 1998; Shao et al., 2002]. Chromosomal breaks and rearrangements have been previously associated with autism [Dykens, Sutcliffe, & Levitt, 2004; Tentler et al., 2001; Vincent et al., 2000, 2008]. Chromosomal rearrangements of 7q in connection with autism have not been observed frequently, but breaks in this region have been observed in patients with other disorders involving language components. Family 7543 is a unique resource in that it has two affected male autistic siblings, a female sibling with expressive language disorder and developmental delay, and a mother with avoidant personality disorder. It has been proposed that 7q is a primary contributor to the language component of autism [Folstein & Mankoski, 2000]. Specific language impairment has been associated to 7q31 [O'Brien, Zhang, Nishimura, Tomblin, & Murray, 2003]. 7q31 breaks have been observed in individuals with Tourette syndrome [Baron-Cohen et al., 1999; Petek et al., 2001] and language impairments [Lennon et al., 2007; Tyson, McGillivray, Chijiwa, & Rajcan-Separovic, 2004; Warburton et al., 2000; Zeeman et al., 2006] as well as in the autism family used in this study.

In the present study, the minimum proximal breakpoint region contains *ZNF498*, an uncharacterized gene with similarity to zinc-finger proteins, and *CYP3A5*, which encodes a member of the Cytochrome P450 family. Three other P450 coding genes are in a cluster with *CYP3A5* and could be affected if the breakpoint extends outside the minimum predicted region, has a distal effect, or has an associated duplication or deletion. Coding or regulation of *ZNF498* or any of several cytochrome P450 genes could be disrupted near the proximal breakpoint. The lack of association with this region in the larger autism data set around the proximal breakpoint suggests that this area does not contain a common variant; however, this does not rule out the possibility that rare variants may exist in a subset of these families or that there is an autistic causative aberration at the breakpoint in Family 7543.

The minimum region predicted to contain the distal breakpoint does not contain any characterized genes or transcripts, being over 500 kb from the nearest well annotated gene. In fact, the distal breakpoint is in an approximately 2Mb (~108–110 Mb) gene desert. The distal inversion breakpoint does, however, show significant association with multiple SNPs in the large autism data set and could exert an effect on an uncharacterized gene, a non-coding transcript, or be involved in the regulation of a distant gene. Interestingly, positive association with *LAMB1*, proximal to the distal breakpoint, has been shown in some studies [Bonora et al., 2005; Huthcheson et al., 2004]. Furthermore, a recent study has implicated the *IMMP2L* and *DOCK4* genes located distal to the distal breakpoint in autism through both association and the identification of deletions and duplications in affected individuals [Maestrini et al., 2009]. Separation of the samples into singleton and multiplex families

generated greater significance for singleton families, indicating that the bulk of the positive association was found in families with presumed sporadic cases of autism.

Using custom CGH chip-based iso-temp oligo-nucleotides, we examined the breakpoint regions for small cytogenetic abnormalities such as deletions or duplications, known collectively as CNVs, which might be associated with the inversion breakpoints. Such CNVs could be an actual causative mutation, and disrupt functional regions, or several genes, alone or in combination with breakpoint alterations. Although several CNVs were detected, none could be directly associated with the putative breakpoint regions, so their status remains unclear. It should be pointed out that it is not possible to examine repeat regions using this approach, so if the breakpoints cause a loss of repeats this would not have been detected in these studies.

This 7q region has been the focus of much research in autism. The convergence of this inversion and the association data reported here with previous linkage, candidate genes studies, and language phenotypes for 7q31 provide support for an autism gene (or genes) in this area. More specifically, the present study suggests further investigations in and around the distal breakpoint region of 108–109Mb on 7q may prove fruitful and resolve some of the problems faced to date in fine mapping the autism locus (loci) on chromosome 7q. In addition to performing larger association studies, more information on the frequency and type of disruptions in the span of 7q22.1-31.2 in autistic individuals would be useful in determining its potential importance in autism. However, until recently, detecting smaller inversions or other cytogenetic abnormalities have been quite difficult with only subtle changes in chromosome banding and few deletions or duplications large enough to be easily detectable. The advent of new approaches involving CGH and oligo-nucleotide arrays for the detection of copy number variants may profoundly affect our understanding of autism [Sebat et al., 2007]. Based on the findings in this region, a thorough investigation of this region of chromosome 7q in a large data set using both association and CGH may lead to a better understanding of 7q alterations and how these alterations may connect to language development.

Acknowledgments

Grant sponsor: The National Institute of Neurological Disorders and Stroke; Grant number: 7P01 NS 26630-19.

We are grateful to many patients with autism and their families for their participation, thus making our research possible. A subset of the participants was ascertained while several of the authors (M. L. C., M. A. P. V., and J. R. G.) were at Duke University. We also acknowledge S. Schwartz and L. Christ for technical assistance with FISH studies.

References

- Ashley-Koch AE, Wolpert CM, Menold MM, Zaeem L, Basu S, et al. Genetic studies of autistic disorder and chromosome 7. *Genomics*. 1999; 61:227–236. [PubMed: 10552924]
- Baron-Cohen S, Mortimore C, Moriarty J, Izaguirre J, Robertson M. The prevalence of Gilles de la Tourette's syndrome in children and adolescents with autism. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*. 1999; 40:213–218.
- Barrett S, Beck JC, Bernier R, Bisson E, Braun TA, et al. An autosomal genomic screen for autism. collaborative linkage study of autism. *American Journal of Medical Genetics*. 1999; 88:609–615. [PubMed: 10581478]

- Bayou N, M'rad R, Belhaj A, Daoud H, Ben Jemaa L, et al. De novo balanced translocation t (7;16) (p221; p11.2) associated with autistic disorder. *Journal of Biomedicine & Biotechnology*. 2008; 2008:231904. [PubMed: 18475318]
- Boghossian-Sell L, Comings DE, Overhauser J. Tourette syndrome in a pedigree with a 7;18 translocation: Identification of a YAC spanning the translocation breakpoint at 18q22.3. *American Journal of Human Genetics*. 1996; 59:999–1005. [PubMed: 8900226]
- Bonora E, Lamb JA, Barnby G, Sykes N, Moberly T, et al. Mutation screening and association analysis of six candidate genes for autism on chromosome 7q. *European Journal of Human Genetics: EJHG*. 2005; 13:198–207. [PubMed: 15523497]
- Conrad DF, Andrews TD, Carter NP, Hurler ME, Pritchard JK. A high-resolution survey of deletion polymorphism in the human genome. *Nature Genetics*. 2006; 38:75–81. [PubMed: 16327808]
- Cuccaro ML, Shao Y, Bass MP, Abramson RK, Ravan SA, et al. Behavioral comparisons in autistic individuals from multiplex and singleton families. *Journal of Autism and Developmental Disorders*. 2003; 33:87–91. [PubMed: 12708583]
- de Smith AJ, Tsalenko A, Sampas N, Scheffer A, Yamada NA, et al. Array CGH analysis of copy number variation identifies 1284 new genes variant in healthy white males: Implications for association studies of complex diseases. *Human Molecular Genetics*. 2007; 16:2783–2794. [PubMed: 17666407]
- Dykens EM, Sutcliffe JS, Levitt P. Autism and 15q11-q13 disorders: Behavioral, genetic, and pathophysiological issues. *Mental Retardation and Developmental Disabilities Research Reviews*. 2004; 10:284–291. [PubMed: 15666333]
- Folstein SE, Mankoski RE. Chromosome 7q: Where autism meets language disorder? *American Journal of Human Genetics*. 2000; 67:278–281. [PubMed: 10889044]
- Hamilton SP, Woo JM, Carlson EJ, Ghanem N, Ekker M, Rubenstein JL. Analysis of four DLX homeobox genes in autistic probands. *BMC Genetics*. 2005; 6:52. [PubMed: 16266434]
- Hutcheson HB, Bradford Y, Folstein SE, Gardiner MB, Santangelo SL, et al. Defining the autism minimum candidate gene region on chromosome 7. *American Journal of Medical Genetics*. 2003; 117B:90–96. [PubMed: 12555242]
- Hutcheson HB, Olson LM, Bradford Y, Folstein SE, Santangelo SL, et al. Examination of NRCAM, LRRN3, KIAA0716, and LAMB1 as autism candidate genes. *BMC Medical Genetics*. 2004; 5:12. [PubMed: 15128462]
- International Molecular Genetic Study of Autism Consortium. A full genome screen for autism with evidence for linkage to a region on chromosome 7q. *Human Molecular Genetics*. 1998; 7:571–578. [PubMed: 9546821]
- Kidd JM, Cooper GM, Donahue WF, Hayden HS, Sampas N, et al. Mapping and sequencing of structural variation from eight human genomes. *Nature*. 2008; 453:56–64. [PubMed: 18451855]
- Korbel JO, Urban AE, Grubert F, Du J, Royce TE, et al. Systematic prediction and validation of breakpoints associated with copy-number variants in the human genome. *Proceedings of the National Academy of Sciences of the United States of America*. 2007; 104:10110–10115. [PubMed: 17551006]
- Lennon PA, Cooper ML, Peiffer DA, Gunderson KL, Patel A, et al. Deletion of 7q31.1 supports involvement of FOXP2 in language impairment: Clinical report and review. *American Journal of Medical Genetics Part A*. 2007; 143A:791–798. [PubMed: 17330859]
- Livak KJ, Marmaro J, Todd JA. Towards fully automated genome-wide polymorphism screening. *Nature Genetics*. 1995; 9:341–342. [PubMed: 7795635]
- Lord C, Rutter M, Goode S, Heemsbergen J, Jordan H, et al. Autism diagnostic observation schedule. A standardization observation of communicative and social behavior. *Journal of Autism and Developmental Disorders*. 1989; 19:185–212. [PubMed: 2745388]
- Lord C, Rutter M, Le Couteur A. Autism diagnostic interview-revised: A revised version of a diagnostic interview for caregivers of individuals with possible pervasive developmental disorders. *Journal of Autism and Developmental Disorders*. 1994; 24:659–685. [PubMed: 7814313]
- Ma D, Salyakina D, Jaworski JM, Konidari I, Whitehead PL, et al. A genome-wide association study of autism reveals a common novel risk locus at 5p14.1. *Annals of Human Genetics*. 2009; 73:263–273. [PubMed: 19456320]

- Maestrini E, Pagnamenta AT, Lamb JA, Bacchelli E, Sykes NH, et al. High-density SNP association study and copy number variation analysis of the AUTS1 and AUTS5 loci implicate the IMMP2L-DOCK4 gene region in autism susceptibility. *Molecular Psychiatry*. 2009
- Martin ER, Bass MP, Gilbert JR, Pericak-Vance MA, Hauser ER. Genotype-based association test for general pedigrees: The genotype-PDT. *Genetic Epidemiology*. 2003; 25:203–213. [PubMed: 14557988]
- Martin ER, Monks SA, Warren LL, Kaplan NL. A test for linkage and association in general pedigrees: The pedigree disequilibrium test. *American Journal of Human Genetics*. 2000; 67:146–154. [PubMed: 10825280]
- McCarroll SA, Hadnott TN, Perry GH, Sabeti PC, Zody MC, et al. Common deletion polymorphisms in the human genome. *Nature Genetics*. 2006; 38:86–92. [PubMed: 16468122]
- Nawa H, Kotani H, Nakanishi S. Tissue-specific generation of two preprotachykinin mRNAs from one gene by alternative RNA splicing. *Nature*. 1984; 312:729–734. [PubMed: 6083453]
- Neitzel H. A routine method for the establishment of permanent growing lymphoblastoid cell lines. *Human Genetics*. 1986; 73:320–326. [PubMed: 3017841]
- O'Brien EK, Zhang X, Nishimura C, Tomblin JB, Murray JC. Association of specific language impairment (SLI) to the region of 7q31. *American Journal of Human Genetics*. 2003; 72:1536–1543. [PubMed: 12721956]
- Perry GH, Ben-Dor A, Tsalenko A, Sampas N, Rodriguez-Revenga L, et al. The fine-scale and complex architecture of human copy-number variation. *American Journal of Human Genetics*. 2008; 82:685–695. [PubMed: 18304495]
- Persico AM, D'Agruma L, Maiorano N, Totaro A, Militeri R, et al. Reelin gene alleles and haplotypes as a factor predisposing to autistic disorder. *Molecular Psychiatry*. 2001; 6:150–159. [PubMed: 11317216]
- Petek E, Windpassinger C, Vincent JB, Cheung J, Boright AP, et al. Disruption of a novel gene (IMMP2L) by a breakpoint in 7q31 associated with tourette syndrome. *American Journal of Human Genetics*. 2001; 68:848–858. [PubMed: 11254443]
- Pinkel D, Straume T, Gray JW. Cytogenetic analysis using quantitative, high-sensitivity, fluorescence hybridization. *Proceedings of the National Academy of Sciences of the United States of America*. 1986; 83:2934–2938. [PubMed: 3458254]
- Purcell, S.; Cherny, SS.; Sham, PC. Genetic Power Calculator. Homepage of Harvard University. 2008. [Online]. Available from: <http://pngu.mgh.harvard.edu/~purcell/gpc/>
- Redon R, Ishikawa S, Fitch KR, Feuk L, Perry GH, et al. Global variation in copy number in the human genome. *Nature*. 2006; 444:444–454. [PubMed: 17122850]
- Sebat J, Lakshmi B, Malhotra D, Troge J, Lese-Martin C, et al. Strong association of de novo copy number mutations with autism. *Science*. 2007; 316:445–449. [PubMed: 17363630]
- Shao Y, Wolpert CM, Raiford KL, Menold MM, Donnelly SL, et al. Genomic screen and follow-up analysis for autistic disorder. *American Journal of Medical Genetics*. 2002; 114:99–105. [PubMed: 11840513]
- Skaar DA, Shao Y, Haines JL, Stenger JE, Jaworski J, et al. Analysis of the RELN gene as a genetic risk factor for autism. *Molecular Psychiatry*. 2005; 10:563–571. [PubMed: 15558079]
- Tentler D, Brandberg G, Betancur C, Gillberg C, Anneren G, et al. A balanced reciprocal translocation t(5;7)(q14;q32) associated with autistic disorder: Molecular analysis of the chromosome 7 breakpoint. *American Journal of Medical Genetics*. 2001; 105:729–736. [PubMed: 11803521]
- Tuzun E, Sharp AJ, Bailey JA, Kaul R, Morrison VA, et al. Fine-scale structural variation of the human genome. *Nature Genetics*. 2005; 37:727–732. [PubMed: 15895083]
- Tyson C, McGillivray B, Chijiwa C, Rajcan-Separovic E. Elucidation of a cryptic interstitial 7q31.3 deletion in a patient with a language disorder and mild mental retardation by array-CGH. *American Journal of Medical Genetics. Part A*. 2004; 129A:254–260.
- Vincent JB, Choufani S, Horike S, Stachowiak B, Li M, et al. A translocation t(6;7)(p11-p12;q22) associated with autism and mental retardation: Localization and identification of candidate genes at the breakpoints. *Psychiatric Genetics*. 2008; 18:101–109. [PubMed: 18496206]

- Vincent JB, Herbrick JA, Gurling HM, Bolton PF, Roberts W, Scherer SW. Identification of a novel gene on chromosome 7q31 that is interrupted by a translocation breakpoint in an autistic individual. *American Journal of Human Genetics*. 2000; 67:510–514. [PubMed: 10889047]
- Wang K, Zhang H, Ma D, Bucan M, Glessner JT, et al. Common genetic variants on 5p14.1 associate with autism spectrum disorders. *Nature*. 2009; 459:528–533. [PubMed: 19404256]
- Warburton P, Baird G, Chen W, Morris K, Jacobs BW, et al. Support for linkage of autism and specific language impairment to 7q3 from two chromosome rearrangements involving band 7q31. *American Journal of Medical Genetics*. 2000; 96:228–234. [PubMed: 10893502]
- Xu J, Zwaigenbaum L, Szatmari P, Scherer SW. Molecular cytogenetics of autism. *Current Genomics*. 2004; 5:347–364.
- Zeesman S, Nowaczyk MJ, Teshima I, Roberts W, Cardy JO, et al. Speech and language impairment and oromotor dyspraxia due to deletion of 7q31 that involves FOXP2. *American Journal of Medical Genetics Part A*. 2006; 140:509–514. [PubMed: 16470794]

Figure 1. Pedigree of Family 7543. This multiplex family demonstrates that the inv7q22.1–q31.2 chromosome was inherited maternally to all three children—two boys and one girl.

Figure 2.
FISH with BAC RP11-316A24 as probe, showing a single signal on one chromosome and a split signal on the other.

Figure 3. Overlapping BACs showing definite maximum and potential minimum regions containing the inversion breakpoints in the proximal (**A**) and distal (**B**) regions.

Figure 4. Linkage disequilibrium (LD) plot showing the r^2 values for the SNPs tested by the Illumina Human 1M beadchip in the distal inversion breakpoint. Three LD blocks were observed.

Figure 5. LD plot showing the r^2 values for the SNPs tested by TaqMan in the distal inversion breakpoint. One LD block of 40 kb was observed, with a single SNP falling outside of this block.

Table I

SNPs Typed in the Genome-Wide Association Study in the Distal Region

SNP	Location	MAF	<i>P</i> -hwe	<i>P</i> -value
rs1525021	108633167	0.45	0.35	0.0387
rs2894534	108633957	0.44	0.44	0.0200
rs6952038	108641773	0.31	0.95	0.0986
rs17459383	108646239	0.32	0.95	0.0732
rs1525029	108646886	0.44	0.38	0.0136
rs1464813	108650173	0.13	0.80	0.7784
rs6947253	108671047	0.43	0.26	0.0072
rs6945785	108676304	0.45	0.32	0.0088
rs2396076	108683705	0.45	0.20	0.0098
rs6975162	108693525	0.45	0.24	0.0076
rs1858897	108695618	0.29	1.00	0.0283
rs6466273	108699280	0.31	0.79	0.0262
rs980400	108704453	0.45	0.24	0.0087
rs12705563	108705797	0.16	0.75	0.0030
rs9969305	108714907	0.45	0.27	0.0089
rs6971422	108721308	0.31	0.89	0.0349
rs848839	108732977	0.45	0.26	0.0099
rs848896	108747265	0.12	0.11	0.0508
rs10257229	108756291	0.11	0.66	0.2829

P-values in dark grey have a value of less than 0.01, *P*-values in light grey have a value of less than 0.05, MAF stands for minor allele frequency.

Table II

PDT *P*-values for SNPs Typed by TaqMan in the Distal Region

SNP	Location	MAF	<i>P</i> -hwe	PDT		
				Joint	Singleton	Multiplex
rs764113	108467533	0.29	0.70	0.869	0.025	0.246
rs4730365	108679058	0.43	0.75	0.105	0.004	0.922
rs2106442	108687109	0.43	0.87	0.033	0.002	0.512
rs2188443	108703639	0.43	0.57	0.039	0.004	0.509
rs1541338	108717585	0.13	0.22	0.245	0.014	0.702
rs995027	108720020	0.42	0.75	0.092	0.007	0.711

P-values in dark grey have a value of less than 0.01, *P*-values in light grey have a value of less than 0.05, MAF stands for minor allele frequency.