



Published in final edited form as:

Mol Psychiatry. 2010 October ; 15(10): 1006–1015. doi:10.1038/mp.2009.42.

Genome-wide linkage in Utah autism pedigrees

K Allen-Brady¹, R Robison¹, D Cannon¹, T Varvil², M Villalobos¹, C Pingree¹, MF Leppert², J Miller¹, WM McMahon¹, and H Coon¹

¹Utah Autism Research Project, Department of Psychiatry, University of Utah, Salt Lake City, UT, USA

²Department of Human Genetics, University of Utah, Salt Lake City, UT, USA

Abstract

Genetic studies of autism over the past decade suggest a complex landscape of multiple genes. In the face of this heterogeneity, studies that include large extended pedigrees may offer valuable insight, as the relatively few susceptibility genes within single large families may be more easily discerned. This genome-wide screen of 70 families includes 20 large extended pedigrees of 6–9 generations, 6 moderate-sized families of 4–5 generations, and 44 smaller families of 2–3 generations. The Center for Inherited Disease Research (CIDR) provided genotyping using the Illumina Linkage Panel 12, a 6K single nucleotide polymorphism (SNP) platform. Results from 192 subjects with an Autism Spectrum Disorder (ASD), and 461 of their relatives revealed genome-wide significance on chromosome 15q, with three possibly distinct peaks: 15q13.1-q14 (HLOD=4.09 at 29,459,872bp); 15q14-q21.1 (HLOD=3.59 at 36,837,208bp); and 15q21.1-q22.2 (HLOD=5.31 at 55,629,733bp). Two of these peaks replicate previous findings. There were additional suggestive results on chromosomes 2p25.3-p24.1 (HLOD=1.87), 7q31.31-q32.3 (HLOD=1.97), and 13q12.11-q12.3 (HLOD=1.93). Affected subjects in families supporting the linkage peaks found in this study did not reveal strong evidence for distinct phenotypic subgroups.

Keywords

Autism Spectrum Disorder; genetic linkage; extended pedigrees; chromosome 15

Introduction

Autism Spectrum Disorders (ASDs; MIM 209850) are characterized by impairments in social interaction and communication as well as by repetitive and stereotyped behaviors and interests. ASDs include Autistic Disorder, Asperger Disorder, and Pervasive Developmental Disorder – Not Otherwise Specified (PDD-NOS). Prevalence estimates for ASDs have recently been reported to be approximately 1 in every 150 children in the general population.¹ In families with an autistic child, recurrence rates are estimated to be greater

Users may view, print, copy, and download text and data-mine the content in such documents, for the purposes of academic research, subject always to the full Conditions of use:http://www.nature.com/authors/editorial_policies/license.html#terms

Correspondence should be sent to: Hilary Coon, Utah Autism Research Project, Department of Psychiatry, 650 Komar Drive, Suite 206, Salt Lake City, UT 84108, Telephone: 801-585-3068, Fax: 801-585-5723, hilary.coon@hsc.utah.edu.

than 15% that an additional offspring will also have autism.^{2, 3} ASDs present a serious public health concern that warrants identification of risk factors with great urgency.

Genetic factors play a substantial role in ASDs (See Review⁴). However, recent gene discovery efforts for ASDs show that genetic heterogeneity is the rule, rather than the exception. While known genes now explain as much as 20% of ASD cases, each known genetic cause explains only a small fraction of the total.⁴ In addition, genome-wide linkage and association studies done over the past decade have implicated multiple regions. Such heterogeneity increases the value of studies that include large extended pedigrees. Many autism studies have focused on small families (sibling pairs, or two parents and an affected offspring) to try to localize autism predisposition genes. These collections of small families may include cases with many different susceptibility loci. Subjects affected with ASDs who are members of a large extended family may be more likely to share the same genetic causes through their common ancestors. Within such families, autism may be more genetically homogeneous.^{5, 6} Additionally, these family members are more likely to share similar environmental exposures, facilitating possible future analyses of gene by environment interaction effects.

We recently published a genome-wide, 10K single nucleotide polymorphism (SNP) linkage analysis using one of the largest known extended ASD pedigrees, with seven ASD cases distributed in a six-generation pedigree, all of whom were descendants of a single founding couple.⁷ In this previous study, we identified three chromosomal regions meeting genome-wide linkage significance criteria at 3q13.2-13.31, 3q26.31-q27.3, and 20q11.21-q13.12; and two additional regions meeting criteria for suggestive significance at 7p14.1-p11.22 and 9p24.3. All five identified regions replicated previous findings, and for some of the regions, we were able to narrow the boundaries. However, our results again illustrated the complex nature of ASD as we found genetic heterogeneity even within a single pedigree.

In our current study, we present genome-wide 6K SNP linkage results for 70 multiple generation pedigrees, ranging in size from 2 to 9 generations. Ascertainment of these 70 multi-generation pedigrees represents over 20 years of autism research in Utah.⁸ These pedigrees show modest positive scores in two of our previously identified regions (9p24.3 and 20q11.21-q13.12), and provide convincing evidence for new ASD susceptibility loci on chromosome 15.

Materials and Methods

Subjects

Subjects were members of 70 pedigrees having at least two family members with an ASD. A total of 653 subjects were genotyped, 192 of whom were defined as having either strictly defined Autistic Disorder or a more broadly defined ASD (Table 1 about here). Table 1 shows the characteristics of these families, which include 20 large extended pedigrees of 6–9 generations, 6 moderately-sized families of 4–5 generations, and 44 smaller families of 2–3 generations. The 20 extended pedigrees have an average of 23 deceased connecting relatives in the upper generations who were not genotyped. These families were identified or confirmed using the Utah Population Database (UPDB), a computerized genealogy database

that contains family history information for over 6.5 million individuals who are, for the most part, descendants of the nineteenth century Utah pioneers (www.hci.utah.edu/groups/ppr/). Previous studies have shown low rates of inbreeding within the UPDB.^{9, 10} Using the UPDB, we were able to identify many distant family relationships between the individuals with ASD that were not known to the subjects or their families. These relationships were kept confidential. This study has ongoing approval from the University of Utah Institutional Review Board.

Phenotyping

Families interested in participating were asked to give questionnaire consent, then to give initial information regarding possible exclusion criteria and to complete the Social Communication Questionnaire (SCQ).¹¹ The SCQ was developed as a parent report measure based on the Autism Diagnostic Interview-Revised.¹² A cutoff score of 15 on the SCQ has been shown to have good sensitivity (0.85) and specificity (0.75) for identifying ASDs from other diagnoses.^{11, 13} Subjects were contacted for possible inclusion if their SCQ score was above 15, or if they had a previous diagnosis of ASD. Subjects were excluded if they reported medical conditions known to be associated with autism (tuberous sclerosis, Fragile X, neurofibromatosis, congenital rubella, or PKU) or evidence of brain injury. If subjects were eligible for the study, they were asked to sign informed consent for DNA and additional assessments. When possible, all subjects with a suspected ASD were then given both the ADI-R¹² and the Autism Diagnostic Observation Schedule-Generic (ADOS-G),¹⁴ and study diagnoses were made using these assessments. For cases where assessments could not be obtained, diagnoses were made according to DSM-IV criteria by a psychologist trained in autism assessment (JM) using all available information (clinical records, other behavioral data, other questionnaire and interview information). Items from the ADI (“age of first words” and “age of first phrases”) were used to assess language delay. For parents who indicated normal onset, but who could not remember exact ages, values were set to 23 months for words and 32 months for phrases (acquiring language after these ages is considered abnormal on the ADI-R). For parents who indicated delayed onset, but could not remember exact ages, values were set to 1-½ standard deviations above the mean. For subjects who never acquired language, values were set to 3 standard deviations above the mean. Qualitative language delay was coded for any subject whose onset was ≥ 24 months for words, or ≥ 33 months for phrases.

IQ was measured in subjects with ASDs using assessments appropriate for age and developmental level. The Wechsler Intelligence Scale for Children WISC-III or Wechsler Adult Intelligence Scale WAIS-III,^{15, 16} Differential Abilities Scale (DAS),¹⁷ and the Mullen Scales of Early Development¹⁸ were used. The DAS is appropriate for children ages 2-1/2 to 18 years with either typical or delayed development, and the General Conceptual Ability Score from the DAS correlates well with the Full Scale IQ score of the Wechsler (WISC-III and WAIS-III) scales.^{19, 20} If we were unable to obtain a valid score on a subject under 68 months on the DAS, we used the Mullen, a standardized measure of cognitive function in young children. For those administered the Mullen, the Early Learning Composite *t*-score (mean=50, sd=10) was converted to a standard score (mean=100, sd=15) as a measure of overall IQ.²¹

Genotyping

Genotyping services were provided by the Center for Inherited Disease Research (CIDR) using the 6K single nucleotide polymorphism (SNP) linkage panel. Originally, 703 samples from the pedigrees were sent for genotyping, in addition to 32 blind duplicates of these pedigree samples for QC, for a total of 735 samples. QC genotyping also included internal controls used by CIDR. The genotyping platform was the Illumina Linkage Panel 12, which includes 6090 single nucleotide polymorphism (SNP) markers, with an average genetic coverage of 0.65 cM. Illumina BeadStudio software was used to evaluate all genotypes using the quantitative GenCall score, which is an indicator of how well a DNA sample performed over all released SNP assays. A total of 55 samples were not released due to one or more of the following reasons: 1) poorly defined clusters, 2) excessive replicate and/or Mendelian errors, 3) more than 50% missing data, or 4) a higher than expected missing data rate for markers on the X chromosome, suggesting a possible mosaic 46XX or 46XO karyotype. Five of these 55 unreleased samples were blind duplicate pairs and the rest were pedigree subjects. There were therefore a total of 680 successfully genotyped subjects, of which 653 were pedigree members and 27 were blind duplicates for QC. Three of the smaller families were left with only one affected case after this QC step, so there were effectively 67 informative families in the sample.

Of the 6090 total SNPs possible, 6044 were released. Loci were dropped if atypical clustering patterns were found. A total of 4,309,372 genotypes were released with a missing data rate of 0.064% and a Mendelian consistency rate of 99.96%. SNPs with Mendelian errors were subsequently zeroed using PedCheck.²² The 27 blind duplicate pairs were checked for consistency between pairs using the file cleaned by PedCheck. Within these cleaned genotypes, duplicate reproducibility was 100%.

Analyses

We used the genetic map provided by CIDR based on the deCODE genetic map.²³ Base pair positions were obtained from the March 2006 human reference sequence (hg18) assembly. Analysis was done using the multipoint linkage software MCLINK, a Markov chain Monte Carlo (MCMC) method that allows for multilocus linkage analysis on large extended pedigrees.²⁴ Using blocked Gibbs sampling, MCLINK generates inheritance matrices from haplotype chains for the markers being analyzed, and performs an approximate calculation of the log-likelihood function linkage statistics. Internally, MCLINK runs the analysis five times to ensure a consistent solution. MCLINK has been used previously to identify candidate genomic regions for a number of complex diseases.^{25, 26} Allele frequencies for the MCLINK analysis were estimated using all of the observed data.

We performed a nonparametric analysis and a general parametric model-based analysis using simple dominant and recessive model parameters that reproduced the reported population frequency of ASDs.¹ For the recessive analysis, the disease allele frequency was set at 0.05 with penetrances for each of the three genotypes of 0.0014, 0.0014, and 0.8. For the dominant analysis, the disease allele frequency was set at 0.0025, with penetrances of 0.0014, 0.8, and 0.8. Only affected cases were assumed to have known phenotypes. While nonparametric methods provide a standard analytic approach for complex psychiatric

disorders, the parametric approach is well suited to the analysis of a complex trait such as autism, particularly when using large extended pedigrees. Parametric models, which provide assumptions about the genotype-phenotype relationship, simplify the parameter space and allow for more powerful and efficient analyses without leading to false positive results.^{27–29} The multipoint heterogeneity lod score (HLOD) allows for unlinked pedigrees and variation in the recombination fraction. The HLOD provided by MCLINK is robust to model misspecification, and may reflect the true position of linkage peaks more accurately under conditions of appreciable heterogeneity (as is the case with ASD).^{26, 28, 30} HLOD scores have been shown to be more powerful than homogeneity LOD scores or model-free methods under these conditions.^{31, 32} The HLOD has been shown to produce scores consistent with other published methods.^{33, 34}

Eliminating linkage disequilibrium (LD) between markers in linkage studies has been strongly recommended, as false positive results can occur in the presence of LD.^{35–39} As these studies have shown, increases in type I error are most dramatic when LD is high and when there are missing genotypes for parents and other connecting relatives. Recommended thresholds of acceptable LD vary, but a pair-wise r^2 value of 0.05 between SNPs has now been supported with extensive simulation studies.³⁷ Linkage peaks achieving at least suggestive linkage evidence (HLOD = 1.86) were reanalyzed after screening for LD. SNPs within a 1-lod drop of each peak were screened using the PLINK software package,⁴⁰ which recursively removes SNPs within a sliding window. We set a window size of 50 SNPs, shifted the window by 5 SNPs at each step, and used a Variance Inflation Factor (VIF) of 1.5, which is equivalent to a multiple correlation coefficient R^2 of 0.33 regressed simultaneously over all SNPs in the selected window. This R^2 considers not only the correlations between SNPs but also between linear combinations of SNPs,⁴⁰ and corresponds in our data to a pair-wise r^2 of approximately 0.05. Also, as part of the validation procedure, we removed rare SNPs with a minor allele frequency less than 0.10. Our screening deleted 63 of the 209 SNPs across all 10 of the re-analyzed regions. Only four SNPs within these screened regions had a pair-wise $r^2 > 0.05$; the maximum pair-wise r^2 was 0.165. We also checked for SNPs out of Hardy-Weinberg Equilibrium using PLINK, and one additional SNP was deleted for this reason.

Results

Table 2 describes the diagnostic information for affected subjects (Table 2 about here). Of the 192 total affected subjects, 166 had data on both the ADI-R and ADOS-G. Of these 166 subjects with complete information, 122 met criteria for strictly defined Autistic Disorder on both assessments, and 44 met criteria for an ASD, having closely missed the cut-off scores for strictly defined autism on one or both measures. The other 27 cases were missing one (N=22) or both (N=5) assessments due to testing difficulties and/or unavailability of a reliable informant. There was a 6.7:1 male/female ratio among the subjects with strictly defined autism, which fell to 3.3:1 among the subjects with an ASD. For all subjects combined, the male/female ratio was 5.4:1. Subjects with ASDs were older than subjects with strictly defined autism at entry into the study (mean age: 15.9 vs. 12.0 years), though the difference was not significant ($t=1.74$, $p=0.09$). As expected, ADI-R scores were significantly higher for the autism group compared to the ASD group ($t=9.78$, $p<0.0001$ for

social; $t=8.30$, $p<0.0001$ for verbal; $t=3.86$, $p=0.0002$ for restricted interests/repetitive behaviors). The nonverbal total cannot be compared because while 28 autistic subjects were nonverbal, only one ASD subject was nonverbal. Means of ADI items indicating delay of first words and phrases also differed by diagnostic group in the expected direction (76.23% in autism vs. 52.38% in ASD; $p=0.003$). Quantitative scores on the ADOS are not compared because of differences between modules, but more individuals in the autism group were administered Modules 1 and 2 compared to the ASD group. IQ was obtained for 172 of the 192 affected subjects. Of these, 112 (65.12%) had $IQ>70$. Not surprisingly, significantly fewer subjects with strictly defined autism had $IQ>70$ (57.14%) compared to the percentage of ASD subjects with $IQ>70$ (86.96%; $p<0.0001$).

Figure 1 shows genome-wide linkage results, and Table 3 gives scores for regions with at least suggestive evidence for linkage ($HLOD = 1.86$) (Figure 1 and Table 3 about here).⁴¹ Each of these regions was screened for possible inflation due to LD as described above. Significant evidence of linkage ($HLOD = 3.3$)⁴¹ was found on chromosome 15, and initially on chromosome 21, though the chromosome 21 peak vanished when SNPs were eliminated because of LD. There were 10 closely spaced SNPs within this region that were deleted in the LD screen.

The chromosome 15 scores in Table 3 represent three possibly distinct regions, as shown in more detail in Figure 2 (Figure 2 about here). Note that the scores between these three maxima do not drop substantially. Using a 1-LOD drop to define regions, the approximate boundaries of these three regions are: 27,440,000 bp – 32,790,000 bp; 32,790,000 bp – 43,260,000 bp; and 50,770,000 bp – 56,800,000 bp. Nonparametric evidence was also found in these regions. Of note, linkage evidence was provided by multiple pedigrees, both large and small. Maximum scores for individual large pedigrees were not large enough to suggest complete sharing across all affected cases within any pedigree. In fact, the highest score for an individual pedigree within the three chromosome 15 peaks was a LOD of 2.27 under the 15q21.1-q22.2 peak in a 7-generation family with nine affected cases.

We investigated characteristics of the autism phenotype for cases in the families supporting the three chromosome 15 linkage peaks in our sample. For the pedigrees that achieved nominal point-wise significance (i.e., $LOD > 0.588$ for an individual pedigree, $p=0.05$) within the three peaks, the proportion of cases with ASD diagnoses was 30.0%, 31.0%, and 25.5% respectively, not significantly different from the overall proportion of autism cases in the entire sample (27.1%). Similarly, the proportion of male affected pedigree members was 75.0%, 79.3%, and 82.4%, not significantly different from the overall total male percentage of 84.4%. The proportion of affected subjects with $IQ>70$ was 66.7%, 65.2%, and 54.6%, not significantly different from the overall proportion of 65.1%. Subjects supporting these peaks also did not show significant differences in quantitative means of ADI domain scores or in language delay as measured by the ADI as compared to the full sample. See Supplemental Table 1 for a summary of these results.

As shown in Table 3 and Figure 2, regions of interest were also initially found on chromosomes 2, 6, 7, 13, 19, and 21. Suggestive linkage evidence was observed on chromosome 2p25.3-p24.1, from about 2,960,000 bp to about 10,660,000 bp that remained

after LD screening. We again tested for phenotypic differences among affected subjects in families supporting this peak. No differences were found for gender, ADI scores, language delay, or IQ>70 (see Supplemental Table 1). However, most of the affected subjects in these families had autistic disorder, with significantly fewer affected ASD subjects than observed in the full sample (1/23=4.35% vs. 52/192=27.08%, Fishers exact test p=0.02).

Two regions on chromosome 6 were initially identified, but one (6q27) was completely eliminated when SNPs were deleted for LD and rare allele frequency (seven close SNPs were dropped from this region). The 6q2-32-q24.1 just missed the suggestive linkage evidence threshold (HLOD=1.81).

The relatively broad chromosome 7q31.31-q32.3 peak was in a region implicated by several previous studies.⁴²⁻⁴⁵ and maintained suggestive linkage evidence even after LD screening. No significant phenotypic differences were found among affected subjects in families supporting this peak (see Supplemental Table 1).

The narrower chromosome 13q12.11-q12.3 peak also exceeded the suggestive linkage evidence threshold even after eliminating SNPs in LD. For this peak, we again did not find evidence for significant phenotypic subgroups (Supplemental Table 1). The peaks on chromosomes 19q13.43 and 21q22.12-q22.13 both disappeared when SNPs in LD were eliminated (10 SNPs from each location).

In a previous linkage screen of a single large Utah family, our group previously reported regions with at least suggestive linkage evidence on chromosomes 3q13.2-q13.31, 3q26.31-q27.3, 7p14.1-p11.22, 9p24.3, and 20q11.21-q13.12.⁷ The current sample includes this original pedigree, but these regions were not strongly supported by the new pedigrees. Within these 70 pedigrees, our strongest signals in any of these previous regions were on chromosomes 9p24.3 (HLOD=1.30 at 2,202,546 bp) and 20q11.21-q13.12 (HLOD=1.23 at 40,018,916 bp), but both of these regions were below the suggestive linkage threshold. The seven affected subjects in this original pedigree were notably heterogeneous across phenotypes, though all were male. An investigation of pedigrees supporting the peaks in the current sample on chromosomes 9 and 20 showed 13% females in families supporting chromosome 9, and 21% females in families supporting chromosome 20. These results suggest that sex is not a defining phenotype of these linkage regions in our families.

Discussion

We have performed a genome-wide linkage scan on a large collection of families with two or more cases of ASD. The 70 families studied here, which included 26 large extended pedigrees and 44 smaller 2- or 3-generation pedigrees, represents 20 years of recruiting and ascertaining ASD cases in Utah. Our study is noteworthy for the significant linkage evidence observed on chromosome 15 in possibly three distinct regions, and suggestive evidence on chromosomes 2, 7, and 13. Although we observed genetic heterogeneity within the families we studied, our large extended pedigree design offers a unique resource for finding linked chromosomal regions. Nonparametric scores showed consistency with the model-based analyses, though the magnitudes of the peaks were reduced. This reduction in

significance is not surprising, given multiple lines of evidence confirming the reduced statistical power of NPL methods.^{27–32} The consistency of the NPL results supports our choice of the HLOD statistic, which is robust to model misspecification.^{28, 30}

Our most significant finding was on chromosome 15q, which may comprise three separate peaks. Our 15q13.1-q14 linkage peak overlaps a region of chromosomal anomalies at 15q11-q13 implicated in autism across numerous studies.^{4, 45} However, genes of primary interest reported in these other studies are primarily centromeric (20,500,000 bp to 25,000,000 bp) to the start of our linkage peak at 27,440,000 bp. A recent multi-center study by Liu et al. of subjects with ASD who had IQ scores ≥ 70 reported a LOD score of 4.01 at chromosome 15q13.3–14,⁴⁶ which overlaps the reported linkage findings in our sample. In our sample, the proportion of IQ scores > 70 from the affected subjects contributing to this peak did not differ from the proportion > 70 in our full sample. While Utah collaborated in the Liu et al. multi-center study,⁴⁶ there was no overlap of subjects between that study and the current study.

The chromosome 15q21.2-q22.1 region in the current study replicates linkage findings reported by Szatmari et al. on a large sample of over 1,400 multiplex families with ASDs.⁴⁷ Szatmari et al. identified an extended region from 56,600,000 bp to 87,500,000 bp, with two maxima: LOD=3.19 at rs1372828 (69,302,220), and LOD=3.41 at rs1433452 (85,395,878). These findings were strongest when subjects with copy number variations (CNVs) were removed from the analysis. The International Molecular Genetic Study of Autism Consortium (IMGSAC) also found evidence for linkage on chromosome 15q21–22 when male-male sib pairs were omitted.⁴⁸ In contrast to Lamb et al., the sex ratio of our affected subjects supporting this peak did not differ from the sex ratio in our overall sample.

Consistent with the complex landscape of autism genetics, our linkage peaks on chromosome 15 showed within- and across-family heterogeneity. No individual pedigree LOD score in any large extended pedigree suggested sharing of a single haplotype across all affected cases, reinforcing the concept that genetic heterogeneity exists within individual pedigrees. Pedigrees providing the most support for the chromosome 15 peaks did not differ from the overall sample in terms of diagnostic severity, sex, IQ, language delay, or ADI domain score. Further characterization of individual cases who share common haplotypes within and across pedigrees may reveal additional insights into ASD predisposition genes in the region.

Other regions with suggestive positive results in our families occurred on chromosomes 2, 7, and 13. Our chromosome 2p25.3-p24.1 findings are more telomeric than the commonly reported chromosome 2p15-p16 autism region.^{4, 49, 50} In particular, the neurexin 1 gene (NRXN1), which has shown association with autism,^{47, 50} is at approximately 50,000,000 bp, well downstream of our peak. Our chromosome 7q31.31-q32.3 peak overlaps previous linkage findings reported in many studies^{42, 51–54} Several interesting gene associations with autism can be found in this region, including reelin (RELN), Ca²⁺-dependent activator protein for secretion 2 (CADPS2), and met proto-oncogene (MET).^{55–61} Finally, our chromosome 13q12.11-q12.3 peak does not appear to overlap with the 13q21 peak for autism and specific language impairment reported in several studies.^{62–64} The neurobeachin

gene (NBEAL2), reported to be disrupted in autism,⁶⁵ is located at 34,633,887 bp, somewhat downstream of our linkage region (21,190,000 bp – 30,950,000 bp).

Our expanded pedigree sample did not strongly support regions we previously reported in a single extended Utah autism pedigree. Our best signals in our current data within these previous regions were on chromosome 9p24.3 (HLOD=1.30) and 20q11.21-q13.12 (HLOD=1.23), but neither of these peaks reached suggestive evidence for linkage. Affected subjects in families supporting these peaks did not share phenotypic similarity with the original family. However, families need not be phenotypically similar, even if they share the same predisposing alleles. Pleiotropy, effects of other susceptibility loci, effect of genetic background, and environmental factors could all modify phenotype. The phenotypic heterogeneity and lack of strong support of original peaks is perhaps not surprising. Even in our original study, heterogeneity (genetic and phenotypic) was found within the single pedigree. Simulation analyses in our original study demonstrated those findings were not likely to be false positives, but our new results suggest these predisposing alleles are not common among our other pedigrees.

In conclusion, our results implicate a region on chromosome 15 that replicates results published by others and may harbor one or more susceptibility genes for autism. While this region overlaps with peaks from previous studies, specific genes have not yet been identified, and will require more detailed molecular work. To date, our investigation of phenotypes in families supporting the linkage evidence have not revealed consistent subtypes, though these analyses have only focused on very broad aspects of the ASD phenotype. More detailed analyses of additional aspects of phenotype, including presence of seizures, verbal ability, and other aspects of the clinical phenotype are planned. These results can be used to direct future gene localization efforts.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

This work was supported by R01 MH069359, the Utah Autism Foundation, the Carmen B. Pingree School for Children with Autism, and GCRC M01-RR00064 from the National Center for Research Resources. Partial support for all datasets within the Utah Population Database (UPDB) was provided by the University of Utah Huntsman Cancer Institute. We thank our staff whose countless hours of work have made this study possible. We also greatly appreciate the time and effort given by the family members who participated in this study.

References

1. Prevalence of autism spectrum disorders--autism and developmental disabilities monitoring network, 14 sites, United States, 2002. *MMWR Surveill Summ.* 2007; 56:12–28. [PubMed: 17287715]
2. Landa RJ, Holman KC, Garrett-Mayer E. Social and communication development in toddlers with early and later diagnosis of autism spectrum disorders. *Arch Gen Psychiatry.* 2007; 64:853–864. [PubMed: 17606819]
3. Landa RJ. Diagnosis of autism spectrum disorders in the first 3 years of life. *Nat Clin Pract Neurol.* 2008; 4:138–147. [PubMed: 18253102]

4. Abrahams BS, Geschwind DH. Advances in autism genetics: on the threshold of a new neurobiology. *Nat Rev Genet.* 2008; 9:341–355. [PubMed: 18414403]
5. Blangero J, Williams JT, Almasy L. Novel family-based approaches to genetic risk in thrombosis. *J Thromb Haemost.* 2003; 1:1391–1397. [PubMed: 12871272]
6. Terwilliger JD, Haghighi F, Hiekkalinna TS, Goring HH. A bias-ed assessment of the use of SNPs in human complex traits. *Curr Opin Genet Dev.* 2002; 12:726–734. [PubMed: 12433588]
7. Allen-Brady K, Miller J, Matsunami N, et al. A high-density SNP genome-wide linkage scan in a large autism extended pedigree. *Mol Psychiatry.* 2008
8. Ritvo ER, Mason-Brothers A, Jenson WP, et al. A report of one family with four autistic siblings and four families with three autistic siblings. *J Am Acad Child Adolesc Psychiatry.* 1987; 26:339–341. [PubMed: 3597289]
9. Jorde LB. Inbreeding in the Utah Mormons: an evaluation of estimates based on pedigrees, isonymy, and migration matrices. *Ann Hum Genet.* 1989; 53(Pt 4):339–355. [PubMed: 2624429]
10. Jorde LB. Consanguinity and prereproductive mortality in the Utah Mormon population. *Hum Hered.* 2001; 52:61–65. [PubMed: 11474206]
11. Berument, SK.; Rutter, M.; Lord, C.; Pickles, A.; Bailey, A. Autism Screening Questionnaire. Los Angeles, CA: Western Psychological Services; 1999.
12. 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. *J Autism Dev Disord.* 1994; 24:659–685. [PubMed: 7814313]
13. Baranek, GT.; Bodfish, JW.; Gordon, AM.; Houser, MB.; Poe, MD. Concurrent validity of the ADI-R and SCQ in high functioning autism. Collaborative Programs of Excellence in Autism/ Studies to Advance Autism Research and Treatment Annual Meeting; Bethesda, MD. 2004. 2004.
14. Lord C, Risi S, Lambrecht L, et al. The autism diagnostic observation schedule-generic: a standard measure of social and communication deficits associated with the spectrum of autism. *J Autism Dev Disord.* 2000; 30:205–223. [PubMed: 11055457]
15. Wechsler, D. Manual for the Wechsler Intelligence Scale for Children-Third Edition. San Antonio, TX: The Psychological Corporation; 1991.
16. Wechsler, D. Wechsler Adult Intelligence Scale - Third Edition. San Antonio, TX: The Psychological Corporation; 1997.
17. Elliott, C. Differential Ability Scales. San Antonio, TX: The Psychological Corporation; 1990.
18. Mullen, E. Mullen Scales of Early Learning, AGS Edition. Circle Pines, MN: American Guidance Service; 1995.
19. Dicerbo KEBA. A convergent validity study of the differential ability scales and the Wechsler Intelligence Scale for Children-Third Edition with Hispanic Children. *J Psychoed Assess.* 2000:344–352.
20. Dumont RCC, Price L, Whelley P. The relationship between the Differential Ability Scales (DAS) and the Wechsler Intelligence Scale for Children - Third Edition (WISC-III) for students with learning disabilities. *Psychology in the Schools.* 1996:203–209.
21. Sattler, J. Assessment of Children: Cognitive Applications. La Mesa, CA: Jerome M. Sattler, Publisher, Inc.; 2001.
22. O'Connell JR WD. PedCheck: A program for identifying genotype incompatibilities in linkage analysis. *Am J Hum Genet.* 1998:259–266. [PubMed: 9634505]
23. Kong A, Gudbjartsson DF, Sainz J, et al. A high-resolution recombination map of the human genome. *Nat Genet.* 2002; 31:241–247. [PubMed: 12053178]
24. Thomas A, Gutin A, Abkevich V, Bansal A. Multipoint linkage analysis by blocked Gibbs sampling. *Statistics and Computing.* 2000:259–269.
25. Coon H, Matsunami N, Stevens J, et al. Evidence for Linkage on Chromosome 3q25-27 in a Large Autism Extended Pedigree. *Hum Hered.* 2006; 60:220–226. [PubMed: 16391490]
26. Christensen GB, Camp NJ, Farnham JM, Cannon-Albright LA. Genome-wide linkage analysis for aggressive prostate cancer in Utah high-risk pedigrees. *Prostate.* 2007; 67:605–613. [PubMed: 17299800]

27. Terwilliger JD, Goring HH. Gene mapping in the 20th and 21st centuries: statistical methods, data analysis, and experimental design. *Hum Biol.* 2000; 72:63–132. [PubMed: 10721614]
28. Goring HH, Terwilliger JD. Linkage analysis in the presence of errors I: complex-valued recombination fractions and complex phenotypes. *Am J Hum Genet.* 2000; 66:1095–1106. [PubMed: 10712220]
29. Greenberg DA, Abreu P, Hodge SE. The power to detect linkage in complex disease by means of simple LOD-score analyses. *Am J Hum Genet.* 1998; 63:870–879. [PubMed: 9718328]
30. Abkevich V, Camp NJ, Gutin A, Farnham JM, Cannon-Albright L, Thomas A. A robust multipoint linkage statistic (tlod) for mapping complex trait loci. *Genet Epidemiol.* 2001; 21(Suppl 1):S492–S497. [PubMed: 11793725]
31. Goldin LR. Detection of linkage under heterogeneity: comparison of the two-locus vs. admixture models. *Genet Epidemiol.* 1992; 9:61–66. [PubMed: 1634107]
32. Abreu PC, Greenberg DA, Hodge SE. Direct power comparisons between simple LOD scores and NPL scores for linkage analysis in complex diseases. *Am J Hum Genet.* 1999; 65:847–857. [PubMed: 10441591]
33. Hunt SC, Abkevich V, Hensel CH, et al. Linkage of body mass index to chromosome 20 in Utah pedigrees. *Hum Genet.* 2001; 109:279–285. [PubMed: 11702208]
34. Christensen GB, Cannon-Albright LA, Thomas A, Camp NJ. Extracting disease risk profiles from expression data for linkage analysis: application to prostate cancer. *BMC Proc.* 2007; 1(Suppl 1):S82. [PubMed: 18466585]
35. Goode EL, Jarvik GP. Assessment and implications of linkage disequilibrium in genome-wide single-nucleotide polymorphism and microsatellite panels. *Genet Epidemiol.* 2005; 29(Suppl 1):S72–S76. [PubMed: 16342185]
36. Huang Q, Shete S, Amos CI. Ignoring linkage disequilibrium among tightly linked markers induces false-positive evidence of linkage for affected sib pair analysis. *Am J Hum Genet.* 2004; 75:1106–1112. [PubMed: 15492927]
37. Levinson DF, Holmans P. The effect of linkage disequilibrium on linkage analysis of incomplete pedigrees. *BMC Genet.* 2005; 6(Suppl 1):S6. [PubMed: 16451672]
38. Chen WV, Amos CI, Etzel CJ, Shete S, Gregersen PK. Comparison of genome-wide single-nucleotide polymorphism linkage analyses in Caucasian and Hispanic NARAC families. *BMC Proc.* 2007; 1(Suppl 1):S97. [PubMed: 18466601]
39. Allen-Brady K, Horne BD, Malhotra A, Teerlink C, Camp NJ, Thomas A. Analysis of high-density single-nucleotide polymorphism data: three novel methods that control for linkage disequilibrium between markers in a linkage analysis. *BMC Proc.* 2007; 1(Suppl 1):S160. [PubMed: 18466506]
40. Purcell S, Neale B, Todd-Brown K, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet.* 2007; 81:559–575. [PubMed: 17701901]
41. Lander E, Kruglyak L. Genetic dissection of complex traits: guidelines for interpreting and reporting linkage results. *Nat Genet.* 1995; 11:241–247. [PubMed: 7581446]
42. Trikalinos TA, Karvouni A, Zintzaras E, et al. A heterogeneity-based genome search meta-analysis for autism-spectrum disorders. *Mol Psychiatry.* 2006; 11:29–36. [PubMed: 16189507]
43. Schellenberg GD, Dawson G, Sung YJ, et al. Evidence for multiple loci from a genome scan of autism kindreds. *Mol Psychiatry.* 2006; 11:1049–1060. 979. [PubMed: 16880825]
44. Marshall CR, Noor A, Vincent JB, et al. Structural variation of chromosomes in autism spectrum disorder. *Am J Hum Genet.* 2008; 82:477–488. [PubMed: 18252227]
45. Christian SL, Brune CW, Sudi J, et al. Novel submicroscopic chromosomal abnormalities detected in autism spectrum disorder. *Biol Psychiatry.* 2008; 63:1111–1117. [PubMed: 18374305]
46. Liu XQ, Paterson AD, Szatmari P. Genome-wide linkage analyses of quantitative and categorical autism subphenotypes. *Biol Psychiatry.* 2008; 64:561–570. [PubMed: 18632090]
47. Szatmari P, Paterson AD, Zwaigenbaum L, et al. Mapping autism risk loci using genetic linkage and chromosomal rearrangements. *Nat Genet.* 2007; 39:319–328. [PubMed: 17322880]
48. Lamb JA, Barnby G, Bonora E, et al. Analysis of IMGSAC autism susceptibility loci: evidence for sex limited and parent of origin specific effects. *J Med Genet.* 2005; 42:132–137. [PubMed: 15689451]

49. Rajcan-Separovic E, Harvard C, Liu X, et al. Clinical and molecular cytogenetic characterisation of a newly recognised microdeletion syndrome involving 2p15-16.1. *J Med Genet.* 2007; 44:269–276. [PubMed: 16963482]
50. Kim HG, Kishikawa S, Higgins AW, et al. Disruption of neurexin 1 associated with autism spectrum disorder. *Am J Hum Genet.* 2008; 82:199–207. [PubMed: 18179900]
51. International Molecular Genetic Study of Autism Consortium. A full genome screen for autism with evidence for linkage to a region on chromosome 7q. *Hum Mol Genet.* 1998; 7:571–578. [PubMed: 9546821]
52. International Molecular Genetic Study of Autism Consortium. A genomewide screen for autism: strong evidence for linkage to chromosomes 2q, 7q, and 16p. *Am J Hum Genet.* 2001; 69:570–581. [PubMed: 11481586]
53. Schellenberg GD, Dawson G, Sung YJ, et al. Evidence for multiple loci from a genome scan of autism kindreds. *Mol Psychiatry.* 2006; 11:1049–1060. [PubMed: 16880825]
54. Further characterization of the autism susceptibility locus AUTS1 on chromosome 7q. *Hum Mol Genet.* 2001; 10:973–982. [PubMed: 11392322]
55. Persico AM, D'Agruma L, Maiorano N, et al. Reelin gene alleles and haplotypes as a factor predisposing to autistic disorder. *Mol Psychiatry.* 2001; 6:150–159. [PubMed: 11317216]
56. Bonora E, Beyer KS, Lamb JA, et al. Analysis of reelin as a candidate gene for autism. *Mol Psychiatry.* 2003; 8:885–892. [PubMed: 14515139]
57. Skaar DA, Shao Y, Haines JL, et al. Analysis of the RELN gene as a genetic risk factor for autism. *Mol Psychiatry.* 2005; 10:563–571. [PubMed: 15558079]
58. Serajee FJ, Zhong H, Mahbulul Huq AH. Association of Reelin gene polymorphisms with autism. *Genomics.* 2006; 87:75–83. [PubMed: 16311013]
59. Sadakata T, Washida M, Iwayama Y, et al. Autistic-like phenotypes in *Cadps2*-knockout mice and aberrant *CADPS2* splicing in autistic patients. *J Clin Invest.* 2007; 117:931–943. [PubMed: 17380209]
60. Campbell DB, D'Oronzio R, Garbett K, et al. Disruption of cerebral cortex MET signaling in autism spectrum disorder. *Ann Neurol.* 2007; 62:243–250. [PubMed: 17696172]
61. Campbell DB, Sutcliffe JS, Ebert PJ, et al. A genetic variant that disrupts MET transcription is associated with autism. *Proc Natl Acad Sci U S A.* 2006; 103:16834–16839. [PubMed: 17053076]
62. Barrett S, Beck JC, Bernier R, et al. An autosomal genomic screen for autism. Collaborative linkage study of autism. *Am J Med Genet.* 1999; 88:609–615. [PubMed: 10581478]
63. Bartlett CW, Flax JF, Logue MW, et al. A major susceptibility locus for specific language impairment is located on 13q21. *Am J Hum Genet.* 2002; 71:45–55. [PubMed: 12048648]
64. Bartlett CW, Flax JF, Logue MW, et al. Examination of potential overlap in autism and language loci on chromosomes, 2, 7, and 13 in two independent samples ascertained for specific language impairment. *Hum Hered.* 2004; 57:10–20. [PubMed: 15133308]
65. Castermans D, Wilquet V, Parthoens E, et al. The neurobeachin gene is disrupted by a translocation in a patient with idiopathic autism. *J Med Genet.* 2003; 40:352–356. [PubMed: 12746398]

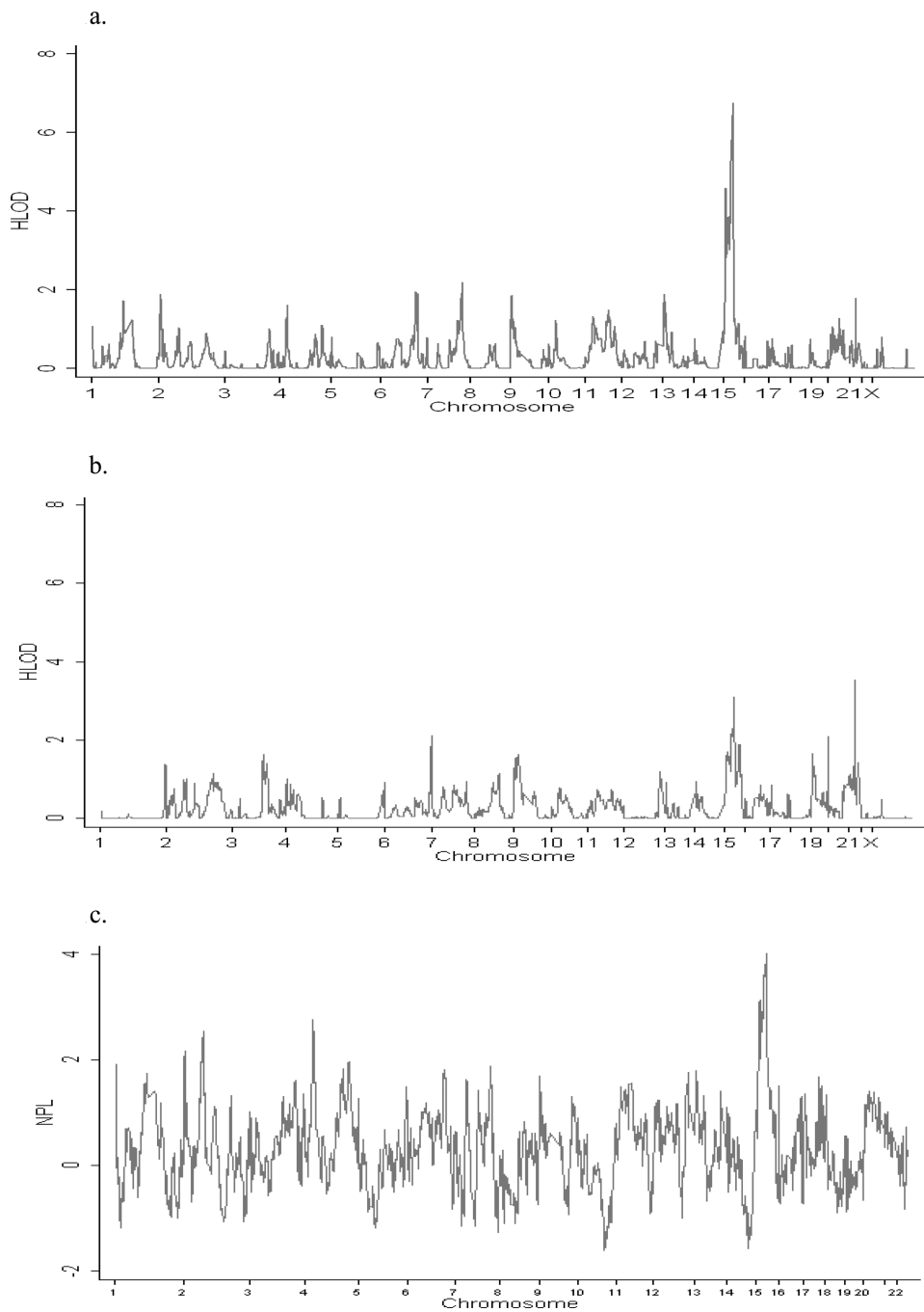
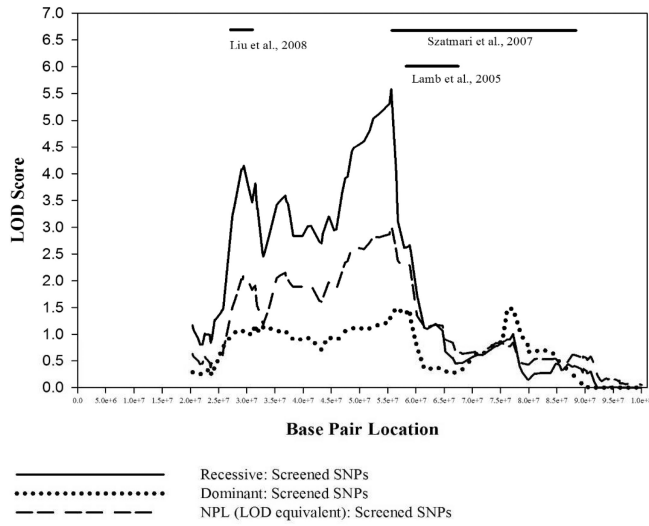


Figure 1. Genome-wide linkage results using all SNPs before screening for possible inflation due to linkage disequilibrium.

- a.** recessive model
- b.** dominant model
- c.** non-parametric analysis

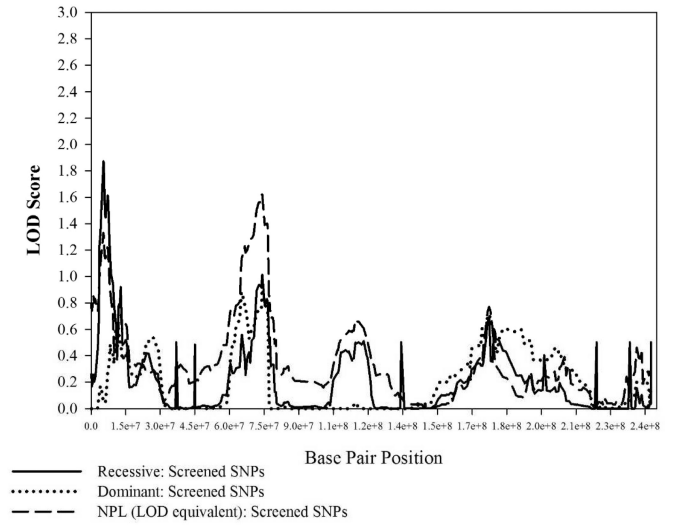
a.

Chromosome 15 Linkage Results



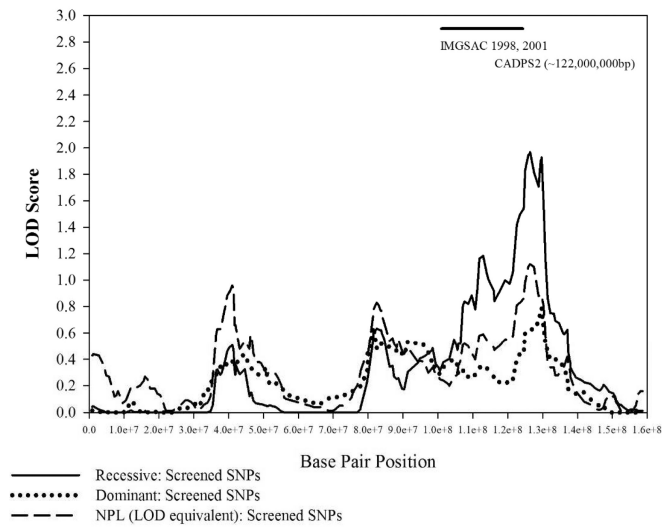
b.

Chromosome 2 Linkage Results



c.

Chromosome 7 Linkage Results



d.

Chromosome 13 Linkage Results

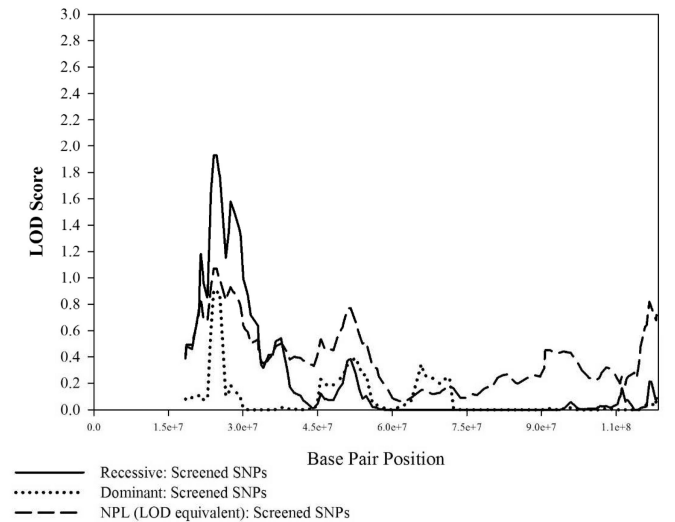


Figure 2.

Detailed results of SNPs screened for linkage disequilibrium for chromosomes.

- a. chromosome 15, significant results
- b. chromosome 2, suggestive results
- c. chromosome 7, suggestive results
- d. chromosome 13, suggestive results

Table 1

Description of the Utah ASD families

Type of pedigree	N of pedigrees	Avg # generations; SD (range)	Total subjects	Avg subjects per pedigree; SD (range)	Total ASD subjects	Avg ASD subjects per pedigree; SD (range)
Large (6–9 generations)	20	7.96; 0.69 (6 to 9)	331	17.21; 12.89 (5 to 50)	82	5.22; 2.54 (2 to 9)
Moderate (4–5 generations)	6	4; 0.00 (4)	85	14.17; 11.34 (6 to 32)	21	4.00; 3.39 (2 to 9)
Small (2–3 generations)	44	2.5; 0.43 (2 to 3)	237	5.39; 2.37 (2 to 11)	89	2.04; 0.60 (1 to 3)
FULL SAMPLE	70		653		192	

Table 2

Diagnostic information for 192 subjects with strictly defined autism or ASD

Diagnostic group	N	Male: Female	Mean Age (SD)	IQ>70 (%)	Language Delay (%)	Mean ADI Domain Scores (SD)					ADOS: N subjects given each module
						Social	Verbal	Non-verbal	Restricted/Repetitive	Module 1; 2; 3; 4	
Autistic Disorder	140	122:18	12.0 (9.6)	72/126 (57.14%)	93/122 (76.23%)	22.2 (6.0; N=122)	17.6 (3.9; N=94)	12.4 (2.3; N=28)	6.8 (2.5; N=122)	38; 31; 30; 34	
Autism Spectrum Disorder	52	40:12	15.9 (15.3)	40/46(86.96%)	22/42 (52.38%)	12.8 (5.7; N=44)	10.9 (6.3; N=43)	12.0 (N=1)	5.2 (2.7; N=44)	4; 10; 20; 16	
All affected subjects	192	162:30	13.1 (11.6)	112/172 (65.12%)	116/166 (69.88%)	19.7 (7.2; N=166)	15.5 (5.7; N=137)	12.4 (2.2; N=29)	6.4 (2.7; N=166)	42; 41; 50; 50	

Note. Of the 140 subjects with autism, 122 had both ADI and ADOS; the 13 who were missing ADI and 7 who were missing ADOS were diagnosed with autism using DSM-IV criteria. Of the 52 subjects with an ASD, 44 had both ADOS and ADI; the 8 who were missing ADI and 6 who were missing ADOS were diagnosed with an ASD using DSM-IV criteria.

Table 3 Chromosomal regions achieving at least suggestive linkage (HLOD 1.86) evidence using the original, unscreened SNP set.

Chromosome location	SNP at maximum HLOD (basepair)	Original HLOD (model)	HLOD after LD screening (model)	Original NPL score	NPL score after LD screening
2p25.3-p24.1	rs792065 (5,434,974)	2.03 (rec)	1.87 (rec)	2.47	2.18
6q22.32-q24.1	rs1570056 (137,101,370)	1.98 (rec)	1.81 (rec)	1.84	1.79
6q27	rs909475 (170,655,714) [screened: rs9295417 (170,734,025)]	2.11 (dom)	0.00 (dom)	0.82	0.61
7q31.31-q32.3	rs1990790 (129,820,866) [screened: rs1419437 (126,447,341)]	2.45 (rec)	1.97 (rec)	2.12	1.88
13q12.11-q12.3	rs6490970 (24,132,738)	1.88 (rec)	1.93 (rec)	1.83	1.81
15q13.1-q14	rs8033248 (29,459,872)	5.01 (rec)	4.09 (rec)	3.43	3.11
15q14-q21.1	rs723049 (36,837,208)	4.05 (rec)	3.59 (rec)	3.31	3.14
15q21.2-q22.1	rs11856 (55,629,733)	6.59 (rec)	5.31 (rec)	4.17	4.02
15q21.1-q22.2	rs383902 (56,821,466) [screened: rs725463 (57,930,371)]	3.10 (dom)	1.49 (dom)	3.37	3.21
19q13.43	rs4801273 (63,692,085) [screened: rs964795 (63,029,177)]	2.09 (dom)	0.01 (dom)	0.09	0.17
21q22.12-q22.13	rs2032088 (37,399,200) [screened rs1016694 (38,156,688)]	3.52 (dom)	0.01 (dom)	1.28	0.72
21q22.12-q22.13	rs2835667 (37,501,784) [screened: rs1012959 (36,983,492)]	2.06 (rec)	0.10 (rec)	1.27	0.63

Note. SNPs were screened and omitted if the minor allele frequency was < 0.10, and for linkage disequilibrium.